



Benefit-risk Assessment of Cladribine Using Multi-criteria Decision Analysis (MCDA) for Patients With Relapsing-remitting Multiple Sclerosis

Patrick Vermersch, MD, PhD¹; Vittorio Martinelli, MD²; Claudia Pfleger, MD, PhD³; Peter Rieckmann, MD^{4,5}; Lucia Alonso-Magdalena, MD⁶; Andrew Galazka, MD⁷; Fernando Dangond, MD⁸; and Lawrence Phillips, PhD⁹

¹Department of Neurology, CHU Lille, LIRIC-INSERM U995, FHU Imminent, University of Lille, Lille, France; ²Neurological Division, San Raffaele Hospital, Milan, Italy; ³Department of Neurology, Aalborg Universitets Hospital, Aalborg, Denmark; ⁴Department of Neurology, Hospital for Nervous Diseases Medical Park Loipl, Loipl-Bischofswiesen, Germany; ⁵Department of Neurology, University of Erlangen, Erlangen, Germany; ⁶Neurology Department, Skåne University Hospital, Malmö, Sweden; ⁷Global Medical Excellence, Merck, Aubonne, Switzerland; ⁸Global Clinical Development, Neurology, EMD Serono, Inc, Billerica, Massachusetts, USA; and ⁹Department of Management, London School of Economics and Political Science, London, United Kingdom

ABSTRACT

Purpose: We applied Multi-Criteria Decision Analysis (MCDA) methods in a structured benefit–risk assessment of cladribine and newer approved disease-modifying drugs (DMDs) for patients with relapsing-remitting multiple sclerosis (RRMS).

Methods: Decision conferencing with clinical neurologists as decision makers was used to create an MCDA model that incorporated available evidence on DMDs for RRMS and clinical judgments about the relevance of the evidence. Benefit–risk assessments were conducted for DMDs in both patients with RRMS and patients with RRMS with high disease activity (HDA; defined as ≥ 2 relapses in the previous year). Treatment options included cladribine and recently approved DMDs available in European Union countries at the time of assessment (December 2015): alemtuzumab, dimethyl fumarate, fingolimod, natalizumab, and teriflunomide. To account for the relative importance of DMD effects, scores for the MCDA model were weighted to ensure that the most clinically important attributes carried more weight in the final benefit–risk calculation. The neurologists weighted different efficacy and safety profile attributes without any reference to individual

DMDs to disassociate the assessment of weights with any specific DMD. The neurologists did not do direct comparisons between DMDs.

Findings: The highest overall weighted preference value for the RRMS model was for dimethyl fumarate (63) followed closely by cladribine (62). For patients with RRMS and HDA, cladribine had the highest overall weighted preference value (76), followed by alemtuzumab (62) and natalizumab (61). The benefit–risk balance of cladribine in patients with RRMS and specifically patients with RRMS who exhibited HDA characterized by high relapse activity (≥ 2 relapses in the previous year) was more favorable than the other DMDs included in the model.

Implications: The balance of high efficacy and the safety profile makes cladribine an important treatment option to consider, both in patients with RRMS and patients with HDA. Regular, single-country meetings could be organized to explore how differences in cultural values (scores and weights) and

Accepted for publication December 28, 2018

<https://doi.org/10.1016/j.clinthera.2018.12.015>

0149-2918/\$ - see front matter

© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

updated input data might affect the usefulness of MCDA in different, real-world, dynamic clinical settings. (*Clin Ther.* 2019;41:249–260) © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: cladribine, Multi-criteria Decision Analysis, multiple sclerosis, relapsing-remitting, risk assessment.

INTRODUCTION

Relapsing-remitting multiple sclerosis (RRMS) is the main phenotype of multiple sclerosis (MS).^{1,2} Prognostic factors such as relapse rates³ and/or magnetic resonance imaging (MRI) activity⁴ in patients with RRMS have been associated with greater risk of disability progression in the future. Increased availability of new treatment options for RRMS over recent years has helped to reduce relapses and new lesions, but worsening disability is still an issue for many patients with MS, regardless of treatment.⁵ Balancing efficacy versus the risks associated with therapy is essential to the selection of a treatment option for every patient with RRMS and particularly for those patients at increased risk of relapses and disability progression.^{6,7}

A traditional view of MS therapy is that, although higher efficacy treatments should be considered for patients at higher risk, these therapies are associated with additional safety profile concerns in some patients.⁷ Basic understanding of the efficacy and safety profiles of available disease-modifying drugs (DMDs) does not provide a complete appreciation of their benefit–risk balance. The experienced clinical judgment of a treating neurologist about the clinical relevance of the underlying favorable and unfavorable effects of treatment options is required, as well as the relative trade-off between the two, for which it would be important to consider the patient's view when deciding on a treatment plan.

Recommendations from the Benefit–Risk Project at the European Medicines Agency (EMA)^{8,9} and from the Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) project of the European Commission's Innovative Medicines Agency¹⁰ suggest that multi-criteria decision analysis (MCDA) within a structured benefit–risk assessment process could clarify and

facilitate communication about drug decision making. In the work described here, we tested these recommendations by applying them to a benefit–risk assessment of cladribine (MAVENCLAD*) and 5 other newer approved DMDs for treatment of both RRMS and patients with high disease activity (HDA). To do this, we applied the group workshop approach of decision conferencing, guided by the EMA/PROTECT framework to create an MCDA model that incorporated both the available evidence for the DMDs and clinical judgments from neurologists about the relevance of that evidence.^{9,10}

METHODS

One of the authors (L. Phillips) facilitated the construction of an initial MCDA model at a 2-day decision conference in October 2015 with Merck KGaA staff providing support to gather evidence from public domain sources. A 1-day decision conference in December 2015, attended by the Merck staff and 5 independent European neurologists (authors P. Vermersch, V. Martinelli, C. Pflieger, P. Rieckmann, and L. Alonso-Magdalena) completed the model and confirmed its validity. The steps in creating, exploring, and finalizing the MCDA model are detailed in [Supplemental Section 1](#) (in the online version at doi:10.1016/j.clinthera.2018.12.015). Here, we briefly describe those stages, the first of which, establishing the decision context, was described in the Introduction above.

Alternatives

Medicinal products evaluated included cladribine (under regulatory consideration at the time of the assessment) and newer approved DMDs available at the time of the assessment: alemtuzumab, dimethyl fumarate, fingolimod, natalizumab, and teriflunomide.

Criteria

Seven favorable and 4 unfavorable effects criteria developed at the October decision conference were introduced at the December 2015 meeting to the panel of neurologists, who discussed and agreed on all 7 favorable effects and extended the 4 original unfavorable effects to 11, the result of a discussion that considered study end points, posology, and

* Trademark: MAVENCLAD (Merck KGaA, Darmstadt, Germany).

established or potential risks associated with the DMDs. Neurologists decided to be overinclusive in the number of criteria included in the final model. All criteria were organized as a hierarchical effects tree.

Neurologists agreed definitions of the effects criteria, recognizing that some flexibility was required to accommodate inconsistencies in the literature. For example, in a population of patients with RRMS exhibiting HDA, the available information reported that for most agents, data for patients having experienced ≥ 2 relapses in the previous year were available most consistently as a definition of HDA. When data for this definition of HDA were not available, but similar definitions of HDA were used (eg, patients with ≥ 2 relapses in the previous year plus at least 1 T1 gadolinium-enhanced [Gd+] lesion), this alternative definition was used (see [Supplemental Section 2](#) for more details). It was also recognized that relevant absolute reductions would be the most informative data to characterize favorable effects. However, it was not possible to obtain data consistently for all of the DMDs. Therefore, relative risk data were used for relapse rate, number of T1 Gd+ and T2 lesions, and time to 3- and 6-month confirmed Expanded Disability Status Scale (EDSS) progression in the model reported. Similarly, the model used the percentage of patients who experienced grade 4 lymphopenia, because data for other grades were not reported for all of the DMDs.

Scoring

Before the decision conference, a hierarchical search strategy obtained treatment effect data for the DMDs by identifying relevant studies in the following order: (1) European Union regulatory approval documents, (2) US regulatory approval documents, (3) reports from post-marketing surveillance, and (4) peer-reviewed publications (pivotal trials; secondary publications). Details of data and sources used in this model, including assumptions, are given in [Supplemental Section 2](#). For both favorable and unfavorable effects, if the relevant data were not found in the first data source, then the second data source was searched and so on down the hierarchy until relevant data were identified and the search was stopped. Only 1 set of data was reported for each effect for each DMD. Only trials in which the DMD was given as monotherapy were chosen, and the highest values for favorable effects attained in any trials were included in the model (ie, the highest efficacy

result). For alemtuzumab, in which the clinical program did not include placebo as a comparator, arbitrary high values were assigned by the neurology experts on MRI and clinical efficacy measures. Data on unfavorable effects selected for inclusion in the model were based on the important identified and potential risks associated with the DMDs to be assessed. Generally, data were captured at the MedDRA System Organ Class level because the terminology across the different sources varied. For the most part data were taken from 2-year studies. Further refinement of the data continued throughout the 2 decision conferences, with data from a few new sources added after the meetings. Definitions of the criteria and all input data were summarized in an effects table ([Table 1](#)).

