Ponesimod for Relapsing Multiple Sclerosis [ID1393]:

A Single Technology Appraisal

Produced by
Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors
Caroline Farmer[1]
Brian O'Toole[1]
David Packman[1]
Amanda Brand[1]
Sophie Robinson[1]
Fraizer Kiff[1]
Olga Ciccarelli[2]
Carl Counsell[3]
Louise Crathorne[1]
G.J. Melendez-Torres[1]
[1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter
[2] Department of Neuroinflammation, Institute of Neurology, Queen Square, University College London (UCL)
[3] Institute of Medical Sciences, University of Aberdeen

Correspondence to
Caroline Farmer
3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; c.farmer@exeter.ac.uk

Date completed
27/05/2021

Source of funding
This report was commissioned by the NIHR Systematic Reviews Programme as project number 13/33/13.

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Declared competing interests of the authors

None

Acknowledgments

The authors acknowledge the administrative support provided by Mrs Sue Whiffin and Ms Jenny Lowe (both PenTAG). In addition, the ERG model was independently appraised for errors by Madhusubramanian Muthukumar (PenTAG).

Rider on responsibility for document

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

Farmer, O'Toole, Packman, Brand, Robinson, Kiff, Ciccarelli, Counsell, Crathorne, Melendez-Torres. 0BPonesimod for Relapsing Multiple Sclerosis [ID1393]: A Single Technology Appraisal. Peninsula Technology Assessment Group (PenTAG), 2021.

Author Contributions:

Caroline Farmer	Project lead, critical appraisal of the company submission, writing and editorial input
Brian O'Toole	Lead for the ERG's appraisal of the economic evidence, drafted economic sections of the report, writing and editorial input
David Packman	Critical appraisal of the economic evidence, checked and re-analysed the economic model, carried out further scenario analyses, and drafted economic sections of the report
Amanda Brand	Critical appraisal of the clinical evidence, conducted additional clinical work for the submission, and drafted sections of the report
Sophie Robinson	Critical appraisal of the literature search strategies, conducted additional literature searching, and editorial review
Fraizer Kiff	Critical appraisal of the clinical evidence and drafted sections of the report
Olga Ciccarelli	Clinical advice and review of draft report
Carl Counsell	Clinical advice and review of draft report
Louise Crathorne	Critical appraisal of the company submission, writing and editorial input, and co-supervised the final report

Author Contributions:	
G.J. Melendez-Torres	Critical appraisal of the company submission, writing and editorial input, and co-supervised the final report. Guarantor of the report

Table of Contents

List of tables

List of Figures

[redacted] 217

[redacted] 218

List of key issues

Abbreviations

A and E	Accident and Emergency
AE	adverse event
ALT	alanine aminotransferase
ARR	annualised relapse rate
AST	aspartate aminotransferase
BSC	best supportive care
CDA	confirmed disability accumulation
CEAC	cost-effectiveness acceptability curve
CI	confidence interval
CNS	central nervous system
CRD	Centre for Reviews and Dissemination
CSR	clinical study report
CS	company submission
CUAL	combined unique active lesions
CV	Cardiovascular
DIC	deviance information criterion
DIS	dissemination in space
DIT	dissemination in time
DMT	disease modifying treatment
DP	decision problem
EDSS	Expanded Disability Status Scale
EQ-5D	EuroQol five dimension
ERG	Evidence Review Group
FSIQ-RMS	Fatigue Symptoms and Impacts Questionnaire: Relapsing Multiple Sclerosis
Gd+	gadolinium-enhancing
HA	highly active
HR	hazard ratio
HRQoL	health-related quality of life
HTA	health technology assessment
ICER	incremental cost-effectiveness ratio
IFNB	interferon beta
IQR	Interquartile range
ITC	indirect treatment comparison

ITT	intention-to-treat
LS	least squared
MD	mean difference
MRI	magnetic resonance imaging
MS	multiple sclerosis
MSFC	multiple sclerosis functional composite measure
MTR	magnetisation transfer ratio
NA	not applicable
NEDA	no evidence of disease activity
NHS	National Health Service
NHSE	National Health Service England
NICE	National Institute for Health and Care Excellence
NMA	network meta-analysis
NR	not reported
OR	odds ratio
OWSA	one-way sensitivity analysis
PAS	patient access scheme
PML	progressive multifocal leukoencephalopathy
PSA	probabilistic sensitivity analysis
QA	quality assessment
QALY	quality adjusted life year
RCT	randomised controlled trial
RES	rapidly evolving severe
RR	relative risk
RRMS	relapsing-remitting multiple sclerosis
RSS	risk-sharing scheme
SD	standard deviation
SF-36	Short Form (36) health survey
SLR	systematic literature review
SoT	suboptimally treated
SPMS	secondary progressive multiple sclerosis
TA	technology appraisal
TEAE	treatment-emergent adverse events
TP	treatment period
UTI	urinary tract infection

Vs	Versus
WTP	willingness to pay

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision-making. It also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

- Section 1.1 provides an overview of the key issues and the differences in the assumptions of the company and the ERG in economic analysis.
- Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER.
- Sections 1.3 to 1.5 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main ERG report.
- Sections 1.6 and 1.7 provide an overview of the ERG's preferred base case and sensitivity analyses undertaken by the ERG.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1. Overview of the ERG's key issues

A brief overview of the key issues identified by the ERG in their appraisal of the company submission (CS) is provided in Table 1. Further detail of the issues is provided in Sections 1.3, 1.4, and 1.5.

Broadly speaking, the key issues related to uncertainty surrounding the clinical effectiveness estimates for ponesimod and its comparators. This uncertainty has implications for the cost-effectiveness of ponesimod in both the active RRMS population and for people with highly active disease (HA RRMS), and for understanding the most appropriate positioning of ponesimod in the treatment pathway. Furthermore, the company's economic evaluation of ponesimod did not fully represent the 'clinical reality' treatment pathway in RRMS, which is often characterised by treatment sequencing, and there is uncertainty about subsequent treatment assumptions after progress to secondary progressive multiple sclerosis (SPMS).

Table 1: Summary of key issues

ID	Summary of issues	Report sections
Key Issue 1	Uncertainty in the evidence base for the rapidly evolving severe (RES) RRMS population	2.3
Key Issue 2	Uncertainty in the clinical efficacy of ponesimod and its comparators	3.3, 3.4, and 3.5
Key Issue 3	Insufficient comparative evidence for the safety of ponesimod	3.4.1, 3.5.3, and 3.5.4
Key Issue 4	Uncertainty surrounding use of 3 month CDA as the primary measure of disease progression in the economic model	1.5 and 6.1.1.1
Key Issue 5	Uncertainty surrounding the assumption that 100% of people who convert to SPMS will receive BSC	1.5 and 6.1.1.2

Abbreviations: BSC, best supportive care; CDA, confirmed disability accumulation; RES, rapidly evolving severe; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

In the economic analysis, the ERG's preferred assumptions vary from the company's in the following ways:

- In the company's base case analysis, the 3-month confirmed disability accumulation (CDA) was chosen as the primary measure of disease progression, which did not align with the preferences of the NICE committees in previous technology appraisals (TAs) (see section 1.5 and 6.1.1.1). The ERG considered that 6-month CDA should be used to estimate disease progression in the model for both the intention-to-treat (ITT) and the HA RRMS highly active populations

- The company assumed that 100% of people who convert to SPMS receive best supportive care (BSC; i.e. largely symptom management). However, the ERG noted that siponimod (TA656)[1] was recommended by NICE in 2020 for the treatment of people with SPMS, and therefore, the analysis should account for some uptake of siponimod in this population. See section 1.5 and 6.1.1.2.

1.2. Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length of life (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Delaying disease progression. The key driver of clinical effectiveness and associated QALY gain for ponesimod (versus most comparators in both the ITT and HA RRMS populations) was due to improved efficacy for CDA. In the model, a higher proportion of people receiving ponesimod remained in lower RRMS Expanded Disability Status Scale (EDSS) health states, relative to most comparator disease modifying treatments (DMTs). A higher proportion of people on 'less efficacious' treatments transitioned to higher EDSS states, where they experience lower health-related quality of life (HRQoL).

- Avoiding higher mortality multipliers, in higher EDSS states, associated with the risk of mortality from multiple sclerosis (MS). As such, higher efficacy DMTs (including ponesimod), resulted in incremental life years gained vs. moderately effective treatments.

In order to do this the technology is modelled to affect costs by:

- Keeping more people in lower EDSS states (0-6) where disease management costs are significantly less than higher states (7-9). Due to the modelled treatment efficacy, people receiving ponesimod had lower disease management costs versus most comparators.

- Ponesimod was also considered to have lower drug acquisition costs, monitoring and administration costs compared to some comparators. Please note, the company's base case analysis did not include confidential patient access scheme (PAS) discounts for the comparators.

The modelling assumptions that have the greatest effect on the ICER are:

- Using six-month CDA for EDSS progression in the model, rather than 3-month CDA (ITT population)

- Using a positioning-based approach to estimate treatment effect (ITT and HA RRMS populations)

- Using an alternative set of annual conversion probabilities, from RRMS to SPMS (ITT population)

- No waning in treatment effect (HA RRMS population)

1.3. Summary of the key issues regarding the decision problem

The ERG reviewed the approach of the company to addressing the NICE decision problem for this appraisal, and identified a key issue relating to the specific inclusion or relevance of different RRMS phenotypes.

The original submission provided by the company did not include evidence for two potential comparators to ponesimod that were under appraisal at the time of submission, however the company presented evidence for these comparators at clarification. While the standard of the evidence presented for these comparators was limited by the timeframe available to the company between submission and their response to clarification, the ERG was satisfied that the evidence presented was sufficiently comparable to other comparators.

Key Issue 1: Uncertainty over the evidence base for the rapidly evolving severe (RES) RRMS population

Report sections	**2.3**
Description of issue and why the ERG has identified it as important	The NICE scope for this appraisal specifies people with RES RRMS as a separate population group; however, in its response to the DP, the company stated that people with RES RRMS were included within its definition of highly active (HA) RRMS, and that no separate subgroup analysis for this population would be presented. The broader HA+RES data was used in the company's base case NMAs, and in the company's economic evaluation. The ERG was unclear whether evidence from a combined HA population could be used to inform a recommendation for the RES population. The ERG understood that while there may be some similarities in presentation between people with HA and RES RRMS in terms of the speed of disease progression, there are differences in the populations: specifically, HA RRMS is disease that progresses despite treatment ('breakthrough disease'), and RES is a separate, rare phenotype of the disease. It is unclear whether relative treatment effects (though often stable across different populations), are comparable in the HA and RES populations. The ERG noted that relative treatment effects in the company's model varied between the ITT and HA population. In addition, the ERG considered that the absolute outcomes and costs for RES RRMS may differ from HA RRMS, which may affect the cost effectiveness of ponesimod versus other available treatments. There has been some uncertainty in previous appraisals about whether recommendations can be generalised across population groups. At clarification the company presented subgroup data for people with RES RRMS from their pivotal trial, though the sample was small, and the comparator treatment (teriflunomide) is not recommended in the NHS for people with RES RRMS. The company's subgroup NMAs

Report sections	2.3
	considered RES within the definition of HA only. The ERG noted that natalizumab is currently recommended in the NHS for RES RRMS (and not HA RRMS), and that while this treatment was included in the company's NMAs, the results were not reported, and natalizumab was not considered as a comparator in the company's economic model.
What alternative approach has the ERG suggested?	The ERG did not believe that the evidence presented by the company is sufficient to evaluate the effectiveness of ponesimod in the RES RRMS population; however further clinical input and evidence may help to resolve this issue.
What is the expected effect on the cost-effectiveness estimates?	The results of the company's economic evaluation vary between the ITT and HA population, though it is unclear whether differences would be seen between the HA and RES populations. Without seeing the results for natalizumab, it is unclear whether ponesimod would be cost-effective against this comparator.
What additional evidence or analyses might help to resolve this key issue?	Evidence to demonstrate that treatment effects for ponesimod are stable across baseline risk, and/or across the different populations of RRMS would provide confidence in generalising evidence to the RES population. Clinical evidence should also be presented for the comparison between ponesimod and natalizumab, as well as all other treatments available for people with RES RRMS. In addition, altered modelling assumptions for the RES population may be needed, in order to evaluate whether ponesimod is cost effective in this population.

Abbreviations: DP, decision problem; ERG, Evidence Review Group; HA, highly active; ITT, intention-to-treat; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis; RES, rapidly evolving severe; RRMS, relapsing-remitting multiple sclerosis

1.4. Summary of the key issues in the clinical effectiveness evidence

The ERG reviewed the clinical effectiveness and safety evidence presented in the CS, and identified the following two key issues for consideration by the committee:

Key Issue 2. Uncertainty in the clinical efficacy of ponesimod and its comparators

Report sections	3.3, 3.4, and 3.5
Description of issue and why the ERG has identified it as important	The clinical effectiveness evidence for ponesimod and its comparators was highly heterogeneous, and there was a paucity of evidence for most of the comparisons in the company's NMAs. Clinical experts to the ERG also noted that the outcomes reported in the included trials were frequently short-term, and that these may be unable to capture meaningful change in disease course. These follow-up durations also varied widely across trials. Treatment effects for all outcomes varied widely between groups treated with placebo, highlighting the extent of the heterogeneity and its impact on treatment effects. Relative treatment effects derived from the NMAs have wide confidence intervals, and there is a

Report sections	3.3, 3.4, and 3.5
	high degree of uncertainty about the true magnitude of the effects reported. The evidence was particularly limited for analyses in the highly active population.
What alternative approach has the ERG suggested?	The ERG appraised the company's NMAs, and validated the methodology and results against previous appraisals, and found that these were consistent. The ERG therefore considered that the methods used by the company were appropriate in the context of the available evidence, and that uncertainty surrounding the clinical effectiveness estimates was principally due to the limitations of the evidence base.
What is the expected effect on the cost-effectiveness estimates?	The ICER was highly sensitive to even small variations in treatment efficacy.
What additional evidence or analyses might help to resolve this key issue?	The ERG was satisfied that the evidence presented by the company is representative of the known treatment effects for ponesimod and its comparators. Until further evidence is available (more direct head-to-head trials of ponesimod, trials with longer follow-up, and evidence identifying whether treatment effects vary according to the sources of heterogeneity in the evidence base), uncertainty surrounding the treatment effects of DMTs is a key issue in appraisals of treatments for RRMS. The ERG has conducted some scenario analyses to demonstrate the sensitivity of the ICER to variation in the treatment effect of ponesimod (see Section 6.1).

Abbreviations: DMT, disease-modifying treatment; ERG, evidence review group; ICER, incremental cost effectiveness ratio; NMA, network meta-analysis; RRMS, relapsing remitting multiple sclerosis

Key Issue 3. Insufficient comparative evidence for the safety of ponesimod

Report sections	3.2.4.3, 3.5.3, 3.5.4
Description of issue and why the ERG has identified it as important	Treatment decisions for RRMS frequently involve a trade-off between the efficacy and safety of DMTs, in addition to consideration of individuals' preferences (towards routes of administering treatment and typical adverse events). Understanding the relative safety of ponesimod is therefore necessary for understanding its likely positioning in the treatment pathway, and its most relevant comparators. The company's main trial, OPTIMUM, compared the safety of ponesimod with teriflunomide, a moderate-safety, first-line DMT. However, no NMA evaluating the relative safety of ponesimod was reported. The company reported annualised rates of adverse events, obtained from included trials, for ponesimod and each comparator DMT. This approach relies upon a naïve comparison of rates that does not take account of the heterogeneity between the included trials (including variations in sample eligibility criteria, healthcare setting, and the measurement and follow-up of safety outcomes). Trial data also lacks external validity when measuring AEs, and trials of DMTs are frequently too small and/or short to reliably measure the incidence of rare, serious AEs.

Report sections	3.2.4.3, 3.5.3, 3.5.4
What alternative approach has the ERG suggested?	The ERG compared the rates of AEs for ponesimod and its comparators, and on the basis of this evidence drew tentative conclusions that ponesimod may be acceptably safe, including in respect to elevated liver enzymes and infections when compared to comparators in the first and second line. With regards to rare serious adverse events, it was uncertain whether ponesimod provides an improved safety profile due to the lack of data in a large enough group of participants. From these data, the ERG drew a comparison between the rates reported for ponesimod and fingolimod. This comparison was chosen as the company posited that ponesimod may be considered a safer alternative to fingolimod, and clinical experts advised that a comparison of the safety of these treatments would aid understanding of the appropriate positioning of ponesimod in the treatment pathway. The evidence did not satisfactorily demonstrate that ponesimod was associated with a lower risk of AEs, including AEs related to liver toxicity. The ERG conducted a further naïve comparison of AE rates reported by the company from the OPTIMUM trial with those reported for fingolimod in its appraisal by NICE in 2012. This comparison was intended to identify rates of cardiac events, macular oedema and treatment discontinuations due to adverse events, which were not reported in the CS for comparators to ponesimod. Based on these data, ponesimod appeared to be an acceptable alternative to fingolimod for macular oedema; however, treatment discontinuations were higher among participants treated with ponesimod. No cardiac data was available from the NICE appraisal of fingolimod.
What is the expected effect on the cost-effectiveness estimates?	The data appeared to suggest that ponesimod is a moderate-safety treatment; however, the quality of safety evidence is poor, and further evidence would inform its most appropriate positioning in the treatment pathway, and therefore the identification of its most relevant comparators in cost-effectiveness evaluations. The risk of rare serious adverse events manifesting over the long-term informs assumptions related to monitoring, as well as healthcare resource use. Increased treatment discontinuations may also affect health resource use. However, the ERG identified that the impact of monitoring has little impact on the ICER.
What additional evidence or analyses might help to resolve this key issue?	A further NMA evaluating the relative risk of discontinuation due to AEs as compared to other available DMTs would contribute to an understanding of the overall safety of ponesimod. While this NMA would also be limited by heterogeneity in the trials, discontinuation gives an overall picture of tolerability, and may be more consistently measured across trials. Moreover, published NMAs of treatments for RRMS often present a graph plotting the relative safety vs. efficacy of all available treatments, which would be useful to aid decision-makers in identifying the most appropriate positioning for ponesimod. Higher quality evidence for the safety of ponesimod, including long-term real-world evidence in

Report sections	3.2.4.3, 3.5.3, 3.5.4
	larger groups of people, would give a more informed insight into the safety of ponesimod, particularly in terms of rare serious adverse events, such as PML. Clinical experts to the ERG also suggested that clearer positioning within the same class of treatment (e.g. if/when to use ponesimod, fingolimod, and siponimod) would be useful to understanding the appropriate positioning of ponesimod.

Abbreviations: DMT, disease-modifying treatment; ERG, Evidence Review Group; HA, highly active; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis; PML, progressive multifocal leukoencephalopathy; RRMS, relapsing-remitting multiple sclerosis

1.5. Summary of the key issues in the cost effectiveness evidence

The ERG reviewed the company health economic evidence and economic evaluation presented in the CS, and identified the following key issues for consideration by the committee:

Key Issue 4. Six-month confirmed disability accumulation (CDA) is considered a more appropriate measure of disease progression

Report sections	4.2.6 and 6.1.1.1
Description of issue and why the ERG has identified it as important	The key driver of clinical effectiveness in the model was treatment effects for 3-month CDA. However the ERG considered 6-month CDA to be a more robust measure of progression. This was following clinical advice to the ERG that 3-month CDA can overestimate progression due to natural fluctuations in the disease. Previous NICE committees have also expressed a preference for 6-month CDA in appraisals of treatments for RRMS (e.g. the NICE appraisal of alemtuzumab, TA312[2]). The company provided additional justification for using 3-month CDA data in the base case (see Section 4.2.6 or their response). However, despite the comparatively lower availability of evidence for 6-month CDA, the ERG considered that this should have been used in the company's base case as it is a more robust measure of progression. The company included an option in their model to use 6-month CDA as the preferred estimate of treatment efficacy.
What alternative approach has the ERG suggested?	The ERG used 6-month CDA estimates in their base case. Results are discussed and reported in Section 6.1.1.1.
What is the expected effect on the cost-effectiveness estimates?	Results were sensitive to using 6-month CDA estimates in the ITT population.
What additional evidence or analyses might help to resolve this key issue?	In the absence of direct head-to-head data, the ERG considered that the use of 6-month CDA data from the NMAs was reasonable. However, 6-month CDA estimates derived from head-to-head studies would increase the validity of these results.

Abbreviations: CDA, confirmed disability accumulation; ERG, Evidence Review Group; ITT, intention-to-treat; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis; RRMS, relapsing-remitting multiple sclerosis

Key Issue 5. The assumption that 100% of people who progress to SPMS receive BSC may not be appropriate

Report sections	4.2.6.1 and 6.1.1.2
Description of issue and why the ERG has identified it as important	In the base case analysis, the company assumed that 100% of people who discontinue treatment go on to receive BSC. Although this is reflective of previous NICE TAs, the ERG were aware that siponimod had recently been accepted by NICE for use in people with SPMS[1], and will soon be available. Clinical advice to the ERG was that some people who have been diagnosed with SPMS will also receive dimethyl fumarate, though this is not considered to be highly efficacious. As siponimod has only recently been approved, there was uncertainty about the rate of uptake in the SPMS population. Based on clinical input to the ERG, the proportion of people who are likely to receive siponimod after converting to SPMS could be approximately 25%; this accounts for a proportion of people who choose not to receive treatment or are ineligible.
What alternative approach has the ERG suggested?	The ERG conducted a scenario analysis that assumed 25% of people who converted to SPMS received siponimod, whilst 75% received BSC. This scenario accounted for the additional costs of managing siponimod in people converting to SPMS, but did not account for the clinical efficacy of siponimod, due to the uncertainty surrounding the expected clinical efficacy
What is the expected effect on the cost-effectiveness estimates?	This scenario analysis did not have a significant impact on the base case results (in either the ITT or HA RRMS populations), however the ERG considered that including this assumption within the base case analysis was likely to better reflect clinical practice.
What additional evidence or analyses might help to resolve this key issue?	Treatment uptake data surrounding siponimod use in the UK (in both the active RRMS and highly active RRMS populations) would help to resolve this issue. The company and ERG model were unable to fully account for the impact of subsequent treatments, and so the potential impact of treatment with siponimod and other DMTs on the cost effectiveness of ponesimod was uncertain.

Abbreviations: BSC, best supportive care; DMT, disease modifying treatment; ERG, Evidence Review Group; HA, highly active; ITT, intention-to-treat; NICE, National Institute for Health and Care Excellence; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis; TA, technology appraisal

1.6. Summary of ERG's preferred assumptions and resulting ICER

The ERG's preferred assumptions for the ITT and HA RRMS are listed in Table 2 and Table 4 below. Results are presented in Table 3 and Table 5; please note that these do not include confidential PAS discounts for comparator treatments. For further details of the exploratory and sensitivity analyses conducted by the ERG, see Section 6.1.

Table 2. ERG preferred assumptions (ITT population)

Preferred assumption	Report Section
Company base-case	*5.1.1*
6 month CDA used to model disease progression	*4.2.6.1 and 6.1.1.1*
25% of people receive siponimod after converting to SPMS, 75% receive BSC	*4.2.6 and 6.1.1.2*

Abbreviations: BSC, best supportive care; CDA, confirmed disability accumulation; ERG, Evidence Review Group; ITT, intention-to-treat; SPMS, secondary progressive multiple sclerosis

Table 3. ERG's preferred case results (ITT population)

Outcomes Ponesimod vs Comparator	ERG base case			Company base case
	Incremental QALYs	Incremental costs (£)	ICER (£/QALY)	ICER (£/QALY)
Teriflunomide 14mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fumarate 240mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30mcg IM	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1b 250mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab 600mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Peginterferon beta-1a 125mcg SC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; QALY, quality adjusted life year

Table 4. ERG preferred assumptions (HA RRMS population)

Preferred assumption	Report Section
Company base-case	5.1.1
6 month CDA used to model disease progression	4.2.6.1 and 6.1.1.1

Preferred assumption	Report Section
25% of people receive siponimod after converting to SPMS, 75% receive BSC	4.2.6 and 6.1.1.2

Abbreviations: CDA, confirmed disability accumulation, ERG, Evidence Review Group; HA, highly active; SPMS, secondary progressive multiple sclerosis

Table 5. ERG's preferred base case results (HA RRMS population)

Outcomes	ERG base case			Company base case
Ponesimod vs Comparator	Incremental QALYs	Incremental costs (£)	ICER (£/QALY)	ICER (£/QALY)
Ocrelizumab 600mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab 12mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine 3.5mg/kg PO	[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod 0.5mg PO	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ERG, Evidence Review Group; HA, highly active; ICER, incremental cost-effectiveness ratio; RRMS, relapsing-remitting multiple sclerosis

1.7. Summary of exploratory and sensitivity analyses undertaken by the ERG

A summary of the ERG's scenario analyses is provided in Table 6 below. For results, please see Section 6.1

Table 6: ERG scenario analyses (ITT population)

Scenario	Report Section
Company base case	*5.1.1*
Scenario 1: 6 month CDA used to model disease progression	*6.1.1.1*
Scenario 2: 25% of SPMS people assumed to receive siponimod and 75% receive BSC	*6.1.1.2*

Scenario	Report Section
Scenario 3: Population characteristics based on UK RSS data	*6.1.1.3*
Scenario 4: Alternative subsequent treatment assumptions	*6.1.1.4*
Scenario 5: No difference in discontinuation rates (assumed 5% for all treatments)	*6.1.1.5*
Scenario 6: No waning in treatment effect (applies to all treatments)	*6.1.1.6*
Scenario 7: Alternative modelled clinical effectiveness parameters	*6.1.1.7*
Scenario 8: Monitoring costs for ponesimod in year 1 assumed to be equal to fingolimod	*6.1.1.8*
Scenario 9: Alternative EDSS health state costs	*6.1.1.9*
Scenario 10: Alternative cost associated with relapse	*6.1.1.10*
Scenario 11: Alternative EDSS health state utilities	*6.1.1.11*
Scenario 12: Alternative annual conversion probabilities (from RRMS to SPMS)	*6.1.1.12*

Abbreviations: BSC, best supportive care; CDA, confirmed disability accumulation; EDSS, Expanded Disability Status Scale; ERG, Evidence Review Group; ICER, incremental cost effectiveness ratio; ITT, intention-to-treat; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

Table 7: ERG scenario analyses (highly active population)

Scenario	Report Section
Company base case	*5.1.1*
Scenario 1: 6 month CDA used to model disease progression	*6.1.1.1*
Scenario 2: 25% of SPMS people assumed to receive siponimod and 75% receive BSC	*6.1.1.2*
Scenario 3: Population characteristics based on UK RSS data	*6.1.1.3*
Scenario 4: Alternative subsequent treatment assumptions	*6.1.1.4*
Scenario 5: No difference in discontinuation rates (assumed 5% for all treatments)	*6.1.1.5*
Scenario 6: No waning in treatment effect (applies to all treatments)	*6.1.1.6*
Scenario 7: Alternative modelled clinical effectiveness parameters	*6.1.1.7*
Scenario 8: Monitoring costs for ponesimod in year 1 assumed to be equal to fingolimod	*6.1.1.8*
Scenario 9: Alternative EDSS health state costs	*6.1.1.9*
Scenario 10: Alternative cost associated with relapse	*6.1.1.10*
Scenario 11: Alternative EDSS health state utilities	*6.1.1.11*
Scenario 12: Alternative annual conversion probabilities (from RRMS to SPMS)	*6.1.1.12*

Abbreviations: BSC, best supportive care; CDA, confirmed disability accumulation; EDSS, Expanded Disability Status Scale; ERG, Evidence Review Group; ICER, incremental cost effectiveness ratio; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

Please note that all page references to the company submission (CS) are using version 2, submitted by the company on 29th March 2021.

2. INTRODUCTION AND BACKGROUND

2.1. Introduction

Multiple sclerosis (MS) is a chronic disease caused by dysfunction of the immune system, which leads to damage to the myelin within the central nervous system. Myelin is an insulating layer surrounding the axons of nerve cells and supports rapid and efficient transmission of electrical impulses along nerve cells. Degradation of this layer leads to neurodegeneration as the electrical impulses transmitted throughout the brain and spinal cord are impeded. Areas where the myelin is damaged are known as lesions, the accumulation of which causes neurological impairment and multifaceted disability.

The symptoms of MS vary between people but can include the following: fatigue; vision issues; numbness or tingling; muscle spasms; stiffness and weakness; mobility issues; pain; issues with cognitive; depression or anxiety; sexual issues; bladder or bowel control issues as well as speech and swallowing difficulties. Public Health England estimates indicate that there are around 105,800 people[3] suffering from all MS forms in the UK. In the general population, MS is twice as common in women as men, although in those aged between 50-59 years the prevalence is three times higher in women[3].

The most common subtype of MS is relapsing remitting MS (RRMS). RRMS is generally diagnosed in when people are in their twenties or thirties, and it accounts for around 85% of those diagnosed with MS[4]. RRMS is characterised by periods of remission interspersed with relapses. A relapse is identified through the presence of new symptoms, or an exacerbation of existing symptoms, lasting over 48 hours. Following a relapse, there will be a period of recovery which may or may not be complete. The recovery from attacks often becomes less complete over time, and residual disability accumulates. The frequency and nature of relapses varies, with natural fluctuation over the disease course, though relapses typically reduce as people age.

People with RRMS will ultimately be considered to have progressed to secondary progressive (SPMS) disease, where they are considered to suffer from fewer attacks but nevertheless show a gradual increase in disability. This is caused by neurodegeneration from existing lesions. SPMS is difficult to diagnose, with the diagnosis often done retrospectively based on a clinical review of symptoms. It is estimated that people with RRMS will progress to SPMS after an average of approximately eight to ten years; this rate has not been shown to change meaningfully since the introduction of disease-modifying treatments (DMTs).

RRMS diagnosis is complex due to the vast range of symptoms and widely varying clinical presentation. Clinicians use the revised McDonald criteria (Thompson et al. 2018[5]), which takes into account the number of relapses and lesions people have, as well as the location of lesions within the central nervous system (CNS), in order to make a judgment. Lesions are detected with magnetic resonance imaging (MRI), of which there are two types used in MS diagnosis; gadolinium (Gd)-enhanced T1 and T2. RRMS can be further categorised by the level of disease activity, as per the categories below. These categories aim to identify those people whose disease will progress more rapidly, in order to inform the choice of treatment.

- **Inactive RRMS** is defined as no relapses and no evidence of new lesions on MRI.
- **Active RRMS** is defined either by up to two relapses per year and/or new MRI activity.
- **Highly active (HA) RRMS** is less easily defined, as there are a range of definitions used internationally. The National Health Service (NHS) defines HA RRMS as: 'People with an unchanged or increased relapse rate or ongoing severe relapses compared with the previous year despite treatment with beta interferon'[6]. Conversely, the definition used in the US is more focused on the radiological burden of MS and rapid disability progression following onset.
- **Rapidly evolving severe (RES) RRMS** can be defined as either two or more disabling relapses in a year and one or more gadolinium-enhanced (Gd+) lesions, or a significant increase in T2 lesion load when compared with an earlier MRI.

People in the UK are currently treated with DMTs according to the NHS treatment algorithm[6]. An MS consultant and a specialist MS nurse will work in conjunction with multi-disciplinary teams from specialist MS centres across the country to determine the optimal treatment course for an individual. Where people have more complex disease, or where clinicians are considering treatment with a DMT with a higher risk of adverse events, such as cladribine or monoclonal antibodies, a meeting is typically held with a specialist team of MS clinicians.