Most of the metrics in the effects table ([Table 1](#)) were expressed as percentage reductions for the favorable effects and incidence, expressed as percentages, for the unfavorable effects. However, 4 effects were expressed differently: (1) rankings based on ease of use; criteria included oral versus intravenous administration, the number of doses required—few or many, whether monitoring is required during administration, and whether co-administration of other drugs is required; (2) durability of treatment effect, expressed as number of months of remaining efficacy after stopping the drug to reflect the acceptable interval of time (per neurologists) to wait after interrupting a therapy to switch to another without risking the appearance of new relapses; (3) malignancy, expressed as number of new cases per 100 patient-years; and (4) progressive multifocal leukoencephalopathy (PML), expressed as number of cases per 1000 patients.

MCDA deals with such differences in metric types by converting all input data to a common metric of preference value in a 2-step process. In the first step the neurologists established a preference scale for each criterion, with low and high values, including all the input data in the effects table ([Table 1](#)), assigned preference values of 0 and 100, respectively. In the second step, the performance metrics for all drugs on each scale were converted linearly to preference scores by using Hiview3 software,[†] which preserved

[†] Hiview3 software, originally developed at the London School of Economics and Political Science and now at Catalyze Ltd, www.catalyze.co.uk, performed the conversion after the ranges were established.

Table 1. Effects Table of input data for the RRMS population and the HDA subset (data shown in parentheses). The low and high figures establish the range of data for which swing weights are assessed. Final normalised weights are shown for the RRMS and HDA models. See also [Supplementary section 2](#) for details and reference sources.

Criteria	Description	Metric	Worst	Best	RRMS weight	HDA weight	Cladribine	Natalizumab	Fingolimod	Dimethyl fumarate	Alemtuzumab	Teriflunomide
	Favourable Effects											
Relapse rate	Relative reduction, compared to the control, in annualised relapse rate at 2 years	%	20	80	9.1	9.3	58 (68)	68 (81)	54 (63)	53 (60)	75 (71)	36 (19)
T2 lesions	Relative reduction in mean number of active T2 lesions per patient per scan over 2 years	%	50	95	8.1	8.8	73	83	74	85	87 (92)	53
T1 Gd+ lesions	Relative reduction in mean number of T1 Gd+ lesions per patient per scan over 2 years	%	80	100	7.0	7.8	86	92	82	94	92 (97)	80
EDSS 3 months	Relative reduction in time to 3-month confirmed EDSS progression over 2 years	%	20	70	8.6	8.3	33 (72)	42 (53)	30 (33)	38 (21)	62 (71)	31 (35)
EDSS 6 months	Relative reduction in the time to 6-month confirmed EDSS progression over 2 years	%	20	85	10.1	9.8	47 (82)	54 (64)	37 (33)	23 (21)	62 (71)	24 (35)
Ease of use	Ranking based on 4 sub-criteria*	1–4	1	4	4.0	2.9	3.5	1.5	2.0	2.5	1.0	3.0
Durability	Number of months of remaining efficacy after stopping the drug	Months	1	12	2.5	3.9	12	2	1	1	12	1
	Unfavourable Effects											
AR infections	Percentage of patients with any infections	%	75	50	2.8	2.7	51.8	73.7	65.1	60.0	70.9	61.7
AR GI effects	Percentage of patients with any GI disorder	%	50	0	3.0	2.9	31.6	0.0	43.0	44.0	49.0	45.3
Liver functions	Percentage of patients experiencing elevated liver enzymes	%	20	0	2.5	2.4	1.5	0.0	10.1	6.0	0.0	15.0
Malignancy	Number of new cases per 100 patient-years	No/100	0.45	0.20	4.0	3.9	0.370	0.320	0.400	0.375	0.370	0.200
Autoimmune disease	Percentage of patients with any autoimmune disease	%	50	0	6.0	5.9	2.0	0.0	0.0	0.0	47.0	0.0
Lymphopenia	Percentage of patients experiencing lymphopenia Grade 4	%	55	0	6.5	6.4	0.7	0.0	18.0	0.13	52.1	0.0
AV block	Percentage of patients with first degree AV block	%	5	0	2.0	2.0	0.0	0.0	4.7	0.0	0.0	0.0
Bradycardia	Percentage of patients with bradycardia	%	4	0	1.5	1.5	0.2	0.0	4.1	0.0	0.3	0.0
Serious infections	Percentage of patients with any serious infection	%	3	1.5	7.0	6.8	2.5	2.4	1.6	2.0	2.7	2.2
Herpetic infections	Percentage of herpetic infections	%	16	0	5.0	4.9	7.9	8.0	9.0	0.0	15.7	0.5
PML	Number of cases of PML per 1,000 patients	No/1000	2	0	10.1	9.8	0.000	2.100	0.104	0.029	0.000	0.000

* (1) oral vs iv, (2) few or many doses, (3) monitoring during administration (Y or N) and (4) Co-administration of other drugs (Y or N).

AR=adverse reaction; AV=atrioventricular; EDSS=expanded disability status scale; Gd+=gadolinium enhanced; GI=gastrointestinal; HDA=high disease activity; iv=intravenous; PML=progressive multifocal leukoencephalopathy.

their relative positions on the input metrics. The linear approach was taken as an appropriate approximation for the criteria under consideration where an effect is measured by proportions of patients, to ensure all patients are treated equally. The number of cases of atrioventricular block, bradycardia, and PML are so small they would all appear on a linear portion of the value function. Thus, for the relative reduction percentages of the favorable effects, a larger input metric received a higher preference score, while for the unfavorable effects, a higher percentage or incidence was assigned a lower preference score. Consequently, higher numbers for unfavorable effects represent better safety. The resulting preference values therefore represent the strength of preference for the performance of each of the 6 drugs on each of the 18 effect scales.

Weighting the Effects Criteria

The differences in preference between 0 and 100 on some of the scales are clearly more clinically relevant than on others, so in MCDA the scales are weighted to ensure the comparability of a unit of preference from one scale to the next. This enables weighted preference values to be summed to give an overall benefit–safety figure for each DMD. The process of doing this is called swing weighting, because it requires the swing in clinical desirability from the 0-to-100 point on one scale to be compared with that on another. This process of weighting requires consideration of how large the range is on a given criterion, and an assessment of how clinically important that range is. The process of weighting was helped by Hiview's projected vertical graphic scales indicating only the worst and best values of the input data. These displays enabled neurologists to judge the added clinical value of moving from the worst to the best scale points on one criterion compared with another, a paired-comparison approach that minimizes the cognitive task of considering too many criteria at once.

The neurologists weighted different efficacy and safety profile attributes without any reference to individual DMDs to disassociate the assessment of weights with any specific DMD. The neurologists did not do direct comparisons between DMDs.

The objective element of a weight is given by the size of the metric's range; that difference is interpreted subjectively and expressed as a number. Differences

between ranges are compared from one criterion to the next, always assigning an arbitrary weight of 100 to the largest range within a set of criteria and then comparing the weighted criteria across the sets.

In a hierarchical model such as the present model, the weighting proceeds from comparing criteria under a grouping to comparisons between groups, through to the final comparison that assesses the swing between the most favorably weighted criterion and the most unfavorably weighted criterion. These last swing weights represent the trade-off between the most beneficial favorable effect and the unfavorable effect associated with most risk.

Generating Results for the Model

With scoring and weighting completed and all data and weights entered into the computer model, the Hiview3 software calculated the weighted average of the preference scores on the 18 criteria and summed the scores to give a single overall weighted preference value for each drug. In addition, the benefit–risk balance of each DMD was compared with that for cladribine. That process followed established priorities for assessing the DMDs and reflected the contributions that individual effects make to the overall weighted preference value for each.

Sensitivity Analyses

Two approaches were taken in sensitivity analyses. First, the criterion weights between the favorable and unfavorable effects were varied to explore the effects of uncertainty and differences in judgments on the final ordering of the drugs. Second, the effects of different input data for some drugs on individual criteria were varied (see [Supplemental Section 3](#) for more detail). These 2 types of analyses were performed for both the RRMS model and the HDA model.

RESULTS

The effects tree of [Figure 1](#) shows the hierarchical model. The effects table ([Table 1](#)) provides the names and definitions of the criteria, the data metric, the worst to best ranges of the data, the weights assessed for the RRMS and HDA models, and the input data for both models.

[Figure 2](#) shows the overall weighted preference values, which take into account available data for the favorable and unfavorable effects, as well as

judgments of the neurologists about the clinical relevance of the effects for the RRMS and HDA populations. Note that the bottom section of each histogram indicates “safety” of the DMDs (higher values represent better safety) and the top section represents “benefits” (higher values represent more benefit). The overall weighted preference values are therefore a sum of the “safety” and “benefits” for each DMD.

The highest overall weighted preference value for the RRMS model was for dimethyl fumarate (63) followed closely by cladribine (62). For patients with HDA, cladribine had the highest overall weighted preference value (76), followed by alemtuzumab (62) and natalizumab (61). The effects with the largest swing-weights were relative reduction in the time to 6-month confirmed EDSS progression over 2 years and incidence of PML (number of cases per 1000 patients).

The neurologists found that the trade-off weight between the most favorable and most unfavorable effect to be the most difficult judgment, so the sum of the benefits against the sum of the safety profile effects, before making that trade-off assessment, was plotted for each model ([Figure 3](#)). This revealed interesting relationships among the DMDs; for example, that cladribine and dimethyl fumarate were both judged to be more beneficial and safer than fingolimod, whatever the relative weights between benefits and safety. In the RRMS model, the 4 DMDs that lie on the line from alemtuzumab, judged to be the most beneficial, down to dimethyl fumarate, judged to be the most safe (ignoring teriflunomide, which is also safe but much less beneficial), established a set of 4 drugs for which more benefit means more risk. Cladribine appeared slightly more beneficial and a little less safe than dimethyl fumarate. Natalizumab appeared at about the half-way point on the line, more beneficial than cladribine but less safe.

For the HDA subgroup model, cladribine moved to an upper-right position, between alemtuzumab and dimethyl fumarate in the RRMS model to a position that would favour the drug over all others over a wide range of the trade-off weights between benefits and safety. Only a modest increase in benefits would be expected from alemtuzumab, compared with cladribine, at a considerable decrease in safety.

Sensitivity Analyses

In the RRMS model, changing the weight on the unfavorable effects node found that dimethyl fumarate was the most preferred drug over a range about 49 to 96. Below this, alemtuzumab was most preferred; above this teriflunomide was most preferred. Over a similar range, cladribine was second to dimethyl fumarate. Cladribine became most preferred with small-to-modest increases of the weight on any of the individual criteria of 6-month confirmed EDSS progression, ease of use, durability, infections, and liver function. See [Supplemental Section 3](#) for details.

In the HDA model, cladribine was the most preferred drug on unfavorable effects, over a range from 19 to 87. Below this, alemtuzumab was most preferred; above this, dimethyl fumarate was most preferred until high weight when teriflunomide was most preferred. Increasing or reducing weights of individual criteria changed the most preferred DMD, depending on that criteria (see [Supplemental Section 3](#)).

Changes in input data, for 3- and 6-month confirmed EDSS progression and the incidence of PML, grade 4 lymphopenia, and malignancy (see [Table 1](#) of [Supplemental Section 3](#)), found that cladribine had a robust benefit–risk balance compared with other DMDs. Despite the hypothetical input data for favorable effects assigned to comparators, cladribine retained its place in terms of benefit–risk, as the most preferred agent for HDA and its second position (after dimethyl fumarate) for RRMS. With respect to unfavorable effects, the only sensitivity analysis for which the relative preference for cladribine compared with other DMDs was not maintained was based on increased input data (0.7%–52.1%) for the grade 4 lymphopenia criterion. For the HDA model, cladribine remained the most preferred option in all these sensitivity analyses.

DISCUSSION

The results of the present model found that in the context of HDA, cladribine had the most favorable benefit–risk balance compared with the other DMDs evaluated and was a close second to dimethyl fumarate for the overall RRMS population (overall weighted preference value for cladribine was 62

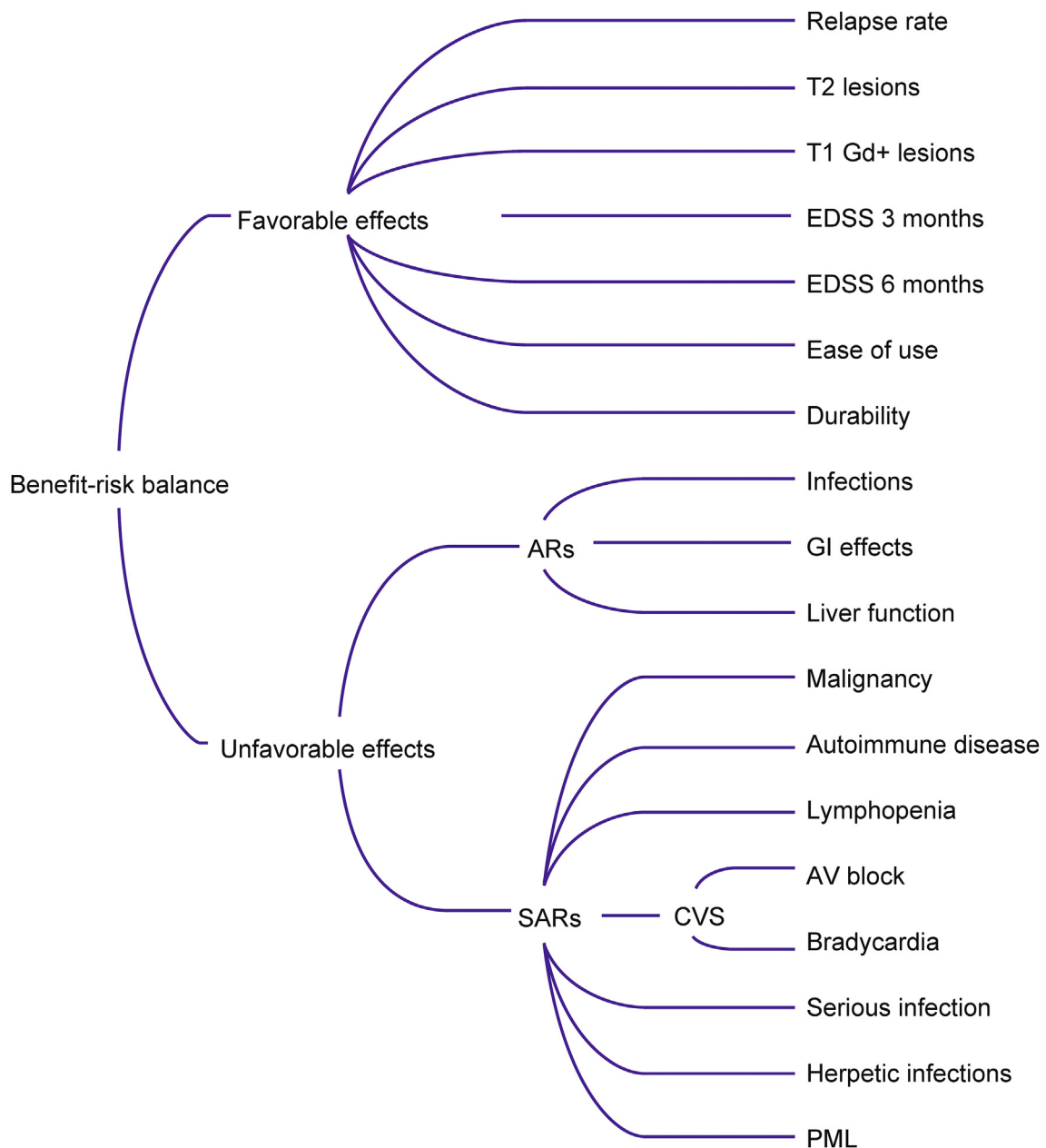


Fig. 1. The effects tree for comparing the 6 disease-modifying drugs (DMDs) in the Multi-Criteria Decision Analysis (MCDA) model. (AR = adverse reaction; AV = atrioventricular; CVS = cardiovascular safety; EDSS = Expanded Disability Status Scale; Gd+ = gadolinium-enhanced; GI = gastrointestinal; PML = progressive multifocal leukoencephalopathy; SAR = serious adverse reaction.)