2.2. Background

2.2.1. Current treatment for RRMS

There are a variety of DMTs currently used to treat RRMS in the UK. The company provided an overview of the NHS England (NHSE) treatment algorithm for DMTs[6], with first-line treatments positioned according to disease features, such as relapse frequency. The ERG considered that the pathway presented by the company accurately represented the NHSE pathway; however

understood that in practice, distinctions between first- and second-line treatments may be an over-simplification as people may receive several lines of therapy within the categories proposed in the NHSE pathway. The choice of a treatment is determined based on a balance of efficacy and safety, while also taking into consideration personal preference with regards to the mode of administration and risk of serious side-effects. Clinicians may choose either an escalation or an induction approach: the former involves administering a first-line, moderate-efficacy, high-safety treatment, with subsequent switching to a second-line treatment more-effective, lower-safety drug after the disease progresses (NHS algorithm [6]; Thompson et al. (2018)[7]); the induction approach involves first administering a highly effective, typically second line drug, to attain rapid remission of highly active MS (two or more severe relapses per year) and prevent rapid disability accumulation (NHS algorithm[6]; Thompson et al. (2018)[7]). Currently, trials to determine which of these approaches are most effective are being conducted (Coyle 2020[8]). People following the escalation approach may receive one or more 'first line' treatments, according to their disease severity, and the individual's and their clinicians' preference. The reasons for switching between first-line DMTs also include inadequate response not fulfilling criteria for second line treatment, adverse reactions or problems with tolerability, or justifiable lateral switches (e.g. low-dose to high-dose interferon beta, or vice versa)[9]. The treatment pathway is therefore highly varied between individuals, and first and second lines are broadly used to offer therapies as a proportion of people show a response to first line therapies and do not need to go to a second line therapy, which are riskier and more costly (NHSE 2019[6], Thompson 2018[7]).

DMTs are intended for use early in the disease course, when CNS inflammation is greatest. This 'window of opportunity' for treatment with DMTs continues until the onset of SPMS, at which point the disease is characterised as a chronic and progressive neurodegenerative process, and DMTs are considered to have little effect in slowing or stopping it (Díaz, Zarco, Rivera 2019[10]). At present, there are only two DMTs available for people with SPMS. Siponimod (TA656)[1] has recently been approved in the UK and is yet to be widely prescribed, while interferon beta (IFNB)-1b (TA527)[11] was approved in the UK in 2018. .

At the time of appraisal, both ozanimod (GID-TA10299)[12] and ofatumumab (TA699)[13] were both under appraisal by NICE as treatments for both first and second line RRMS, and it was not clear where in the treatment pathway these treatments would be positioned if recommended.

2.2.2. The technology

Ponesimod is a sphingosine 1-phosphate type 1 ($S1P_1$) receptor modulator that sequesters lymphocytes in lymph nodes by blocking S1P signalling. It can, therefore, be classified as an immunosuppressant drug in the same class as fingolimod (TA254[14]; second line treatment for RRMS/HA RRMS), ozanimod (GID-TA10299[12]; currently under appraisal for first and second line RRMS) and siponimod (TA656[1]; for the treatment of SPMS). However, these drugs are less specific, with fingolimod binding to S1P Type 1 as well as Types 3 to 5, while ozanimod and siponimod bind to S1P Types 1 and 5[15]. The off-target interactions with other S1P types are thought to cause undesirable effects as these receptor types are found in various cells, including tissues of the heart muscle and smooth arterial muscle. These effects range from cardiomyopathy and high blood pressure generally to bradyarrythmias, macular oedema and varicella-zoster viral infections with fingolimod specifically (Chaudhry 2017[15], Gajofatto 2015[9]). As a result of its increased specificity for $S1P_1$, ponesimod is proposed by the company to have fewer adverse effects than others in its class, however as with other DMTs, infections are still a potential concern due to its immuno-suppressive effects.

The company proposed that ponesimod may be used to treat people with active or highly active RRMS, and therefore could be considered as either a first- or second-line treatment for RRMS. As the line of treatment received by people with RRMS is guided by the balance in efficacy and safety shown by treatments, the appropriate positioning for ponesimod will be informed by clinicians' views towards its performance relative to existing treatments. The company further suggest that ponesimod may be preferred by people who prefer an oral treatment and/or a treatment with a shorter half-life. While covered under the licence, the company have not presented evidence for the use of ponesimod to treat people with SPMS, as few participants with SPMS were included in the trials of ponesimod. The ERG was unclear whether the company intended to position ponesimod towards people with RES RRMS: while people with RES RRMS were included in the company's clinical trials, and covered under the company's chosen definition of HA RRMS, the company excluded evidence for one of the treatments currently used to treat RES RRMS in the NHS (natalizumab).

Generally, the ERG considered that there may be a role for ponesimod to treat people with RRMS; however, there is no fixed position for ponesimod in the treatment pathway, due to variation in the pathway between people with RRMS, and the need to identify the relative balance of efficacy and safety of ponesimod. Clinical experts to the ERG stressed that DMT for HA RRMS need to show high efficacy, as there are efficacious treatments already available and

clinicians typically prefer an early, high efficacy treatment for people with this faster progressing disease course.

The ERG was aware that the treatment pathway for RRMS has changed within the context of the SARS-CoV 2 coronavirus pandemic, following updated guidelines from the Association of British Neurologists (ABN)[16]. As all DMTs interact with the immune system, the guidance aims to identify and prioritise those DMTs that pose a lower risk of infection or where the risk of lymphocyte rebound is greater than the risk of infection. The recommendations state that it is safe to start or continue on all NHSE first line treatments with the exception of ocrelizumab, as these DMTs pose a small risk of infection. Fingolimod poses a moderate risk of infection, but the risk of lymphocyte rebound is considered a larger risk. Alemtuzumab, cladribine and ocrelizumab are not recommended due to significantly heightened risk of viral infection. As ponesimod belongs to the same drug class as fingolimod, and is reported as having lower lymphocyte rebound, it is likely to pose a small to moderate risk of infection and would probably be considered safe in the pandemic context. There is uncertainty about when these guidelines will change, though clinical experts advised the ERG that some of the changes (for example around the frequency of monitoring) may be retained on a long-term basis.

2.3. Critique of company's definition of decision problem

The ERG's critique of the company's definition of the decision problem is provided in Table 8.

Table 8: Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	People with relapsing MS	People with RRMS (limited to people with active RRMS and people with highly active RRMS)	The decision problem is focused on a sub-population of people with MS because there is limited evidence available for ponesimod in SPMS for health technology evaluation. The evidence presented in the submission is based on a RCT (OPTIMUM) that evaluated ponesimod compared to teriflunomide in people with RMS. At study entry, most people in the trial were diagnosed with RRMS (97.4%). The trial included only a small proportion of people with SPMS (2.6%). Phase 3 data for people with RRMS is more robust in people with active RRMS and highly active RRMS (35% of trial population) and so the submission focuses on these two subgroups i.e. not in people with RES RRMS.	The company positioning of ponesimod has been adjusted since the NICE scope to focus on the treatment of people with active and highly active RRMS, and to exclude people with SPMS. This means that the intended use of ponesimod following this appraisal is narrower than the product licence for ponesimod. The ERG agrees that the available evidence for ponesimod is strongest in these populations, and it would not be possible for the ERG to evaluate the clinical efficacy of ponesimod in the SPMS population. There is no internationally standard definition of highly active RRMS, and all definitions rely on the judgement of the treating clinician. This creates heterogeneity in the evidence base, and some uncertainty in generalising evidence to the UK HA population. The company's definition of highly active varies from the definition used by NHS England, and includes people with RES. At clarification, the company presented a post-hoc subgroup analyses of data

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
				from their main trial in the RES population.
Intervention	Ponesimod	As per scope	N/A	The intervention in the company's main trial, OPTIMUM, matches the scope and licence for ponesimod. The company's Phase 2 trial compared the licensed dose of ponesimod with a higher and lower dose; the ERG appraisal of this trial is restricted to the licensed dose.
Comparator(s)	For people with active RRMS: • beta-interferon • dimethyl fumarate • glatiramer acetate • teriflunomide • ocrelizumab • peginterferon beta-1a • ozanimod (subject to ongoing NICE appraisal) • ofatumumab (subject to ongoing NICE appraisal) For people with highly active RRMS despite previous treatment: • alemtuzumab • cladribine • fingolimod • ocrelizumab (only if alemtuzumab is	For people with active RRMS (disease activity and treatment naïve): • beta-interferon • dimethyl fumarate • glatiramer acetate • teriflunomide • ocrelizumab • peginterferon beta-1a For people with highly active RRMS (i.e. disease activity whilst on 1st line therapy) • alemtuzumab • cladribine • fingolimod • ocrelizumab (only if alemtuzumab is contraindicated or otherwise unsuitable)	At the time of submission, ozanimod and ofatumumab have not been recommended by NICE as treatment options for MS and cannot be considered as standard of care within the NHS. Therefore, they not been considered in the submission. The OPTIMUM trial included only [redacted] SPMS population, therefore it was deemed that there is insufficient evidence for this population In line with previous clinical trials in MS, the definition of highly active RRMS employed in the OPTIMUM trial was broad, and thus also incorporates people with RES RRMS as defined by NHS England [6,17,18]. As a result, separate subgroup analyses of people with	At the time of writing, the ERG understood that ozanimod and ofatumumab were still under consideration by NICE. Previous appraisals of technologies for RRMS have included evidence for technologies currently under appraisal by NICE, and it was the view of the ERG and NICE that the company should have therefore included these comparators in their evidence base and economic model. At clarification the company provided this evidence, however within the timeframe, the company stated that their updated submission would be less rigorous (e.g. less comprehensive searching, and limitations in the way these treatments were added to the model). The ERG nevertheless

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	contraindicated or otherwise unsuitable) • ozanimod (subject to ongoing NICE appraisal) • ofatumumab (subject to ongoing NICE appraisal) For people with RES RRMS • alemtuzumab • cladribine • natalizumab • ocrelizumab (only if alemtuzumab is contraindicated or otherwise unsuitable) • ozanimod (subject to ongoing NICE appraisal) • ofatumumab (subject to ongoing NICE appraisal) For people with active SPMS (evidenced by continuing relapses) • established clinical management, including IFN-beta or other DMTs used outside their marketing authorisations • siponimod (subject to ongoing NICE appraisal)		RES RRMS were not part of the prespecified analysis.	considered the updated submission to be sufficient. The ERG agreed with the exclusion of siponimod as a direct comparator to ponesimod, due to the low numbers of people with SPMS included in the available trials. However, as SPMS health states were included in the company model, the ERG considered that evidence for siponimod should have been included in the company model (no treatment effects or costs for siponimod were included). The ERG was uncertain as to whether the company wish to position ponesimod for the treatment of people with RES RRMS; if so, the ERG considered that the company should have presented data for the relative efficacy of ponesimod to natalizumab.
Outcomes	The outcome measures to be considered include: • relapse rate • severity of relapse	The outcome measures to be considered include: • relapse rate • ARR	The outcomes captured by the OPTIMUM clinical trial of ponesimod are relevant for people with active RRMS or highly active RRMS and are	The outcomes reported by the company for the trial OPTIMUM are relevant to the NICE scope, and clinically meaningful for evaluating the efficacy of

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	• disability (for example, EDSS) • disease progression • symptoms of MS (such as fatigue, cognition and visual disturbance) • freedom from disease activity (for example lesions on MRI scans) • mortality • adverse effects of treatment • HRQoL	• Time to first confirmed relapse • disability • change from baseline in EDSS score • disease progression • 12-week CDA • 24-week CDA • symptoms of MS • change from baseline in FSIQ-RMS score • freedom from disease activity • CUAL • NEDA-3 • NEDA-4 • adverse effects of treatment • mortality • HRQoL • Change from baseline in SF-36 score • Change from baseline in MSFC Z-score	representative of current clinical practice in England. Outcomes such as severity of relapse and mortality could not be included in the pharmacoeconomic analyses due to the absence of comparative trial data. The OPTIMUM trial did not formally measure severity of relapse, which is difficult to measure in trials for MS. The OPTIMUM trial captures new Gd+ T1 lesions plus new or enlarging T2 lesions, which can indirectly denote disease severity. OPTIMUM trial outcomes are in line with outcome measures in previous MS trials appraised by NICE.	treatments for RRMS. The ERG agreed that measuring relapse severity is challenging, though was aware that the importance of distinguishing the severity of relapse has been noted previously by NICE. In addition to the outcomes noted by the company, the ERG noted that the company also measured additional markers of severity, including duration of relapse and relapses requiring hospitalisation (the latter was retrieved from the trial CSR)[19]. The ERG noted that most outcomes were only comprehensively measured and/or reported for OPTIMUM, and only a subset of the outcomes were reported for the extension phase of OPTIMUM and the company's placebo-controlled Phase 2 trial.
Economic analysis	Cost utility analysis	As per the scope, a cost utility analysis has been presented, whereby QALYs were used to capture the health benefits of ponesimod and comparator treatments. Costs were considered from an NHS and Personal Social Services perspective. Carer disutility has been included in the company's base case.	N/A	The ERG considered that the cost utility analysis was appropriate and matched the analysis outlined by the company in the scope.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Subgroups	Highly active RRMS	As per scope	N/A	No comment
Special considerations including issues related to equity or equality	None	The company did not identify any equity or equality concerns in the scope	N/A	The ERG agreed that there are no equity or equality concerns to be considered in this appraisal.

Abbreviations ARR, annualised relapse rate; CDA, confirmed disability accumulation; CSR, clinical study report; CUAL, combined unique active lesions; DMT, disease modifying therapy; EDSS, expanded disability status scale; ERG, Evidence Review Group; FSIQ-RMS, Fatigue Symptoms and Impacts Questionnaire-relapsing multiple sclerosis; Gd+, gadolinium-enhancing; HA, highly active; HRQoL, health-related quality of life; IFN, interferon; MRI, magnetic resonance imaging; MS, multiple sclerosis; MSFC, multiple sclerosis functional composite measure; NA, not applicable; NEDA, no evidence of disease activity; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life years; RCT, randomised controlled trial; RES, rapidly evolving severe; RMS, relapsing multiple sclerosis; RRMS, relapsing-remitting multiple sclerosis; SF-36, 36-item short form survey; SPMS, secondary progressive multiple sclerosis

3. CLINICAL EFFECTIVENESS

3.1. Critique of the methods of review(s)

The Company undertook a single systematic literature review (SLR) to identify evidence for ponesimod (summarised in Section 3.2) and to identify evidence for comparators to ponesimod to inform their indirect treatment comparison (Section 3.3 and 3.4). An overview of the methods used in the SLR is provided in Table 9 below.

Table 9: Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix D	The searches are thorough and well constructed. Searches have been run in three Ovid databases at once, and the results for each database have been extracted from the total results. The searches are therefore difficult to interpret or replicate but appear to be correctly executed. Suitable RCT filters have been used[20,21]. Search strategies for supplementary searches (e.g. in clinical trials registries) are not given, so it is not possible to determine how comprehensive these are. The ERG carried out some additional searches for multiple sclerosis NMAs in Medline and Embase from 2016 onwards (Appendix A) and found 1,044 papers. The company did not carry out any additional searches for adverse effects. Because the clinical effectiveness searches were limited to RCTs, any additional safety data not in RCTs may not have been found by the searches. The ERG carried out some additional searches for adverse effects for ponesimod in Medline and Embase (Appendix A) and found 148 papers, 30 of these were considered eligible following full-text screening.
Inclusion criteria	Appendix D	The ERG considered that that inclusion criteria used by the company in their review were broadly appropriate. However, the ERG disagreed with the

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		company's decision to exclude phase 4 trials from the NMA. The company rationale for this exclusion was due to variability in the methods used in phase 4 trials, however the ERG considered that problematic methods could have been accounted for in specific exclusion criteria. The ERG noted that these criteria led to the exclusion of several RCTs that have been included in previous NMAs of DMTs for RRMS, and could have expanded the available body of evidence for the company's analyses. However, the effect estimates for these comparators were not expected to alter greatly if the trials were included, and therefore the ERG did not investigate this further.
Screening	Appendix D	Conducted appropriately
Data extraction	Appendix D	Not described
Tool for quality assessment of included study or studies	TBA	Risk of bias assessment of OPTIMUM in the main body of the CS was reported according to the CRD tool, while the Cochrane risk of bias tool (version 1) was used to evaluate all RCTs included in the company's ITC. The Phase 2 trial and all trials included in the company's NMA were evaluated using the Cochrane risk of bias tool v.1. Both methods are appropriate for evaluating the quality of RCTs though the updated Cochrane v2 tool is generally preferred. No risk of bias assessment was reported for either of the long-term trial extensions to OPTIMUM or the Phase 2 trial.
Evidence synthesis	TBA	No synthesis of the ponesimod trials was conducted, as there is only one trial per comparison available. The company conducted several (number uncertain) NMAs to evaluate the comparative efficacy of ponesimod with other available treatments. Separate NMAs were conducted for trial-specified RRMS (ITT population, including both active and HA participants) and HA RRMS participants analysed in separate subgroup analyses. The ERG considered that further outcomes could have been evaluated in the NMAs, although as the company did not report their feasibility assessment in full, it is

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		not possible to determine if these outcomes were considered but found not feasible for analysis. The methods used in the NMAs were appropriate, though the ERG highlighted concerns about heterogeneity in the networks and the paucity of evidence, which both contributed to uncertainty in the results. The ERG also noted that several key outputs of the NMAs were not reported in the CS.

Abbreviations: CRD, Centre for Reviews and Dissemination; CS, Company submission; DMT, disease modifying therapy; ERG, Evidence Review Group; HA, highly active; ITT, intention-to-treat; NMA, network meta-analysis; RCT, randomised controlled trial; RRMS, relapsing-remitting multiple sclerosis

3.2. Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

The company presented evidence for ponesimod from one head-to-head Phase 3 randomised controlled trial (RCT; OPTIMUM) and one Phase 2 placebo-controlled dose-finding trial in participants with RRMS (B202). Each of these studies were followed by an extension phase evaluating ponesimod only. An overview of the methods used in these studies is presented across the following sections (Sections 3.2.1 to 3.2.4).

3.2.1. Study design

The company's primary evidence for ponesimod is derived from OPTIMUM, a randomised, double-blind, parallel-group, multicentre Phase 3 trial of ponesimod 20 mg vs. teriflunomide 14 mg in participants with RRMS. The trial measured a broad range of clinical efficacy and safety outcomes up to 108 weeks. OPTIMUM is a well-designed RCT, and the ERG agreed with the company approach to place the evidence from this trial in greater prominence than the earlier Phase 2 trial. However, clinical advisors to the ERG cautioned that the trial follow-up may be too short to evaluate meaningful disease progression. This may lead to some uncertainty surrounding disability estimates, including impact on conversion to SPMS (where levels of disability are most pronounced. It was also noted that the sample size of OPTIMUM may be too small to identify the risk of rare, but serious adverse events.

The double-blind phase of OPTIMUM was followed by AC-058B303, a single-arm extension phase for those participants who completed the double-blind phase, and wished to continue on ponesimod or switch to ponesimod from teriflunomide. Follow-up of the extension was up to 132

weeks following the double-blind phase. The CS contains a subset of the clinical efficacy and safety outcomes measured for OPTIMUM for the extension phase, and a full clinical study report (CSR) for the extension phase was not provided by the company. However, despite reporting data at a longer follow-up than the core trial, treatment was open-label and uncontrolled, and is therefore of a lower evidence quality.

The Phase 2 trial, AC-058B202, was a randomised dose-finding trial of ponesimod, which compared three doses of ponesimod with each other and with placebo. The trial lasted 24 weeks, after which point all people receiving placebo were offered ponesimod. The extension phase lasted 552 weeks and consisted of three phases, over which groups were randomised to different doses of ponesimod until in the final phase all people received a 20 mg dose of ponesimod only (the current licensed dose). As differences in efficacy and safety were noted across the doses, for the purposes of this appraisal the ERG focused on the subset of people who received the 20 mg dose continually across all phases of the trial (n=147)

An overview of the trial designs is provided in Table 10.

Table 10: Overview of ponesimod trial designs

Study name and acronym	Study design	Phase	Intervention / Comparator	Study Objectives	Population
OPTIMUM; AC-058B301 [NCT02425644]	Randomised, double-blind, active-controlled parallel trial Follow-up: 108 weeks	3	Ponesimod 20 mg once daily / Teriflunomide 14 mg once daily	Efficacy and safety	N = 1,133 Participants with active RRMS who were treatment naïve or have received previous treatment with interferons, glatiramer acetate, natalizumab, or dimethyl fumarate. Participants were ambulatory, with EDSS score 0-5.5 at screening and baseline. Subgroup analyses were conducted in highly active RRMS.
OPTIMUM-LT; AC-058B303 [NCT03232073]	Single-group, open-label, non-comparative long-term	3	Ponesimod, gradually up-titrated over day 1 to 14 until a maintenance	Long-term safety and control of RMS	N = 877. Extension in participants who completed up to week 108 of the OPTIMUM trial.

Study name and acronym	Study design	Phase	Intervention / Comparator	Study Objectives	Population
	extension of OPTIMUM Follow-up: up to 132 weeks		dose of 20 mg is reached on day 15 / No comparator		
AC-058B201[a] [NCT01006265] (Olsson et al. 2014[22])	Randomised, double-blind, placebo-controlled dose-finding study Follow-up: 24 weeks	2b	Ponesimod 10, 20, or 40 mg once daily / Matching unspecified placebo once daily	Efficacy, safety and tolerability of ponesimod at various doses	N = 237. Participants with RRMS (per revised 2005 McDonald criteria[23]) with ≥ 1 documented relapse(s) within 12-months before screening, ≥ 2 relapses within 24 months before screening, or at least one T1-weighted Gd+ lesion on brain MRI at screening. EDSS score 0-5.5.
AC-058B202[a] [NCT01093326]	Randomised, double-blind, multiple-dose, uncontrolled long-term extension of AC-058B202 Follow-up: 528 weeks	2b	Ponesimod 10, 20, or 40 mg once daily / No comparator	Long-term efficacy, safety and tolerability of ponesimod at various doses	N = 147. Extension in participants who completed the dose-finding study AC-058B201.

Abbreviations: EDSS, Expanded Disability Status Scale; MRI, magnetic resonance imaging; N, number; RRMS, relapsing-remitting multiple sclerosis

Notes: [a] Number of people reported are the total of those randomised to ponesimod 20 mg and placebo only

3.2.2. Trial populations

Population eligibility and characteristics are outlined in this section, including comparability of the trials and trial arms, and generalisability of the trial samples to the target population.

3.2.2.1. Eligibility criteria

Key inclusion and exclusion criteria used in the two included trials are summarised in Table 11 below. The trials identified participants according to the McDonald 2010[24] (OPTIMUM) and McDonald 2005[23] (Phase 2) criteria; while these criteria were most recently updated in 2017, the earlier versions are appropriate for this appraisal, as the update mainly affects those earlier in the disease course who would not normally be considered for DMT. The trials sought to exclude

people with progressive MS, including both primary and secondary progressive MS; however, OPTIMUM did include a small minority of people with SPMS in their final results. The most likely explanation for this is that the diagnosis of SPMS is often done retrospectively, and so participants may have received a diagnosis following inclusion in the trial. The age and EDSS inclusion criteria for participants in the trials were appropriate for the target population.

Both treatment-naïve and previously treated people were included in the trials, which aligns with the proposed positioning of ponesimod as either a first or second line treatment. Where appropriate, the previous DMT was required to have washed out prior to the start of the trial, and no previous treatment with cladribine or ocrelizumab was permitted.

The trials excluded people with certain cardiovascular (CV) comorbidities and abnormal liver diagnostics; this may have been a precaution as both are known risks with fingolimod treatment. These exclusion criteria were broadly comparable with the contraindications outlined in the licence for ponesimod, though the ERG noted that people who had experienced macular oedema in the past were still eligible for inclusion (macular oedema is also a known risk of treatment with S1P modulators. The exclusion of people at risk of these outcomes may also be an obstacle in identifying similarities in the safety profile of ponesimod and other S1P modulators.

For the long-term extensions, all participants who completed the core phases of each trial and were willing to continue were eligible for inclusion. However, those participants who discontinued ponesimod for any reason, including for adverse events (AEs) or lack of efficacy, would not have been included in the long-term trial extensions. This is generally reflective of likely use in UK practice since people who do not tolerate ponesimod for any reason will not continue on treatment for any extended period.

Table 11: Eligibility for the included trials

Study	Inclusion criteria	Exclusion criteria
OPTIMUM	Aged 18-55 MS with relapsing course from onset (2010 revised McDonald[24] criteria): • 1+ attacks with onset within 12-1 months prior to baseline EDSS or; • 2+ attacks with onset within 24-1 months prior to baseline EDSS or; • 1+ (Gd+) lesions on an MRI within 6 months prior to baseline EDSS Treatment-naïve or previously treated with IFN beta-1a, IFN beta-1b, glatiramer acetate, natalizumab, or dimethyl fumarate Ambulatory with EDSS of 0-5.5 Agreed to use an accelerated elimination for teriflunomide after study	Lactating/pregnant women Progressive MS Significant medical conditions or receiving therapies for such conditions Unlikely to comply
B201	Aged 18-55 Presented with RRMS as defined by revised McDonald criteria (2005) At least one of the following characteristics of RRMS: • 1+ relapse within 12 months prior to screening • 2+ relapses within 24 months prior to screening • 1+ Gd+ lesion Ambulatory with EDSS 0-5.5 No exacerbation last 30 days	Progressive MS Treatment with the following medications within 30 days prior to randomisation: Systemic corticosteroids or adrenocorticotropic hormone Beta-blockers, diltiazem, verapamil or digoxin or QT-prolonging drugs Pregnancy; or women breast-feeding Treatment with certain DMTs and immunosuppressive agent within 3-6 months of trial start Treatment with the following medications at any time prior to randomization: Cyclophosphamide, mitoxantrone or cladribine

Study	Inclusion criteria	Exclusion criteria
		Lymphocyte-depleting biologic agents
		Autoimmune disorder other than MS
		Ongoing bacterial, viral or fungal infection (with the exception of onychomycosis and dermatomycosis), positive hepatitis B surface antigen or hepatitis C antibody tests
		Certain current infections
		History or presence of malignancy
		Poorly controlled type I or type II diabetes and associated complications
		History of clinically significant drug or alcohol abuse
		People with certain CV or pulmonary conditions
		Abnormal LFTs
		Abnormal blood test results
		Known allergy to any of the study drug excipients
		Any other condition which would put the person at risk by participating in the study
		Unlikely to comply

Abbreviations: CV, cardiovascular; DMT, disease modifying therapy; EDSS, expanded disability status scale; Gd+ gadolinium-enhancing; IFN, interferon; LFT, liver function test; MRI, magnetic resonance imaging; MS, multiple sclerosis; QT, start of the Q wave to end of the T wave on electrocardiogram; RRMS, relapsing-remitting multiple sclerosis

3.2.2.2. Baseline characteristics

The baseline characteristics of the participants in the included trials are summarised in Table 12 alongside comparative characteristics of the UK risk-sharing scheme (RSS) population. No separate population characteristics were reported for the HA populations included in the included trials. In the following sections, the ERG summarised the comparability of the trial arms in the included trials, as well as the relevance of the trial populations to the NHS target population.

Table 12: Baseline characteristics of the intention-to-treat populations of the included trials, and their comparability with UK risk-sharing scheme populations

Characteristic	OPTIMUM		Phase 2 trial (B202)		UK RSS[25]
	Ponesimod	Teriflunomide	Ponesimod 20mg	Placebo	
Age (SD)	36.7 (8.74)	36.8 (8.74)	35.5 (8.5)	36.6 (8.6)	39.4(9.05)
Female	64%	65.7%	67.5%	70.2%	74.2%
Received 1+ prior DMT	[redacted]	[redacted]	35.1%	39.7%	
DMT received in 2 years prior to randomisation	37.6%	37.3%			
EDSS (Median (Q1-Q3))	[redacted] Range: [redacted]	[redacted] Range: [redacted]	2.0 (1.5-3.0) Range: 0.0-5.5	2.0 (1.5-3.0) Range: 0.0-5.5	3.5 (2.0-5.0)
Years since first symptoms at randomisation (SD)	7.63 (6.781)	7.65 (6.782)	7.3(6.25)	6.9(5.7)	8.8(7.47)
Mean relapses within year prior to study	1.2 (0.61)	1.3 (0.65)			

Characteristic	OPTIMUM		Phase 2 trial (B202)		UK RSS[25]
entry (SD)					
Mean months since last relapse (SD)	[redacted]	[redacted]	5.1(5.51)	5.6(4.53)	
Disease subtype	97% RRMS 3% SPMS	98% RRMS 2% SPMS			86.2% RRMS 13.8% SPMS
Presence of Gd+ T1 lesions	39.9%	45.4%	40%	47.4%	
Number of T2 lesions	[redacted]	[redacted]			
Mean volume of T2 lesions (mm^3 (SD))	8301.4 (10346.28)	9489.2 (11265.42)	7747(10,005)	6125(8988)	
Mean BMI kg/m^2 (SD)	[redacted]	[redacted]			
Geographic region	[redacted]	[redacted]			
Mean FSIQ-RMS weekly symptom	31.9 (20.4)	32.8 (19.1)			

Characteristic	OPTIMUM		Phase 2 trial (B202)		UK RSS[25]
s score (SD)					
% of people 'highly active'	35.6%	35.3%			
% of people with RES	[redacted]	[redacted]			
White race	97.2%	97.7%	98.2%	94.2%	
Number of relapses in last 24 months	NR	NR	0 – 1.8% 1 – 43% 2+ - 55.3%	0 – 0.8% 1 – 40.5% 2+ - 58.7%	3 (2-3) Median (quartiles)
Mean relapses in last year (SD)	1.2 (0.61)	1.3 (0.65)	1.2 (0.62)	1.3 (0.68)	
Mean number of Gd+ T1 lesions (SD)	NR	NR	2.5 (6.61)	1.7 (3.31)	

Abbreviations: BMI, body mass index; DMT, disease modifying therapy; EDSS, expanded disability status scale; EU, European Union; FSIQ-RMS, Fatigue Symptoms and Impacts Questionnaire-relapsing multiple sclerosis; Gd+ gadolinium-enhancing; Q1, quartile 1; Q3, quartile 3; RES, rapidly evolving severe; RoW, rest of the world; RRMS, relapsing-remitting multiple sclerosis; RSS, risk-sharing scheme; SD, standard deviation; SPMS, secondary progressive multiple sclerosis

Comparability of trial arms

The baseline characteristics of participants in the ITT population of the included studies were balanced across arms. Randomisation had been stratified by EDSS score at baseline and prior DMT in the previous two years. Baseline characteristics were not reported separately for the HA population, and so it was not possible to determine if characteristics were also balanced for the company's subgroup analyses.

Relevance of trial populations to the target population

Based on the data reported, the ITT population characteristics in both included trials investigating ponesimod appear broadly similar to people in the UK population who are eligible for first or second line DMTs; this was a view shared by clinical advisors to the ERG. The EDSS scores in both ponesimod trials appear marginally lower than in the RSS population[25], suggesting that people in the trials had lower disability than the target population; however this is likely due to a higher proportion of people with SPMS in the RSS population, and because people in the RSS population generally had a longer disease course without early access to DMT.

However, the definition of HA RRMS used in OPTIMUM included people with RES RRMS, which varies from the definition used in the NHS. Overall, [redacted] of the people in OPTIMUM had RES, equating to [redacted] of the highly active population. People diagnosed with RES are at a higher risk of disease progression, and therefore absolute clinical outcomes may vary from the active and highly active RRMS populations. It is unclear whether treatment efficacy may also vary in people with RES, though they may be treated with different, more efficacious treatments earlier in the disease course (and typically not with teriflunomide). The variation in the definition of HA reflects the international nature of the OPTIMUM trial, given that there is no universally accepted definition of 'highly active' RRMS (see Section 2.1 and Table 13 below for a comparison of these definitions). The generalisability of evidence to different RRMS populations is an area of uncertainty within this appraisal.

The ERG noted that participants in OPTIMUM had on average been symptomatic for over seven years, and that [redacted] were treatment naïve, with the remaining [redacted] having had at least one DMT previously. Clinical advice to the ERG was that use of DMT within the first two years of the disease is associated with better outcomes, though the ERG was aware that many people with RRMS choose not to receive DMT. Amongst participants who had previously

received DMT, previous treatments were generally consistent with those prescribed in the NHS, though as to be expected with an international trial, some minor differences were noted. Notably, the inclusion of participants with HA RRMS in OPTIMUM is an alteration from the NHS treatment pathway, as teriflunomide is not used to treat HA RRMS in the UK.