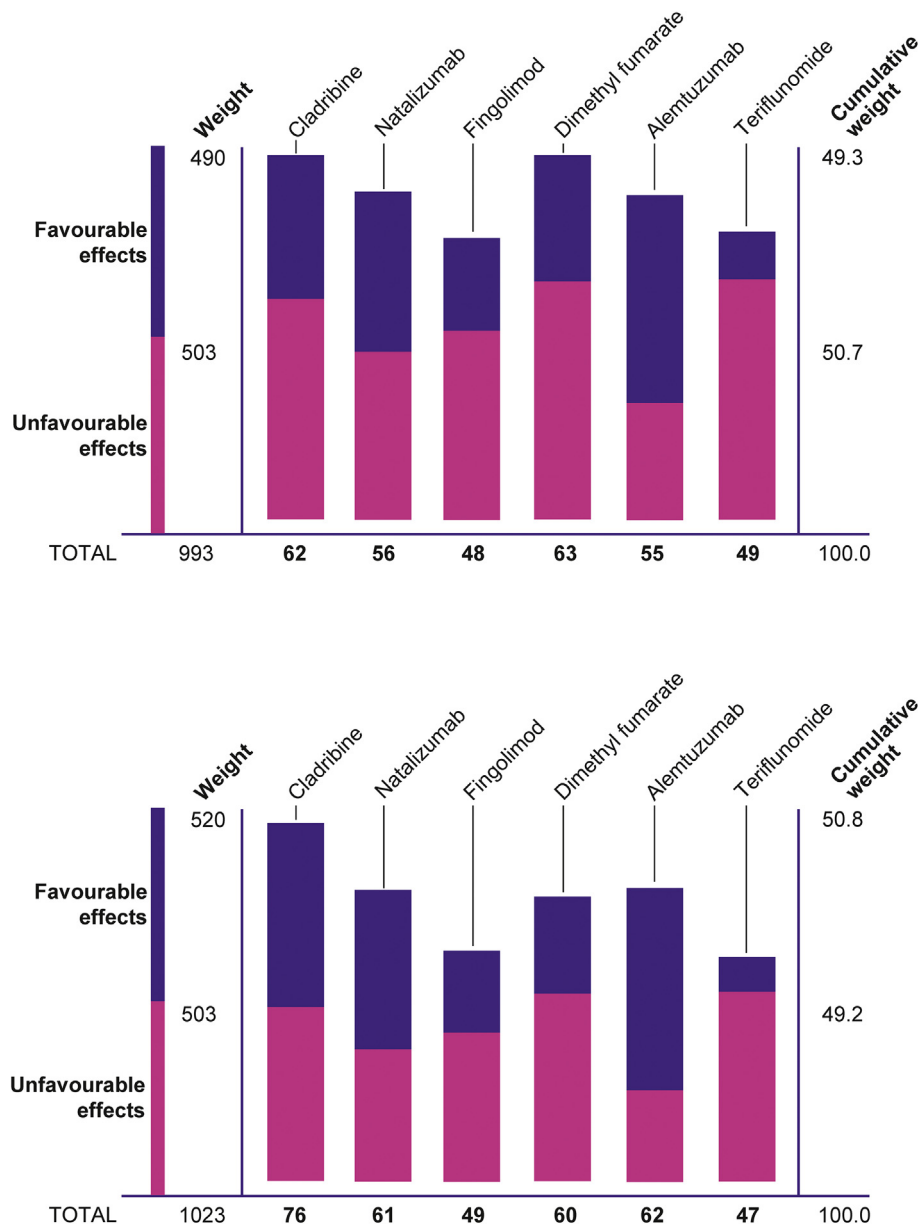
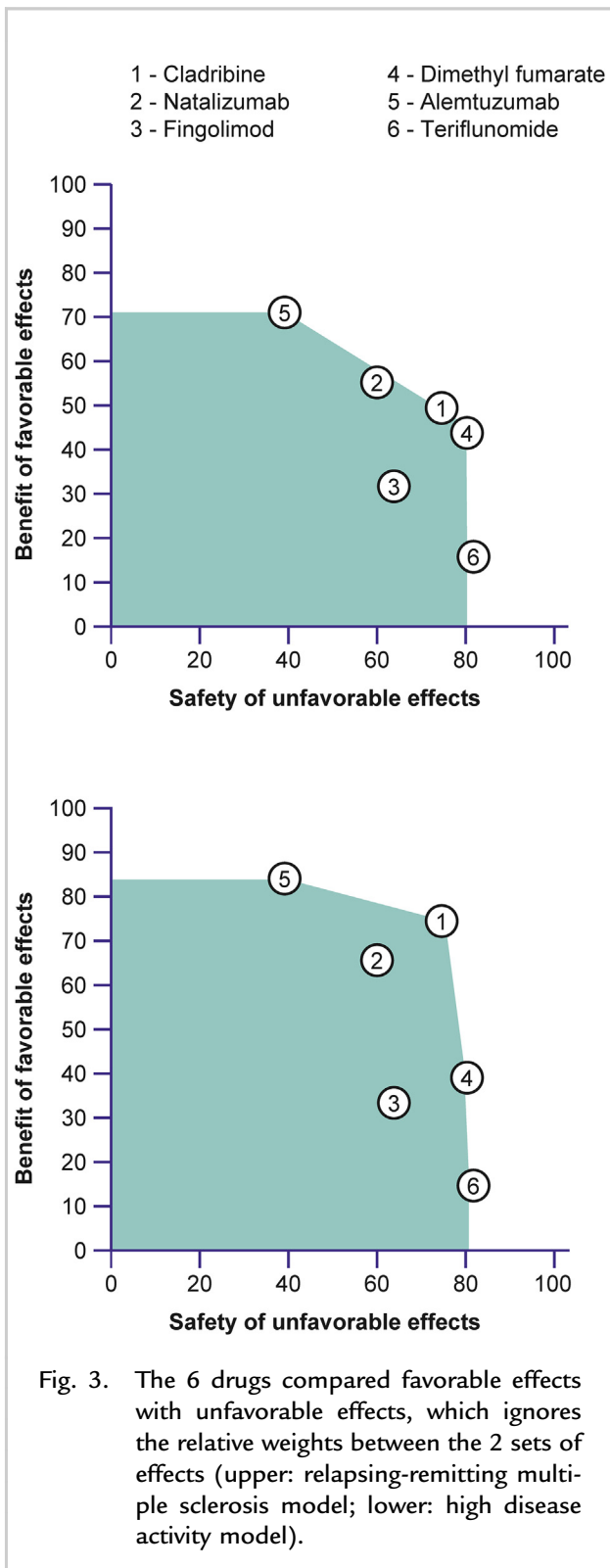


Fig. 2. The overall weighted preference values for the 6 drugs (upper: relapsing-remitting multiple sclerosis model; lower: high disease activity model) shown in the Total row. Note that the bottom section of each histogram indicates “safety” of the DMDs (higher values represent better safety) and the top section represents “benefits” (higher values represent more benefit). The overall weighted preference values are therefore a sum of the “safety” and “benefits” for each DMD. Weights shown in the white field are the sums of the non-normalised cumulative weights, and the normalised values are given in the right column.



versus 63 for dimethyl fumarate). The work considered the judgments of neurologists on the favorable and unfavorable effects of the DMDs included in the model, according to evidence available in December 2015. The outcome indicated that the balance of high efficacy and the safety profile makes cladribine an important treatment option to consider, both in patients with RRMS and patients with HDA.

It is important to recognize that the overall purpose of MCDA is to provide guidance on a specific problem, based on the decisions of a defined group. It does not set out to provide the right answer. MCDA methods may help prescribers by informing more objective decisions about the appropriate DMD to recommend to an individual patient. In this MCDA model, the judgments were made by neurologists with specific expertise in the therapeutic profiles of DMDs and who understood the importance of different aspects of MS and its treatment, with an objective view of the size of possible treatment effects. The neurologists' preferences were relevant to the decision problem, and a direct contribution from patients or caregivers was not included in the MCDA. Nevertheless, individual patient preferences and lifestyle must also be considered carefully, discussed fully, and included in overall treatment decisions to encourage long-term adherence and to ensure that the optimum balance of benefit–risk is most likely to be positive in specific circumstances. Many other approaches aimed at supporting patients in their decisions about treatment options are often based on assumptions or patient reports of preference only and are therefore less systematic methods than MCDA.^{11–16}

Similarly, cost-effectiveness of treatments for MS is an important consideration but was outside the scope of the exercise reported here because cladribine tablets were not commercially available at the time the work was undertaken. However, a cost-effectiveness analysis has recently been published, reporting that cladribine is a cost-effective alternative to alemtuzumab and natalizumab in the treatment of HDA from the perspective of the National Health Service in England.¹⁷

Over recent years the range of DMDs available for MS has grown, with each drug having characteristic

efficacy and safety profiles. The favorable and unfavorable effects identified for the present model were selected to be inclusive and therefore cover the most important features of all the DMDs included. To achieve this meant that the present model had a relatively large number of effects, particularly unfavorable effects, which were needed to encompass the known or potential risks associated with the range of DMDs included.

Uncertainty because of imprecision of data was also a problem in the present model which used relative risk reductions for relapse rate, number of T1 Gd+ and T2 lesions, and time to 3- and 6-month confirmed EDSS progression rather than absolute reductions because it was not possible to obtain consistent data across all the DMDs included. Clinical, MRI, and safety profile data used for the effects table (Table 1) were collected from different clinical studies conducted in different patient populations, and definitions of HDA in reported data were not identical for every DMD. We also acknowledge that selecting the most favorable set of data, rather than an alternative such as the most conservative, a blended value, or according to the quality of evidence, will influence the results of the MCDA model. It should also be noted that in clinical practice, the benefits and risks of natalizumab for an individual patient in terms of PML risk must also take account of John Cunningham virus status.^{18,19} However, sensitivity analyses in the present work did explore potential disagreements and imprecision of the effects criteria. More information about the methods used to identify input data for the effects table (Table 1) and the basis for sensitivity analyses is shown in [Supplemental Information Sections 2 and 3](#). For the present model, input data in the effects table (Table 1) were the most robust available at the time the exercise was conducted, and changes in the input data for 3- and 6-month confirmed EDSS progression and the incidence of PML and malignancy did not change the relative preference for cladribine compared with other DMDs.

Sensitivity analyses of the trade-off weights between the favorable and unfavorable effects for the RRMS model found that dimethyl fumarate and cladribine remained in first- and second-place preference over a wide range of weights, which may well accommodate the preferences of many clinicians and patients. For

the HDA model, cladribine remains most preferred because of its high scores for both benefits and safety, so the trade-off is almost irrelevant to decision making.

Finally, an important caveat of this model is the limited number of neurologists who provided clinical judgments based on available evidence. The neurologists involved were representing their own views and not those of any specific group or organization's viewpoint when making decisions. Caution is required when interpreting the data which may not be representative of the wider, MS specialist community. Reflecting this, the authors suggest that regular, single-country meetings could be organized to explore how differences in cultural values (scores and weights) and updated input data might affect the usefulness of MCDA in different, real-world, dynamic clinical settings.

CONCLUSIONS

The MCDA model described in this report was developed by using decisions from 5 independent European neurologists, who weighted different efficacy and safety profile attributes without any reference to individual DMDs. The resulting model suggests that the benefit–risk balance of cladribine, in patients with RRMS and specifically patients with RRMS who exhibit HDA characterized by high relapse activity (≥ 2 relapses in the previous year), was favorable compared with the other DMDs included in the model.

ACKNOWLEDGMENTS

This decision conferencing and subsequent analysis was sponsored by EMD Serono, Inc, a business of Merck KGaA, Darmstadt, Germany (in the United States) and Merck Serono SA – Geneva, an affiliate of Merck KGaA Darmstadt, Germany (ROW). Merck KGaA staff searched for and provided data used in construction of the Multi-Criteria Decision Analysis (MCDA) model.

The neurologist decision makers who participated in the structured MCDA benefit–risk assessment also served as authors on this manuscript (P. Vermersch, V. Martinelli, C. Pfleger, P. Rieckmann, and L. Alonso-Magdalena). L. Phillips was engaged as an independent specialist in MCDA and group processes to facilitate the decision conference. P. Vermersch, V.

Martinelli, C. Pflieger, P. Rieckmann, L. Alonso-Magdalena, and L. Phillips received honoraria from Merck KGaA for their role in the 2015 workshop but not for their role in developing this subsequent manuscript. Each of the authors contributed equally to the writing, reviewing, revision and approval of this manuscript. Medical writing assistance was provided by Mark O'Connor of inScience Communications, Springer Healthcare, Chester, UK, and was funded by Merck KGaA, Darmstadt, Germany.