The ERG was unclear to what extent evidence from this population would generalise across populations at different lines of treatment; the company did not report any subgroup analyses according to line of treatment, and little is known about how treatment effects vary according to the previous treatments people with RRMS have received. Clinical advice to the ERG was that evidence from people who have stopped treatment due to a lack of efficacy may represent people with more active disease, and therefore subgroup analyses in the HA population may identify if treatment effects vary as compared to the main ITT population. The ERG recognised the broad inclusion criteria of OPTIMUM as an attempt to evaluate ponesimod across a broad RRMS population; however, the trial was potentially not large enough for comprehensive subgroup analyses to explore variation in treatment effects across variability in the trial population. As little is known about effect modifiers in the broader RRMS literature, there is some uncertainty about the generalisability of evidence from the included trials to the target NHS populations.

Table 13: Currently used definitions of highly active disease

Source	Definition of highly active population	Includes RES
OPTIMUM	Any DMT for MS received within 12 months prior to randomisation and one or both of the following: - ≥1 relapse within 1 year prior to study entry and the baseline MRI read centrally showed either ≥1 Gd+ T1 lesion and/or ≥9 T2 lesions. - Number of relapses within 1 year prior to study entry ≥ number of relapses between 2 and 1 year prior to study entry, for people with at least one relapse within 2 years prior to study entry. ≥2 relapses within the 1 year prior to study entry and baseline EDSS score >2 and baseline MRI read centrally showed ≥1 Gd+ T1 lesion.	Yes
NHS	People with an unchanged or increased relapse rate or ongoing severe relapses compared with the previous year despite treatment with beta interferon.	No
TA254[14] (25/04/2012)	People with high disease activity despite treatment with a beta-interferon. A treatment failure is defined as a lack of response to a full and adequate course of beta interferon (normally at least one year of treatment). People 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 people with an unchanged or increased relapse rate or ongoing severe relapses compared to the previous year	No
TA312[2] (28/05/2014)	Adults with high disease activity despite treatment with a beta interferon (normally at least one year of treatment). People have at least 1 relapse in the previous year while on therapy, and have at least 9 T2-hyperintense lesions in cranial MRI or at least 1 gadolinium-enhancing lesion; OR unchanged or increased relapse rate or ongoing severe relapses, as compared to the previous year.	Yes
TA533[26] (25/07/2018)	Treated with interferon or glatiramer acetate for ≥1 year and had: (1) ≥1 relapse in the previous year; (2) ≥1 gadolinium-enhancing T1 lesion at baseline; (3) ≥9 hyperintense T2 lesions at baseline.	No
TA616[27] (19/12/2019)	The NICE committee considered that the sub optimally treated (SoT) group in the company submission best reflected the UK HA population. SoT was defined as at least 1 relapse in the previous year while the person was on disease-modifying therapy, and at least 1 T1 gadolinium-enhancing lesion or 9 T2 lesions	Yes

Abbreviations: DMT, disease modifying therapy; EDSS, Expanded Disability Status Scale; HA, highly active; MRI, magnetic resonance imaging; MS, multiple sclerosis; RES rapidly evolving severe; SoT, suboptimally treated

3.2.3. Intervention characteristics

The intervention characteristics delivered during the included trials are summarised in Table 14 below. Ponesimod is delivered as an oral treatment taken as one 20 mg tablet each day. This dose was selected following the company's Phase 2 dose-finding trial, which also evaluated a lower (10 mg) and higher (40 mg) dose of ponesimod. This trial showed that the higher dose of ponesimod resulted in an increased risk of adverse events without a commensurate benefit for efficacy. Interestingly, a recent analysis reported that a 40 mg of ponesimod resulted in the worst rate of discontinuations due to adverse events as compared to other DMTs in the active population (Tong 2021 et al.[28]). No reductions or increases in dose were permitted during OPTIMUM, and none are specified in the licence for ponesimod.

The company recommends a period of up-titration for ponesimod, which they stated in section B.2.50 of the CS is to avoid cardiac adverse events such as those associated with fingolimod. Different up-titration protocols were used in the two trials, with a longer (two weeks) period used in OPTIMUM compared to the Phase 2 trial (one week).

Many concomitant therapies were used by participants in OPTIMUM to manage the symptoms of RRMS and adverse events experienced during the trial. Their use was broadly comparable between the ponesimod and teriflunomide arms; however, the ERG noted lower use of corticosteroids in the ponesimod arm (31.4% of the ponesimod group used corticosteroids, compared to 43.1% of those in the teriflunomide arm). Corticosteroids have an established safety profile, though side effects were considered unlikely to alter the efficacy of treatments in the trial.

Table 14: Intervention characteristics of the included trials

Trial		Treatment
OPTIMUM	Ponesimod	Up titration at initiation from 2 mg to 10 mg over first 14 days 20 mg daily from Day 15 onwards 108 weeks
	Teriflunomide	Mock up-titration of 14 mg for first 14 days 14 mg taken daily 108 weeks
OPTIMUM Extension		As above, up to 240 weeks
Phase 2 trial (B201)[a]		10 mg on days 1-7 Up-titrated to 20 mg on Day 8

Trial		Treatment
		24 weeks
B201 Extension	10mg Group	10 mg up to 96 weeks (core study and TP1) 10 mg during TP2 Increased to 20 mg for TP3
	20 mg Group	20 mg up to 432 weeks (TP1, 2 and 3)
	40 mg Group	40 mg up to 96 weeks (core study and TP1) Randomised 1:1 to 10 mg and 20 mg for TP2 All received 20 mg in TP3
	Placebo	Once daily placebo for 24 weeks Placebo populations switched to one of the above treatment regimens for long term extension

Abbreviation: TP, treatment period

Notes: [a] As a dose-finding study, Olsson et al.[22] also treated groups with 10 mg and 40 mg, the ERG has excluded these groups here as they are outside the licensed dose.

3.2.4. Clinical effectiveness results

3.2.4.1. Outcome measurement

As noted previously, the choice of DMT for RRMS frequently involves a trade-off between efficacy and safety (see Section 2.2). Clinical advice to the ERG was that the clinical efficacy of DMTs is firstly demonstrated by a reduced risk of relapse, including neurological evidence that disease progression is delayed (e.g. reduced number and size of lesions). Reduced disability and impact on HRQoL are also important outcomes, and clinical advice was that reducing the relapses may lead to benefits in these outcomes. DMTs are not expected to reverse disease progression or disability, and therefore efficacy is demonstrated by stability or slower disease progression at follow-up.

Specific safety concerns associated with DMTs for RRMS include infection, due to the immune-suppressive mechanisms of the treatments, hypertension and cardiac events, liver disorders, malignancy, and macular oedema. The ERG noted that fingolimod, also a sphingosine 1-phosphate (S1P) receptor modulator, has been associated with an increased risk of liver and cardiac events[29], which means that some people are ineligible for treatment, and increased monitoring for adverse effects is required during treatment.

The company reported a range of absolute and relative effect estimates to evaluate the efficacy of ponesimod. The clinical efficacy outcomes reported by the company can be grouped into measures of the *risk of relapse*, *neurological/radiological outcomes*, and measures of *disability and HRQoL*. In addition, the company reported *safety* based on the risks of treatment-emergent adverse events and discontinuation due to adverse events. The company also reported *additional outcomes*, including rates of NEDA, which is the rate of people demonstrating an absence of disease activity as a composite of several clinical outcomes, and the rate of all-cause discontinuation, which represents discontinuations due to either efficacy or tolerability (or trial attrition). The bulk of these outcomes were only measured and reported for OPTIMUM, with a subset only report for the long-term phase of OPTIMUM and for the Phase 2 trial. An overview of outcome definitions and their measurement is provided below. These descriptions also capture limitations with measurements in the included trials.

Relapse

Clinical advice to the ERG was that the company's definition of relapse was broadly appropriate: the company defined relapse as new, worsening, or recurrent neurological symptoms occurring ≥30 days following the onset of a prior relapse and sustained ≥24 hours without fever or infection (CS Document B, p. 45). However, clinical advice to the ERG was that this definition may include an exacerbation of symptoms caused by anxiety or stress that is not a relapse. This difficulty highlights the subjective nature of measuring relapse, which requires the judgement of the person with RRMS and their clinicians.

The primary outcome of OPTIMUM was annualised relapse rate (ARR), which represents the number of reported relapses per patient-year. The average relapse rate for people receiving ponesimod at baseline was [redacted]

The company reported a variety of further measures to characterise the efficacy of ponesimod on relapse rates, including: time to first confirmed relapse; proportion of participants with ≥1 relapse; duration of relapse; and rates of relapse requiring corticosteroids. The ERG also identified rates of relapse resulting in hospitalisations and A&E admission from the trial CSR[30]. As discussed in Section 2.3, previous NICE appraisal committees have highlighted the importance of distinguishing variation in the severity of relapses experienced by people. The

severity of relapse is challenging to define, though relapse resulting in hospitalisation is sometimes used.

Neurological/radiological outcomes

The company reported a range of neurological and radiological outcomes, including the proportion of new or enlarging lesions across various definitions, and magnetisation transfer ratio (MTR) values. These outcomes are typically challenging to interpret, due to reliability issues in MRI measures and uncertainty about the relationship of the measures with disease progression. However, clinical advice to the ERG was that the rate of combined unique active lesions (CUALs) and loss in brain volume are both considered to be useful markers of disease progression. At clarification, the company noted that measurement of CUAL may vary across trials, thus making any evaluation challenging.

Disability

The principal measure of disability used in evaluations of DMT for RRMS is the time to confirmed disability accumulation (CDA), which is a measure of sustained, meaningful change in disability. The company definition is consistent with previous appraisals; i.e. an increase of ≥1.5 in expanded disability status scale (EDSS) score for people with a baseline EDSS score of 0.0, an increase of ≥1.0 for people with a baseline score of 1.0 to 5.0, or an increase of ≥0.5 for people with a baseline score of ≥5.5. To account for natural fluctuation in RRMS, a change in disability is considered to have occurred if the change in EDSS score is maintained for a prolonged period. The company evaluated CDA confirmed at 12 weeks (CDA at 3 months, or CDA-3) or at 24 weeks (CDA at 6 months, or CDA-6). While these time periods are consistent with those evaluated in previous appraisals of RRMS treatments, committees have commented that these time periods may be too short to evaluate a meaningful change in disability. These concerns were echoed by clinical advice to the ERG. The company also separately reported change in participants' EDSS scores.

Health-related quality of life and participant-reported outcomes

Health-related quality of life (HRQoL) was measured by the SF-36 (domain and composite scores); however, these data were not reported in the CS, apart from some categorised data of the proportion of people who considered their health to be 'much better' during the trial. The ERG considered the latter data to be highly limited, and the absence of HRQoL data in the CS

was considered to be a major omission. These data were therefore retrieved from the trial CSR[30].

Additional participant-reported outcome data was available from the Multiple Sclerosis Functional Composite (MSFC) scale. The MSFC combines three separate measures to assess lower extremity, upper extremity, and cognitive function. People are asked to complete a series of tasks, which are then rated by a trained observer. For each measure, participants' scores are standardised into a z-score using a reference population (e.g. representing the standard deviation from baseline scores for the trial population), which are then combined to give an overall measure of function across the three measures. Higher positive scores were associated with improvement, while negative scores were associated with deterioration. It has been suggested that a change of 15-20% can be considered clinically meaningful; a threshold chosen in part because lower thresholds may reflect natural fluctuations in functioning[31]. The company did not report a threshold to interpret the results of the MSFC, and data were not reported as a percentage change.

The Fatigue Symptoms and Impacts Questionnaire: Relapsing Multiple Sclerosis (FSIQ-RMS) is a new scale developed by the company[32] to measure fatigue, which can significantly affect the lives of people with RRMS. The company proposed that this scale better represents the symptoms of fatigue in RRMS than other available measures as it evaluates both cognitive and physical symptoms. The FSIQ-RMS consists of two scales, one measuring symptoms and one measuring the impact of symptoms. On both scales, higher scores represent more fatigue or impact. As the FSIQ-RMS is a new tool, it has not been evaluated in previous appraisals or research, and the associated publication did not report a threshold for what change or difference in scores would be considered clinically meaningful.

No HRQoL or PRO outcomes were considered in the company's ITC.

Safety

The ERG noted that ascertainment of AEs was conducted through voluntary reporting or non-directed interviewing of participants, and considered this approach to be reasonable. Safety assessments for post-treatment follow-up, both for those entering the extension of OPTIMUM and those who did not, as well as post-treatment observation, for those people who discontinued the study prematurely, as reported in the CSR for OPTIMUM[19] appear reasonable.

From the Phase 2b core study publication (Olsson 2014[22]), the details of these assessments appear similar to those for OPTIMUM.

S1P modulators such as fingolimod and ozanimod have known safety concerns, i.e. cardiovascular, immune, ophthalmologic, pulmonary and hepatic effects (Novartis 2019[29], Gajofatto & Benedetti 2015[9], Swallow 2020[33]). The coverage of safety assessments for people treated with ponesimod, as reported by the company in the CSR for OPTIMUM[19] and for the Phase 2b core study (Olsson 2014[22]), seemed reasonable.

With regards to the handling of data, the company reported receiving scientific advice approving of the pooling strategy of safety data across the Phase 2b and OPTIMUM trials, as well as their extensions, with consideration given to differences in characteristics of the trials.

Clinical advice to the ERG indicated that the length of follow-up of the included trials for ponesimod may not be sufficient to detect rare, serious AEs; as has been the case in the NICE appraisal of fingolimod (TA254) in 2012. Following approval, cases of progressive multifocal leukoencephalopathy (PML) have occurred in the post-marketing context. The duration of both trials assessing direct comparisons of fingolimod in this appraisal were 12 and 24 months, [redacted]). The company reported [redacted] in the Phase 2b study and its extension, indicating that some rare serious AEs, were they to have occurred, may have manifested by the time of submission, though the sample size of this study is very limited (n=[redacted]).

Other outcomes

The company also reported NEDA, representing the absence of disease activity according to several levels of criteria. The company cited references proposing that NEDA-3 (the absence of confirmed relapse, Gd+ T1 lesions, new or enlarging T2 lesions and 12-week CDA) is considered to be a valuable treatment goal of DMT, as it may have a stronger association with long-term outcomes as compared to single measures. However, the ERG understood that there is uncertainty about whether the criteria appropriately measure disease progression, and to what extent this outcome is able to predict further progression. The company reported data from OPTIMUM for NEDA-3 as well as NEDA-4 (NEDA-3 criteria plus absence of brain atrophy). Neither outcome was considered in the company's NMA.

The company also reported data for the rate of all-cause discontinuation. This outcome could be considered to represent a composite of discontinuation due to either efficacy or safety, though it

could also include discontinuation due to trial attrition. As DMTs for RRMS often involve a compromise of efficacy and safety, comparisons of all-cause discontinuation will derive very different results from analyses restricted to discontinuation due to either efficacy or safety.

3.2.4.2. Results

Clinical efficacy

Key clinical efficacy results for the ITT populations in the OPTIMUM trial and its extension, and the Phase 2 trial and its extension, are summarised in Table 15. The company did not report clinical effectiveness data specifically from the Phase 2 placebo-controlled trial of ponesimod (B201), opting instead to report limited clinical efficacy data from across the core and long-term phases of the trial; however the ERG identified select data points from the trial CSR[19]. Limited data only were provided by the company for the long-term extension of OPTIMUM in the CS, and no full CSR was provided to the ERG.

Overall, the results showed [redacted]. Measures of brain volume loss, CUALs, and NEDA also suggested that participants receiving ponesimod [redacted]. The ERG noted that [redacted]. Both OPTIMUM and the Phase 2 trial showed [redacted] in the ponesimod arm, [redacted]

However, [redacted]

[redacted]

No data for these outcomes were reported for the Phase 2 trial.

The company reported that ponesimod was associated with [redacted]

Data from OPTIMUM suggested that approximately [redacted]. A similar breakdown was not available in the Phase 2 trial, though overall rates of discontinuation were [redacted].

Table 15: Clinical effectiveness results for ponesimod (ITT population; OPTIMUM and Phase 2 trial)

Outcome	Outcome measurement	OPTIMUM Follow-up 108 weeks		OPTIMUM Extension. Follow-up 132 weeks	B201 Follow-up 24 weeks		B201 Extension Follow-up 432 weeks
Treatment		Ponesimod	Teriflunomide	Ponesimod	Ponesimod[#]	Placebo	Ponesimod[#]
ITT sample		567	566	877	116	121	145
Relapse	Total relapses (n)	242	344	NR	[redacted]	[redacted]	[redacted]
	ARR (mean)	0.202 (95%cl 0.173, 0.235)	0.290 (95%cl 0.254, 0.331)	[redacted]	[redacted]	[redacted]	[redacted]
	ARR (relative rate)	0.695 (95%cl 0.570, 0.848)[¥]		-	[redacted]		-
	Population with ≥1 relapse (%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Time to first relapse (HR)	[redacted]		-	[redacted]		-
	Median (IQR) duration of relapse (days)	[redacted]	[redacted]	NR	[redacted]	[redacted]	[redacted]
	Relapses requiring corticosteroid treatment	[redacted]	[redacted]	NR	[redacted]	[redacted]	[redacted]
	Relapses requiring hospitalisation	[redacted]	[redacted]	NR	[redacted]	[redacted]	[redacted]
	Relapses requiring A&E admission	[redacted]	[redacted]	NR	NR	NR	NR
3-month CDA	Rate (%)	[redacted]	[redacted]	[redacted]	NR	NR	[redacted]

Outcome	Outcome measurement	OPTIMUM Follow-up 108 weeks		OPTIMUM Extension. Follow-up 132 weeks	B201 Follow-up 24 weeks		B201 Extension Follow-up 432 weeks
	HR	0.83 (95%cl 0.58, 1.18)		-	NR		-
6-month CDA	Rate	8.1%	9.9%	[redacted]	NR	NR	[redacted]
	Risk reduction	0.84 (95%cl 0.57, 1.24)		-	NR		-
Trial discontinuation	All-cause	[redacted]	[redacted]	NR	[redacted]	[redacted]	[redacted]
	Rate due to safety or tolerability	[redacted]	[redacted]	NR	NR	NR	[redacted]
	Rate due to efficacy	[redacted]	[redacted]	NR	NR	NR	[redacted]
CUALs	Mean (annualised)	1.405	3.164	NR	NR	NR	[redacted]
	RR (95%CI)	0.44 (0.364, 0.542)		-	NR		-
Brain volume loss	LS mean Δ	-0.91%	-1.25%	NR	NR	NR	NR
	LS mean difference (95%CI)	0.34% (0.17, 0.50)		-	NR		-
	Rate of populations with annual brain volume decrease ≥0.4% from baseline	33%	42%	NR	NR	NR	NR
Fatigue	FSIQ-RMS LS mean Δ from baseline	-0.01	3.56	NR	NR	NR	NR
	LS MD	-3.57 (95%cl -5.83, -1.32)		-	NR		-

Outcome	Outcome measurement	OPTIMUM Follow-up 108 weeks		OPTIMUM Extension. Follow-up 132 weeks	B201 Follow-up 24 weeks		B201 Extension Follow-up 432 weeks
	OR for improvement or stable response (Δ≤6.3 from baseline)^	[redacted]		-	NR		-
mFIS	Mean Δ from baseline	[redacted]	NR	NR	[redacted]	[redacted]	NR
EDSS	Mean Δ from baseline	[redacted]	[redacted]	NR	[redacted]	[redacted]	[redacted]
	LS Mean diff	[redacted]		-	NR		-
NEDA	NEDA-3 (rate)	[redacted]	[redacted]	NR	NR	NR	NR
	NEDA-3 (OR)	1.70 (95%cl 1.27, 2.28)		-	NR		-
	NEDA-4 (rate)	[redacted]	[redacted]	NR	NR	NR	
	NEDA-4 (OR)	1.85 (95%cl 1.24, 2.76)		-	NR		-
MSFC	LS mean change in z-score	[redacted]	[redacted]	NR	NR	NR	NR
	LS mean difference	[redacted]		-	NR		-
SF-36	Physical component mean (SD)	[redacted]	[redacted]	NR	NR	NR	NR
	Mental component mean (SD)	[redacted]	[redacted]	NR	NR	NR	NR

Abbreviations: A & E, Accident and Emergency; ARR, annualised relapse rate; CDA, confirmed disability accumulation; CI, confidence interval; CUAL, combined unique active lesions; EDSS, Expanded Disability Status Scale (scale 0-10; higher is poorer outcome); FSIQ-RMS, Fatigue Symptoms and Impacts Questionnaire: Relapsing Multiple Sclerosis; HR, hazard ratio; IQR, interquartile range; ITT, intention-to-treat; LS, least squared; MD, mean difference; mFIS, Modified fatigue impact scale (scale 0 – 84, higher is poorer outcome); MSFC, multiple sclerosis functional composite measure; NR, not reported, NEDA, no evidence of disease activity; OR, odds ratio; RR, relative risk; SD, standard deviation; SF-36, short-form-36 health survey; TEAE, treatment-emergent adverse event

Source: CS, Document B and the trial CSR[19]; All-cause discontinuation data is from the company's clarification response

[#] Figures reported are for populations who received 20mg throughout the trial. $from baseline of OPTIMUM through to end of follow-up period ¥Adjusted for EDSS strata (≤3.5 vs >3.5), DMT in 2 years prior to trial, and number of relapses in year prior to trial (≤1 vs ≥2). [ǂ]Effects for ARR are reported in the per protocol population.

Subgroup analyses

The company reported subgroup analyses of the OPTIMUM trial data for HA participants (pre-planned definition that included participants with RES; [redacted]), HA participants according to the NICE definition (post-hoc analysis excluding RES participants; [redacted]), for the RES population (post-hoc analysis; n=[redacted]), and for the ITT population excluding participants with SPMS (pre-planned analysis; [redacted]). Few outcomes were reported for each of the subgroup analyses, and as 95% confidence intervals were proportionally wider for each analysis, it was difficult to draw conclusions about whether population was an effect modifier. As to be expected, the absolute rates of relapse and disability were [redacted] in both arms of OPTIMUM as compared to the ITT population, at follow-up, [redacted]

In general, relative treatment effects are stable across baseline risk, and clinical advisors to the ERG were unaware of any reason why treatment efficacy would vary across the difference RRMS subgroups. A comparison with the results for the ITT population showed [redacted] for both CDA-3 and CDA-6 as compared to teriflunomide in the HA and RES groups; [redacted] The ERG identified evidence from the CSR of OPTIMUM(OPTIMUM trial CSR[19]) [redacted]

- CDA-3: [redacted]
- ARR: [redacted]

Unsurprisingly, [redacted]. Population subgroup analyses were not reported for the Phase 2 trial, due to there being a lack of statistical power.

Table 16: Population subgroup analyses from OPTIMUM (HA, HA excluding RES, RES, and ITT excluding SPMS)

Outcome	Measurement	OPTIMUM HA		OPTIMUM HA (NICE definition)		OPTIMUM RES		OPTIMUM ITT excluding SPMS	
Treatment		Ponesimod	Teriflunomide	Ponesimod	Teriflunomide	Ponesimod	Teriflunomide	Ponesimod	Teriflunomide
ITT sample		202	200	177	172	34	40	552	552
Relapse	ARR (mean, 95%cl)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ARR (rate ratio, 95%cl)	[redacted]		[redacted]		[redacted]		[redacted]	
3-month CDA	Rate (%)	[redacted]	[redacted]	NR	NR	[redacted]	[redacted]	[redacted]	[redacted]
	HR	[redacted]		[redacted]		[redacted]		[redacted]	
6-month CDA	Rate (%)	[redacted]	[redacted]	NR	NR	[redacted]	[redacted]	[redacted]	[redacted]
	Risk reduction (95%cl)	[redacted]		NR		[redacted]		[redacted]	
CUALs	Mean	[redacted]	[redacted]	NR	NR	NR	NR	[redacted]	[redacted]
	RR	[redacted]		NR	NR	NR	NR	NR	
Fatigue	FSIQ-RMS LS mean Δ from baseline	[redacted]	[redacted]	NR	NR	NR	NR	[redacted]	[redacted]
	LS MD (95%cl)	[redacted]		NR	NR	NR	NR	[redacted]	

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; CI, confidence interval; CUAL. combined unique active lesions; FSIQ-RMS, Fatigue Symptoms and Impacts Questionnaire: Relapsing Multiple Sclerosis; HA, highly active; HR, hazard ratio; ITT, intention-to-treat; LS, least squared; MD, mean difference; NICE, National Institute for Health and Care Excellence; RES, rapidly evolving severe; RR, relative risk; SPMS, secondary progressive multiple sclerosis

Adverse effects

The company reported direct safety evidence for ponesimod from OPTIMUM and the Phase 2 core study, as well as a long-term safety set pooling evidence from all participants receiving ponesimod during OPTIMUM, its extension (OPTIMUM-LT), the Phase 2 trial, or its extension. Safety evidence from a sample of all randomised participants in the OPTIMUM trial who received a dose of either ponesimod 20 mg or teriflunomide 14 mg resulted in a comparative safety set of 1,131 participants. Only two participants who should have, but did not, receive ponesimod 20 mg were excluded from this analysis. No separate comparison of AEs was reported for different population subgroups.

Results provided by the company for overall treatment-emergent adverse events (TEAEs) in OPTIMUM are presented in Table 17, and showed that the vast majority of participants experienced one or more TEAE.

Table 17: Participants with at least one treatment-emergent adverse event in the OPTIMUM trial

Person with at least one:	Ponesimod 20 mg n=565 (%)	Teriflunomide 14 mg n=566 (%)
AE	502 (88.8)	499 (88.2)
Severe AE	[redacted]	[redacted]
AE leading to study discontinuation	49 (8.7)	34 (6.0)
Serious AE	49 (8.7)	46 (8.1)

Abbreviations: AE, adverse events

The company reported similar overall TEAEs in the ponesimod 20 mg (88.8%) and teriflunomide 14 mg (88.2%) groups, though higher rates of treatment discontinuations due to adverse events (AEs) were observed in the ponesimod group (8.7%) when compared to the teriflunomide group (6.0%). The proportion of participants with serious TEAEs are similar across treatment groups: 8.7% of participants in the ponesimod group and 8.1% of participants in the teriflunomide group experienced a serious TEAE; though no TEAEs in either group were fatal. Two fatalities occurred in the teriflunomide group but were considered unrelated to teriflunomide by the study investigator; no fatalities occurred in people treated with ponesimod in the OPTIMUM trial. Clinical advice to the ERG suggested that this rate of TEAE is broadly consistent with other DMTs, though noted that the sample size and length of follow-up in the trials may not yet have identified rare serious side effects.

The company reported the rates of AEs experienced by ≥5% of people in the CS (CS appendices, page 185); the ERG has summarised TEAEs of particular interest in Table 18.

Table 18: Incidence of key treatment-emergent adverse events in the OPTIMUM trial

Safety set	Ponesimod 20 mg n=565 (%)	Teriflunomide 14 mg n=566 (%)
People with ≥1 TEAE, n (%)	502 (88.8)	499 (88.2)
Infections[a]	[redacted]	[redacted]
ALT increased	[redacted]	[redacted]
AST increased	[redacted]	[redacted]
Nasopharyngitis	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]
UTI	[redacted]	[redacted]

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; TEAE, treatment-emergent adverse event; UTI, urinary tract infection

Note:
a Composite number of people with infections comprising nasopharyngitis, upper respiratory tract infection, urinary tract infection; total number of all infections may therefore be greater

From these data, the ERG noted that hepatobiliary disorders and liver test abnormalities occurred more frequently in the ponesimod arm, but a lower proportion were serious as compared to the teriflunomide arm. It was unclear from the data presented by the company whether ponesimod posed a higher risk for cardiac disorders when compared to teriflunomide over the course of treatment, but the evidence indicated that ponesimod may lead to an increased risk of cardiovascular effects initially than teriflunomide.
[redacted] of TEAEs related to [redacted] ([redacted] versus [redacted] in the ponesimod and teriflunomide groups, respectively), before becoming more comparable over the full course of the study ([redacted] in the ponesimod and [redacted] in the teriflunomide group). However conversely,
[redacted].

Ponesimod was also associated with
[redacted]

[redacted]).

[redacted]

The paper by Olsson et al. (2014)[22] reported the safety results of the Phase 2 core study. The results were similar to those reported for OPTIMUM, though marginally smaller proportions of participants experienced TEAEs and liver abnormalities when compared to participants who received ponesimod 20 mg in the OPTIMUM trial. Lower occurrences would be expected due to the shorter follow-up of 24 weeks (compared to 108 in OPTIMUM); though the similarity in the proportions suggested the possibility that most TEAEs with ponesimod have an early onset. All AEs related to heart rate and rhythm were also reported as occurring on Day 1 of treatment. No fatalities were reported in the ponesimod 20 mg group, or any other trial arms. The ERG summarised key TEAEs from the Phase 2 core study in Table 19.

Table 19: Key treatment-emergent adverse events in the Phase 2b core trial

Event	Ponesimod 20 mg n=114 (%)	Placebo n=121 (%)
People with ≥ 1 TEAE, n (%)	88 (77.2)	90 (74.4)
Infections[a]	36 (31.6)	47 (38.8)
Bronchitis	4 (3.5)	2 (1.7)
Gastroenteritis	3 (2.6)	4 (3.3)
Influenza	3 (2.6)	2 (1.7)
Nasopharyngitis	11 (9.6)	17 (14.0)
Sinusitis	5 (4.4)	5 (4.1)
Upper respiratory tract infection	9 (7.9)	11 (9.1)
UTI	1 (0.9)	6 (5.0)
ALT increased	7 (6.1)	1 (0.8)
AST increased	-	-

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; TEAE, treatment-emergent adverse event; UTI, urinary tract infection

Notes: [a] Composite number of people with infections comprising bronchitis, gastroenteritis, influenza, nasopharyngitis, sinusitis, upper respiratory tract infection, urinary tract infection; total number of all infections may therefore be different

The company also reported safety evidence from a long-term pooled safety analysis, which included all [redacted] participants who received ponesimod 20 mg in either the OPTIMUM or Phase 2 trial (representing [redacted] patient-years of exposure from the Phase 2 and [redacted] patient-years of exposure from OPTIMUM, with data cut-off for both extensions at [redacted]).

These long-term data showed similar proportions of people with at least one TEAE, treatment discontinuation due to TEAE and at least one serious TEAE when compared to the ponesimod group in OPTIMUM. The proportion of participants with elevated ALT and AST levels were [redacted], respectively) than in OPTIMUM; [redacted] (Table 25, pp.93-94 of the CS) in the pooled set. The company did not report proportions of participants experiencing effects on heart rate and rhythm, or macular oedema, but Table 25, pp.93-94 in the CS showed that [redacted] and [redacted], in participants on ponesimod. [redacted] were reported as part of the long-term safety analysis.

3.2.5. [redacted] Quality assessment of the included trials

The company used two different quality appraisal tools to appraise the quality of the OPTIMUM trial (CRD tool, CS Document B p.53-54 and Cochrane Risk of Bias Tool v.1, CS appendices p.160-161), whereas the Phase 2 trial was evaluated by the Cochrane Risk of Bias Tool v.1 tool only. While both tools are acceptable for evaluating risk of bias in RCTs, in 2019 the Cochrane tool was updated and would have been a preferable tool. For the Cochrane tool, the company evaluated an additional domain under the 'other' category of the tool, which they titled 'balance of dropouts and baseline traits'. No explanation of this domain was provided, and the ERG was unclear whether this double counted for differential attrition already covered within the attrition domain of the Cochrane tool, or assessed something different.

The company appraised the core phases of both trials to be at low risk of bias; this assessment was made at the trial level, with no differential ratings given across outcomes. The ERG agreed with the assessments made by the company according to the domains of the tools used, though noted that outcome measurement in both trials was subject to some limitations. The clinical outcomes of the trials may be subject to some measurement error, and the short-term evaluation of outcomes may not provide a reliable measure of changes in disability. In particular, clinical advice to the ERG was that CDA-3 may be likely to over-estimate disability due to natural fluctuations in the condition, and therefore CDA-6 is a more reliable measure (see Key Issue 4). Clinical advice to the ERG was also that the samples of both trials are likely too

small to identify rare serious adverse events associated with treatment. These issues were expected to apply equally to both arms.

No quality assessment or commentary about risk of bias was provided for the long-term extensions of either trial. The ERG considered both to be at a high risk of bias. The extension to OPTIMUM was uncontrolled, meaning that it is not possible to determine to what extent clinical outcomes were determined by treatment or by natural changes in the disease course or chance adverse events. It was also open-label, meaning that all outcomes that required a degree of subjectivity in measurement (particularly relapse rate, CDA, and PROs, but to some extent also neurological/radiological outcomes) are at a high risk of bias. All arms of the Phase 2 extension received ponesimod, and therefore comparisons can be made between doses of ponesimod only. While the different doses were blinded to participants, all were nevertheless aware that they were receiving an active treatment.