CONFLICTS OF INTEREST

Patrick Vermersch has received honoraria or consulting fees from Biogen, Sanofi-Genzyme, Bayer, Novartis, Merck KGaA, Celgene, Roche, and Almirall and research support from Biogen, Sanofi-Genzyme, Bayer, and Merck KGaA. Vittorio Martinelli has received honoraria for speaking and/or for consultancy and support for travel expenses and participation in Congresses from Bayer, Biogen, Merck-Serono, Novartis, Genzyme, and Teva. Claudia Pflieger has received honoraria for advisory council meetings from Novartis and travel expenses for attending meetings from Novartis, Biogen Idec, Roche, Genzyme, and Merck Serono. The Department of Neurology, Aalborg Hospital has received compensation for participation in industry-sponsored clinical trials from Merck Serono, Roche, Novartis, Genzyme, and Biogen Idec. Peter Rieckmann has received honoraria for lectures/steering committee meetings from Merck, Biogen Idec, Bayer Schering Pharma, Boehringer-Ingelheim, Sanofi-Aventis, Genzyme, Novartis, Teva Pharmaceutical Industries, and Serono Symposia International Foundation. Lucia Alonso-Magdalena has received honoraria from Merck KGaA for her role in the 2015 workshop but not for her role in developing this subsequent manuscript. Andrew Galazka is an employee of Merck, Aubonne, Switzerland, a division of Merck KGaA, Darmstadt, Germany. Fernando Dangond is an employee of EMD Serono Research & Development Institute Inc, a business of Merck KGaA, Darmstadt, Germany. Lawrence Phillips has received consultancy fees from Amgen, GE Healthcare, Merck KGaA, Pfizer, Sanofi, and, indirectly, Reckitt Benckiser. The authors have indicated that they have no other conflicts of interest regarding the content of this article.

REFERENCES

1. Lublin FD, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*. 2014;83:278–286.
2. Comi G, Radaelli M, Soelberg Sorensen P. Evolving concepts in the treatment of relapsing multiple sclerosis. *Lancet*. 2017;389:1347–1356.
3. Fox RJ, et al. MS disease activity in RESTORE: a randomized 24-week natalizumab treatment interruption study. *Neurology*. 2014;82:1491–1498.
4. Cohen JA, et al. Fingolimod versus intramuscular interferon in patient subgroups from TRANSFORMS. *J Neurol*. 2013;260:2023–2032.
5. Miller AE, Rhoades RW. Treatment of relapsing-remitting multiple sclerosis: current approaches and unmet needs. *Curr Opin Neurol*. 2012;25:S4–S10.
6. Sorensen PS. Balancing the benefits and risks of disease-modifying therapy in patients with multiple sclerosis. *J Neurol Sci*. 2011;311(Suppl 1):S29–S34.
7. Soelberg Sorensen P. Safety concerns and risk management of multiple sclerosis therapies. *Acta Neurol Scand*. 2017;136:168–186.
8. EMA. *Benefit-risk Methodology Project. Work Package 2 Report: Applicability of Current Tools and Processes for Regulatory Benefit-Risk Assessment*; 2010. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2010/10/WC500097750.pdf.
9. EMA. *Benefit-risk Methodology Project. Work Package 4 Report: Benefit-Risk Tools and Processes*; 2012. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2012/03/WC500123819.pdf.
10. Hughes D, et al. Recommendations for benefit–risk assessment methodologies and visual representations. *Pharmacoepidemiol Drug Saf*. 2016;25:251–262.
11. Reen GK, Silber E, Langdon DW. Interventions to support risk and benefit understanding of disease-modifying drugs in Multiple Sclerosis patients: a systematic review. *Patient Educ Couns*. 2017;100:1031–1048.
12. Kopke S, et al. Information provision for people with multiple sclerosis. *Cochrane Database Syst Rev*. 2014;(4):CD008757.
13. Colligan E, Metzler A, Tiryaki E. Shared decision-making in multiple sclerosis. *Mult Scler*. 2017;23:185–190.
14. Cocco E, et al. Perception of risk and shared decision making process in multiple sclerosis. *Expert Rev Neurother*. 2017;17:173–180.
15. Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. *Med Decis Making*. 2015;35:114–131.
16. Raju GK, Gurumurthi K, Domike R. Benefit-risk analysis for decision-making: an approach. *Clin Pharmacol Ther*. 2016;100:654–671.

17. Hettle R, Harty G, Wong SL. Cost-effectiveness of cladribine tablets, alemtuzumab, and natalizumab in the treatment of relapsing-remitting multiple sclerosis with high disease activity in England. *J Med Econ.* 2018;21:676–686.
18. Major EO, Yousry TA, Clifford DB. Pathogenesis of progressive multifocal leukoencephalopathy and risks associated with treatments for multiple sclerosis: a decade of lessons learned. *Lancet Neurol.* 2018;17:467–480.
19. Mills EA, Mao-Draayer Y. Understanding progressive multifocal leukoencephalopathy risk in multiple sclerosis upatients treated with immunomodulatory therapies: a bird's eye view. *Front Immunol.* 2018;9:138.

Address correspondence to: Lawrence Phillips, PhD, Department of Management, London School of Economics and Political Science, London, UK. E-mail: larry_phillips@msn.com

APPENDIX A. SUPPLEMENTARY INFORMATION

Section 1

Steps in creating and exploring the MCDA model¹

Step	Task
CONTEXT	
1. Establish the decision context	• Review the current landscape of MS treatment options. • Recognize the unmet medical need, severity and morbidity of condition, affected population, patients' and physicians' concerns, time frame for health outcomes. • Describe the comparators
ALTERNATIVES	
2. Identify the options.	
CRITERIA	
3. Identify and define the criteria for assessing the effects of each alternative.	• Select the favourable effects (eg, endpoints, clinical outcomes, durability, and ease of use). • Select the unfavourable effects (eg, adverse events, serious adverse events, infections, serious infections, major cardiac events, malignancies, progressive multifocal leukoencephalopathy).
Represent these in an Effects Tree (Value Tree).	
SCORING	
4. Describe how the alternatives perform for each of the criteria and show how to convert input data into preference values (ie, assess value functions).	• Gather available data. Present data summaries and confidence intervals. • Provide data summaries in an 'Effects Table' with alternatives in columns and criteria in rows. • Assess linear or non-linear value functions, usually direct (more means better) for favourable effects, and inverse (more means worse) for unfavourable effects. • Define each effect's measurement scale and its units (eg, proportions, incidence) and determine upper and lower limits that encompass a plausible range for the data. In most cases minimum and maximum defined the range. • Assess swing-weights to represent the clinical relevance of the swing from lower to upper limit on each scale.
WEIGHTING	
5. Establish a measurement scale for each criterion and assess the relative importance of the scales.	• Multiply preference values and criterion weights and sum the products to obtain overall value (carried out by Hiview3 software). • Construct preference-value bar graphs for favourable and unfavourable effects, and for individual effects. • Calculate difference displays for pairs of alternatives. • Vary individual weights over their entire range from 0 to 1.0; display the overall results graphically. • Change input data in order to check the impact of changing values, which are important but are associated with a high level of uncertainty. • Examine the overall benefit-risk balance under possible future scenarios by changing input data and criteria weights.
RESULTS	
6. Calculate results and provide graphical displays.	
SENSITIVITY ANALYSES	
7. Explore effects of uncertainty on the benefit-risk balance.	

(continued on next page)

(Continued)

Step	Task
RECOMMENDATION	
8. Formulate recommendations.	• Judge the relative importance and effect of the decision maker's risk tolerance for this product (eg, unmet medical need, risk management plan). • Consider how this decision is consistent with similar past decisions, might set a precedent, or make similar decisions in the future easier or more difficult.

MCDA, Multi-criteria decision analysis.

Section 2

Rationale for the favourable and unfavourable effects of DMDs including in the MCDA models

This document describes the approach used to identify source data and populate the effects table used for the benefit-risk assessment of oral cladribine compared with other disease modifying drugs (DMDs) in both relapsing remitting multiple sclerosis (RRMS) patients and those demonstrating high disease activity (HDA).

Identification of data and sources for each of the favourable effects

Cladribine 3.5 mg/kg

Annualized relapse rate (ARR) was the primary endpoint in the Phase III CLARITY trial. For the analysis, data for the cladribine 3.5 mg/kg dose group from the CLARITY trial has been used.² In the CLARITY trial, the primary endpoint, ARR, was met, with cladribine 3.5 mg/kg reducing ARR to 0.14 compared to an ARR of 0.33 in placebo patients, showing a relative risk reduction of 58%.² Standard measures of disability progression are time to sustained disability progression, measured by a 3-month or 6-month confirmed change in expanded disability status scale (EDSS) score. Both timepoints for assessment of change in EDSS are widely accepted as demonstrative of disability progression; and time to 3-month sustained disability progression was chosen as a study endpoint in the cladribine studies. In the CLARITY trial cladribine 3.5 mg/kg led to a lower risk of 3-month confirmed EDSS progression (hazard ratio [HR] of 0.67, 95% confidence interval [CI] of 0.48–0.93), a relative risk reduction of 33%.²

An analysis of 6-month confirmed EDSS progression in patients from CLARITY was included in the combined summary of efficacy. In this analysis, cladribine 3.5 mg/kg led to a relative risk reduction of 6-month confirmed EDSS progression (HR of 0.53, 95% CI of 0.36–0.79), of 47% (data on file). In the CLARITY study, there was strong suppression of T1 Gd+ lesions (least squares [LS] mean number of lesions 0.12 for cladribine 3.5 mg/kg, and 0.91 for placebo; a relative risk reduction of 86%) and active T2 lesions (LS mean number of lesions 0.38 for cladribine 3.5 mg/kg, and 1.43 for placebo; a relative reduction of 73%).²

High disease activity subset

In the HDA group relevant to this analysis (patients with 2 or more relapses in previous year, regardless of previous treatment status), cladribine 3.5 mg/kg reduced ARR to 0.16 compared to an ARR of 0.50 in placebo patients, showing a relative risk reduction of 68%. Cladribine 3.5 mg/kg reduced the risk of 3-month confirmed EDSS progression by 72% and reduced the risk of 6-month confirmed EDSS progression by 82% (data on file).