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

3.3.1. Search strategy

A single search strategy was used to identify RCTs evaluating the efficacy and safety of ponesimod and comparators for RRMS for the company submission; the methods are described in Section 3.1.

3.3.2. Feasibility assessment

The company did not clearly state whether they conducted a feasibility assessment to inform the analyses for this appraisal. It is therefore not possible for the ERG to evaluate the scope of any assessment, and appraise the rigour and rationale of decision-making for the company's NMAs. The company did report that several outcomes they considered were not "feasible". At clarification, the company reported that the choice of outcomes was based on the outcomes needed to populate the economic model, however it's unclear to the ERG why the company did not conduct NMAs for relative safety (discontinuation due to adverse events) or HRQoL, which could have informed both the clinical and economic evaluations of ponesimod. The ERG further noted that some analyses were stated to have been conducted but the results not reported in the CS, and so overall there was a lack of certainty about the analyses planned, conducted, and found not to be feasible.

The company stated that analyses restricted to the active RRMS population were not possible, due to the lack of available comparator data. Therefore, the company base case analyses are conducted with the ITT populations of the included trials. Evidence in the HA population is still more sparse, and the company reported that data from the ITT population were needed to complete the networks for the HA population, and that an NMA evaluating all-cause discontinuation was not feasible in this population. Without a comprehensive report of any feasibility assessment, it is unclear when company decisions to 'flex' inclusion criteria were deemed appropriate to complete networks, and when not. All networks were unadjusted for effect modifiers, and it is unclear whether the company explored this as an option but found that it was not feasible.

Tables presenting limited details about the included studies were provided, though the ERG considered that these did not fully reflect key factors that may create heterogeneity in the network. The ERG was aware that the evidence base for treatments of RRMS is highly heterogeneous, in study design, population characteristics/definitions, intervention delivery, and outcome follow-up and measurement. While to some extent these issues are unavoidable for these appraisals, a rigorous and transparent feasibility assessment would nevertheless have added trust to the analyses.

3.3.3. Study selection criteria

The selection criteria used by the company are described in the CS appendices, with a summary presented in Table 2 of Section D.3 (p.14-16). The ERG considered the selection criteria used by the company to be broadly appropriate.

As stated in Section 3.3.2, the company stated that it was not feasible to conduct analyses only in the active RRMS population, which would have been most pertinent to the decision problem. Instead, the company base case analyses were conducted in the ITT populations of the included trials, provided that at least 80% of trial samples should be people with RRMS (an arbitrary threshold based on IQWiG guidance). The ERG considered this to be a reasonable, pragmatic approach. The company further conducted subgroup NMAs using the OPTIMUM-definition of HA, which includes a small proportion of RES participants. In general, relative treatment effects are stable across baseline disease severity, though the ERG was unclear if this had been established in the RRMS population. Furthermore, the ERG was aware that different treatment recommendations are used in the NHS for people with differing RRMS

disease severity. Accordingly, the ERG considered that the generalisability of evidence across people with different disease severity was unclear.

Other selection criteria were judged to be appropriate, or to likely have minimal impact on the effect estimates. Notably, interventions included in the analysis were restricted to those recommended for each population (active and HA), and at licensed doses used in the NHS, which the ERG accepted.

The company chose to exclude phase 4 trials, which the ERG did not consider appropriate, since any problematic variation in methods between trials (the company's given rationale) could be more appropriately managed through more specific exclusion criteria. These criteria led to the exclusion of several trials that the ERG considered should have been included in the company's analyses; however, a comparison of treatment effects between the company's NMAs and those previously published that contained the excluded studies did not demonstrate major differences in reported effects, and therefore the ERG did not consider this to be a major concern for the analyses.

The company implemented several exclusion criteria following the completion of screening, which is generally considered to be a risk of bias. However, the ERG considered all the criteria implemented (e.g. excluding trials with fewer than 10 people in any treatment arm, and trials with zero events) were reasonable.

3.3.4. Included studies

The ERG found the flow of studies identified for the NMAs to be unclearly reported in the CS, and the descriptions contained some discrepancies in numbers; however, this lack of clarity was aided by information provided by the company at clarification. Following the inclusion of evidence for ofatumumab, the company reported that 41 RCTs were identified for inclusion in the ITT analyses, and 12 RCTs were included in the HA analyses. 42 trials reported discontinuation in the ITT population. However less than half of the trials reported CDA (three month CDA n=22; six-month CDA n=20 [note that all trials reporting six-month CDA also reported three-month CDA]).

The majority of trials were placebo-controlled (n=26), though 15 trials included a head-to-head comparison (not including trials that compared different doses of the same treatment). Included RCTs for each of the comparator treatments were as follows: beta-interferons n=18; glatiramer acetate n=9; fingolimod n=5; teriflunomide n=5; ozanimod n=3; dimethyl fumarate n=4;

alemtuzumab n=3; ocrelizumab n=3; natalizumab n=2; ponesimod n=2; peginterferon beta-1a n=1; cladribine n=1, and placebo n=26). The trials included 4 extensions to other included trials[36-39].

Enrollment periods for the included trials ranged from 1993 to 2020 (as reported in table 6 of the company's clarification response; question A4). The trials were conducted across a range of different geographic areas and healthcare settings. Most trials were conducted across multiple countries (n=33), with other trials conducted in the US (n=3), Japan (n=2), Iran (n=1), and Russia (n=1) and Italy (n=1). The median follow-up, based on the company's clarification response, was 96 weeks (range of 24-144 weeks).

Table 7 of the CS appendices (p. 128) reported the population eligibility criteria for the included studies (for ofatumumab these were reported in the company's clarification response). The table showed further variation in the diagnostic criteria and definition of active and highly active RRMS used within the trials. While this variation introduces some uncertainty into the analysis, clinical advice to the ERG was that these differences are unlikely to have a major impact on the comparability of the trials. Since the earliest trials, there have been various changes to the diagnostic criteria of RRMS, however clinical experts also considered that this is unlikely to undermine the analysis; the changes to diagnosis may have led to earlier diagnosis of RRMS, though the most impact will be for people not eligible for DMTs.

3.3.5. Quality assessment of studies included in indirect treatment comparison

The company reported using the Cochrane risk of bias tool (version 1) to assess the quality of trials included in the ITC. The ERG noted that the domains used in the assessments were appropriate for Cochrane risk of bias. The judgements are summarised in a colour-coded table in the appendices to the CS (Appendix D.7). Overall, the company reported that studies included in the NMA were generally at low risk of selection, attrition and reporting bias, with greater variability reported with regards to performance bias and other bias. The company did not, however, provide justifications for their quality judgments. This made it difficult to assess whether these judgments were reasonable, in particular for the composite 'other bias' domain, described as both a balance of baseline characteristics and drop-outs. It was also not stated whether these were done independently in duplicate, making it difficult for the ERG to assess whether these judgments were unbiased.

Within the timeframe of this appraisal it was not feasible for the ERG to independently assess the risk of bias for all trials included in the ITC. However, the ERG compared the judgments in the company submission with those reported in other NICE RRMS appraisals, finding that there was a good level of agreement.

In general, several trials included in the NMA had some uncertainty around selection bias, but few of these had issues around the balance of baseline characteristics; indicating few trials with serious problems regarding randomisation or allocation concealment. A considerable number of included trials were at high risk of performance bias, and less posed a risk of detection bias. Given the nature of the outcomes, which requires the individual's involvement in identifying relapses and disability, it is difficult to assess the impact of these biases on trial results. The ERG noted that very few trials had issues related to attrition or reporting bias, but nearly half of the included trials had high risk related to imbalances in baseline characteristics and/or attrition.

3.4. Critique of the indirect comparison and/or multiple treatment comparison

The following sections contain the ERG's appraisal of the company's NMA methods and results. Overall, the ERG considered that the choice of analyses could have been more comprehensive towards the decision problem; for example, analyses comparing treatment discontinuation due to adverse events and HRQoL would have been informative, as well as further analyses in populations specific to the NHS treatment pathway. However, it is possible that further analyses were not feasible, due to a paucity of evidence across other comparisons. The ERG identified a number of limitations with the NMAs, particularly for the analyses conducted in the HA population, which significantly undermine the validity of the results. These limitations were generally due to the paucity and quality of evidence for ponesimod and comparator treatments, and not because of the company's methods for selecting and analysing evidence.

3.4.1. Summary of analyses undertaken

The ERG was unclear how many NMAs the company conducted in total, though this included eight NMAs in the ≥80% RRMS population (random- and fixed-effects models of ARR, CDA-3, CDA-6, and all-cause discontinuation); six in the HA population (random- and fixed-effects models of ARR, CDA-3, CDA-6) and three in the RRMS only population (ARR, CDA-3, CDA-6). The company also stated that additional NMAs were conducted to explore the impact of informed priors (CS Document B p.70) and to replace HA subgroup data for two teriflunomide trials with the ITT data (CS appendix p.148); however, it was not clear which outcomes were

subjected to these sensitivity analyses, and while model fit statistics were reported for one of the analyses using informed priors, the priors used and the remaining results were not reported. The CS appendix also reported the results of an NMA of effect estimates for trials with long-term follow-up of ARR, which at clarification the company stated included trials with comparative follow-up data beyond the core trial period

The NMAs were conducted using a Bayesian framework, based on a Markov Chain Monte Carlo simulation. Consistent with NICE Decision Support Unit (DSU) guidance, vague prior distributions assuming no pre-existing information on the values of treatment effects, trial baselines, and common regression terms were used in the base case analyses. Model fit was assessed using the residual deviance (ResDev), deviance information criterion (DIC), and estimated between-study SD. The posterior mean deviance (of individual data points for ARR and treatment discontinuation and individual studies for three- and six-month CDA) was used to investigate consistency. The company did not report estimates separately for direct and indirect evidence, and did not comment on consistency of the networks. The company also did not state how heterogeneity would be evaluated: between-study SD was stated to inform model selection, though it was not stated if this would be used to investigate heterogeneity, and no further measures (e.g. I^2, Cochran's Q, chi-square) were reported.

For ARR, the company used a Poisson model with log link to generate relative rates, while HRs were derived for three- and six-month CDA using log HRs and a Normal model with identity link. A binomial model with logit link was used to calculate ORs for all-cause treatment discontinuations. The analyses were conducted in R and JAGS, and the full code used was provided in the CS appendix for the main (fixed- and random-effects) analyses (Section D5). The code was consistent with the analyses described, and appeared to contain no errors. The company stated that they calculated the probability of being best, the probability that ponesimod is better than other interventions in the network, and the Surface Under the Cumulative Ranking curve (SUCRA); however, only the 'rank' of ponesimod against other treatments based on SUCRA was reported, and this data was not accompanied by confidence intervals: this is a limitation of the analysis, given that ranking data such as SUCRA are very sensitive to uncertainty in the relative treatment effects, which is a concern for the analyses in this submission.

All analyses were unadjusted for covariates, and at clarification the company confirmed that only unadjusted rates were used from the included trials. Previous NMAs in this field have also

selected unadjusted rates, due to variation in the covariates used to adjust treatment effects across trials

The company conducted both random- and fixed-effects models, and reported the findings of both in the CS. The company selected fixed-effects models for ARR and three- and six-month CDA for the ≥80% RRMS population on the basis that the DIC criterion suggested a better fit to the data. The ERG considered that DIC is an estimate of model fit rather than of heterogeneity in the network, and therefore did not agree with the rationale for selecting fixed effects models on this basis. Rather, in recognition of the high degree of heterogeneity in the studies included in the network, the ERG considered that a random-effects approach should have been taken for all analyses. The principal difference between random- and fixed- effects models were the certainty of the effect estimates, and some of the differences reported between treatments were no longer statistically significant when using the random effects analyses.

3.4.2. Critique of assumptions used in the indirect treatment comparison

The company's analyses proceeded despite known heterogeneity in the evidence base. At clarification, the company outlined their approach to selecting the effect estimates from the included trials; all of these decisions appeared reasonable, though they demonstrated the complexity of an evidence base characterised by varying population definitions, trials conducted in different international healthcare settings across a span of decades, and where disease outcome measures are not standard and involve some measurement subjectivity/error. The impact of this heterogeneity was evident in the wide variation of placebo effects: the input data used for the company's NMAs, provided at clarification, showed that ARR ranged from 0.18 to 1.73 (n=26; for context, the ERG noted that the differences in ARR between ponesimod and comparator treatments were all <0.1), and the rate of treatment discontinuation ranged from 0% to 62.8% (n=25), without this variation being explained from length of follow-up only. Due to the paucity of evidence for each comparison in the networks, it was not possible to fully evaluate the range in effects in the CDA networks and for other treatments.

The company used unadjusted effects from each of the included trials, which they stated was due to variation in the adjustments made within each trial, and the company did not calculate effects using meta-regression: in effect, therefore, the company have assumed homogeneity in the trial evidence, despite evidence that this is not the case. The ERG was aware that previous appraisals of treatments for RRMS have required the acceptance of heterogeneity in networks to generate indirect treatment effects, due to the lack of direct head-to-head evidence. In all

cases, concerns about the impact of this heterogeneity have been noted as significantly undermining the validity of the treatment effects due to the differential effects of known or unknown effect modifiers (Klawiter 2009[40]; Jansen 2011[41]).

Finally, the ERG considered it a limitation of the company's analyses that the analyses do not represent the line and sequencing of treatments that would be expected in practice: all treatments available within each population are compared, no matter the line they would be received in practice. As in practice people would not be 'at risk' of every treatment, this undermines the transivity assumption of the analyses (Rouse 2017[42]). Moreover, as participants in the included trials were treated at varying lines of treatment, it's unclear to what extent effects are generalisability to the target population.

3.4.3. Relevance to the target population

As described above, the company's analyses are pragmatic and do not fully represent the populations and treatment pathways present within the NHS. While analyses were restricted to treatments available within the NHS, the analyses involve a comparison of treatment effects across participants with varying disease severity and on various lines of treatment. There is a lack of evidence for treatment effect modifiers in RRMS, though it is known that treatment efficacy varies widely between individuals, and discontinuing treatment is dependent on previous treatment history, and several demographic, radiological and clinical characteristics[43]. It is therefore unclear to what extent the mixed evidence base in the company's NMAs is generalisable to the target UK population.

3.4.4. Results of the indirect treatment comparison

3.4.4.1. RRMS participants (trial ITT populations)

A summary of the results from the company's updated base case NMAs is provided in Table 20 and Table 21 below (updated from the CS to include ofatumumab).

In the company's base case analyses, 95% credible intervals around the effects comparing ponesimod and the other comparators were extremely wide for all outcomes, indicating a high degree of uncertainty in the true effects. This was particularly the case for the CDA outcomes and for all-cause treatment discontinuation. This was likely due in part to the distance between ponesimod and the other comparators in the network, as well as the paucity and heterogeneity of the evidence for all treatments. To aid interpretation, the ERG have used colouring in the table to highlight both statistically significant differences and large numerical differences (i.e.

outside thresholds of 0.80 – 1.25) that were not statistically significant. However, the ERG acknowledged that there is greater uncertainty in determining the latter of these, and that smaller differences may nevertheless be clinically meaningful. In addition, effects estimated for all comparators as compared to placebo are summarised in Table 21, where the effects are more precise due to the weight and proximity of evidence for placebo relative to all treatments.

The results suggested that ponesimod was [redacted] the risk of relapse in people with active RRMS than interferon beta 1-a (all doses), interferon beta 1-b, glatiramer acetate (all doses), and teriflunomide.

[redacted] Clinical advice to the ERG was also that these treatments are used less in clinical practice, due to a lack of clinical efficacy.

[redacted] Rank data suggested that ponesimod was [redacted] for ARR, three-month CDA, six-month CDA, and all-cause treatment discontinuations, respectively; as noted earlier, no confidence intervals around the ranks were reported.

Table 20: NMA outcomes for ponesimod vs. comparator in ≥80% RRMS population (company base case)

Comparator	Dose	ARR, Rate ratio (95% CrI)[a]	3-month CDA[a]	6-month CDA[a]	All-cause discontinuation[b]
interferon beta-1a	22SC TIW	[redacted]	[redacted]	[redacted]	[redacted]
	44SC TIW	[redacted]	[redacted]	[redacted]	[redacted]
	30 IM QW	[redacted]	[redacted]	[redacted]	[redacted]
glatiramer acetate	20QD	[redacted]	[redacted]	[redacted]	[redacted]
	40 TIW	[redacted]	[redacted]	[redacted]	[redacted]
peginterferon beta-1a		[redacted]	[redacted]	[redacted]	[redacted]

Comparator	Dose	ARR, Rate ratio (95% CrI)[a]	3-month CDA[a]	6-month CDA[a]	All-cause discontinuation[b]
ocrelizumab		[redacted]	[redacted]	[redacted]	[redacted]
interferon beta 1b		[redacted]	[redacted]	[redacted]	[redacted]
dimethyl fumarate		[redacted]	[redacted]	[redacted]	[redacted]
teriflunomide		[redacted]	[redacted]	[redacted]	[redacted]
alemtuzumab		[redacted]	[redacted]	[redacted]	[redacted]
Cladribine		[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod		[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod		[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab		[redacted]	[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; CrI, credible interval; IM, intramuscular; NMA, network meta-analysis; QD, once a day; QW, weekly; RRMS, relapsing-remitting multiple sclerosis; SC, subcutaneous; TIW, three times weekly

Notes: [a] fixed effects NMA; [b] random effects NMA. Darker coloured cells represent statistically significant differences: green cells are in favour of ponesimod, red cells are in favour of the comparator. Lighter shading is used to represent large numerical differences in outcome (≥0.80 – 1.25) that were not statistically significant.

Table 21: NMA outcomes for all treatments vs. placebo in ≥80% RRMS population (company base case)

	[redacted]	[redacted]	[redacted]	[redacted]
Ponesimod	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fumarate	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate 20	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22 µg	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1b	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]	[redacted]	[redacted]
Pegylated interferon beta-1a	[redacted]	[redacted]	[redacted]	[redacted]
Teriflunomide	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; HR, hazard ratio; NMA, network meta-analysis; OR, odds ratio; RRMS, relapsing-remitting multiple sclerosis

Notes: Darker coloured cells represent statistically significant differences: green cells are in favour of ponesimod, red cells are in favour of the comparator. Lighter shading is used to represent large numerical differences in outcome (≥0.80 – 1.25) that were not statistically significant

3.4.4.2. Highly active subgroup

An overview of the company results from the highly active networks is provided in Table 22 and Table 23 below. Networks evaluated ARR and 3- and 6-month CDA only; no analysis was conducted to evaluate relative effects for treatment discontinuation due to a lack of evidence for this outcome in the HA population.

Across the clinical outcomes, the data suggested that ponesimod performed better than interferon beta 1a and teriflunomide, although neither of these treatments are currently recommended for treating people with HA RRMS. There [redacted].

The results were comparable with those in the company's ≥80% RRMS base case analysis, although there was [redacted].

Table 22: NMA outcomes for ponesimod vs. comparator in the highly active population

Comparator	Dose	ARR, Rate ratio (95% CrI)[a]	3-month CDA[a]	6-month CDA[a]
interferon beta-1a	44SC TIW	[redacted]	[redacted]	[redacted]
	30 IM QW	[redacted]	[redacted]	[redacted]
Ocrelizumab		[redacted]	[redacted]	[redacted]
Teriflunomide		[redacted]	[redacted]	[redacted]
Alemtuzumab		[redacted]	[redacted]	[redacted]
Cladribine		[redacted]	[redacted]	[redacted]
Fingolimod		[redacted]	[redacted]	[redacted]
Ofatumumab		[redacted]	[redacted]	[redacted]
Ozanimod		[redacted]	[redacted]	[redacted]
Placebo		[redacted]	[redacted]	[redacted]

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; CrI, credible interval; NMA, network meta-analysis; QW, weekly; TIW, three times weekly

Notes: Darker coloured cells represent statistically significant differences: green cells are in favour of ponesimod, red cells are in favour of the comparator. Lighter shading is used to represent large numerical differences in outcome (≥0.80 – 1.25) that were not statistically significant

Table 23: NMA outcomes for all treatments vs. placebo in the highly active population

	ARR	CDA-3	CDA-6
Ponesimod	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22 mcg	[redacted]	0.62 (0.43 - 0.90)	0.62 (0.43 - 0.90)
Interferon beta-1a 30 mcg	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44 mcg	[redacted]	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]	[redacted]
Ofatumumab	[redacted]	[redacted]	[redacted]
Ozanimod	[redacted]	[redacted]	[redacted]
Teriflunomide	[redacted]	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; NMA, network meta-analysis

Notes: Darker coloured cells represent statistically significant differences: green cells are in favour of ponesimod, red cells are in favour of the comparator. Lighter shading is used to represent large numerical differences in outcome (≥0.80 – 1.25) that were not statistically significant

3.4.4.3. Additional sensitivity analyses

Additional sensitivity analyses reported by the company were random- (ARR, CDA-3, and CDA-6) and fixed- (all-cause treatment discontinuation) effects analyses, restricted inclusion to long-term follow-up data (definition not provided; ARR only), and inclusion of ITT data for the teriflunomide trials in the highly active population (CDA-3 and CDA-6). The analyses revealed little that was pertinent to the appraisal: partly because the analyses do not address the key uncertainties with the company's analyses, and partly because wide confidence intervals in all analyses meant that it was not possible to detect whether differences across analyses conveyed meaningful effect modifiers.

3.4.5. Conclusions on the indirect treatment comparison

The ERG appraised the company's methods for the NMAs as pragmatic and appropriate in context of the available evidence. The ERG considered that a broader range of outcomes, to include the relative safety and impact on HRQoL of ponesimod, would have been informative to the appraisal; though as the company did not report their feasibility assessment, it was unclear whether these outcomes were not considered or were not feasible. There was a paucity of evidence across treatments for RRMS that could be used to inform these analyses; many

parameters in the networks relied on one or two studies only, which is particularly problematic in RRMS where both the condition and the available trials are heterogeneous in nature. However, the ERG considered that the company should have presented further outcome data from their NMAs, in addition to further exploration of heterogeneity and inconsistency in the networks.

Overall, the ERG considered the company's base case analyses to suggest that ponesimod could be considered as a moderate efficacy treatment for active RRMS amongst the treatments available, in terms of relapse rate and CDA of 3- and 6-months. However, clinical advice to the ERG was that the treatments that ponesimod out-performed were [redacted] Overall, the company's NMAs were associated with a high degree of uncertainty: the ERG considered that the true magnitude of any treatment effects in the analyses were uncertain, due to major limitations in the available evidence base. Finally, the ERG did not consider the company to have presented evidence of the relative efficacy of ponesimod in the RES population.

3.5. Additional work on clinical effectiveness undertaken by the ERG

The ERG conducted additional work to validate the company's NMAs and address uncertainty in treatment effects, and to address gaps in the evidence base for the safety of ponesimod. Specifically, the ERG:

- Conducted additional literature searches to identify (a) previous NMAs conducted in RRMS, with a particular focus on people with HA RRMS and (b) additional evidence of the safety of ponesimod (Section 3.5.1).
- Compared the methods used in previous TA appraisals for defining HA RRMS and for comparing treatments for HA (Section 3.5.2.1).
- Validated the treatment effects in the company's NMAs by comparing these with previous TAs/published NMAs (Section 3.5.2.2).

- Appraised the adverse event rates for all comparators, as identified by the company's SLR of RCT data, to evaluate the comparative safety of ponesimod (Section 3.5.3).
- Conducted a naïve comparison of the safety of ponesimod with fingolimod using evidence from the NICE appraisal (TA254[14])(Section 3.5.4).

An overview of this work is provided in the following sections, with supplementary information in the appendices.

3.5.1. Additional searches

The ERG carried out some additional searches for multiple sclerosis NMAs in Medline and Embase from 2016 onwards (Appendix A) and found 1,044 papers. This was a partial (modified) update of the searches used in Melendez-Torres (2018)[44], limited to papers published in 2016 onwards. These searches informed additional work conducted by the ERG to validate the methods and results of the company's NMAs.

In addition the ERG carried out some additional searches for adverse effects for ponesimod in Medline and Embase (Appendix A) and found 148 papers. This search used the broad adverse effects expert search filter from Ovid (Adverse Effects - Medline – Broad[45]) without any study type filter, in order to find any additional (non RCT) papers reporting safety data. Safety evidence measured within clinical trials can lack external validity (e.g. due to restrictive population eligibility criteria, and treatment use that may not reflect real world use). The search was also translated into Embase using the equivalent Ovid search filter (Adverse Effects – Embase – Broad[45]). This search was used to inform additional work conducted by the ERG to evaluate the relative safety of ponesimod.

Within the timeframe of this appraisal, it was not possible for the ERG to fully appraise the results of this search; though a single reviewer screened the results: 30 papers were found eligible, of which 20 papers were related to the included Phase 2 and OPTIMUM trials and their extensions. The remaining records were safety studies in healthy volunteers, and therefore were outside the scope of this appraisal. Within the timeframe of this appraisal, the ERG were unable to consider these papers in detail, to identify whether the evidence could meaningfully impact on this appraisal. However, based on the results of the search, the ERG concluded that the company included all available safety evidence for ponesimod.

3.5.2. Validation of the company's NMAs

Within the timeframe of this appraisal, it was not feasible for the ERG to conduct a comprehensive review and comparison of methods and effect estimates across previous NMAs. However, in order to validate the findings of the company's NMA, particularly for the HA subgroup where there is a high degree of uncertainty in the estimates, the ERG sought to compare the company's NMAs with those previously conducted. To this end, the ERG screened and selected published NMAs from targeted searches (described in 3.5.1) to identify previous NMAs evaluating treatments for RRMS.

3.5.2.1. Comparison of methods

The previous NMAs conducted in the HA RRMS population identified by the ERG, and a brief overview of the included trials and methodology used, are provided in Table 58 in the appendix. As with the company's NMAs, all required a broad definition of HA, to account for the various definitions used in the available trials. These analyses also always required the inclusion of indirect evidence to complete the network; either from indirect populations and/or treatments. The analyses were all associated with more uncertainty than analyses in the RRMS population. Based on the evidence accessible to the ERG, the impact of the assumptions used in the analyses were not investigated, with the exception of a meta-regression conducted by the company for TA616 (NICE evaluation of cladribine), which adjusted treatment effects for baseline disease severity. Unfortunately, the results from this analysis were considered by the ERG to show that effects were also affected by additional effect modifiers, which undermined the validity of the results.

Previous NICE committees have accepted the variability in HA population definitions in NMAs presented by companies, and have further accepted the inclusion of indirect evidence to complete networks as pragmatic. However, it is clear that all NMAs in the HA population include highly heterogeneous data, with unknown impacts on effect estimates, which cannot easily be resolved through statistical techniques.

A review of previous appraisals highlighted ongoing uncertainty in whether effect estimates could be generalised between the active HA and RES populations. Notably, for TA699[13] the committee heard from clinical experts who proposed that definitions of HA and RES may not be used in practice, in favour of classifications based on relapse severity and line of treatment, and in this case the committee concluded that recommendations could be made for the HA and RES populations based on evidence from a broad RRMS population. Conversely, within TA10299[12]

the committee considered that they could not make a recommendation for ozanimod in the 2nd line population as the company had not presented evidence specific to these people in its submission. The ERG considered that these discrepancies in opinion may be inevitable in a disease where population definitions are not standardised, and where there is a lack of evidence for treatment effect modifiers.

Overall, the ERG concluded that the methods used by the company to evaluate the relative efficacy of treatments in the HA RRMS population were broadly consistent with previous appraisals, and pragmatic according to the available evidence. The uncertainties in this analysis were considered to be related to the quality of the available evidence, and the ERG considered it unlikely that these uncertainties could have been adequately resolved by the company in their submission.

3.5.2.2. Comparison of relative effects in the intention-to-treat versus the highly active populations

The NMA in the HA subgroup had very sparse data for all clinical effectiveness outcomes. To determine whether data from the base case in the ITT population could be used to form a more complete network in the HA population, the ERG compared the relative effects for clinical outcomes between these populations using the effects reported in the company NMA. Relative effects were extracted from the league tables presented in the appendices to the company submission for ARR (Figures 2 and 5), CDA at 3 months (Figures 6 and 9) and CDA at 6 months (Figures 10 and 13) and tabulated to enable a comparison. The data used in this comparison are summarised in Tables C1-3 of Appendix C.

The ERG found differences in effects between the HA and ITT populations across ARR, CDA at 3 months and CDA at 6 months. Thresholds of 0.1 and 0.2 were used to identify differences in relative effects, both of which are within the bounds found to have a meaningful impact on the ICER; higher ARRs and hazard ratios for CDA, both at 3 and 6 months, were more frequently observed in the ITT population. Differences in nominal significance between results were low for ARR and CDA at 3 months, and larger for CDA at 6 months. Results from the ITT population were more frequently significant in these cases. While these comparisons are not conclusive, due to the wide confidence intervals reported around the effects, the ERG considered there to be some uncertainty in the use of ITT data to complete networks in the HA subgroup, given the frequency of significant and less favourable findings in the ITT population.

The ERG conducted further comparative analyses to establish whether there is a differential treatment effect for DMTs in the ITT versus the HA populations. To do this, the ERG calculated the ratio of the relative effect in the HA group to the relative effect in the ITT group for all DMTs compared to placebo. These ratios are available in the far right columns of Tables C1-3 in Appendix C. Using the approach from Cochrane guidance[46] for interpreting the importance of relative measures, these ratios were classified as 'inappreciably' or 'appreciably' lower or higher in the HA group using the cut-offs of 0.75 and 1.25. Inappreciably higher or lower ratios were considered as a comparable effect of treatments in the two populations on the outcome of interest. A summary of these conclusions for ARR, CDA at 3 months and CDA at 6 months is presented in Table 22.

Table 24: ERG conclusions on the estimated relative efficacy of disease modifying treatments in the highly active population compared to the intention-to-treat population

DMT	ARR	CDA at 3 months	CDA at 6 months
Alemtuzumab	[redacted]	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]
IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]
IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]	[redacted]
Ponesimod	[redacted]	[redacted]	[redacted]
Teriflunomide	[redacted]	[redacted]	[redacted]

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; DMT, disease modifying treatment; ERG, Evidence Review Group; HA, highly active; IFNB, interferon beta

[redacted]

.

3.5.3. Trial adverse event rates for ponesimod and its comparators

The company reported the rates of specific AEs for comparator treatments to ponesimod using annualised safety data obtained from trials identified by their SLR. These rates were reported in

Tables 42 to 45 of the CS (Document B, p 122 - 125). Please note that safety data was not reported in this table for ofatumumab or ozanimod, as AE data for these comparators were not submitted by the company during clarification. As the company's NMAs did not include an indirect comparison of safety between ponesimod and comparator treatments, the ERG reviewed the reported rates to inform a judgement on the relative safety of ponesimod to other available treatments. Reported rates of key AEs, with potentially large implications for healthcare resource use and/or safety, are reported in Table 24, and rates of these AEs that were serious are reported in Table 25.

The ERG noted that these rates are subject to a high degree of uncertainty, as all are based on trial data, which lacks external validity for estimating the risk of AEs. In addition, the trials were highly heterogeneous, with variations in health setting and country, population eligibility criteria, definition and measurement of safety outcomes, and length of follow-up. The ERG therefore considered that the rates reported may be indicative of the comparative safety of ponesimod, but that they should be interpreted with caution. Using these data, a naïve comparison of adverse event rates between ponesimod and fingolimod is summarised in Section 3.5.3.1, and between ponesimod and all other comparators in Section 3.5.3.2.

Table 25: Incidence of key adverse events reported in trials of ponesimod and its comparators

Treatment	Elevated ALT	Elevated AST	Infections [a]	Non-fatal PML	Fatal PML
Ponesimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fumarate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22 µg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pegylated interferon beta-1a	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Teriflunomide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Treatment	Elevated ALT	Elevated AST	Infections [a]	Non-fatal PML	Fatal PML
Natalizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Best supportive care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; PML, progressive multifocal leukoencephalopathy

Note:

[a] Composite percentage of participants with infections comprising nasopharyngitis, upper respiratory tract infection, urinary tract infection; total number of all infections may therefore be greater

Source: CS Document B, p.122-123

Table 26: Proportions of key adverse events that were serious for ponesimod and its comparators

Treatment	Elevated ALT	Elevated AST	Infections [a]	Non-fatal PML [b]	Fatal PML [b]
Ponesimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fumarate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22 µg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1b	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pegylated interferon beta-1a	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Teriflunomide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Natalizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Best supportive care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; N/A, not applicable; PML, progressive multifocal leukoencephalopathy

Note:

[a] Composite percentage of participants with infections comprising nasopharyngitis, upper respiratory tract infection, urinary tract infection; total number of all infections may therefore be greater

[b] Due to its serious nature, all PML events are considered serious adverse events. Cells with N/A reflect DMTs with no reported incidence of PML and, therefore, no calculable proportions of serious PML

Source: CS Document B, p. 124-125

3.5.3.1. Naïve comparison of AE rates for ponesimod vs. fingolimod

The company posited that ponesimod may be a safer alternative to fingolimod, due to its increased specificity on the $S1P_1$ receptor. A comparison between the rates of AEs reported for ponesimod and fingolimod suggested [redacted]. The ERG noted that [redacted]. In the absence of larger participant samples in trials of ponesimod and considering the rarity of these events, however, it is uncertain whether [redacted].

The rates of cardiac events and macular oedema (both known AEs of S1P modulators) from the trials of comparator treatments were not reported, and therefore the ERG was unable to evaluate whether the risk of these events was lower for ponesimod as compared to fingolimod. To address this, the ERG conducted a naïve comparison between the ponesimod trials and the evidence base for fingolimod considered in its appraisal by NICE (see Section 3.5.4)

Overall, as the relative safety evidence for ponesimod and fingolimod relies on a naïve comparison between heterogeneous trials, the ERG considered that it was not possible to draw firm conclusions about whether ponesimod does present a reduced risk of AEs due to its increased specificity on the $S1P_1$ receptor. Clinical advisors to the ERG considered that, pending further safety evidence, the monitoring of people receiving ponesimod should be comparable to that used for fingolimod.

3.5.3.2. Naïve comparison of AE rates for ponesimod vs. other comparators

Ponesimod showed [redacted] compared to other comparators: [redacted]

[redacted].

[redacted].

[redacted] (discussed in Section 3.5.3.1).

Overall, as noted above, the ERG considered the relative safety data to be highly limited, and that conclusions about the relative safety of ponesimod should be interpreted with caution, due to the heterogeneity in the trial methods.

[redacted] however, the ERG maintained that a NMA comparing discontinuation due to AEs would further inform the relative safety of ponesimod. There were outstanding uncertainties [redacted] (such as PML), and further safety data exploring these outcomes would inform the appropriate positioning of ponesimod in the treatment pathway, as well as the frequency of monitoring.

3.5.4. Naïve comparison of macular oedema rates and treatment discontinuation due to adverse events between ponesimod and fingolimod

The ERG noted that the company did not provide any data on the risk of cardiac events, macular oedema or treatment discontinuation due to AEs for comparators to ponesimod. As cardiac events and macular oedema are considered important AEs related to S1P modulators, and treatment discontinuations are a useful marker of overall tolerability, the ERG conducted a naïve comparison of these outcomes for ponesimod versus fingolimod, using safety data from the OPTIMUM trial for ponesimod and from the NICE technology appraisal for fingolimod (TA254[14]). Fingolimod was prioritised for this comparison as it is in the same drug class as ponesimod (S1P modulators) but is thought to have a less specific action on S1P receptors than ponesimod. Ponesimod is, therefore, posited by the company to have an improved safety profile. A limitation of this comparison was that additional safety evidence for fingolimod has been published since its appraisal by NICE[47] in 2012, including evidence that has highlighted

concerns about liver toxicity[48]. It is therefore feasible that the data appraised by NICE does not present a full picture of other serious AEs. However, in the timeframe of this appraisal, the ERG was unable to review the full evidence base for fingolimod, and this comparison should therefore be considered indicative, but interpreted with caution.

The results showed that the rate of treatment discontinuations due to AEs [redacted] than in either of the trials of fingolimod included in the NICE appraisal: 2.2% to 3.1% of people discontinued due to AEs in the FREEDOMS and TRANSFORMS trials, respectively, [redacted] of people treated with ponesimod in the OPTIMUM trial.

No cardiac event data were reported in the NICE appraisal of fingolimod, and therefore the ERG was unable to comment on whether ponesimod is safer for these events. The risk of macular oedema was [redacted] the NICE appraisal of fingolimod (0.4%, as reported from the SmPC), though, the Phase 2 trial of ponesimod reported a higher rate (2/114, 1.8%)[22].

Based on the evidence reviewed by the ERG, [redacted]. This comparison is limited, and has the same limitations due to trial heterogeneity as the comparison of AE rates in Section 3.5.3 (and in the company's NMAs). However overall, the ERG did not consider that [redacted].

3.6. Conclusions of the clinical effectiveness section

The clinical evidence presented by the company suggested that there may be a place for ponesimod in the current treatment pathway for people with active RRMS: based on the evidence available, ponesimod demonstrated [redacted] The ERG also considered the shorter half-life of ponesimod and its use as an oral treatment as potential benefits to people with RRMS. However, the ERG considered that weaknesses in the collective evidence base meant that the magnitude of clinical benefits relative to other comparators were uncertain, and combined with the paucity of reliable comparative safety evidence, this created some uncertainty as to the most appropriate positioning of ponesimod in the current treatment

pathway. The uncertainty was most evident in the HA RRMS population, where uncertainty in clinical effects was greatest, and there was no relevant direct head-to-head comparison (as teriflunomide is not recommended for the treatment of HA RRMS).

[redacted]. It may be reasonable to consider ponesimod as an alternative to fingolimod, particularly if the increased specificity of ponesimod to the $S1P_1$ receptor results in an improved safety profile, as posited by the company. However, the ERG did not consider that the company had demonstrated this in the evidence provided. Finally, the ERG did not consider that sufficient evidence had been presented to consider ponesimod for the treatment of RES RRMS.

4. COST-EFFECTIVENESS

4.1. ERG comment on company's review of cost-effectiveness evidence

The company carried out a SLR, using a single search strategy, to identify existing cost-effectiveness evidence, HRQoL evidence, and cost and resource use evidence for ponesimod in multiple sclerosis. A summary of the ERG's critique of the methods implemented by the company to identify relevant evidence is presented in Table 27.

Table 27. Summary of ERG's critique of the methods implemented by the company to identify health economic evidence

Systematic review step	Section of CS in which methods are reported			ERG assessment of robustness of methods
	Cost-effectiveness evidence	HRQoL evidence	Cost and resource use evidence	
Searches	Appendix G	Appendix G	Appendix G	The same search strategy was used for all three searches and was an update of the searches for TA624. It only covered Nov 2018 to July 2020. The cost-effectiveness/HRQoL/Costs searches were carried out as one search. The strategy did not use a recognised search filter to identify relevant publications such as those by SIGN[21] or CADTH[49]. The search strategy did not include any search terms for siponimod, ozanimod, ofatumumab or ponesimod. Therefore few or no papers will have been identified for these interventions. In clarification the company agreed to carry out some additional searches for ofatumumab and the results were shared with the ERG. The ERG carried out additional searches for the additional technologies in Medline and Embase (Appendix A) and found 105 papers.

Systematic review step	Section of CS in which methods are reported			ERG assessment of robustness of methods
	Cost-effectiveness evidence	HRQoL evidence	Cost and resource use evidence	
Inclusion criteria	Appendix G.1.3	Appendix G.1.3	Appendix G.1.3	The inclusion criteria were appropriate.
Data extraction	Appendix G.1.4 and 1.5	Appendix G.1.4 and 1.5	Appendix G.1.4 and 1.5	Methods for screening and data extraction were clearly described, and were considered appropriate.
Quality appraisal	NA	NA	Appendix G.1.6, and I	Quality appraisal of economic evaluations was conducted using the Drummond[50] checklist, which was appropriate. The evidence submitted by the company was consistent with the NICE reference case.

Abbreviations: CADTH, Canadian Agency for Drugs and Technologies in Health; CS, company submission; ERG, Evidence Review Group; HRQoL, health-related quality-of-life; NA, not applicable; SIGN, Scottish Intercollegiate Guideline Network

4.2. Summary and critique of company's submitted economic evaluation by the ERG

4.2.1. NICE reference case checklist

Table 28: NICE reference case checklist

Attribute	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for participants or, when relevant, carers	QALYs were estimated for participants and carer disutilities were included in the company's base case.
Perspective on costs	NHS and PSS	NHS and PSS as appropriate.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The company submitted a cost utility analysis.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	A 50 years time horizon was used in the base case analysis. The ERG considered the base case time horizon to be appropriate.
Synthesis of evidence on health effects	Based on systematic review	For the active and HA RRMS populations, clinical effectiveness data pertaining to ARR, CDA and treatment discontinuation were based on

Attribute	Reference case	ERG comment on company's submission
		NMAs conducted by the company. Treatment efficacy in the economic model is based on the relative risk vs. natural history
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs were used as appropriate.
Source of data for measurement of health-related quality of life	Reported directly by participants and/or carers	Utility values were derived from published literature.(Orme 2007[51]) The ERG considered this to be an appropriate source, however for completeness an alternative source has been tested in a scenario analysis.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Dolan et al.[52] as appropriate.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	There were no equity concerns.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	NHS reference costs and PSSRU were used as appropriate. Resource use estimates were based on previous NICE MS appraisals including ocrelizumab (TA533[53]) and peginterferon beta 1a (TA624[54]).
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Costs and outcomes were discounted at 3.5% as appropriate.

Key: EQ-5D, EuroQol 5 dimension; HA, highly active; HRQoL: health-related quality of life; NHS, National Health Service; NMA, network meta-analysis; PSS, Pseronal Social Services; QALY: quality-adjusted life year; RRMS, relapsing-remitting multiple sclerosis; TA: technology appraisal

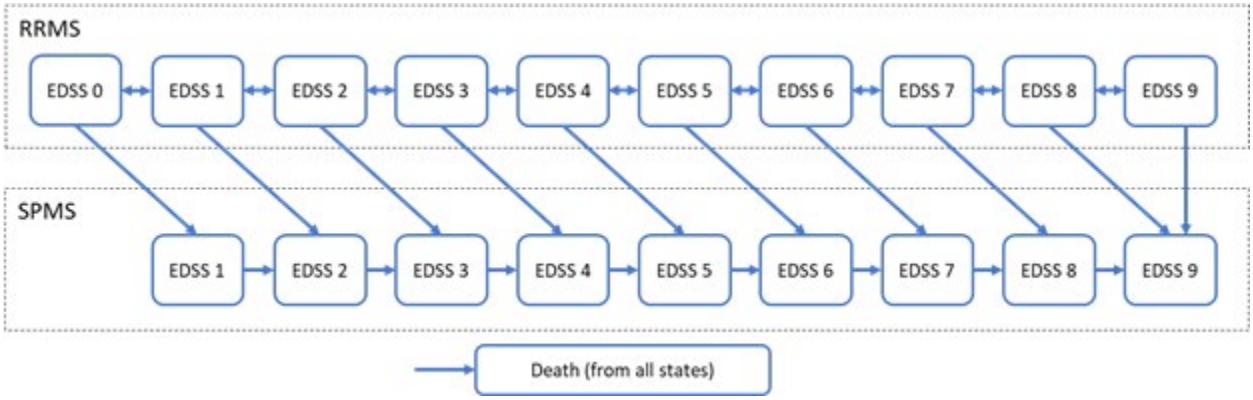
4.2.2. Model structure

The company submitted a Markov model consisting of 20 health states, based on EDSS scores (EDSS 0-9 for RRMS, EDSS 1-9 for SPMS and death, which was assumed to be equivalent to EDSS 10 for both RRMS and SPMS). People moved through EDSS health states based on treatment transition probabilities, which were derived from natural history data and adjusted to

account for treatment effect. See Section 4.2.6 for further detail surrounding the estimation of transition probabilities.

Whilst in the RRMS part of the model, people were capable of improving (moving to lower EDSS states) or getting worse (moving to higher EDSS states), upon progression into the SPMS part of the model, people were only able to move to higher EDSS states (see p.98 for further detail surrounding the probability of converting to SPMS and treatment discontinuation assumptions).The ERG acknowledged that the model structure was broadly in line with models used in previous NICE MS appraisals including fingolimod (TA254)[14], teriflunomide (TA303)[55], alemtuzumab (TA312)[2], dimethyl fumarate (TA320)[56], beta interferons and glatiramer acetate (TA527)[11] and peginterferon beta-1a (TA624)[54], which were based on 21 EDSS health states (previous models had included a EDSS 0 health state for SPMS). The company justified removing this on the basis that the conversion assumption (which assumes people who convert to SPMS move into an EDSS score of +1), had been used previously in a study by Mauskopf et al.[57] and previous NICE TAs including ocrelizumab (TA533)[53].

Figure 1: Model structure



4.2.3. Population

The company presented cost effectiveness results for two RRMS populations: the ITT population, which reflected the ITT population from OPTIMUM, used to represent people with active RRMS; and the HA RRMS population, which reflected the subgroup population of OPTIMUM, including people with highly active or RES RRMS (see Document B, p.110).

For the ITT population, people entered the model based on their baseline EDSS distribution in the OPTIMUM study. As outlined in Table 29, approximately [redacted] of participants had an EDSS

score of three or less, with [redacted] in EDSS 4 and 5. For the HA subgroup ([redacted]), comparatively fewer participants had a baseline EDSS score of three or less ([redacted]), whilst a higher proportion of participants had an EDSS score of 4 and 5 ([redacted]).

Table 29: Baseline EDSS distribution of participants within the economic model

	Baseline EDSS distribution (ITT population)	Baseline EDSS distribution (HA RRMS)
EDSS 0	[redacted]	[redacted]
EDSS 1	[redacted]	[redacted]
EDSS 2	[redacted]	[redacted]
EDSS 3	[redacted]	[redacted]
EDSS 4	[redacted]	[redacted]
EDSS 5	[redacted]	[redacted]

Abbreviations: EDSS, expanded disability status scale; HA, highly active; ITT, intention-to-treat; MS, multiple sclerosis; RRMS, relapsing remitting multiple sclerosis

The ERG noted that OPTIMUM was a global multi-centre study that included relatively few participants from the UK. However, based on clinical input to the ERG, participant characteristics from OPTIMUM for both the active and HA populations were considered to be broadly generalisable to the UK. Therefore the ERG considered these characteristics to be appropriate for use in the model. For completeness the ERG conducted a scenario analysis that used population characteristic data from the UK RSS dataset for the ITT population; however, while this population is based in the UK, it also included people with SPMS, and people who had a longer disease duration without access to DMTs, and may therefore not be highly generaliseable to the target population. It is worth noting that using UK RSS population data in the model did not have a material impact on the base case results (see Section 6.1.1.3).

Finally, in Document B, p107, the ERG noted that a relatively small proportion of participants in OPTIMUM had SPMS; i.e. [redacted]. Due to the small proportion of participants with SPMS, the ERG considered that the inclusion of this group was unlikely to be a key concern. In support of this, subgroup analyses from OPTIMUM removing these participants showed comparable findings to the ITT population.

4.2.4. Interventions and comparators

In the ITT population, the company initially compared ponesimod to teriflunomide, dimethyl fumarate, pegylated interferon beta-1a, glatiramer acetate, interferon beta-1a (22 mcg, 44 mcg),

interferon beta-1a (30mcg), interferon beta-1b and ocrelizumab. The company stated that the comparators were selected based on approved first line treatments for RRMS, as per the NHSE treatment algorithm. Based on clinical expert opinion to the ERG, the comparators appeared appropriate; however, two treatments (ofatumumab and ozanimod), which are currently under NICE review, were not included as part of the clinical or economic analyses. During the clarification stage, the ERG asked the company to update the NMA's and economic model to incorporate evidence for both treatments and provide updated results. This was subsequently provided, though the ERG acknowledged that it was unclear if these treatments would be recommended for the treatment of active RRMS.

For the HA RRMS subgroup, the company compared ponesimod to alemtuzumab, cladribine, fingolimod and ocrelizumab. Clinical input to the ERG confirmed that these treatments are widely used to treat people with HA RRMS in the UK (see Section 5.1.1.2 for results). At clarification the company also provided clinical and cost effectiveness analyses comparing ponesimod to ozanimod and ofatumumab in the HA RRMS population. Again, the ERG acknowledged that at the time of writing, it was unclear whether ozanimod and ofatumumab would be recommended by NICE for the treatment of HA RRMS.

4.2.5. Time horizon, perspective and discounting

A 50-year (lifetime) horizon was used in the company's base case. As MS is considered to be a progressive, lifelong condition the ERG considered that 50 years was sufficiently long enough to capture the differences in costs and effects between treatments. Furthermore, a 50-year time horizon has been used and accepted in previous MS submissions to NICE including fingolimod (TA254)[14], teriflunomide (TA303)[55], alemtuzumab (TA312)[2], dimethyl fumarate (TA320)[56], beta interferons and glatiramer acetate (TA527)[11], ocrelizumab (TA533)[53] and peginterferon beta-1a (TA624)[54]. Overall, the ERG considered the modelled time horizon to be reasonable.

The cycle length used in the model was one year. In the CS (Document B, p.108) the company stated that this was selected in order to be consistent with MS natural history data, as reported by Palace (2014)[25] and Mauskopf (2016)[57]. The ERG considered this justification to be reasonable and acknowledged the appropriateness of a 1 year cycle length in the model, but noted that the model did not allow for the cycle length to be varied.

There were no concerns surrounding discounting. Costs and benefits were discounted at 3.5% which reflects NICE guidance. All costs and outcomes were estimated from an NHS and PSS perspective, as appropriate.

4.2.6. Treatment effectiveness and extrapolation

4.2.6.1. Modelled treatment efficacy based on 3-month CDA

As noted in Section 4.2.3, people entered the model according to their OPTIMUM baseline EDSS score and moved through the EDSS health states via treatment-specific transition probabilities. Transition probabilities were estimated using clinical data from the company's NMA outlined in Tables Table 30 and Table 31 below; i.e. 3-month CDA hazard ratios vs. placebo were applied to natural history data from the British Columbia MS dataset[25] (see Document B, p112 for the transition matrix). Based on the results from the company's 3-month CDA NMA for the ITT population, ponesimod was associated with a lower risk of 3-month CDA than many of the other DMTs, with the exception of alemtuzumab, ofatumumab and ocrelizumab. For the HA subgroup, ponesimod was less effective for reducing the risk of 3-month CDA than cladribine, ofatumumab and ocrelizumab.

The ERG were uncertain why the company used three-month CDA as the primary outcome measure for disease progression, when six-month CDA estimates from the NMAs were also available for all but one comparator (interferon beta 1a SC22). The ERG opined that the six-month CDA was a more appropriate measure of disease progression on the basis of clinical advice, which noted that three-month CDA may potentially overestimate progression due to natural fluctuations in the disease. Furthermore, six-month CDA was considered as NICE's preferred measure of disease progression in previous MS TAs, including alemtuzumab (TA312)[2].

During the clarification stage the company was asked to comment on why six-month CDA was not used in the base case analysis to derive treatment effect estimates. The company commented that there were a larger number of closed loops in the three-month CDA and that it was considered more robust, stating that the six-month CDA was defined more frequently as a secondary outcome in the networks. Furthermore, the company noted that in the NICE appraisal of ocrelizumab (TA533)[53] and ofatumumab (TA699)[13], committee members had identified concerns surrounding the inclusion of the ADVANCE study for peginterferon beta-1a, as it produced clinically implausible six-month CDA results. The ERG acknowledged the potential limitations surrounding the six-month CDA highlighted by the company, noting that

heterogeneity and a lack of robust evidence is a significant cause of uncertainty across all of the company's NMAs. However on balance, the ERG still considered six-month CDA to be a more valid measure of disease progression, and preferred this outcome measure in the its preferred base case.

Clinical efficacy data used in the economic model

With respect to the key clinical efficacy data used in the company's economic model, the ERG considered the robustness of the NMAs to be a key area of concern (see Section 3.4.5). Clinical effectiveness estimates (based on 3-month CDA) used to derive transition probabilities, relapse rates and treatment discontinuation rates, were all associated with a high degree of uncertainty, and were surrounded by relatively wide confidence intervals. As a means of testing uncertainty surrounding modelled treatment effect estimates, the ERG conducted a further scenario analysis which derived DMT clinical effectiveness estimates by grouping treatments according to their positioning and using the median effect estimate to parameterise the model (see section 6.1.1.7 for further detail).

Natural history progression

In terms of the natural progression data used within the model, the ERG considered the British Columbia dataset used by the company to be an appropriate source for the active RRMS population (see CS Document B, p112 for the transition matrix). This Canadian observational study, which followed 898 people with RRMS and SPMS over 15 years, has also been accepted in previous NICE RRMS appraisals, including the appraisal of cladribine (TA493)[27], beta interferon and glatiramer acetate (TA527)[11], ocrelizumab (TA533)[53] and peg interferon beta 1a (TA624)[54]. The ERG was aware of an alternative natural history dataset (London Ontario), which could have been used to estimate base case transition probabilities for the active RRMS population; however previous NICE TAs, including teriflunomide (TA303)[55] and alemtuzumab (TA312)[2], have noted limitations in the use of this dataset, given that the study did not collect data on people whose disease had improved.

As a means of exploring the impact of using an alternative set of natural history transition probabilities in the model, the company conducted a scenario analysis using a combination of data from the placebo arm of DEFINE[58], a dimethyl fumarate trial, and the London Ontario dataset. Transition probabilities for EDSS states 0-7 were therefore derived from DEFINE whilst transitions between EDSS states 8-9 were taken from the London Ontario dataset. The

company stated that this analysis represented RRMS progression in a controlled environment, though the ERG were unclear on the company's rationale for selecting DEFINE for this analysis. As outlined in Document B, p171, results were not overly sensitive to this analysis. The ERG acknowledged that the scenario of using alternative natural history transition probabilities was useful, however considered the British Columbia dataset to be a better representation of real world disease progression.

For the HA RRMS subgroup, the natural history transition matrix was based on a previous NICE appraisal for ocrelizumab (TA533)[53], which reflected progression of participants in the placebo arm of the AFFIRM trial for natalizumab (for EDSS 0-6). For EDSS 7-9 the company used values from the British Columbia database (Document B, p119). Given that NICE had previously critiqued the use of the London Ontario data to model natural disease progression for the HA population in its appraisal of alemtuzumab (TA312)[2], the ERG considered the company's approach to be reasonable.

For people who progressed to SPMS, people were assumed to transition through health states based on the London Ontario dataset.

Table 30: Modelled CDA (ITT population)

Treatment	3 month CDA (hazard ratio vs. placebo)	6 month CDA (hazard ratio vs. placebo)
Ponesimod	[redacted]	[redacted]
Teriflunomide	[redacted]	[redacted]
Dimethyl fumarate	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]
Interferon beta-1a 22mcg	[redacted]	[redacted]
Interferon beta-1a 30mcg	[redacted]	[redacted]
Interferon beta-1a 44mcg	[redacted]	[redacted]
Interferon beta-1b 250mcg	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]
Peginterferon beta-1a 125mcg	[redacted]	[redacted]
Ofatumumab	[redacted]	[redacted]
Ozanimod	[redacted]	[redacted]

Abbreviations: CDA, confirmed disability accumulation; ITT, intention-to-treat

Table 31: Modelled CDA (HA RRMS group)

Treatment	3 month CDA (hazard ratio vs. placebo)	6 month CDA (hazard ratio vs. placebo)
Ponesimod	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]

Abbreviations: CDA, confirmed disability accumulation; HA, highly active; RRMS, relapsing-remitting multiple sclerosis

Annualised relapse rates

The company's model captured the impact of relapse associated with RRMS via the inclusion of annualised relapse rates for each treatment. When a person experienced a relapse, they incurred a utility decrement associated with relapse and incurred a specific relapse cost. See Section 4.2.7.1 and 4.2.8.2 for further detail on modelled disutility and cost per relapse).

For people with active RRMS, default annual relapse rates (or natural history rates) associated with each EDSS health state were derived from published literature (Mauskopf et al.[57]). Annualised relapse rates were then derived by applying treatment-specific rate ratios from the NMA to these natural history data (see Table 32 below). For the HA subgroup, the company derived average annual relapse rates from the placebo arm of the AFFIRM trial from natalizumab (TA127)[59]. The company stated that ARRs in the HA RRMS population are approximately 1.98 times higher compared with the ITT population. The ERG noted that the company did not provide rationale for selecting to use AFFIRM as a means of estimating annualised relapse rates for people with HA RRMS. As such there may be some uncertainty surrounding modelled ARR estimates for people with HA RRMS.

Overall, the ERG identified a number of limitations with the results of the company's NMAs, which increased uncertainty surrounding modelled relapse rates (see Section 3.4.5). The company conducted one-way sensitivity analyses that varied the rate ratio in relapse for DMTs, however this did not have a material impact on the results. Whilst the ERG acknowledged that relapse rates are not the key efficacy driver within the company's model, differences in relapse rates between treatments are expected to impact on the incremental costs and QALYs when varied.

Table 32: Relapse rates used in the company's model for ITT and HA RRMS

Treatment	Rate ratio for relapse vs. placebo (ITT population)	Rate ratio for relapse vs. placebo (HA RRMS)
Ponesimod	[redacted]	[redacted]
Teriflunomide	[redacted]	[redacted]
Dimethyl fumarate	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]
Interferon beta-1a 22mcg	[redacted]	[redacted]
Interferon beta-1a 30mcg	[redacted]	[redacted]
Interferon beta-1a 44mcg	[redacted]	[redacted]
Interferon beta-1b 250mcg	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]
Peginterferon beta-1a 125mcg	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]
Ofatumumab	[redacted]	[redacted]
Ozanimod	[redacted]	[redacted]

Abbreviations: RRMS, relapsing remitting multiple sclerosis

Progression from RRMS to SPMS

The modelled annual EDSS baseline probability of progressing from RRMS to SPMS was derived from a US study by Mauskopf (2016)[57], which estimated the cost effectiveness of delayed release dimethyl fumarate for the treatment of RRMS. Annual SPMS conversion probabilities were based on the London Ontario natural history study, which was considered to be an appropriate data source. Upon progressing to SPMS the company assumed that EDSS would increase by 1. Although the base case conversion rates used by the company were considered reasonable, the ERG noted that these were higher than those used in the submission for peginterferon (TA624)[54]. These values appear to have been derived from hazards presented in the appraisal of daclizumab (TA441)[60], which has recently had its marketing authorisation withdrawn. A comparison of these probabilities is provided in below.

Table 33 Annual probability of converting from RRMS to SPMS

EDSS state	Mauskopf et al.[57]	Peginterferon (TA624)[54]
EDSS 0	0.000	0.004
EDSS 1	0.003	0.002
EDSS 2	0.032	0.029

EDSS state	Mauskopf et al.[57]	Peginterferon (TA624)[54]
EDSS 3	0.117	0.097
EDSS 4	0.210	0.181
EDSS 5	0.299	0.225
EDSS 6	0.237	0.168
EDSS 7	0.254	0.211
EDSS 8	0.153	0.064
EDSS 9	1.000	0.154

Abbreviations: EDSS, expanded disability status scale; RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

The company explored uncertainty surrounding this parameter via probabilistic sensitivity analysis and did not conduct one way sensitivity or scenario analyses. As an exploratory analysis, the ERG conducted a scenario analysis using the SPMS annual conversion probabilities reported in pegylated interferon (TA624)[54]. The ERG noted that this scenario analysis had a material impact on the ITT analysis results, though not those in the HA RRMS subgroup. See Section 6.1.1.12 for results and further discussion.

Modelled treatment discontinuation rates

Within the model, people are capable of discontinuing treatment for the following three reasons

1. When a person's EDSS score equals or exceeds 7
2. When a person progresses from RRMS to SPMS
3. When a person discontinues prematurely for any reason (e.g. lack of efficacy, due to adverse events).

As outlined in CS Document B, p.118, discontinuation assumptions 1 and 2 above have been used in previous NICE TAs, including natalizumab (TA127)[59], fingolimod (TA254)[14], alemtuzumab (TA312)[2], cladribine (TA493)[27], ocrelizumab (TA533)[53] and peginterferon beta 1a (TA624)[54], and were considered appropriate. However, there was some uncertainty surrounding assumption 3, which involved estimating annual treatment discontinuation rates using odds ratios for all-cause discontinuation from the ITT population NMA (for ponesimod versus each comparator). To derive annual discontinuation rates, the relative risk of discontinuation for each treatment was then multiplied by the annual discontinuation rate for ponesimod (see annual discontinuation rates in Table 34 below). The annual discontinuation rate for ponesimod was calculated from pooled data from the OPTIMUM and the Phase 2 trial of ponesimod. As noted

previously in Sections 3.4.2 and 3.6, there is a high degree of uncertainty surrounding estimates derived from the company's NMAs. Furthermore, due to the lack of all-cause discontinuation data reported by trials for the HA RRMS population, no NMA was conducted for this outcome and the company assumed that discontinuation rates from the ITT population would be generalisable to the HA population and its relevant comparators. The ERG considered the lack of robust treatment discontinuation data for the HA population to be an area of uncertainty. More broadly, while the company's decision to use all-cause discontinuation in the model may have been pragmatic, the ERG noted that the definition included discontinuation due to trial attrition. Notably, in the OPTIMUM trial, less than half of trial discontinuations were due to efficacy or safety issues (see Table 15). It was therefore unclear to what extent this outcome could be generalised to clinical practice, and how variation in trial methodology and outcome measurement created heterogeneity in the evidence base.

Table 34. Modelled treatment discontinuation rates

Treatment	Annual discontinuation rates (%)
Ponesimod	[redacted]
Teriflunomide	[redacted]
Dimethyl fumarate	[redacted]
Glatiramer acetate	[redacted]
Interferon beta-1a 22mcg	[redacted]
Interferon beta-1a 30mcg	[redacted]
Interferon beta-1a 44mcg	[redacted]
Interferon beta-1b 250mcg	[redacted]
Ocrelizumab	[redacted]
Peginterferon beta-1a 125mcg	[redacted]
Alemtuzumab	[redacted]
Cladribine	[redacted]
Fingolimod	[redacted]
Ofatumumab	[redacted]
Ozanimod	[redacted]

To explore uncertainty in discontinuation rates, the company included an option in the model to apply a common discontinuation rate to all treatments (5%), for both the active RRMS and HA RRMS populations. Whilst this scenario was considered useful for determining the impact of discontinuation rates on the base case results, the ERG noted that in peginterferon beta 1a (TA624)[54], NICE preferred the use of treatment specific discontinuation rates. The ERG acknowledged that assuming a flat discontinuation rate of 5% for all treatments was simplistic

and may not reflect clinical practice, given that each treatment is associated with a specific adverse event and efficacy profile.

The ERG noted that the company's scenario analysis that applied a 5% discontinuation rate to all treatments resulted in increased total costs and QALYs for all treatments; however it did not have a material impact on the base case ICERs. For completeness, the ERG conducted a further scenario analysis that applied a 5% discontinuation rate to all treatments and incorporated the model changes made by the company during the clarification stage. See Section 6.1.1.5 for description and results.

Treatment waning assumptions

The ERG noted that there was a lack of long term clinical effectiveness data for ponesimod and comparator DMTs (input data for the NMAs were generally derived from endpoints under 3 years; range 24-144 weeks, median = 96 weeks). Therefore, there is uncertainty surrounding the maintenance of treatment effects for disease progression and relapse rates over time. In the base case analysis (for both the active RRMS and HA RRMS populations) the company applied the same treatment waning assumption to all DMTs; i.e. a 25% decrease in treatment efficacy was applied from years 2 to 5, followed by a 50% decrease in efficacy applied from year 6 onwards. The ERG noted that this assumption had previously been used in NICE appraisals of dimethyl fumarate (TA320)[56] and peginterferon beta 1a (TA624)[54]. In the appraisal of peginterferon beta 1a (TA624), the committee acknowledged that DMTs are likely to have different waning assumptions in practice, however in the absence of evidence, the same waning assumptions should be applied to all treatments.