Natalizumab 300mg

Natalizumab was evaluated in two randomized, double-blind, placebo-controlled trials in patients with MS. Both studies enrolled patients who experienced at least one clinical relapse during the prior year and had an EDSS score between 0 and 5.0. Study 1 (monotherapy study) enrolled patients who had not received any IFN- β or glatiramer acetate for at least the previous 6 months; approximately 94% had never been treated with these agents.³ Study 2 (add-on study) enrolled

patients who had experienced one or more relapses while on treatment with IFN- β -1a IM once-weekly during the year prior to study entry; patients continued to receive IFN- β -1a 30 μ g IM once-weekly during the study and were randomly assigned to natalizumab or placebo.³ In Study 1, there was a placebo-adjusted annualized relapse rate (ARR) of 68% and a 42% reduction in relative risk of disability progression. In Study 2, there was a placebo-adjusted ARR of 56% and a 24% reduction in relative risk of disability progression. So, for the effects table, the higher values from Study 1 were used for the ARR (68%) and 3-month confirmed EDSS risk reduction (42%).⁴ The sensitivity analysis of progression of disability that was sustained for 24 weeks yielded a 54% risk reduction in the natalizumab group (hazard ratio, 0.46; 95% CI, 0.33 to 0.64),⁴ so this value (54%) was used for the 6-month confirmed EDSS risk reduction.

Natalizumab led to an 83% reduction in the accumulation of new or enlarging hyperintense lesions, as detected by T2-weighted MRI, over two years (mean numbers of lesions, 1.9 with natalizumab and 11.0 with placebo). There were 92% fewer lesions (as detected by gadolinium enhanced MRI) in the natalizumab group than in the placebo group at both one and two years.⁴

High disease activity subset

For the high disease activity subset (at least two relapses in the previous year, or at least two relapses in the previous year and at least 1 T1 Gd+ lesion), data were taken from the Tysabri EPAR.⁵ ARR was 0.282 in the natalizumab group and 1.455 in the placebo group, a reduction of 81%. A 3-month confirmed EDSS HR of 0.47 (95% CI 0.24, 0.93) and a 6-month confirmed EDSS HR of 0.36 (95% CI 0.17, 0.76) were reported, and the reduction in risk of 53% and 64%, respectively were used in the effects table.

Fingolimod 0.5mg

Efficacy of fingolimod was demonstrated in two studies that evaluated once daily doses of fingolimod 0.5 mg and 1.25 mg in patients with RRMS, one placebo-controlled study and one active-controlled study. Study 1 was a 2-year randomized, double-blind, placebo-controlled study in patients with RRMS who had not received any interferon-beta (IFN- β) or glatiramer acetate for at least the previous 3 months and had not received any natalizumab for

at least the previous 6 months.⁶ In this study, the ARR was significantly lower in patients treated with fingolimod 0.5 mg (0.18) than in patients who received placebo (0.40), a relative reduction of 54% (entered in the effects table). Time to onset of 3-month confirmed disability progression was significantly delayed with fingolimod 0.5 mg treatment compared to placebo (HR, 0.70, 95% CI 0.52, 0.96) and regarding disability progression that was confirmed after 6 months, the risk was also reduced with fingolimod 0.5 mg over the 24-month study period (HR, 0.63, 95% CI 0.44, 0.90). Therefore, the 3-month and 6-month EDSS risk reductions were entered as 30% and 37%, respectively.

The mean number of new or newly enlarging T2 lesions over 24 months was 2.5 in the fingolimod 0.5 mg group and 9.8 in the placebo group, a reduction of 74%. The mean number of T1 Gd+ lesions at 24 months was 0.2 in the fingolimod 0.5 mg group and 1.1 in the placebo group, a reduction of 82%.⁶

High disease activity subset

For the high disease activity subset (two or more relapses in the previous year), data were taken from the EPAR.⁷ The ARR rate ratio was 0.37 (95% CI 0.27, 0.51) so a relapse rate reduction of 63% was entered in the effects table. The 3-month confirmed EDSS risk reduction was 33% (HR of 0.67, 95% CI 0.40, 1.11). The same value was used for the 6-month confirmed EDSS risk reduction in the effects table because this analysis was not reported.

Alemtuzumab 12 mg

Efficacy of alemtuzumab in MS was established in two open-label, rater-blinded, active-comparator (IFN- β -1a) Phase III trials: one in treatment-experienced RRMS patients (CARE MS II, also known as CAMMS 324) and a second in treatment-naïve RRMS patients with early, active disease (ie, as a first-line treatment) (CARE MS I, also known as CAMMS 323). There are no placebo-controlled studies of alemtuzumab. In both pivotal studies, alemtuzumab was administered by intravenous infusion once daily over a 5-day course, followed one year later by intravenous infusion once daily over a 3-day course. The clinical outcome measures in both studies were the ARR over 2 years and the time to confirmed disability progression.⁸

Alemtuzumab demonstrated a statistically significant IFN- β -1a-adjusted effect on ARR reduction in treatment-experienced RRMS patients (49%)⁹ and in treatment-naïve RRMS patients (55%).¹⁰ The 6-month confirmed EDSS risk reduction was 42% in treatment-experienced RRMS patients (Coles et al 2012) and 30% in treatment-naïve RRMS patients (Cohen et al 2012). For the purposes of the benefit-risk analysis the higher relapse rate of 55% was chosen and arbitrarily increased by 20% on the assumption that the effect would have been greater relative to placebo. Therefore, the ARR value entered into the effects table for alemtuzumab for the RRMS analysis was 75%. Similarly, for 6-month confirmed EDSS risk reduction the higher value was chosen (42%) and arbitrarily increased by 20% to give a value of 62% which was entered in the effects table. The same value was used for the 3-month confirmed EDSS risk reduction in the effects table because this analysis was not reported.

In treatment refractory RRMS patients the proportion of patients with new or enlarging T2-hyperintense lesions at 24 months was 68% in the IFN- β -1a group and 46% in the alemtuzumab group.⁹ Corresponding data for treatment-naïve RRMS patients were 58% and 48%, respectively.¹⁰ Similarly, the proportion of refractory RRMS patients with gadolinium enhancing lesions at 24 months was 23% in the IFN- β -1a group and 9% in the alemtuzumab group,⁹ while for naïve patients the proportions were 19% and 7%, respectively.¹⁰ These publications did not provide data on mean number of active T2 lesions or mean number of T1 Gd+ lesions. This information was also not found in the Lemtrada EPAR.¹¹

For the purposes of the benefit-risk analysis it was concluded that if alemtuzumab had been compared with placebo, MRI efficacy would have been extremely high. Consequently, arbitrary values of 87% and 92% reduction were included in the effects table for T2 lesions and T1 Gd+ lesions, respectively, as a conservative approach.

High disease activity subset

In the CARE MS I (CAMMS 323) study the ARR risk ratio for “patients with >2 relapses” was 0.53, consistent with a 47% risk reduction and the risk ratio for the “highly active population” was 0.49 consistent with a 51% risk reduction.¹¹ Similarly, in

the CARE MS II (CAMMS 324) study the ARR risk ratio for “patients with >2 relapses” was 0.51, consistent with a 49% risk reduction and the risk ratio for the “highly active population” was 0.49 consistent with a 51% risk reduction.¹¹ Since these results were obtained in a trial against IFN- β -1a and not against placebo, an arbitrary ‘premium’ was assigned and a value of 71% ARR risk reduction was entered in the effects table.

In the CARE MS I (CAMMS 323) study the 6-month confirmed EDSS risk reduction for “patients with >2 relapses” and for the “highly active population” was 17% and 13%, respectively (risk ratios of 0.83 and 0.87, respectively).¹¹ Similarly, in the CARE MS II (CAMMS 324) study the 6-month confirmed EDSS risk reduction for “patients with >2 relapses” and for the “highly active population” was 23% and 51% respectively (risk ratios of 0.77 and 0.49, respectively).¹¹ The of 20% was added to give a value of 71%. This value was entered in the effects table for both the 3-month confirmed EDSS risk reduction and the 6-month confirmed EDSS risk reduction.

As for the overall RRMS population, it was concluded that had alemtuzumab been compared with placebo, MRI efficacy would have been extremely high and slightly higher in the HDA population than the overall population. Consequently, arbitrary values of 92% and 97% reduction were included in the effects table for T2 lesions and T1 Gd+ lesions, respectively.

Dimethyl fumarate

Efficacy of dimethyl fumarate in MS was established in two randomized, double-blind, placebo-controlled studies in RRMS^{12,13}; one of the studies also included a glatiramer acetate reference arm. In each study the primary endpoint was the proportion of patients relapsed at 2 years. Additional endpoints at 2 years included the number of new or newly enlarging T2 hyperintense lesions, number of new T1 hypointense lesions, number of Gd+ lesions, ARR, and time to confirmed disability progression. In the study that included the glatiramer acetate arm, statistical significance against placebo was not confirmed for either active treatment on the secondary endpoint of sustained disability progression but statistical significance against placebo was confirmed for all other endpoints.¹⁴ In the other

study, statistical significance against placebo was confirmed for all endpoints.¹⁴ An improvement in disability outcomes was statistically significant in a pooled analysis of the entire Phase III program.¹⁵ As the benefits of dimethyl fumarate were better in Study 1 (also known as DEFINE and 109MS301) than Study 2 (also known as CONFIRM and 109MS302), the data from Study 1 have been used in the effects table.

In Study 1 the ARR at 2 years was 0.17 in the dimethyl fumarate 240 mg bid group compared with 0.36 in the placebo group, representing a relative reduction of 53%.¹³ The estimated proportion of patients with 3-month confirmed EDSS was 16% in the dimethyl fumarate 240 mg bid group and 27% in the placebo group, a relative risk reduction of 38% (HR 0.62; 95% CI, 0.44 to 0.87). Data for the 6-month confirmed EDSS were not included in the publication¹³ but are provided in Tecfidera EPAR.¹⁶ The HR is given as 0.77 (95% CI, 0.52 to 1.14), a relative risk reduction of 23%. In Study 1, as compared with placebo, dimethyl fumarate 240 mg bid reduced the number of new or enlarging hyperintense lesions on T2-weighted images at 2 years by 85% and decreased the mean number of T1 Gd+ lesions from 1.8 to 0.1, a reduction of 94%.¹³ These values were used in the effects table (for the RRMS analysis).

High disease activity subset

For the high disease activity subset (2 or more relapses and at least 1 T1 Gd+ lesion), pooled data from Study 1 and Study 2 were presented in the Tecfidera EPAR.¹⁶ In this subset, the ARR at 2 years was 0.23 in the dimethyl fumarate 240 mg bid group compared with 0.58 in the placebo group, representing a relative reduction of 60% (HR 0.40, 95% CI 0.22, 0.71). The 3-month confirmed EDSS was 26% in the dimethyl fumarate 240 mg bid group and 33% in the placebo group, a relative risk reduction of 21%. The same value was used for the 6-month confirmed EDSS risk reduction in the effects table because this analysis was not reported.

Teriflunomide

The efficacy of teriflunomide in patients with RRMS was demonstrated in two similarly designed, Phase III, randomized, double-blind, placebo-controlled, parallel-group studies (TOWER and TEMSO) and one Phase III, randomized, single-blind, parallel-group, study comparing teriflunomide with IFN- β -1a.

In the two placebo-controlled studies, teriflunomide significantly reduced the ARR and significantly reduced the risk of sustained accumulation of disability (at the licensed dose of 14 mg once daily). Although both placebo-controlled studies showed a relative risk reduction for the primary outcome of ARR of >30%, the adjusted ARR in both the placebo and active treatment groups was low in both studies, so that the absolute risk reduction in ARR was small (<0.2).^{17,18}

In the TOWER study, teriflunomide 14 mg significantly reduced the ARR (adjusted rates of 0.50 [95% CI 0.43, 0.58] for placebo versus 0.32 [95% CI 0.27, 0.38] for teriflunomide 14 mg), corresponding to relative rate reduction of 36%.¹⁹ Similar results were seen in the TEMSO study with a relative rate reduction of 31%.²⁰ For the purpose of the benefit-risk analysis, the higher value from the TOWER study was used.

In the TOWER study a 31.5% reduction in the risk of sustained accumulation of disability was observed with teriflunomide 14 mg compared with placebo.¹⁹ Similar results were seen in the TEMSO study with a 29.8% reduction.²⁰ A value of 31% was used in the benefit-risk analysis for 3-month confirmed EDSS.

In an integrated analysis of the TOWER and TEMSO studies, the relative reduction in the 6-month confirmed EDSS was 24% (HR 0.76 95% CI 0.57, 1.01).²¹

In the TEMSO study, patients in the teriflunomide 14 mg group had fewer T1 Gd+ lesions per scan than those in the placebo group. Relative risk versus placebo was 0.2 (95% CI 0.12, 0.32), indicating an 80% reduction.²⁰

The effect of teriflunomide 14 mg on MRI activity was also demonstrated in a Phase II, randomized, double-blind, placebo-controlled clinical trial of multiple sclerosis patients with relapses.²² MRI was to be performed at baseline, 6 weeks, 12 weeks, 18 weeks, 24 weeks, 30 weeks, and 36 weeks after treatment initiation. Secondary endpoints in this study were based on MRI scans including, average number of new T2 lesions per scan. The mean number of new or newly enlarging T2 lesions was 1.52 in the placebo group and 0.71 in the teriflunomide 14 mg group, indicating a reduction of 53%.

High disease activity subset

In the TEMSO study a subgroup analysis of patients with high disease activity (2 or more relapses and at

least 1 T1 Gd+ lesion at baseline; N = 127) showed an ARR relative reduction of 19% (RR 0.81, 95% CI 0.51, 1.28) and 3-month confirmed EDSS relative reduction of 35% (HR 0.65, 95% CI 0.26, 1.59).²¹ The same value (35%) was used for the 6-month confirmed EDSS risk reduction in the effects table because this analysis was not reported.

Durability

Data on durability of the effect of cladribine and the comparators are estimates inferred from the mode of action of the drugs assuming that the drug has already shown efficacy and that the treatment interruption occurs for reasons other than efficacy: it is what physicians generally consider to be an acceptable interval of time to wait after interrupting a therapy to switch to another without risking the appearance of new relapses.

Ease of use

Patients with MS who have low adherence to therapy are at greater risk of relapse than more adherent patients.^{23,24} Conversely, patients with MS who are highly adherent are less likely to have a relapse or require hospitalization.^{24–27} Furthermore, because of the low frequency of exacerbations, patients have difficulties perceiving the beneficial impact of being adherent to a drug administered on a frequent basis. This is a contributing factor to the “treatment fatigue” phenomenon, where MS patients become averse to repeated parenteral or oral dosing over prolonged periods of time, despite the potential effectiveness of the therapy. Ease of use was therefore included in the benefit-risk comparison.

Four criteria were considered when ranking the six drugs for “ease of use”. These were:

- (1) Oral versus intravenous administration
- (2) The number of doses required (few or many)
- (3) Whether monitoring was required during administration and
- (4) Whether co-administration of other drugs was required.

The six drugs were ranked for overall ease of use based on these criteria with a score of 1.0 being assigned to the drug considered to be the most difficult to use with increments of 0.5 up to the

highest score of 3.5 for the drug perceived to be easiest to use.

Identification of data for each of the unfavourable effects

Data on unfavourable effects selected for inclusion in the model were based upon the important identified and potential risks associated with cladribine and the comparator products. Generally data were captured at the MedDRA System Organ Class (SOC) level as the terminology across the different sources varied. For the most part the comparators data were taken from 2-year studies.

The data sources for the unfavourable effects are described in the following sections.

Cladribine 3.5 mg/kg

In the integrated analysis of AEs in the placebo-controlled double-blind cohort, 1458 patients were exposed to cladribine (all doses) and 745 patients were exposed to placebo. The placebo controlled, double-blind cohort without the ONWARD study was used for the cladribine data (January 2016). This was compared to the CLARITY clinical study report (CSR) for validation.

In the Infections and infestations system organ class (SOC), in the cladribine group AEs were reported for 755/1458, (51.8%) patients (data on file). This value was entered in the effects table. For comparison, the incidence rate in the placebo group was 348/745 (46.7%).

Serious adverse events (SAEs) in the Infections and infestations SOC were reported for 37/1458, (2.5%) patients in the cladribine group and 12/745 (1.6%) in the placebo group (data on file). The value of 2.5% was entered in the effects table for cladribine.

Herpetic infection AEs were reported for 115/1458, (7.9%) patients in the cladribine group and 24/745 (3.2%) in the placebo group (data on file). The value of 7.9% has been entered in the effects table for cladribine.

In the Gastrointestinal disorders SOC, AEs were reported for 460/1458 (31.6%) patients in the cladribine group and 233/745 (31.3%) patients in the placebo group (data on file). The value of 31.6% was entered in the effects table for gastrointestinal (GI) effects.

In the Hepatobiliary disorders SOC, AEs were reported for 22/1458 (1.5%) patients in the

cladribine group and 12/745 (1.6%) in the placebo group (data on file). The value of 1.5% was entered in the effects table for liver function.

In the Immune system disorders SOC, AEs were reported for 29/1458 (2.0%) patients, in the cladribine group and 11/745 (1.5%) patients in the placebo group (data on file). The value of 2.0% was entered in the effects table for autoimmune disease.

Bradycardia was reported for 3/1458 (0.2%) patients in the cladribine group and was not recorded for any patients in the placebo group (data on file).

There were no events of atrioventricular block in either treatment group (data on file).

There were no cases of progressive multifocal leukoencephalopathy (PML) reported in the studies (data on file).

For the analysis of malignancy, the “All exposed” cohort was used, comprising all Phase II/III studies with any formulation of cladribine. It contained the maximum safety data available and was used to assess cladribine safety profile over time and in comparison to placebo. This cohort includes patients from the ONWARD study who were on background IFN- β therapy. The All exposed cohort comprised 1976 patients in the cladribine group (all doses) and 802 patients in the placebo group, with 8650 and 2361 patient years at risk, respectively. In this cohort, 32 patients in the cladribine group and 4 patients in the placebo group had at least one malignant tumor, giving an incidence per 100 PY of 0.37 and 0.17, respectively (data on file). The value of 0.37 cases per 100 PY was entered in the effects table.