The company explored uncertainty surrounding treatment efficacy waning by conducting two scenario analyses using alternative assumptions; i.e. no treatment waning and a further analysis which applied a 50% decrease in treatment effectiveness to all DMTs at 10 years. As outlined in (Document B, p171-174), results in both the ITT and HA populations were not considered sensitive to these scenarios. For completeness, the ERG conducted a scenario analysis which assumed 100% treatment efficacy for all DMTs; i.e. no waning over time. This analysis was based on the company's updated NMAs, which included ozanimod and ofatumumab, as well as alternative monitoring assumptions for ponesimod. See Section 6.1.1.6 for description and results.

Subsequent treatment assumptions

RRMS population

In the base case analysis, the company assumed that all people with active and HA RRMS people who stop treatment will go on to receive BSC. The company justified this approach on the basis that it allows the analysis to highlight differences in treatment effects for the initial phase of treatment and is consistent with previous appraisals including alemtuzumab (TA312)[2], ocrelizumab (TA533)[53], dimethyl fumarate (TA320)[56], beta interferons and glatiramer acetate (TA527)[11]. Although the ERG largely accepted the company's justification and acknowledged that there is precedent for using BSC as the primary treatment option post discontinuation, clinical advice to the ERG outlined that people are highly likely to receive a further DMT in practice (with the choice of subsequent DMT dependent on the rationale for discontinuation; e.g. lack of response or tolerability, or treatment break for pregnancy). The ERG considered conducting scenario analyses using assumptions for subsequent treatments suggested by clinical experts, however given that the choice and probability of subsequent treatment use will differ due to the reasons for discontinuing, the scenario was considered to introduce additional complexity and uncertainty. Furthermore, the ERG was unable to identify any prescribing data, which could inform subsequent treatment use in the model. As a result, the ERG accepted the company's base case assumption; however, acknowledged that it is unlikely to reflect clinical practice.

The company explored the impact of subsequent treatment use on the base case results via scenario analyses: the company assumed that 100% of the ITT population who discontinued went on to receive cladribine, whilst 100% of the HA population who discontinued went on to receive natalizumab (see Document B, p125-126 outlining the company's justification for selecting these as subsequent treatments). This is a simplifying approach, given that, as noted above, the choice of subsequent treatment will depend on the rationale for stopping treatment. It should be noted that, for these scenarios, the company included the clinical effectiveness of subsequent treatments (based on the NMA results). Due to the limitations surrounding the clinical effectiveness estimates, the ERG considered that modelling subsequent treatment effects introduced additional uncertainty.

As an exploratory analysis the ERG conducted a scenario using alternative subsequent treatments for both the ITT and HA RRMS populations. Both of these scenario analyses applied additional costs of subsequent treatments, but did not account for the clinical efficacy of these

treatments. These analyses were therefore considered to evaluate the impact of altered costs of subsequent treatment on the ICER, but would over-estimate rates of disease progression in those who switch treatment. See section 6.1.1.4 for further description and results.

SPMS group

The ERG noted that siponimod (TA656)[1] has been recommended by NICE for the treatments of people with SPMS, however in the model the company has not included siponimod as a treatment option for people who progress to SPMS; i.e. it is assumed that 100% of people who progress to SPMS will go on to receive BSC as the primary subsequent treatment option. Clinical experts to the ERG also noted that in practice a proportion of people with SPMS will receive treatment with interferon beta (IFNB)-1b. During the clarification stage, the company was asked to comment on why siponimod was excluded from the analysis and responded noting that their approach was consistent with NHS treatment guidelines and previous NICE appraisals (clarification question B4). The ERG confirmed that previous appraisals had not included siponimod or IFNB-1b as treatment for SPMS; however following its recent approval by NICE, clinical advisors to the ERG advised that a proportion of people with SPMS will start to routinely receive this. As siponimod is a new treatment, true rates of treatment uptake in the NHS are yet unknown, however clinical experts to the ERG advised between 12% - 50% of people may receive siponimod. As an exploratory analysis, the ERG conducted a scenario analysis that assumed a proportion of people who progress to SPMS receive siponimod. See Section 6.1.1.2 for description and results.

Mortality

All-cause mortality rates for people with RRMS were included for each EDSS health state, based on age and gender mortality risks for the UK, which were taken from UK life tables[61]. These underlying rates were then adjusted by applying a RRMS specific mortality relative risk to each health state using a linear interpolation approach as reported by Pokorski (1997)[62]l. In (Document B, p126), the company stated that there was a lack of data to inform differentiated mortality risk for each EDSS health state in people with SPMS (as compared to RRMS). Therefore, a simplifying assumption was made whereby people with RRMS and SPMS, in the same EDSS health state, were assumed to have the same relative risk of mortality (see Document B, p126 for EDSS mortality used in the model). Given the paucity of data surrounding SPMS mortality risk according to EDSS state, and the acceptance of this assumption previously

in peginterferon beta 1a (TA624)[54], the ERG considered the company's assumption to be reasonable.

The ERG noted that linear interpolation relative risks of mortality from Pokorski et al. were considered appropriate for use by the committee in the NICE appraisal of peginterferon beta 1a (TA624), as these values better reflected the mortality risk versus the general population as EDSS levels increase when compared to non interpolated values. The ERG noted that the company provided a scenario analysis that used raw mortality rates (without interpolation), however this did not have a material impact on the base case results.

4.2.7. Health-related quality of life

4.2.7.1. Baseline EDSS utility

For both the ITT and highly active populations, baseline EDSS utility values were derived from published literature Orme (2007)[51] (see Table 35 below). Orme et al. is a UK study that estimated the effect of disease, functional status and relapses on the utility of people with RRMS in the UK. Within the study, 12,968 people registered on the MS trust database were sent a postal survey, and utility was assessed using the EQ-5D (note only 15% of responses were used in the analysis due to low response rates). Utilities were estimated via an appropriate UK value set using the time trade off method from Dolan et al.[52]. The ERG acknowledged the strengths of Orme et al. as the primary source of EDSS utility; i.e. values were elicited directly from people with RRMS in the UK (or carers), however several key limitations were identified. The primary concern related to the generalisability of these participants to those within the OPTIMUM study. For instance, participants included in the Orme et al. study were older and had more severe disease at baseline compared to those in OPTIMUM. Mean age in Orme et al.[51] was 51.4 years and 59.6% were distributed across EDSS states 4 – 6 (compared to a mean age of [redacted] and [redacted] distribution of ITT participants across EDSS states 4 – 6). As such, it's feasible that utility values within the model could be underestimated.

Based on the appraisal of peginterferon beta 1a (TA624)[54], the ERG were aware of a more recent study by Thompson (2017)[63], which reported quality of life burden and costs associated with RRMS in a UK population. As an exploratory analysis, the ERG conducted a scenario analysis that used baseline EDSS utility values reported in Thompson et al. (2017)[63]. See Section 6.1.1.11 for further description and results.

Based on a review of previous NICE RRMS appraisals, including fingolimod (TA254)[14], alemtuzumab (TA312)[2], teriflunomide (TA303)[55], and ocrelizumab (TA533)[53], the ERG confirmed that Orme et al.[51] had been accepted as an appropriate source of patient utility. As such, the ERG considered Orme et al. to be a reasonable source for use in the base case analysis. However, the ERG noted that the lack of HRQol data from OPTIMUM in the company's model was a source of uncertainty: while HRQoL was measured in OPTIMUM (SF-36), these data were not mapped to EQ-5D values or used in the model. The company did not provide justification for this.

A final limitation surrounding the modelled utility values is the assumption that people with active RRMS and HA RRMS have the same EDSS utilities, which the ERG considered was unlikely due the impact of more severe disease on the lives of people with HA RRMS.

In the model, a person's baseline EDSS utility was assumed to decrease upon progression, relapse and as a result of adverse events. Disutility associated with a relapse was estimated to be -0.071 (based on Orme et al.[51]). The company conducted a one-way sensitivity analysis that varied disutility associated with relapse using upper and lower bound percentiles, however this did not have a material impact on the base case results. Finally, upon progression to SPMS within the model, a further utility decrement of -0.045 was applied to each baseline EDSS utility value (based on Orme et al.[51]).

Table 35: Modelled EDSS utility values (based on Orme et al.[51])

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

Abbreviations: EDSS, expanded disability status scale; RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

4.2.7.2. Carer disutility

For both the ITT and HA populations, the model captured the HRQoL impact for caregivers based on a published study by Acaster et al. (2013)[64] (see disutilities in Table 36 below).

The ERG acknowledged that the inclusion of caregiver disutility was appropriate in the base case and is preferred by NICE, based on its appraisal of fingolimod (TA254)[14] and natalizumab (TA127)[59]. Acaster et al.[64] was a UK observational study that assessed the HRQoL impact on carers of people with RRMS: an online survey of 200 RRMS carers was conducted and compared to a matched control cohort. Impact on HRQoL was assessed using a number of instruments including the EQ-5D, and utilities were estimated using the UK value set from Dolan et al.[52], as appropriate. Carer disutility was estimated for patient-determined disease steps (PDSS) states, which is a self-assessment scale that assesses functional disability in people with MS.

The ERG noted that although similar, the PDSS and EDSS are not identical assessment measures; i.e. the EDSS is clinician led and offers a more granular assessment of disease. As such, there may be some uncertainty surrounding the assumption that PDSS states translate directly to EDSS states. Additionally, from Acaster et al.[64], it was unclear what proportion (if any) of respondents were carers of people with HA RRMS.

The ERG confirmed that Acaster et al. has been used in previous NICE TAs, including beta interferon 1a and 1b and glatiramer acetate (TA527)[11] and peginterferon beta 1a (TA624)[54]; however noted that Gani[65] was the primary source of carer disutility in NICE appraisals of alemtuzumab (TA312)[2], fingolimod (TA254)[14], and terilunomide (TA303)[55]. The UK study by Gani et al.[65] assessed the cost effectiveness of natalizumab compared to other DMTs for people with HA RRMS (see CS Document B, p131 for further description surrounding the estimation of these values). The company provided a scenario analysis that used carer disutility values reported by Gani et al.[65] for the active RRMS population, however the ERG noted that this did not have a material impact on results (see Document B, p166). For completeness, the ERG assessed the impact of using Gani et al.[65] values for the HA RRMS population. Results were not found to be sensitive to these values.

The ERG noted that carer disutility for SPMS was assumed to be the same as RRMS. This is considered to be a limiting assumption, however, given the paucity of data surrounding carer disutility in SPMS, and the fact that this assumption had been previously used in peginteferon beta 1a (TA624)[54], the ERG considered this to be reasonable.

Table 36: Modelled EDSS carer disutilities

	Base case carer disutilities (Acaster et al.)[64]	Scenario carer disutilities (Gani et al.)[65]
EDSS 0	0.002	0.000
EDSS 1	0.002	0.000
EDSS 2	0.045	0.000
EDSS 3	0.045	0.010
EDSS 4	0.142	0.010
EDSS 5	0.160	0.020
EDSS 6	0.173	0.030
EDSS 7	0.030	0.050
EDSS 8	0.095	0.110
EDSS 9	0.095	0.140

Abbreviations: EDSS, expanded disability status scale

4.2.7.3. Adverse event disutility

Disutility associated with serious and non-serious adverse events were captured in the model. Given that each treatment has a specific adverse event profile, the ERG considered adverse event disutility to be appropriate for inclusion (see CS Document B, p130 for the full list of adverse events and disutilities included in the model). For all treatments, the incidence rates for severe and non-severe events were derived from a SLR conducted by the company. The ERG noted that the incidence of adverse events for ponesimod as based on the company's SLR were lower than the incidence rates based from the long-term pooled analysis set for ponesimod reported in the CS (Document B, p.92; see Table 37 below).

Table 37: Adverse event incidence

Adverse event	Ponesimod (AE incidence long term pooled analysis set)	Ponesimod (modelled AE incidence)
Nasopharyngitis	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]
Headache	[redacted]	[redacted]
Upper respiratory tract infection	[redacted]	[redacted]
Lymphopenia	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]
Backpain	[redacted]	[redacted]
Nausea	[redacted]	[redacted]
Upper urinary tract infection	[redacted]	[redacted]
Aspartate aminotransferase increased	[redacted]	[redacted]
Alopecia	[redacted]	[redacted]
Dizziness	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]

The company conducted one-way sensitivity analysis and probabilistic sensitivity analysis which varied incidence rates and disutilities associated with adverse events for ponesimod and comparator DMT's. However this did not have a material impact on results. The ERG considered that adverse event disutilities in the model were not a key driver of cost effectiveness.

4.2.8. Resources and costs

Medicine acquisition costs were included for all treatments and are outlined in the CS (Document B, p.135). Within the CS, the company presented the annual acquisition cost for each treatment, with the model providing further detail on the calculation of each. Unit costs (price per pack) and dose frequency were primarily derived from the British National Formulary (BNF)[66], which is considered to be an appropriate source. The ERG noted that annual drug acquisition costs in years 1 and 2 were largely in line with previous RRMS apraisals including pegiterferon (TA624) and therefore seemed reasonable.

It should be noted that for alemtuzumab and cladribine treatment acquisition costs, the company assumed that a proportion of people receiving these treatments would require re-treatment, if relapses continued to occur. For alemtuzumab, 28%, 11% and 1% of people were assumed to reinitiate treatment in Years 3, 4 and 5 respectively, and for cladribine this was 9.3%, 4.2% and 3.2% respectively. These rates were derived from the NICE appraisal of cladribine (ID64)[67]. The ERG acknowledged that re-treatment rates for both alemtuzumab and cladribine had been included in previous appraisals of cladribine by NICE (TA493)[27] and the SMC (SMC 1300/18)[68]. In both appraisals, uncertainty surrounding the appropriateness of these rates was outlined (due to the lack of effectiveness evidence on re-exposure). In cladribine (TA493)[27], the ERG conducted an analysis which removed retreatment rates for both treatments, however this did not have a material impact on the base case results. The ERG considered that removing cladribine and alemtuzumab re-treatment rates would result in a decrease in total costs for these treatments, however it was unlikely to have a meaningful impact on results, given the high acquisition costs of both.

4.2.8.1. Administration and monitoring costs

The model included differentiated costs for year one and subsequent years in order to account for differences in monitoring and administration assumptions between treatments. The ERG considered this approach to be consistent with previous NICE TAs for RRMS, and therefore appropriate. Administration costs were included for IV and SC treatments in year 1 and years 2+ (for both treatments in both the ITT and HA RRMS populations). As ponesimod is taken orally (20mg once daily), no administration costs were included in the model in years 1 and 2+. Similarly, administration costs were not included for other oral treatments, including dimethyl fumarate, teriflunomide, ozanimod, and cladribine. The ERG considered that the exclusion of administration costs for oral treatments was reasonable.

For peginterferon beta 1a 125mcg, glatiramer acetate, interferon beta-1a, interferon beta-1b, alemtuzumab and fingolimod, administration costs were estimated based on resource use estimates within NICE (TA624)[54] and ocrelizumab (TA533)[53]. For ofatumumab and ozanimod, resource use estimates were taken from (TA ID1677)[13] and ID1294[12]. Costs were valued using the Personal Social Services Research Unit (PSSRU) and NHS reference costs 2018/19, as appropriate. Overall, the ERG considered the administration costs included in the analysis to be appropriate.

In relation to monitoring costs, annual resource use estimates for each treatment, apart from ponesimod and cladribine, were based on estimates used in the NICE appraisal of ocrelizumab (TA533)[53]. In the company's base case analysis it was assumed that ponesimod would be associated with 30% of the monitoring costs for fingolimod in year 1, and no monitoring required in subsequent years thereafter. With respect to monitoring costs in year 1, the company justified this assumption in the CS (Document B, p134), noting that 30% of participants in OPTIUMUM required monitoring after the first dose, which was based on an estimated 18.5% of participants being at risk of symptomatic bradycardia, then inflated to account for the exclusion of people with certain cardiovascular disorders. The company claimed that the methods for up-titrating ponesimod, and the increased specificity of ponesimod to the $S1P_1$ receptor will result in fewer AEs than fingolimod. Based on clinical input to the ERG and the safety profile of ponesimod reported in Section 3.5.3 and 3.5.4 (which indicated cardio and ophthalmic concerns with ponesimod when compared to teriflunomide), these assumptions were not considered to be have been fully justified. As the data did highlight some concerns of liver toxicity in participants treated with ponesimod, clinical advisors to the ERG suggested that monitoring should match that of fingolimod until further evidence for its safety is available. As an exploratory analysis, the ERG conducted a scenario analysis which assumed fingolimod had identical monitoring costs to fingolimod in year 1. See section 6.1.1.8 for further description and results

During the clarification stage, the company subsequently provided a revised model that updated ponesimod monitoring costs in year 2+, as clinical expert advice to the ERG considered £0 monitoring costs in subsequent years to be inappropriate. The ERG acknowledged that the updated monitoring cost provided by the company (£228.82) was broadly in line with other oral DMTs. Overall, monitoring and administration costs were not a key driver of cost effectiveness results (in either the ITT or HA RRMS populations), given the magnitude of drug acquisition costs and disease management costs for all treatments.

4.2.8.2. **Health state costs**

The model included EDSS health state costs for people with RRMS and SPMS, which represented costs associated with disease management. Costs were derived from a study by Tyas et al. (2007)[69], and included direct health care costs as well as costs for community services i.e. nurse visits, home helper and other major investments (see Document B, p143). Values were inflated to 2019 as appropriate. The ERG noted that indirect costs (e.g. informal care, productivity losses) were excluded. Given that the analysis was conducted from an NHS and PSS perspective, this was considered to be reasonable.

Tyas et al.[69] was a UK study that examined the cost of RRMS according to disease severity. The ERG noted that Tyas et al.[69] had been used previously in NICE appraisals for RRMS including teriflunomide (TA303)[55], alemtuzumab (TA312)[2] and ocrelizumab (TA533)[53]. The ERG noted that the company did not provide results for a scenario analysis basing EDSS disease management costs on alternative literature sources. The ERG was aware of other relevant sources including a relatively recent study by Thompson et al. (2018)[5] that re-examined the financial impact associated with RRMS in the UK. As an exploratory analysis the ERG conducted a scenario analysis using this alternative study to estimate disease management costs. See section 6.1.1.9 further description and results.

In the NICE appraisal of beta interferons and glatiramer acetate (TA527),[11] the assessment group preferred costs used in the appraisal of dimethyl fumarate (TA320),[56] which used costs from the UK MS survey in 2005 (subsequently reported by Tyas et al.[69]). As such, the ERG considered the use of direct costs from Tyas et al.[69] to be an appropriate source for use within the base case analysis.

The ERG noted that disease management costs were the same for both the ITT and HA RRMS populations. From Tyas et al.[69] it was unclear what proportion of participants (if any) had HA RRMS. Clinical advice to the ERG was that disease management costs are likely to be higher for people with HA RRMS, as people will have more relapses. The company conducted a one-way sensitivity analysis that varied disease management costs in RRMS and SPMS. ITT results were not overly sensitive to this analysis, however the ERG noted that in the HA RRMS subgroup, varying disease management costs for SPMS did have a material impact on results. The ERG acknowledged that the lack of robust EDSS disease management costs for HA RRMS (particularly in SPMS) is an area of uncertainty, however in the absence of relevant cost data for this subgroup, the use of Tyas et al. was considered reasonable.

Table 38: Modelled disease management costs

	RRMS	**SPMS**
EDSS 0	£998.74	NA
EDSS 1	£1,039.11	£1,386.86
EDSS 2	£760.70	£1,108.45
EDSS 3	£4,165.75	£4,512.46
EDSS 4	£2,018.19	£2,364.90
EDSS 5	£3,422.64	£3,771.42

	RRMS	SPMS
EDSS 6	£4,569.38	£4,916.10
EDSS 7	£12,027.36	£12,374.08
EDSS 8	£29,293.73	£29,641.48
EDSS 9	£23,439.95	£23,788.74
Relapse costs		
0-9	£2,243.81	£2,243.81

Abbreviations: EDSS, expanded disability status scale; RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

Costs associated with relapse

As noted in Section 4.2.2, in the model people were capable of experiencing relapse whilst in any EDSS state. The source of relapse cost used was the NICE appraisal of peginterferon beta 1a (TA624)[54]. Relapse costs within the appraisal were inflated from previous values published within dimethyl fumarate (TA320);[56] i.e. Tyas et al.[69] The company inflated costs to 2019 levels using the HCHS index and the PSSRU as appropriate. The cost of relapse in the model was estimated to be £2,243.

Although Tyas et al.[69] had been used previously in the NICE appraisal of fingolimod (TA254)[14], ocrelizumab (TA533)[53] and peginterferon beta 1a (TA624)[54], the ERG noted that there was a lack of granularity surrounding the cost of relapse estimated in the study; i.e. it was unclear what proportion of people were assumed to require hospitalisation. To explore uncertainty surrounding the cost of relapse, the company conducted a one-way sensitivity analysis that varied the cost using upper and lower bound percentiles. Base case results were not sensitive to this analysis.

The ERG identified that an alternative source by Dee et al.[70] had been used in several previous NICE appraisals including teriflunomide (TA303)[55] and alemtuzumab (TA312)[2]. Dee et al.[70] was an Irish study that assessed the budget impact of natalizumab. The study, which was conducted from a Health Service Executive (HSE) perspective, included people with RRMS deteriorating on one of the first line DMTs. The average cost of relapse was derived using a database that reported length of stay (LoS) data for neurology bed MS admissions from six large neurology centres. The average LoS for people requiring admission was reported to be 10.71 days. The average cost of relapse was estimated to be €3,696, based on 20% of people requiring an inpatient stay and 80% requiring a day case visit. For completeness, the ERG conducted a scenario analysis using the average relapse cost as reported by Dee et al.[70], inflated to 2020

GBP costs (see Section 6.1.1.10 for discussion and results). However it should be noted that clinical opinion to the ERG noted that the majority of relapses in the UK are treated in an outpatient setting via GP. Therefore resource use data from Dee et al. may overestimate the cost of relapse.

Similar to EDSS health state costs, the company assumed that relapse costs were the same for both people with HA RRMS. Based on clinical input to the ERG, this assumption may be reasonable, however it could be plausible for highly active groups to experience more severe relapses and therefore higher costs. Overall, the ERG considered that the company's base case approach to estimating the cost of relapse was reasonable. Based on sensitivity analysis conducted by the company and the ERG, cost of relapse was not considered to be a key driver of cost effectiveness in the model.

Costs associated with adverse events

The model included costs associated with both non-serious and serious adverse events (see Document B, p138). The ERG considered that adverse event costs were reasonable to include, given that most people receiving DMTs experience AEs, either mild or serious, and that the rates and types of AE vary across each DMT. Resource use estimates were primarily based on previous NICE TAs including ocrelizumab (TA533)[53], with costs reflecting PSSRU 2019 and NHS reference costs, as appropriate. However, as noted in the CS (Document B, p138), the company needed to make several assumptions surrounding resource utilisation with respect to alopecia, diarrhoea, dyspnoea, hypertension and nausea, due to a lack of data.

The ERG acknowledged that the majority of unit costs were relatively minor, with the exception of non-fatal and fatal progressive multifocal leukoencephalopathy (PML). PML was associated with a relatively high cost compared to other modelled adverse events. Adverse event costs ranged from £5.78 (for treatment of nausea) to £19,391 (for treatment of PML). However, PML costs only applied to a small proportion of people receiving natalizumab in the model, as PML incidence rates for all other DMTs were 0%.

From the base case results provided by the company, the ERG noted that there were differences in total adverse event costs between treatments due to variation in modelled incidence rates between treatments, however adverse event costs were not considered a key driver of incremental costs. The company conducted one-way sensitivity analysis which varied

the cost of adverse events using upper and lower bound percentiles, however this did not have a material impact on results.

5. COST-EFFECTIVENESS RESULTS

5.1. Company's cost-effectiveness results

5.1.1. Base case results

The company provided base case cost effectiveness results for both the ITT population and HA RRMS subgroup (see Document B, Sections B.3.27-28).

5.1.1.1. ITT population

The company's base case results are provided in Table 39 below.

[redacted] The ERG noted that incremental savings were largely due to lower drug acquisition costs and disease management costs, whilst the incremental QALY gained associated with ponesimod stemmed primarily from improved relative treatment efficacy. Compared to interferon beta 1a 22 mcg and peginterferon beta 1a 125 mcg, ponesimod resulted in ICER of [redacted] and [redacted] respectively. Compared to ocrelizumab and ofatumumab, ponesimod resulted in [redacted].

Table 39: Company base case results (ITT population)

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained (ICER)
Company deterministic base case					
Ponesimod	[redacted]	[redacted]	-	-	-
Teriflunomide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fumarate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22mcg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30mcg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44mcg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained (ICER)
Interferon beta-1b 250mcg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Peginterferon beta-1a 125mcg	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ITT, intention-to-treat; ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life years

5.1.1.2. HA RRMS subgroup

The results of the company's subgroup analysis, shown in Table 40, showed that ponesimod was [redacted]. Compared to fingolimod and ozanimod, [redacted]. The results further showed that ponesimod resulted in a [redacted].

Table 40: Company base case results (HA RRMS population)

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained (ICER)
Company deterministic base case					
Ponesimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HA, highly active; ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life years; RRMS, relapsing remitting multiple sclerosis

5.2. Company's sensitivity analyses

The company conducted a variety of sensitivity analyses including one-way sensitivity analysis, scenario analyses and probabilistic sensitivity analyses. The results of these analyses are appraised in the following sections (Sections 5.2.1, 1.1.1 and 5.2.3).

5.2.1. One-way sensitivity analysis

In the CS (Document B, Section 3.31), the company provided the results of a one-way sensitivity analysis for comparisons between ponesimod and teriflunomide, in the ITT population, and fingolimod in the HA RRMS population. One-way sensitivity analysis results comparing ponesimod to the remaining comparators were included in an appendix. The results for the twelve most noteworthy parameters are displayed via tornado diagrams in Figure 2 and Figure 3 below. The ICER for the ITT population was relatively robust with respect to most of the model parameters; though it was highly sensitive to the EDSS progression hazard ratio for teriflunomide, and was also sensitive to the EDSS progression hazard ratio and annual discontinuation rate for ponesimod. Varying the baseline conversion to SPMS progression, the annual discontinuation rate for ponesimod, and the relapse rate ratios and EDSS progression hazard ratios for both comparators had the biggest impact on the ICER for the HA RRMS subgroup.

[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; QALYs, quality adjusted life years; EDSS, Expanded Disability Status Scale; OWSA, one-way sensitivity analysis; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

3

Abbreviations: EDSS, expanded disability status scale; HA, highly active; ICER, incremental cost-effectiveness ratio; OWSA, one-way sensitivity analysis; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

5.2.2. Probabilistic sensitivity analysis

The company conducted a probabilistic sensitivity analysis (PSA) to explore the impact of parameter uncertainty when the model parameters were varied as per the respective distributions (CS, Document B, Section 3.10.1). The PSA was run for 5,000 iterations (see PSA results in Document B, p159, Table 62).

The PSA results are presented in Table 41 for the ITT population and in Table 42 for the HA RRMS subgroup, along with the deterministic ICERs (for reference). The cost effectiveness acceptability curves (CEAC) in Figure 5 and Figure 7 indicated that the probability of ponesimod being cost-effective at a £30k threshold was [redacted] for the ITT population and [redacted] for the HA RRMS subgroup. The cost-effectiveness scatterplots in Figure 4 and Figure 6 suggested that there was significant uncertainty around the results, especially for the HA RRMS subgroup.

[redacted].

5.2.2.1. ITT population

Table 41: PSA results (ITT population)

Outcomes Ponesimod vs Comparator	Incremental costs (£)	Incremental QALYs	Probabilistic ICER (£/QALY)	Deterministic ICER (£/QALY)
Teriflunomide 14 mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fuarate 240 mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate 20 mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22 mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30 mcg IM	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44 mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1b 250 mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab 600 mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20 mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0 mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Peginterferon beta-1a 125 mcg SC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; PSA, probabilistic sensitivity analysis; QALYs, quality adjusted life years

[redacted]

Abbreviations: ITT, intention-to-treat; QALYs, quality adjusted life years

5

Abbreviations: ITT, intention-to-treat; QALYs, quality adjusted life years

5.2.2.2. HA RRMS subgroup

Table 42: PSA results (HA RRMS subgroup)

Outcomes Ponesimod vs Comparator	Incremental costs (£)	Incremental QALYs	Probabilistic ICER (£/QALY)	Deterministic ICER (£/QALY)
Ocrelizumab 600mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab 12mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine 3.5mg/kg PO	[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod 0.5mg PO	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI, confidence interval; HA, highly active; ICER, incremental cost-effectiveness ratio; PSA, probabilistic sensitivity analysis; QALYs, quality adjusted life years; RRMS, relapsing remitting multiple sclerosis

6

Abbreviations: HA, highly active; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis

7

Abbreviations: HA, highly active; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis

5.2.3. Scenario analyses

The company conducted a range of scenario scenario analyses for both the ITT and HA RRMS populations (see Table 43 and Table 44), the results of which are reported in Table 45 and Table 46 below. Total costs and QALYs for each treatment are reported in the CS (Document B, p171).

Table 43: Scenario analyses conducted by the company (ITT population)

Number	Parameter	Scenario
S1	Discounting	1.5% for both costs and outcomes
S2	Population characteristics	UK RSS data set
S3	Natural history transition matrix between EDSS health states	Dimethyl fumarate and London Ontario data source
S4	Disease progression to higher EDSS	Treatment effect based on 6 month data
S5	Treatment waning effect	a) No waning effect
		b) 50% loss after 10 years
S6	Care giver disutilities	Disutility based on Gani et al.[65]
S7	Mortality	Pokorski et al.[62] without interpolation
S8	Treatment discontinuation	5% discontinuation for all treatments
S9	Post treatment discontinuation	100% of people move to cladribine

Abbreviations: EDSS, Expanded Disability Status Scale; ITT, intention-to-treat; RSS, risk sharing scheme

Table 44: Scenario analyses conducted by the company (HA RRMS population)

Number	Parameter	Scenario
S10	Population	Highly active RRMS subgroup from OPTIMUM
S11	Disease progression to higher EDSS	Treatment effect based on 6-month data
S12	Treatment waning effect	a) No waning (backed up with Phase 2 long term data)
		b) 50% loss after 10 years
S13	Treatment discontinuation	5% discontinuation for all treatments
S14	Post treatment discontinuation	100% of people move to natalizumab

Abbreviations: EDSS, Expanded Disability Status Scale; HA highly active; RRMS, relapsing remitting multiple sclerosis

5.2.3.1. ITT population

Based on the company's scenario analyses, results for the ITT population were most sensitive to using an alternative EDSS natural history transition matrix (derived from the London Ontario

and dimethyl fumarate dataset), disease progression based on six-month CDA, and a post-treatment discontinuation assumption that assumed that 100% of people received cladribine after first-line treatment.

[redacted]

Table 45: Company scenario analysis results (ITT population)

Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
Base case	Incremental QALYs	[redacted]										
	Incremental Costs	[redacted]										
	ICER	[redacted]										
S1: Discounting	Incremental QALYs	[redacted]										
	Incremental Costs	[redacted]										
	ICER	[redacted]										
S2: UK RSS	Incremental QALYs	[redacted]										

Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
population	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S3: EDSS health states	Incremental QALY	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S4: Disease progression	Incremental QALY	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S5a: Treatment waning (no waning)	Incremental QALY	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S5b: Treatment waning (50% loss at 10 yrs)	Incremental QALY	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
S6: Caregiver disutilities	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S7: Mortality	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S8: Treatment discon	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
tinu-ation	Incremental Costs											
	ICER											
S9: Post-treatment discon tinu-ation	Incremental QALYs											
	Incremental Costs											
	ICER											

Abbreviations: DMF, Dimethyl fumarate 240mg PO; EDSS, Expanded Disability Status Scale; GA, Glatiramer acetate 20mg SC; ICER, incremental cost-effectiveness ratio; IFNB-1a 22 µg, interferon beta-1a 22 µg subcutaneously; IFNB-1a 30 mcg, interferon beta-1a 30 µg intramuscular once weekly; IFNB-1a 44 µg, interferon beta-1a 44 µg subcutaneously three times weekly; ITT, intention-to-treat; OCR, ocrelizumab 600 mg every six months; Ofatumumab 20mg SC; Ozanimod 1.0mg PO; PBO, placebo; PEG, Peginterferon beta-1a 125mcg subcutaneously; PON, ponesimod 20 mg once daily; QALYs, quality adjusted life years; RSS, risk sharing scheme; TER, teriflunomide 14 mg once daily

5.2.3.2. HA RRMS

A complete list of scenario analyses undertaken by the company can be found in the CS (Document B, p168). Based on the company's scenario analyses, results for the HA RRMS population were most sensitive to a scenario that assumed 100% of people would receive alemtuzumab post-treatment discontinuation.

[redacted] **Table 46: Company scenario analysis results (HA RRMS population)**

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
S10: Population	Incremental QALYs	[redacted]					
	Incremental Costs						
	ICER						
S11: Disease progression	Incremental QALYs						
	Incremental Costs						

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S12a: Treatment waning (no waning)	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S12b: Treatment waning (50% loss at 10yrs)	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S13: Treatment	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
discontinu-ation	Incremental Costs						
	ICER						
S14: Post-treatment discontinu-ation	Incremental QALYs						
	Incremental Costs						
	ICER						

Abbreviations: ALE, alemtuzumab 12 mg once daily; CLA, cladribine 3.5 mg/kg once daily; FIN, fingolimod 0.5 mg once daily; HA, highly active; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OCR, ocrelizumab 600 mg every six months; Ofatumumab 20mg SC; Ozanimod 1.0mg PO; PBO, placebo; PON, ponesimod 20 mg once daily; QALYs, quality adjusted life years; RRMS, relapsing remitting multiple sclerosis

5.3. Model validation and face validity check

The ERG did not identify any errors in the company's original model. The results outlined below are based on the company's revised model (submitted during clarification), which included ozanimod and ofatumumab.

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1. Exploratory and sensitivity analyses undertaken by the ERG

As noted throughout the report, the ERG identified a number of uncertainties surrounding the clinical efficacy parameters used in the economic analysis, the model assumptions and choice of literature sources. The ERG conducted scenario analyses in order to explore the potential impact of these uncertainties. See Section 6.1.1 for description of each scenario and Section 6.2 for the impact on the ICER. Please note that the results below incorporate the PAS discount for ponesimod, but do not include PAS discounts for comparator treatments.

6.1.1. ITT and HA RRMS populations

6.1.1.1. Scenario analysis 1: Six-month CDA used to model disease progression

The ERG considered that the use of three-month CDA in the economic model to estimate clinical effectiveness was not appropriate, given that six-month CDA is a more robust measure of disease progression and has been preferred by NICE in previous MS appraisals (see Section 4.2.6.1). This scenario analysis explored the impact of using six-month CDA data from the NMA to estimate hazard ratios and treatment-specific transition probabilities in both the active and highly active RRMS populations.

[redacted]. Overall, results from this analysis indicated that using the six-month CDA had a material impact on base case cost effectiveness results versus three key comparators (see Section 6.2).

In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to the base case results i.e. [redacted] (See Section 6.2)

6.1.1.2. Scenario analysis 2: 25% of SPMS group assumed to receive siponimod and 75% receive BSC

The ERG noted that siponimod had been recommended by NICE for the treatment of people with SPMS (see Section 4.2.6). Clinical experts to the ERG estimated that between 12.5% and 50% of people will receive siponimod after progressing to SPMS. To explore the potential impact of subsequent treatment with siponimod, this scenario analysis assumed that 25% of people who converted to SPMS in the model went on to receive siponimod, whilst 75% received BSC. Please note that for this scenario only the costs for siponimod were considered; i.e. siponimod was not assumed to have a treatment effect. Given the lack of robust long-term clinical effectiveness data for siponimod, this approach was considered to introduce less uncertainty into the analysis and provide indicative results based on treatment cost only.

Based on this analysis, the cost effectiveness of ponesimod improved versus all comparator DMTs in the ITT population.

[redacted]see Section 6.2).

In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to the base case results; i.e.

[redacted]see Section 6.2).

6.1.1.3. Scenario analysis 3: Participant characteristics based on UK RSS population

As noted in Section 4.2.3, clinical opinion to the ERG was that characteristics from participants in OPTIMUM were likely to be generalisable to the UK population. For completeness, this scenario analysis (conducted in the ITT population only) used characteristics from people in the UK RSS dataset; i.e. mean age, sex and EDSS distribution. The ERG noted that the UK RSS dataset included people with RRMS and did not outline the proportion of participants with HA RRMS, therefore the ERG did not consider it appropriate to conduct a scenario analysis for the HA RRMS population based on UK RSS characteristics. The ERG also considered that the UK RSS dataset may not be fully generalisable to the target population, as clinical advice to the ERG was that it included people with SPMS, and participants generally had longer disease duration without access to DMT.

Based on this analysis,

[redacted] (see Section 6.2). The ERG did not consider baseline population characteristics to be a key source of uncertainty in this analysis.

6.1.1.4. Scenario analysis 4: Alternative subsequent treatment assumptions

The ERG accepted the company's base case assumption surrounding the use of BSC as the primary subsequent treatment option for people who discontinued treatment (see Section 4.2.6). Furthermore, the ERG acknowledged the company's attempt to explore uncertainty surrounding the impact of subsequent treatment use by including scenario analyses that assumed 100% of people who discontinue treatment go on to receive cladribine and natalizumab in the ITT and HA RRMS populations, respectively. However based on clinician feedback to the ERG, subsequent treatment will depend primarily on the rationale for stopping first-line treatment; i.e. if a person discontinues due to adverse events then they will likely go on to receive a treatment with a more favourable adverse event profile. As such the ERG considered that the selection of cladribine and natalizumab as the primary subsequent treatments for this scenario (as outlined in CS Document B, p125-126) was overly simplistic.

In this alternative exploratory scenario the ERG opted to use teriflunomide and alemtuzumab as the subsequent treatments in the respective ITT and HA RRMS populations. It should be noted that only the costs associated with these treatments were considered; i.e. drug acquisition costs, administration costs and monitoring costs only, and the efficacy of these treatments for health outcomes were not considered. The ERG opined that including subsequent treatment effects on the basis of the company's NMAs would introduce additional uncertainty, due to the limitations surrounding these results (see section 3.4 and 3.6). The alternative treatments selected by the ERG therefore explored the impact of using subsequent treatments with different acquisition costs.

The ERG noted that ITT population results were not overly sensitive to this scenario analysis. Most notably, ponesimod went from being [redacted] to resulting in a minor ICER of [redacted] and [redacted] compared to glatiramer acetate and interferon beta 1b 250 mcg respectively. Compared to peginterferon beta 1a 125mcg and interferon beta 1a 22mcg, ponesimod became more cost effective resulting in reduced ICERs (see Section 6.2).

In the HA RRMS population, this scenario primarily impacted the cost effectiveness results against alemtuzumab, as ponesimod went from being [redacted], to being [redacted]. The ERG noted that this scenario may lack validity as it assumed that people continued to receive alemtuzumab, despite not responding to alemtuzumab as a first line treatment (see Section 6.2).

6.1.1.5. Scenario analysis 5: No difference in discontinuation rates (assumed 5% for all treatments)

As noted in Section 4.2.6, the use of treatment specific all-cause discontinuation rates in the company's base case was considered to be reasonable. However, the ERG noted that there is uncertainty surrounding the validity of the NMA estimates due to the limitations outlined in Section 3.4. This scenario analysis explored the impact of variation in discontinuation rates by assuming no difference in rates between treatments; i.e. a discontinuation rate of 5% is applied to all treatments.

Based on this analysis [redacted] See Section 6.2.

In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to the base case results; i.e. [redacted] See Section 6.2.

6.1.1.6. Scenario analysis 6: No waning in treatment effect (applies to all treatments)

The ERG considered the company's base case assumption surrounding treatment waning to be broadly acceptable (see Section 4.2.6). However, due to the lack of long term data surrounding treatment efficacy over time, this scenario analysis explored the impact of removing the treatment waning assumption used in the base case for all treatments. Please note that although this scenario analysis was useful in exploring uncertainty, the ERG did not consider it to reflect clinical practice and therefore it may lack plausibility.

Based on this analysis, cost effectiveness results for ponesimod compared to each comparator improved; i.e.

[redacted] See Section 6.2. In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to most base case results; i.e.

[redacted] However, the ERG noted that compared to fingolimod, [redacted]. See Section 6.2.

6.1.1.7. Scenario analysis 7: Alternative modelled clinical effectiveness parameters

In Sections 3.4 and 3.6, the ERG noted there to be uncertainty surrounding the clinical effectiveness estimates used in the economic model, which were derived from the NMAs. This scenario analysis estimated alternative clinical effectiveness estimates by adopting a positioning-based approach; i.e. DMTs were stratified into 3 groups according to their approximate position within the treatment pathway (see Table 47 below). For CDA and ARR, each treatment group was compared to BSC and the median efficacy estimate (hazard ratio and rate ratio) was selected for each. For treatment discontinuation, each group was compared to ponesimod (which was not included in Group B) and the median odds ratio was selected for each. This analysis was considered exploratory in nature, however it helped to demonstrate the sensitivity of base case results to a change in key treatment efficacy parameters.

Table 47: Treatment groups according to positioning

Group A	Group B	Group C
Interferon beta 1a (22mcg, 30mcg, 44mcg)	Ponesimod (except for treatment discontinuation)	Alemtuzumab
Interferon beta 1b	Ozanimod	Fingolimod
Peginterferon beta 1a	Ofatumumab	Cladribine
Glatiramer acetate	Teriflunomide	
Dimethyl fumarate	Ocrelizumab	

Table 48: Median efficacy effect estimates for positioning-based groups (ITT population)

Group	ARR Rate Ratio	3-month CDA HR	6-month CDA HR	Premature Treatment Discontinuation OR
A	[redacted]	[redacted]	[redacted]	[redacted]
B	[redacted]	[redacted]	[redacted]	[redacted]
C	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; HR, hazard ratio; ITT, intention-to-treat; OR, odds ratio

Table 49: Median efficacy effect estimates for positioning-based groups (HA RRMS)

Group	ARR Rate Ratio	3-month CDA HR	6-month CDA HR	Premature Treatment Discontinuation OR
A	[redacted]	[redacted]	[redacted]	[redacted]
B	[redacted]	[redacted]	[redacted]	[redacted]
C	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ARR, annualised relapse rate; CDA, confirmed disability accumulation; HA, highly active; HR, hazard ratio; ITT, intention-to-treat; OR, odds ratio; RRMS, relapsing remitting multiple sclerosis

Based on this analysis, in the ITT population [redacted] In the HA RRMS population, results were sensitive to this scenario analysis; i.e. [redacted] Notably, ponesimod was [redacted] by ozanimod and fingolimod. See Section 6.2.

6.1.1.8. Scenario analysis 8: Increased monitoring costs for ponesimod in Year 1

As noted in Section 4.2.8.1, the ERG highlighted some uncertainty surrounding monitoring costs for ponesimod in year one. In order to explore the impact of increased monitoring costs, this scenario assumed that ponesimod would require monitoring equivalent to that of fingolimod in year 1. The ERG noted that for both the ITT and HA RRMS populations, results were not considered sensitive to this scenario (see Section 6.2).

6.1.1.9. Scenario analysis 9: Alternative EDSS health state costs

Tyas et al.[69] was considered to be an appropriate source for deriving EDSS disease management costs in the base case (see Section 4.2.8.2). This scenario analysis was

conducted to determine the impact of using an alternative literature source to derive EDSS health state costs. Based on direct health care costs, community services costs and investment costs from Thompson et al.[63], the mean annual EDSS disease management cost per person was estimated to be £6,369, £7,994 and £13,325 for EDSS states 0-3, 3-6 and 6-9, respectively. The ERG noted that the RRMS costs reported by Thompson et al.[63] are somewhat limited, given that values were reported for mild, moderate and severe disease (and not individual EDSS health states).

The ERG noted that results for the ITT population were not overly sensitive to this scenario analysis. Most notably, [redacted]. See Section 6.2.

In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to the base case results; i.e. [redacted] See Section 6.2.

6.1.1.10. Scenario analysis 10: Alternative cost associated with relapse

The ERG acknowledged that the cost of relapse used in the company's base case analysis was largely appropriate (see Section 4.2.8.2). This scenario explored the impact of using a higher cost of relapse in the model for both the ITT and HA populations, based on an Irish study by Dee et al.[70]. For this analysis costs were converted from euros into GBP and inflated to 2020 values, resulting in a cost per relapse of £3,451. The ERG accepted that this study may be associated with generalisability concerns given that it is non-UK based and there are likely to be differences in healthcare resource utilisation for RRMS groups between Ireland and the UK.

The ERG noted that results were not overly sensitive to this scenario analysis and slightly improved the cost effectiveness of ponesimod compared to other DMTs in the ITT population. This scenario analysis resulted in minor incremental cost and QALY changes, however base case results remained largely unchanged (see Section 6.2).

In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to the base case results i.e.

[redacted] See Section 6.2.

6.1.1.11. Scenario analysis 11: Alternative EDSS health state utilities.

Overall the ERG considered Orme et al.[51] to be an appropriate source for estimating EDSS health state utilities (see Section 4.2.7.1). To test uncertainty surrounding the utility value source, the company provided a scenario analysis that used an alternative values reported by Gani et al.[65] (for the active RRMS population). Results were not considered overly sensitive to these alternative values (see Section 5.2.3). For completeness this scenario analysis applied utility values from an additional UK study by Thompson et al.[63] (see Table 50 below) to the ITT and HA RRMS populations. It should be noted that in the absence of robust HRQol data, utility values for the SPMS population were estimated by applying the -0.045 utility decrement from Orme et al.[51] to the RRMS values from Thompson et al.[63].

The ERG noted that results were not overly sensitive to this analysis, as utility values were broadly similar to Orme et al.[51]. In the ITT population [redacted] Compared to interferon beta 1a 22 mcg and peginterferon beta 1a 125 mcg, ponesimod resulted in an ICER of [redacted] and [redacted] respectively. Compared to ocrelizumab and ofatumumab, ponesimod resulted in [redacted].

In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to the base case results; i.e. [redacted] See Section 6.2.

Table 50: Utility values from Thompson et al.[63]

	EDSS									
	0	1	2	3	4	5	6	7	8	9
RRMS	0.898	0.787	0.695	0.573	0.605	0.569	0.48	0.373	0.157	-0.111
SPMS	N/A	0.742	0.650	0.528	0.560	0.524	0.435	0.328	0.112	-0.156

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

6.1.1.12. Scenario analysis 12: Alternative annual conversion probabilities (from RRMS to SPMS)

As noted in Section 4.2.6, the ERG was aware that alternative SPMS conversion probabilities had been used previously in the NICE appraisal of peginterferon (TA624)[54]. To explore the sensitivity of results to a change in this modelled parameter, this scenario used the annual EDSS baseline probabilities of converting to SPMS reported in peginterferon (TA624)[54], which were lower than the estimates used by the company in their base case.

The ERG noted that ITT results were highly sensitive to this analysis. Notably the ICER compared to interferon beta 1a 22 mcg increased from [redacted] to [redacted]. Compared to teriflunomide, dimethyl fumarate and ozanimod, ponesimod was no longer dominant, resulting in [redacted]. For the comparison with peginterferon beta 1a 125 mcg, ponesimod went from being the [redacted] treatment to being [redacted]. See Section 6.2.

In the HA RRMS population, this scenario had an impact on the incremental costs and QALYs for all treatments, but did not result in material changes to the base case results; i.e. [redacted] See Section 6.2.

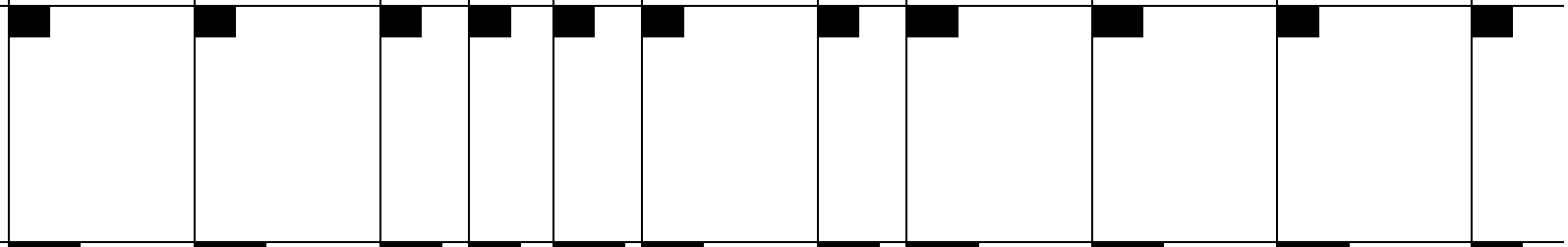
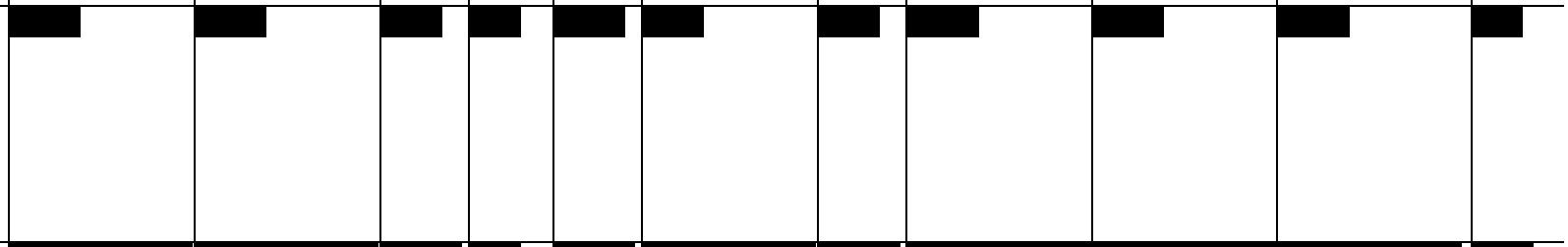
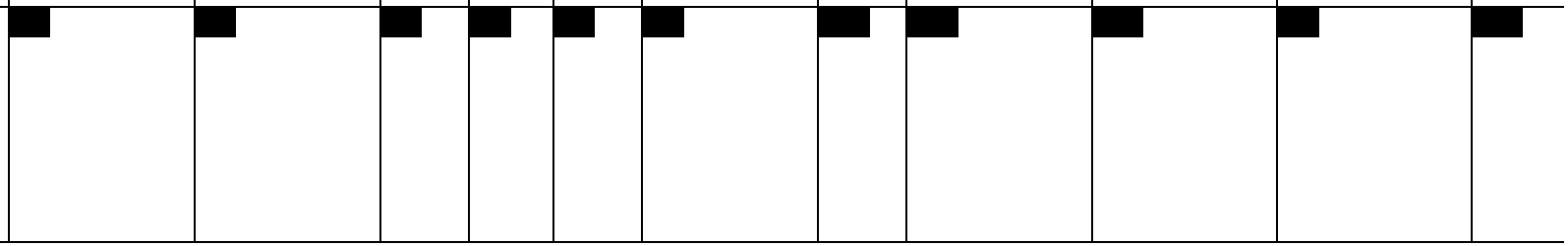
6.2. Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

The results of the ERG's one-way sensitivity analyses are reported in Table 51 (ITT) and Table 52 (HA RRMS). A full description of the analyses undertaken is provided in Sections 6.1.1.1 - 6.1.1.12. The scenarios that had the most impact on the base case results were:

- Using six-month CDA for EDSS progression in the model, rather than 3-month CDA (ITT population)

- Using a positioning-based approach to estimate treatment effect (ITT and HA RRMS populations)
- Using an alternative set of annual conversion probabilities, from RRMS to SPMS (ITT population)
- No waning in treatment effect (HA RRMS population)

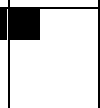
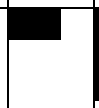
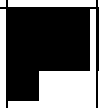
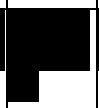
Table 51: ERG scenario analysis results (ITT population)

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
Company base case	Incremental QALYs											
	Incremental Costs											
	ICER											
S1: CDA-6	Incremental QALYs											

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S2: 25% population receive sponimod	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S3: UK RSS population	Incremental QALY	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S4: Alternativ	Incremental	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
e subsequent treatments	QALY											
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S5: 5% discontinue-ation rate	Incremental QALY	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
	Costs											
	ICER											
S6: No treatment waning	Incremental QALY											
	Incremental Costs											
	ICER											

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
S7: Alternative modelled clinical effectiveness parameters	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S8: Monitoring costs	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S9: EDSS health state costs	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S10: Relapse costs	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S11: EDSS healt	Incremental	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
h state utilities	QALYs											
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S12: Conversion to SPMS	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ERG Scenario	Outcome PON vs. comparator	TER	DMF	GA	IFNB-1a 22 mcg	IFNB-1a 30 mcg	IFNB-1a 44 mcg	IFNB-1b	OCR	OFA	OZA	PEG
	Costs											
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: DMF, Dimethyl fumarate 240mg PO; EDSS, Expanded Disability Status Scale; ERG, Evidence Review Group; GA, Glatiramer acetate 20mg SC; ICER, incremental cost-effectiveness ratio; IFNB-1a 22 µg, interferon beta-1a 22 µg subcutaneously; IFNB-1a 30 mcg, interferon beta-1a 30 µg intramuscular once weekly; IFNB-1a 44 µg, interferon beta-1a 44 µg subcutaneously three times weekly; ITT, intention-to-treat; OCR, ocrelizumab 600 mg every six months; Ofatumumab 20mg SC; Ozanimod 1.0mg PO; PBO, placebo; PEG, Peginterferon beta-1a 125mcg subcutaneously; PON, ponesimod 20 mg once daily; QALYs, quality adjusted life years; SPMS, secondary progressive multiple sclerosis; TER, teriflunomide 14 mg once daily

Table 52: ERG scenario analysis results (HA RRMS subgroup)

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
Company HA subgroup analysis	Incremental QALYs						
	Incremental Costs						
	ICER						
S1: CDA-6	Incremental QALYs						
	Incremental Costs						
	ICER						
S2: 25% population	Incremental QALYs						

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
receive siponimod	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S3: UKRSS population	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S4: Alternative subsequent treatments	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
S5: 5% discontinu-ation rate	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S6: No treatment waning	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S7: Alternative modelled clinical effectiveness	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
parameters	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S8: Monitoring costs	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S9: EDSS health state costs	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S10: Relapse costs	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Scenario	Outcome PON vs. comparator	OCR	OFT	OZA	ALE	CLA	FIN
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S11: EDSS health state utilities	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
S12: Conversion to SPMS	Incremental QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Incremental Costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ICER	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ALE, alemtuzumab 12 mg once daily; CLA, cladribine 3.5 mg/kg once daily; EDSS, Expanded Disability Status Scale; ERG, Evidence Review Group; FIN, fingolimod 0.5 mg once daily; HA, highly active; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OCR, ocrelizumab 600 mg every

six months; Ofatumumab 20mg SC; Ozanimod 1.0mg PO; PBO, placebo; PON, ponesimod 20 mg once daily; QALYs, quality adjusted life years; RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

6.3. ERG's preferred assumptions

The ERG considered that several of the company's base case assumptions were inappropriate, and alternatives to these were used in the ERG base case. The ERG's preferred assumptions are outlined in Table 53 for both the ITT and HA RRMS populations. The ICERs presented in Table 54 and Table 55 below incorporate all of the ERG's preferred assumptions.

Table 53: ERG preferred base case assumptions (ITT and HA RRMS)

Preferred assumption	Report Section
Company base-case	*5.1.1*
6 month CDA used to model disease progression	4.2.6.1 and 6.1.1.1
25% of people receive BSC after converting to SPMS, 75% receive Siponimod	4.2.6 and 6.1.1.2

Abbreviations: BSC, best supportive care; CDA, confirmed disability accumulation; ERG, Evidence Review Group; HA, highly active; ITT, intention-to-treat; RRMS, relapsing remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

6.3.1. Deterministic analysis

Table 54: ERG's preferred base case results (ITT population)

Outcomes **Ponesimod vs** **Comparator**	**ERG base case**			**Company base case**
	Incremental QALYs	**Incremental costs (£)**	**ICER (£/QALY)**	**ICER (£/QALY)**
Teriflunomide 14mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fumarate 240mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30mcg IM	[redacted]	[redacted]	[redacted]	[redacted]

Outcomes Ponesimod vs Comparator	ERG base case			Company base case
Interferon beta-1a 44mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1b 250mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab 600mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Peginterferon beta-1a 125mcg SC	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention to treat; QALY, quality-adjusted life year

Table 55: ERG's preferred base case results (HA RRMS population)

Outcomes Ponesimod vs Comparator	ERG base case			Company base case
	Incremental QALYs	Incremental costs (£)	ICER (£/QALY)	ICER (£/QALY)
Ocrelizumab 600mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab 12mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine 3.5mg/kg PO	[redacted]	[redacted]	[redacted]	[redacted]

Outcomes **Ponesimod vs** **Comparator**	**ERG base case**			**Company base case**
Fingolimod 0.5mg PO	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ERG, Evidence Review Group; HA, highly active; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; RRMS, relapsing remitting multiple sclerosis

6.3.2. One-way sensitivity analysis

In order to test the impact of parameter uncertainty on results (based on the ERG's preferred assumptions), one-way sensitivity analyses were conducted, to vary key parameters using low and high values. Tornado diagrams are presented below for comparisons with teriflunomide and fingolimod for the ITT and HA RRMS populations, respectively. Due to the large number of comparators within this appraisal, the rest of the results have been included in Appendix D.

[redacted]

Abbreviations: EDSS, expanded disability status scale; ERG, Evidence Review Group; HA, highly active; ICER, incremental cost-effectiveness ratio; OWSA, one-way sensitivity analysis; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis;

Abbreviations: ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; QALYs, quality adjusted life years; EDSS, expanded disability status scale; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

9

Abbreviations: EDSS, expanded disability status scale; ERG, Evidence Review Group; HA, highly active; ICER, incremental cost-effectiveness ratio; OWSA, one-way sensitivity analysis; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis;

Abbreviations: ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life years; EDSS, expanded disability status scale; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis; HA RRMS, highly active relapsing remitting multiple sclerosis

6.3.3. Probabilistic sensitivity analysis

The ERG's preferred probabilistic ICERs comparing ponesimod to each comparator are presented below (alongside the company's deterministic ICER for each comparison). The ERG's probabilistic results were broadly similar to the company's deterministic results, which seemed reasonable, given that the ERGs base case only included two different assumptions.

6.3.3.1. ITT population

Table 56: ERG PSA results (ITT population)

Outcomes Ponesimod vs Comparator	Incremental costs (£)	Incremental QALYs	Probabilistic ICER (£/QALY)	Deterministic ICER (£/QALY)
Teriflunomide 14 mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Dimethyl fuarate 240 mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Glatiramer acetate 20 mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 22 mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 30 mcg IM	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1a 44 mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Interferon beta-1b 250 mcg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ocrelizumab 600 mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20 mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0 mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Peginterferon beta-1a	[redacted]	[redacted]	[redacted]	[redacted]

Outcomes **Ponesimod vs** **Comparator**	**Incremental costs (£)**	**Incremental QALYs**	**Probabilistic ICER (£/QALY)**	**Deterministic ICER (£/QALY)**
125 mcg SC				

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; PSA, probabilistic sensitivity analysis; QALYs, quality adjusted life years

10 [redacted]

Abbreviations: ERG, Evidence Review Group; ITT, intention-to-treat; QALY, quality-adjusted life year

11

[redacted]

Abbreviations: ERG, Evidence Review Group; ITT, intention-to-treat; QALYs, quality adjusted life years

6.3.3.2. HA RRMS subgroup

Table 57: ERG PSA results (HA RRMS subgroup)

Outcomes Ponesimod vs Comparator	Incremental costs (£)	Incremental QALYs	Probabilistic ICER (£/QALY)	Deterministic ICER (£/QALY)
Ocrelizumab 600mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Ofatumumab 20mg SC	[redacted]	[redacted]	[redacted]	[redacted]
Ozanimod 1.0mg PO	[redacted]	[redacted]	[redacted]	[redacted]
Alemtuzumab 12mg IV	[redacted]	[redacted]	[redacted]	[redacted]
Cladribine 3.5mg/kg PO	[redacted]	[redacted]	[redacted]	[redacted]
Fingolimod 0.5mg PO	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ERG, Evidence Review Group; HA, highly active; ICER, incremental cost-effectiveness ratio; PSA, probabilistic sensitivity analysis; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis

12

Abbreviations: ERG, Evidence Review Group; HA, highly active; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis

13

Abbreviations: ERG, Evidence Review Group; HA, highly active; QALYs, quality adjusted life years; RRMS, relapsing-remitting multiple sclerosis

6.4. Conclusions of the cost-effectiveness section

Based on the ERG's preferred assumptions in the ITT population [redacted] Compared to interferon beta 1a 22 mcg ponesimod resulted in an ICER of [redacted]. Ponesimod resulted in [redacted] compared to ocrelizumab and ofatumumab (the latter currently under appraisal by NICE); [redacted]. ERG preferences had the most impact on the results versus Interferon beta-1b 250mcg SC and peginterferon beta-1a 125mcg, with ponesimod becoming the [redacted] treatment.

In the HA population, using the ERG's preferred assumptions did not have a material impact on the base case results; i.e. ponesimod remained [redacted] when compared to [redacted] Compared to fingolimod and ozanimod, ponesimod [redacted] treatment. As in the company's base case, cladribine [redacted] ponesimod.

The ERG considered that the company had broadly used the best available evidence to inform the data and modelled assumptions, and most modelled parameters and assumptions were informed by sources used and accepted in previous NICE MS appraisals. However, the ERG nevertheless considered that these were subject to a high degree of uncertainty. In most cases the ERG were unable to identify improved sources, though tested the sensitivity of the ICERs to variations in each of the uncertainties. These analyses identified that ICERs were broadly robust to most assumptions, with the exception of clinical efficacy estimates (CDA, ARR, and discontinuation rates). As discussed in Section 3, the ERG identified considerable limitations surrounding NMAs for both the ITT and HA RRMS populations, and the true estimates for each of the included treatments could vary considerably. Sensitivity analyses showed that even small variations in clinical efficacy estimates could materially change the ICERs. This was particularly true in the HA RRMS population.

Finally, NICE should be aware that comparators in both the ITT and HA RRMS populations have patient access scheme discounts (PASs). The inclusion of comparator PAS discounts had a substantial impact on the base case cost effectiveness results (see addendum to this report).

7. END OF LIFE

The ERG considered that ponesimod does not meet NICE end of life criteria as the treatment is not indicated for people with a short life expectancy (normally defined as less than 24 months).