Grade 4 lymphopenia was infrequent throughout the clinical program. In CLARITY, Grade 4 lymphopenia occurred in 13 patients (2.9%) treated with oral cladribine at 5.25 mg/kg and in only 3 patients (0.7%) treated at 3.5 mg/kg (data on file). The value of 0.7% was entered in the effects table for lymphopenia.

Natalizumab 300mg

Natalizumab is generally well tolerated but is associated with a small but confirmed risk of PML, which prompted the drug's temporary withdrawal from the US market in February 2005.

Because of the risk of PML, natalizumab is available in the US only through a restricted program under a REMS called the TOUCH Prescribing Program.³ More than 450 post-marketing cases of PML,

resulting in more than 100 fatalities, have been reported with natalizumab monotherapy; survivors of PML (~80% of MS patients with PML) are left with varying degrees of disability.

The most important of the unfavourable effects in the comparison as far as natalizumab is concerned is therefore PML. A value of 2.1 cases per 1000 patients was included based on the report that as of February 29, 2012, there were 212 confirmed cases of PML among 99,571 patients treated with natalizumab.²⁸

A total of 1617 multiple sclerosis patients in controlled studies received natalizumab, with a median duration of exposure of 28 months.³ Of these 1617 patients, 1192 (73.7%) had an infection and 39 (2.4%) had a serious infection.²⁹ These values (73.7% and 2.4%) are entered in the effects table. It should be noted that in these trials 839/1135 (73.9%) of patients receiving placebo also had an infection and 26/1135 (2.3%) had a serious infection.²⁹

In Study 1 (monotherapy study), which enrolled patients who had not received any IFN- β or glatiramer acetate for at least the previous six months, herpes infection was reported by 8% of patients.³

For the current analysis of malignancy risk data the incidence for natalizumab-treated patients was 0.32/100 patient-years (All exposed) compared to 0.65/100 patient-years for placebo-treated patients.³⁰

For the remaining unfavourable effects in the analysis, (ie GI effects, liver function, autoimmune disease, lymphopenia, AV block, and bradycardia) a value of zero was entered for natalizumab.

Fingolimod 0.5mg

In multiple sclerosis clinical studies the overall rate of infections (65.1%) at the 0.5 mg dose was similar to placebo.³¹ The 65.1% value was entered in the effects table.

FREEDOMS II was a double-blind, randomized, placebo-controlled, Phase III trial that started shortly after the pivotal Phase III FREEDOMS and TRANSFORMS studies had started and was part of the global clinical development program to investigate fingolimod in multiple sclerosis.

In this study, AEs in the gastrointestinal disorders SOC were reported by 155/358 (43%) patients in the fingolimod 0.5 mg group, compared with 143/355

(40%) in the placebo group.³² The 43% value was entered in the effects table.

In clinical trials, elevations 3-fold the upper limit of normal (ULN) or greater in alanine aminotransferase (ALT) occurred in 8.0% of patients treated with fingolimod 0.5 mg compared to 1.9% of placebo patients.

Elevations 5-fold the ULN occurred in 1.8% of patients on fingolimod and 0.9% of patients on placebo.³¹ In the EPAR, the data from three completed, double-blind, controlled MS studies and interim data from two long-term extension studies in MS patients were pooled into five datasets using appropriate cut-offs to accommodate differences between studies in duration of treatment, doses, and comparators. In one of these datasets (24-month treatment; N = 1272), 43/425 (10.1%) patients in the fingolimod 0.5 mg group had an AE of ALT increased compared with 16/418 (3.8%) patients in the placebo group and 50/429 (11.7%) in the fingolimod 1.25 mg group.⁷ The 10.1% value was entered in the effects table.

In the same dataset (24-month treatment; N = 1272), 7/425 (1.6%) patients in the fingolimod 0.5 mg group had a serious adverse event (SAE) in the system organ class (SOC) Infections and infestations compared with 8/418 (1.9%) patients in the placebo group and 11/429 (2.6%) in the fingolimod 1.25 mg group.⁷ The 1.6% value was entered in the effects table for serious infections.

In placebo-controlled trials, the rate of herpetic infections was 9% in patients receiving fingolimod 0.5 mg and 7% in patients receiving placebo.³³ The 9% value was entered in the effects table.

For the current analysis of malignancy risk data the incidence of basal cell carcinoma for fingolimod 0.5 mg-treated patients was 0.4/100 patient-years across all studies (controlled and extensions).³⁴

Within 4–6 hours after the first dose of fingolimod 0.5 mg, the lymphocyte count decreases to approximately 75% of baseline in peripheral blood. With continued daily dosing, the lymphocyte count continues to decrease over a two-week period, reaching a minimal count of approximately 500 cells/mL or approximately 30% of baseline. A total of 18% of patients reached a minimal count below 200 cells/mL (CTC Grade 4) on at least one occasion.³¹ Low lymphocyte counts are maintained with chronic daily dosing. The 18% value was entered in the effects table.

In multiple sclerosis clinical studies first-degree atrioventricular block (prolonged PR interval on ECG) was detected after treatment initiation in 4.7% of patients on fingolimod 0.5 mg, in 2.8% of patients on intramuscular IFN- β -1a, and in 1.6% of patients on placebo. Second-degree atrioventricular block was detected in less than 0.2% patients on fingolimod 0.5 mg. The 4.7% value was entered in the effects table.³¹

In the completed pivotal Phase III MS studies, bradycardial AV conduction abnormalities were reported as AEs by 35/854 (4.1%) patients on fingolimod 0.5 mg and 9/418 (2.2%) patients on placebo.³⁵ The 4.1% value was entered in the effects table.

As of August 2015, according to Novartis' Gilenya Information Center webpage, more than 125,000 patients have been treated with fingolimod and there are 240,000 patient years of exposure in both clinical trials and the post-marketing setting. In the Gilenya risk management plan (RMP) dated 24 February 2015, there had been 11 cases of PML in patients (with prior history of natalizumab use).³⁵ A further 2 cases of PML have been reported during 2015 giving 13 cases in total. Therefore, a value of 0.104 cases per 1000 patients was entered into the effects table.

For autoimmune disease, a value of 0% was entered for fingolimod as this risk is not relevant for this drug.

Alemtuzumab 12 mg

The most common adverse events in the alemtuzumab development program were infusion reactions, reported in 90% of patients in the Phase III studies,^{9,10,36} of which approximately 3% were regarded as serious. There were no cases of anaphylaxis and no infusion reactions resulted in death in the Phase III studies. Infections (eg, respiratory/urinary tract infection, oral herpes, sinusitis) were more common in the alemtuzumab group than in the IFN- β -1a group in the Phase III studies; most were mild to moderate and none were life-threatening. By 24 months, approximately 17% of patients in the alemtuzumab group had thyroid-associated adverse events and 1% had immune thrombocytopenia. Alemtuzumab has not been associated with any cases of PML in studies of patients with MS (although there have been cases in patients treated with alemtuzumab for transplant rejection or for chronic lymphocytic leukemia).^{9,10,36}

In light of the potentially serious autoimmune and other side effects that can occur many months to years after treatment with alemtuzumab, a strict monitoring program is in place requiring monthly complete blood count with differential, urine, and serum creatinine and quarterly thyroid function testing; moreover, according to the US and European labels, patients must be monitored for 48 months following their last infusion.

The AE profile in the dataset including all Phase II/III active-controlled studies, is presented in Table 35 of the Lemtrada EPAR.¹¹ Overall the incidence of infections reported as an AE was 652/919 (70.9%) for alemtuzumab 12 mg and 264/496 (53.2%) for IFN- β -1a. The incidence of serious infections was 25/919 (2.7%) patients for alemtuzumab 12 mg and 5/496 (1.0%) patients for IFN- β -1a over 2 years of follow up in the Phase II/III active-controlled studies (Table 36 of the Lemtrada EPAR).¹¹ The incidence of herpetic infections was 144/919 (15.7%) patients in the alemtuzumab 12 mg group and 15/496 (3.0%) patients in the IFN- β -1a group (Table 66 of the Lemtrada FDA background package).³⁷

AEs in the gastrointestinal disorders SOC were reported by 450/919 (49.0%) patients in the alemtuzumab 12 mg group and by 163/496 (32.9%) patients in the IFN- β -1a group.¹¹

In the Immune system disorders SOC, AEs were reported by 53/919 (5.8%) patients in the alemtuzumab 12 mg group and by 16/496 (3.2%) patients in the IFN- β -1a group.¹¹ However, a real-world study of the long-term (mean follow-up of 6.1 years) efficacy and safety of alemtuzumab treatment in patients with RRMS has recently been reported and has shown a much higher level of autoimmunity than was seen in the trials. One hundred patients were identified with a mean follow-up of 6.1 years (range 1–13). Forty patients were retreated with at least one further treatment cycle. Approximately half (47/100) of the patients included in the study developed secondary acquired autoimmune disease (AID), with 35% of patients developing thyroid AIDs.³⁸ The value of 47% was entered in the effects table for autoimmune disease.

In total, 22 of 1485 (1.5%) alemtuzumab-treated patients had a treatment-emergent adverse event of malignancy, as of 20 April 2013, at which point 5874 person-years of follow-up were available, giving an incidence per 100 PY of 0.374.³⁷ The value of 0.37 cases per 100 PY