References

1. NICE. Siponimod for treating secondary progressive multiple sclerosis. Technology appraisal guidance [TA656] 2020.
2. NICE. Alemtuzumab for treating highly active relapsing remitting multiple sclerosis [TA312]: technology appraisal guidance 2020. Available from: https://www.nice.org.uk/guidance/ta312.
3. Public Health England. Multiple sclerosis: prevalence, incidence and smoking status - data briefing 2020. Available from: https://www.gov.uk/government/publications/multiple-sclerosis-prevalence-incidence-and-smoking-status/multiple-sclerosis-prevalence-incidence-and-smoking-status-data-briefing.
4. MS Society. What is MS? 2021. Available from: https://www.mssociety.org.uk/about-ms/what-is-ms.
5. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018;17(2):162-73.
6. NHS. Treatment Algorithm for Multiple Sclerosis Disease-Modifying Therapies. NHS England Reference: 170079ALG. 2019. Available from: https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2019/03/Treatment-Algorithm-for-Multiple-Sclerosis-Disease-Modifying-Therapies-08-03-2019-1.pdf.
7. Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. Lancet. 2018;391(10130):1622-36.
8. Coyle PK. Commentary: The Multiple Sclerosis Controversy: Is It Escalation or Induction High Efficacy? Neurotherapeutics. 2020;17(3):971-2.
9. Gajofatto A, Benedetti MD. Treatment strategies for multiple sclerosis: When to start, when to change, when to stop? World J Clin Cases. 2015;3(7):545-55.
10. Díaz C, Zarco LA, Rivera DM. Highly active multiple sclerosis: An update. Multiple Sclerosis and Related Disorders. 2019;30:215-24.
11. NICE. Beta interferons and glatiramer acetate for treating multiple sclerosis [TA527]: technology appraisal guidance 2018. Available from: nice.org.uk/guidance/ta527.
12. NICE. Ozanimod for treating relapsing-remitting multiple sclerosis [ID1294]; In development [GID-TA10299]. 2021.
13. NICE. Ofatumumab for treating relapsing multiple sclerosis; Technology appraisal guidance [TA699]. 2021.
14. NICE. Fingolimod for the treatment of highly active relapsing–remitting multiple sclerosis [TA254]: technology appraisal guidance 2012. Available from: nice.org.uk/guidance/ta254.
15. Chaudhry BZ, Cohen JA, Conway DS. Sphingosine 1-Phosphate Receptor Modulators for the Treatment of Multiple Sclerosis. Neurotherapeutics. 2017;14(4):859-73.
16. Coles AL, M; Giovannoni, G; Anderson, P; Dorsey-Campbell, R; Qualie, M. ABN guidance on the use of disease modifying therapies in multiple sclerosis in response to the threat of a coronavirus epidemic Association of British Neurologists; 2020.
17. Giovannoni G, Comi G, Cook S, Rammohan K, Rieckmann P, Soelberg Sorensen P, et al. A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. The New England journal of medicine. 2010;362(5):416-26.
18. Derfuss T, Bergvall NK, Sfikas N, Tomic DL. Efficacy of fingolimod in patients with highly active relapsing-remitting multiple sclerosis. Curr Med Res Opin. 2015;31(9):1687-91.
19. Actelion Pharmaceuticals Ltd. Data on file - Study AC-058B301: OPTIMUM Final Clinical Study Report 2020 5 Feb.
20. Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, M-I; M, et al. Chapter 4: Searching for and selecting studies. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions2021.
21. SIGN. Search filters. Available from: https://www.sign.ac.uk/what-we-do/methodology/search-filters/.
22. Olsson T, Boster A, Fernandez O, Freedman MS, Pozzilli C, Bach D, et al. Oral ponesimod in relapsing-remitting multiple sclerosis: a randomised phase II trial. J Neurol Neurosurg Psychiatry. 2014;85(11):1198-208.
23. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol. 2005;58(6):840-6.
24. 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(2):292-302.
25. Palace J, Bregenzer T, Tremlett H, Oger J, Zhu F, Boggild M, et al. UK multiple sclerosis risk-sharing scheme: a new natural history dataset and an improved Markov model. BMJ Open. 2014;4(1):e004073.
26. NICE. Ocrelizumab for treating relapsing–remitting multiple sclerosis: technology appraisal guidance. 2018.
27. NICE. Cladribine tablets for treating relapsing–remitting multiple sclerosis [TA493]: technology appraisal guidance 2017. Available from: nice.org.uk/guidance/ta493.

28. Tong J, Zou Q, Chen Y, Liao X, Chen R, Ma L, et al. Efficacy and acceptability of the S1P receptor in the treatment of multiple sclerosis: a meta-analysis. Neurol Sci. 2021;42(5):1687-95.
29. Novartis Europharm Ltd. GILENYA: Summary of Product Characteristics 2020.
30. Janssen Pharmaceutical Co. OPTIMUM Clinical Study Report 2019.
31. Krysko K, O'Connor P. Measuring Disability Progression with the Multiple Sclerosis Functional Composite. European Neurological Review. 2011;6:31-5.
32. Hudgens S, Schüler R, Stokes J, Eremenco S, Hunsche E, Leist TP. Development and Validation of the FSIQ-RMS: A New Patient-Reported Questionnaire to Assess Symptoms and Impacts of Fatigue in Relapsing Multiple Sclerosis. Value in Health. 2019;22(4):453-66.
33. Swallow E, Patterson-Lomba O, Yin L, Mehta R, Pelletier C, Kao D, et al. Comparative safety and efficacy of ozanimod versus fingolimod for relapsing multiple sclerosis. J Comp Eff Res. 2020;9(4):275-85.
34. European Medicines Agency. Guideline on clinical investigation of medicinal products 5 for the treatment of Multiple Sclerosis. 2012. Contract No.: EMA/CHMP/771815/2011, Rev 2.
35. Meyer-Moock S, Feng YS, Maeurer M, Dippel FW, Kohlmann T. Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. BMC Neurol. 2014;14:58.
36. Cohen JA, Comi G, Selmaj KW, Bar-Or A, Arnold DL, Steinman L, et al. Safety and efficacy of ozanimod versus interferon beta-1a in relapsing multiple sclerosis (RADIANCE): a multicentre, randomised, 24-month, phase 3 trial. Lancet Neurol. 2019;18(11):1021-33.
37. Hauser SL, Bar-Or A, Comi G, Giovannoni G, Hartung HP, Hemmer B, et al. Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. The New England journal of medicine. 2017;376(3):221-34.
38. Coles AJ, Twyman CL, Arnold DL, Cohen JA, Confavreux C, Fox EJ, et al. Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. Lancet. 2012;380(9856):1829-39.
39. 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. The New England journal of medicine. 2010;362(5):387-401.
40. Klawiter E, Cross A, Naismith R. The present efficacy of multiple sclerosis therapeutics: Is the new 66% just the old 33%? Neurology. 2009;73:984-90.
41. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting Indirect Treatment Comparisons and Network Meta-Analysis for Health-Care Decision Making: Report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 1. Value in Health. 2011;14(4):417-28.
42. Rouse B, Chaimani A, Li T. Network meta-analysis: an introduction for clinicians. Intern Emerg Med. 2017;12(1):103-11.
43. Kalincik T, Brown JWL, Robertson N, Willis M, Scolding N, Rice CM, et al. Treatment effectiveness of alemtuzumab compared with natalizumab, fingolimod, and interferon beta in relapsing-remitting multiple sclerosis: a cohort study. The Lancet Neurology. 2017;16(4):271-81.
44. Melendez-Torres GJ, Armoiry X, Court R, Patterson J, Kan A, Auguste P, et al. Comparative effectiveness of beta-interferons and glatiramer acetate for relapsing-remitting multiple sclerosis: systematic review and network meta-analysis of trials including recommended dosages. BMC Neurology. 2018;18(1):162.
45. Walters Kluwer. Expert Search - Adverse Effects Available from: https://tools.ovid.com/ovidtools/expertsearches.html.
46. Schünemann H, Higgins J, Vist G, Glasziou P, Akl E, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions 2021.
47. Druart C, El Sankari S, van Pesch V. Long-term safety and real-world effectiveness of fingolimod in relapsing multiple sclerosis. Patient Relat Outcome Meas. 2018;9:1-10.
48. Joni SS, Cheshmavar M, Shoureshi P, Zamani Z, Taoosi N, Akbari M, et al. Effects of fingolimod treatments on alanine transaminase and aspartate transaminase levels in patients with multiple sclerosis. Int J Physiol Pathophysiol Pharmacol. 2020;12(3):88-94.
49. CADTH. Strings Attached: CADTH's Database Search Filters 2019. Available from: https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters.
50. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313(7052):275-83.
51. 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(1):54-60.
52. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
53. NICE. Ocrelizumab for treating relapsing–remitting multiple sclerosis [TA533]: technology appraisal guidance 2018. Available from: nice.org.uk/guidance/ta533.
54. NICE. Peginterferon beta-1a for treating relapsing–remitting multiple sclerosis [TA624]: Technology appraisal guidance 2020. Available from: https://www.nice.org.uk/guidance/ta624.
55. NICE. Teriflunomide for treating relapsing–remitting multiple sclerosis [TA303]: technology appraisal guidance 2014. Available from: nice.org.uk/guidance/ta303.

56. NICE. Dimethyl fumarate for treating relapsing-remitting multiple sclerosis [TA320]: technology appraisal guidance 2014. Available from: nice.org.uk/guidance/ta320.
57. Mauskopf J, Fay M, Iyer R, Sarda S, Livingston T. Cost-effectiveness of delayed-release dimethyl fumarate for the treatment of relapsing forms of multiple sclerosis in the United States. J Med Econ. 2016;19(4):432-42.
58. 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. The New England journal of medicine. 2012;367(12):1098-107.
59. NICE. Natalizumab for the treatment of adults with highly active relapsing–remitting multiple sclerosis [TA127]: technology appraisal guidance 2007. Available from: nice.org.uk/guidance/ta127.
60. NICE. Daclizumab for treating relapsing-remitting multiple sclerosis [TA441]: Technology appraisal guidance 2017. Available from: https://www.nice.org.uk/guida nce/ta441.
61. Office for National Statistics. National life tables, United Kingdom, 2017-2019 2020. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables.
62. Pokorski RJ. Long-term survival experience of patients with multiple sclerosis. J Insur Med. 1997;29(2):101-6.
63. Thompson A, Kobelt G, Berg J, Capsa D, Eriksson J, Miller D, et al. New insights into the burden and costs of multiple sclerosis in Europe: Results for the United Kingdom. Mult Scler. 2017;23(2_suppl):204-16.
64. Acaster S, Perard R, Chauhan D, Lloyd A. A forgotten aspect of the NICE reference case: an observational study of the health related quality of life impact on caregivers of people with multiple sclerosis. BMC Health Serv Res. 2013;13:346.
65. Gani R, Giovannoni G, Bates D, Kemball B, Hughes S, Kerrigan J. Cost-effectiveness analyses of natalizumab (Tysabri) compared with other disease-modifying therapies for people with highly active relapsing-remitting multiple sclerosis in the UK. Pharmacoeconomics. 2008;26(7):617-27.
66. British National Formulary. Multiple sclerosis. Available from: https://bnf.nice.org.uk/treatment-summary/multiple-sclerosis.html.
67. NICE. Cladribine tablets for treating relapsing–remitting multiple sclerosis[ID64]: Committee Papers 2017. Available from: https://www.nice.org.uk/guidance/ta616/evidence/committee-papers-for-ta493-pdf-7021081261.
68. Scottish Medicines Consortium. cladribine (Mavenclad) - SMC 1300/18 2018. Available from: https://www.scottishmedicines.org.uk/medicines-advice/cladribine-mavenclad-fullsubmission-130018/.
69. 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(5):386-9.
70. Dee A, Hutchinson M, De La Harpe D. A budget impact analysis of natalizumab use in Ireland. IrJMedSci. 2012(181):199-204.

Appendix A: Additional searches conducted by the ERG

Additional Medline search strategy for multiple sclerosis NMAs

This was a partial (modified) update of the searches used in Melendez-Torres (2018)[44], limited to papers published in 2016 onwards. The search was also translated into Embase.

1. exp Multiple Sclerosis/

2. multiple sclerosis.tw.

3. 1 or 2

4. (metaanalys* or "meta analys*" or "meta-analys*").tw.

5. meta analysis.pt.

6. 4 or 5

7. 3 and 6

8. limit 7 to yr="2016 -Current"

Additional Medline strategy for adverse events

This search uses the broad adverse effects expert search filter from Ovid (Adverse Effects - Medline – Broad[45]) without any study type filter. The search was also translated into Embase using the equivalent Ovid search filter (Adverse Effects – Embase – Broad[45]).

1. exp "Drug-Related Side Effects and Adverse Reactions"/ or adverse.ti,ab,kf. or side effect?.ti,ab,kf. or adverse effects.fs. or exp drug overdose/ or overdos*.ti,ab,kf. or exp drug misuse/ or misus*.ti,ab,kf. or exp Substance-Related Disorders/ or abus*.ti,ab,kf. or exp pregnancy/ or pregnan.ti,ab,kf. or exp pregnancy complications/ or exp lactation/ or exp lactation disorders/ or exp breast feeding/ or (exp milk, human/ and exp secretion/) or exp fertility/ or exp infertility/ or exp reproduction/ or exp fetus/ or exp embryonic structures/ or terat*.ti,ab,kf. or drug efficacy.ti,ab,kf. or therapeutic efficacy.ti,ab,kf. or drug withdrawal.ti,ab,kf. or exp medication errors/ or exp death/ or death*.ti,ab,kf. or fatal*.ti,ab,kf. or exp drug interactions/ or exp carcinogens/ or carcinogen*.ti,ab,kf. or mutagen*.ti,ab,kf. or exp "Off-Label Use"/ or exp occupational exposure/ or toxicity.fs. or toxic*.ti,ab,kf. or pharmacotox*.ti,ab,kf. or

neurotox*.ti,ab,kf. or cardiotox*.ti,ab,kf. or nephrotox*.ti,ab,kf. or immunotox*.ti,ab,kf. or hepatotox*.ti,ab,kf. or cytotox*.ti,ab,kf. or immunocytotox*.ti,ab,kf. or intoxicat*.ti,ab,kf. or exp "Congenital, Hereditary, and Neonatal Diseases and Abnormalities"/ or drug treatment failure.ti,ab,kf. or drug toxicity.ti,ab,kf. or exp case report/ or case report?.ti,ab,kf. or exp environmental exposure/ or treatment contraindication.ti,ab,kf. or exp contraindications, drug/ or exp "Wounds and Injuries"/ or suicid*.ti,ab,kf. or exp poisoning/ or poisoning.fs. or exp drug tolerance/ or exp treatment failure/ or exp drug resistance/ or exp substance-related disorders/

2. Ponesimod/ or (ponesimod$2 or "act 128800" or act128800 or act-128800 or "rg 3477" or rg3477 or 854107-55-4).ti,ab,kw,du,rn.

3. 1 and 2

Additional Medline strategy for cost effectiveness (adding four additional technologies)

This was the search as used in the CS but with the existing drug terms removed and the four missing drug terms (for siponimod, ozanimod, ofatumumab and ponesimod) added. The search was also translated into Embase.

1. exp multiple sclerosis/ or (multiple sclerosis or ((primary or progressive or secondary) and (relapsing or remittent or (relapsing and remitting)) and multiple and sclerosis) or ppms or spms or rrms).tw,kw.

2. ("health utilit$" or "health state utility$" or "utility score$" or "utility valu$").tw,kw.

3. ("standard gamble" or "time trade-off" or "time tradeoff" or tto or "visual analog$ scale$" or "patient preference$" or preference$).tw,kw.

4. (eq-5d or eq5d or euroqol or "health utility$ index" or hui or sf-6d or "short form 6d" or "quality of well-being scale$" or "utility assessment" or qaly$ or "quality adjusted life year$" or utility$).tw,kw.

5. 2 or 3 or 4

6. exp economics/ or exp cost control/ or exp cost of illness/ or exp drug costs/ or exp hospital costs/ or exp health care costs/ or exp socioeconomic factors/ or exp health care economics and organizations or exp fee and charges/ or exp budgets/

7. (fiscal or funding or financ$ or economic$ or pharmacoeconomic$ or pric$).tw,kw.

8. 6 or 7

9. exp patient acceptance of health care/

10. ("health care use" or "healthcare use" or "health service$ use" or "health care utili?ation" or "healthcare utili?ation" or "health resource utili?ation" or "health service$ utili?ation" or "resource use" or "medical leave" or "work disability").tw,kw.

11. exp absenteeism/ or absenteeism.tw,kw.

12. exp retirement/ or retirement.tw,kw.

13. exp sick leave/

14. exp workers' compensation/

15. ("disability absence" or "illness day" or "sick day" or "work absence" or "work day loss" or "work incapacity" or "work loss" or "work time loss" or "workmans compensation" or "workers compensation" or "productivity loss" or "work impairment" or "sickness absence" or "lost days" or "productivity").tw,kw.

16. or/9-15

17. ("cost minimi?ation analys$" or ("cost-minimi?ation" adj analys$)).tw,kw.

18. exp cost benefit analysis/

19. (("cost benefit" or "cost-benefit") adj analys$).tw,kw.

20. (("cost utility" or "cost-utility" or "cost-effective$") adj analys$).tw,kw.

21. exp cost utility analysis/ or exp economic evaluation/

22. (cost adj effective$ adj analys$).tw,kw.

23. or/17-22

24. ((economic or pharmacoeconomic) adj (evaluation or assessment or analys?s or stud$)).tw,kw.

25. (cea or cma or cba or cua or cca).tw,kw.

26. exp decision theory/ or exp decision trees/

27. "decision tree".tw,kw.

28. ((economic or cohort or transition) adj model).tw,kw.

29. (markov or deterministic).tw,kw.

30. ((transition adj probability$) or (health adj stat$) or (sensitivity adj analys$) or (health adj outcome) or (("patient level" or "patient-level" or "discrete event" or "discrete-event") adj simulat$)).tw,kw.

31. (incremental-cost or icer or qaly or daly or wtp or tto).tw,kw.

32. or/24-31

33. 5 or 8 or 16 or 23 or 32

34. Ponesimod/ or (ponesimod$2 or "act 128800" or act128800 or act-128800 or "rg 3477" or rg3477 or 854107-55-4).ti,ab,kw.

35. (siponimod$2 or 1230487-00-9 or 1230487-85-0 or "baf 312" or baf312 or mayzent$2 or nvpbaf312nx).ti,ab,kw.

36. (ozanimod$2 or 1306760-87-1 or 1618636-37-5 or “rpc 1063” or rpc1063 or Zeposia$2).ti,ab,kw.

37. (ofatumumab$2 or arzerra$2 or "gsk 1841157" or gsk1841157 or "humac CD20" or "HuMax CD20" or HuMax-CD20 or HuMaxCD20 or "omb 157" or omb157 or 679818-59-8).ti,ab,kw.

38. or/36-39

39. 1 and 33 and 38

Appendix B: NMA methods used in the HA RRMS population in previous NICE appraisals

A brief overview of the included trials and methodology used to evaluate treatments for RRMS in the HA population in previous appraisals and publications of NMAs is provided in Table 58 below.

Table 58: NMA methods used to evaluate treatments for HA RRMS in previous NICE HTA appraisals

Appraisal (NMA publications)	Included publications (*reporting on HA MS; ^included in NMA)	Included treatments (note that not all treatments were included in all analyses)	Definition of HA	Assumptions
Alemtuzumab [TA312] 2013[2]	FREEDOMS*^ CARE-MS: II*^ TRANSFORMS *^ TENERE TEMSO & TOWER TEMSO & TENERE & TOWER	Terifuonimide 7mg Terifuonimide 14mg Interferon 44mg* Interferon 1a 30mg* Alemtuzumab 12mg* Fingolimod 0.5mg* Fingolimod 1.25mg Placebo	HA despite interferon use, although various trial definitions accepted	In their response to ACD, the company conducted NMAs in the highly active population. Treatments relevant to HA populations were included, along with treatments and evidence in the ITT population that were added to complete the networks where necessary. The ERG noted that heterogeneity in population definitions and the inclusion of indirect evidence increased uncertainty in the effect estimates.
Ocrelizumab [TA533] 2018[53]	CARE-MS II (ALE)*^ CONFIRM & DEFINE (pooled) (DMF) *^ TRANSFORMS (FIN) *^ FREEDOMS & FREEDOMS II (pooled) (FIN) *^ OPERA I & II (pooled) (OCR) *^	IFNB-1b 250 µg SC EOD IFNB-1a 22 µg TIW Glatiramer acetate 20 mg QD IFNB-1a 30 µg IM QW IFNB-1a 44 µg SC TIW Fingolimod 0.5 mg QD Alemtuzumab 12 mg Ocrelizumab 600 mg Daclizumab 150 mg Q4W	Populations treated with INFBs or GA for at least one year with (1) ≥ 1 relapse(s) in previous year, (2) ≥ 1 Gd+ lesion on brain MRI at baseline, or (3) ≥ 9 T2 hyperintense lesions on brain MRI at baseline	Networks for the HA population were disconnected. To connect the networks, ITT data from studies investigating ABCR treatments (IFNB-1a [Avonex], IFNB-1b [Betaferon], glatiramer acetate [Copaxone], and IFNB-1a [Rebif]) were included. In addition, where studies did not report CDA-6 data from CDA-3 was used to complete the network. The ERG suggested that the results of the HA analyses be interpreted with caution. Furthermore, the committee stated a preference for evidence for CDA-3 in the ITT population to be excluded where evidence for CDA-6 in the same comparison was not also available.

Appraisal (NMA publications)	Included publications (*reporting on HA MS; ^included in NMA)	Included treatments (note that not all treatments were included in all analyses)	Definition of HA	Assumptions
	TEMSO & TOWER (pooled) (TER) *^ SELECT (DAC) *^ DECIDE (DAC) *^ N=12*			All-cause discontinuation rates were assumed to be the same as the whole RRMS population. Unlicensed doses and treatment regimens were excluded.
Cladribine [TA616] 2017[67] *Need access to appendix D of CS*	AFFIRM*^ CONFIRM*^ DEFINE *^ FREEDOMS*^ TOWER*^ TRANSFORMS*^ CLARITY*^ CAMMS223*^ CARE-MS*^ PRISMS*^ CARE-MS I*^	Alemtuzumab dimethyl fumarate fingolimod glatiramer acetate 20mg IFN-β-1a 30 µg IFNβ-1a 44 µg Natalizumab teriflunomide 7 mg/14 mg Cladribine	Two definitions were explored: HA (licensed population): 1 relapse in the previous year while on treatment and ≥1 T1 Gd+ lesion or ≥9 T2 lesions OR populations with ≥2 relapses in the previous year whether on treatment or not Sub-optimally treated: Populations with ≥1 relapses in the prior year whether on treatment or not. A 2nd (very limited NMA) is reported for 'sub-optimally treated MS' – relapse despite treatment.	Assumed that subgroups in CLARITY were comparable to those in other trials despite differences in definitions of subgroups from previous NICE guidance. NMA conducted for HA population but not for the sub-optimally treated, due to small number of populations in relevant cladribine trials that met this criteria, and the paucity of evidence available from other trials. In the HA population NMA, it was assumed that outcomes were comparable between trials despite differences in outcome measures in CLARITY and clinical trials for other treatments. While the ERG expressed concerns about the validity of this approach, the committee accepted that the results were sufficiently similar. A meta-regression was conducted to estimate effect sizes for the sub-optimally treated population adjusted by baseline disease severity. This analysis assumed a linear relationship between baseline severity and treatment efficacy. The ERG recognised that this approach was used to address heterogeneity across trials; however, noted that the analysis was still subject to the other limitations associated with the company's NMA. The ERG also flagged indications that

Appraisal (NMA publications)	Included publications (*reporting on HA MS; ^included in NMA)	Included treatments (note that not all treatments were included in all analyses)	Definition of HA	Assumptions
				the reported effect sizes were influenced by other effect modifiers, thus undermining the validity of the analysis.
Fingolimod [TA254] 2011[14]	AFFIRM EVIDENCE FREEDOMS INCOMIN MSCRG IFNB MS Study Group PRISMS TRANSFORMS BEYOND BECOME REGARD Hurwitz 2008 Etemadifar 2006 Wroe 2005 Saida 2005 Johnson 1995 Comi 2001 Bornstein 1987	fingolimod 0.5 mg* natalizumab 300mg* interferon beta-1a* 22mcg interferon beta-1a* 44mcg interferon beta-1a* 30mcg interferon beta-1b* 250mcg glatiramer acetate 20mg placebo Interferon 1b 50mcg Interferon 1b 500mcg	Populations who have an unchanged or increased relapse rate or ongoing severe relapses compared with the previous year despite treatment with beta interferon (including RES RRMS). The company suggested that the populations included in the indirect comparison were populations with RRMS regardless of previous treatment, rather than from those whose disease had a suboptimal response to disease-modifying therapy.	Fingolimod was the first DMT to be recommended specifically in the HA population. The company conducted a NMA using the active RRMS population, though this was not used to inform the economic analysis for the HA population due to indirectness/heterogeneity of the trial poplations. Instead, an indirect comparison between fingolimod and placebo was generated from two of the included trials (FREEDOMS and TRANSFORMS). As fingolimod was the first treatment to be recommended by NICE for the HA population, and so no NMA was conducted in the HA population. The results of the NMA in the active population were not used in the model. Notably for this appraisal, discontinuation due to AEs was evaluated instead of all-cause discontinuation, due to variability in the reasons for exclusion used across trials. Unadjusted data was used, also due to variability in covariates applied in the included trials. For the CDA analysis, the company excluded three trials that didn't report CDA-3, but these were included in a sensitivity analysis.

Appraisal (NMA publications)	Included publications (*reporting on HA MS; ^included in NMA)	Included treatments (note that not all treatments were included in all analyses)	Definition of HA	Assumptions
Ozanimod [TA1294] 2021[12]; note that information reported is based on documents published following AC2 (May 2021)	NA	NA	Those with an unchanged or increased relapse rate, or ongoing severe relapses compared with the previous year despite treatment with at least one DMT	No separate NMA conducted for the HA population, and the company did not present evidence for comparators used in the HA population, as ozanimod was not originally positioned for these populations. In the active population, the ERG did not consider that heterogeneity across the included trials to have a major impact on the results of the analyses. The company conducted an analysis combining CDA-3 and CDA-6, to account for older trials that did not report CDA-6. However, the ERG considered that the assumption of a proportional relationship between the CDA-3 and CDA-6 hazard ratios for ozanimod appeared to have been violated.
Ofatumumab [TA1677][13] 2021; note that information reported is based on documents published following AC1 (April 2021)	NA	NA	-	The company's feasibility assessment concluded that it was not possible to conduct NMAs in the HA or RES populations, due to heterogeneity between the trials and the paucity of data in completing the network. In the ACD, it was reported that clinical experts had suggested that HA and RES definitions may not be used in practice, but rather clinicians would consider treatment and relapse history. On that basis, the committee concluded that it was reasonable to consider the RRMS in full.

Abbreviations: ABCR, immunomodulators; ACD, appraisal consultation document; ALE, alemtuzumab; CDA, confirmed disability accumulation; CS, Company Submission; DAC, daclizumab; DMF, dimethyl fumarate; EOD, every other day; ERG, Evidence Review Group; FIN, fingolimod; GA, Glatiramer acetate; HA, highly active; HTA, health technology assessment; IFNB interferon beta; ITT, intention-to-treat; MRI, magnetic resonance imaging; MS, multiple sclerosis; NA, not applicable; NICE, National Institute for Health and Care Excellence; NMA, network meta analysis; OCR, ocrelizumab; QD, once a day; QW, weekly; RES, rapidly evolving severe; RRMS, relapsing remitting multiple sclerosis; SC subcutaneous; TER, teriflunomide, TIW, three times weekly

Appendix C: Comparison of relative effects in ITT and HA populations

Table 59: Comparison of relative effects on annualised relapse rate between the highly active and intention-to-treat populations in the NMA

Comparison	HA subgroup [a,b]	ITT subgroup [a,b]	Difference [c]	Ratio [d]
OCR vs ALE	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs CLA	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs PON	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs TER	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs CLA	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs PON	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs TER	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs PON	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs TER	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
PON vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
PON vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
PON vs TER	[redacted]	[redacted]	[redacted]	[redacted]
PON vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
PON vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs TER	[redacted]	[redacted]	[redacted]	[redacted]

Comparison	HA subgroup [a,b]	ITT subgroup [a,b]	Difference [c]	Ratio [d]
FIN vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
IFNB 1a 44 vs TER	[redacted]	[redacted]	[redacted]	[redacted]
IFNB 1a 44 vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
IFNB 1a 44 vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
TER vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
TER vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
PBO vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ALE, alemtuzumab 12 mg once daily; CLA, cladribine 3.5 mg/kg once daily; FIN, fingolimod 0.5 mg once daily; HA, highly active; IFNB-1a 30 µg , interferon beta-1a 30 µg intramuscular once weekly; IFNB-1a 44 µg , interferon beta-1a 44 µg subcutaneously three times weekly; ITT, intention-to-treat; NMA, network meta-analysis; OCR, ocrelizumab 600 mg every six months; PBO, placebo; PON, ponesimod 20 mg once daily; TER, teriflunomide 14 mg once daily

Notes:

[a] Data are point estimates of relative risk (lower 95% confidence interval; upper 95% confidence interval)

[b] Cells with grey shading denote significant results

[c] Difference of HA point estimate – ITT point estimate

[d] Ratio of HA point estimate/ITT point estimate

Table 60: Comparison of relative effects on confirmed disability accumulation at 3 months between the highly active and intention-to-treat populations in the NMA

Comparison	HA subgroup [a,b]	ITT subgroup [a,b]	Difference [c]	Ratio [d]
CLA vs OCR	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs PON	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs TER	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs PON	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs TER	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]

Comparison	HA subgroup [a,b]	ITT subgroup [a,b]	Difference [c]	Ratio [d]
OCR vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
PON vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
PON vs TER	[redacted]	[redacted]	[redacted]	[redacted]
PON vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
PON vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
PON vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs TER	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs IFNB 1a-44 µg	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs IFNB 1a-30 µg	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
TER vs IFNB 1a-44 µg	[redacted]	[redacted]	[redacted]	[redacted]
TER vs IFNB 1a-30 µg	[redacted]	[redacted]	[redacted]	[redacted]
TER vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
IFNB 1a-44 µg vs IFNB-1a 30 µg	[redacted]	[redacted]	[redacted]	[redacted]
IFNB-1a 44 µg vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
IFNB-1a 30 µg vs PBO	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CLA, cladribine 3.5 mg/kg once daily; FIN, fingolimod 0.5 mg once daily; HA, highly active; IFNB-1a 30 µg , interferon beta-1a 30 µg intramuscular once weekly; IFNB-1a 44 µg , interferon beta-1a 44 µg subcutaneously three times weekly; ITT, intention-to-treat; NMA, network meta-analysis; OCR, ocrelizumab 600 mg every six months; PBO, placebo; PON, ponesimod 20 mg once daily; TER, teriflunomide 14 mg once daily

Notes:

[a] Data are point estimates of hazard ratios (lower 95% confidence interval; upper 95% confidence interval)

[b] Cells with grey shading denote significant results

[c] Difference of HA point estimate – ITT point estimate

[d] Ratio of HA point estimate/ITT point estimate

Table 61: Comparison of relative effects on confirmed disability accumulation at 6 months between the highly active and intention-to-treat populations in the NMA

Comparison	HA subgroup [a,b]	ITT subgroup [a,b]	Difference [c]	Ratio [d]
CLA vs ALE	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs OCR	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs PON	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs TER	[redacted]	[redacted]	[redacted]	[redacted]

Comparison	HA subgroup [a,b]	ITT subgroup [a,b]	Difference [c]	Ratio [d]
CLA vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
CLA vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs OCR	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs PON	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs TER	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
ALE vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs PON	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs TER	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
OCR vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
PON vs FIN	[redacted]	[redacted]	[redacted]	[redacted]
PON vs TER	[redacted]	[redacted]	[redacted]	[redacted]
PON vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
PON vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs TER	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
FIN vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
TER vs IFNB-1a 44 µg	[redacted]	[redacted]	[redacted]	[redacted]
TER vs PBO	[redacted]	[redacted]	[redacted]	[redacted]
IFNB-1a 44 µg vs PBO	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ALE, alemtuzumab 12 mg once daily; CLA, cladribine 3.5 mg/kg once daily; FIN, fingolimod 0.5 mg once daily; HA, highly active; IFNB-1a 44 µg , interferon beta-1a 44 µg subcutaneously three times weekly; ITT, intention-to-treat; NMA, network meta-analysis; OCR, ocrelizumab 600 mg every six months; PBO, placebo; PON, ponesimod 20 mg once daily; TER, teriflunomide 14 mg once daily

Notes:

[a] Data are point estimates of hazard ratios (lower 95% confidence interval; upper 95% confidence interval)

[b] Cells with grey shading denote significant results

[c] Difference of HA point estimate – ITT point estimate

[d] Ratio of HA point estimate/ITT point estimate

Appendix D: ERG One-way sensitivity analysis

This section contains additional tornado plots displaying the results of one-way sensitivity analyses conducted by the ERG for ponesimod as compared to its comparators. Due to the large number of comparators included within this appraisal, the ERG opted only include the one-way sensitivity analysis results versus teriflunomide and fingolimod in the main report (for the ITT and HA RRMS populations respectively). The remaining results, have been included here for completeness.

ITT population

[redacted]

Abbreviations: EDSS, expanded disability status scale; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

15

[redacted]

Abbreviations: EDSS, expanded disability status scale; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

16

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

17

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]19[redacted]

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

20

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]21

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]22[redacted]

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]23[redacted]

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

HA RRMS subgroup

[redacted]24[redacted]

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]25

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]27[redacted]

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]28

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis

[redacted]29

[redacted]

Abbreviations: EDSS, expanded disability status scale ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OWSA, one-way sensitivity analysis; QALY, quality adjusted life year; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis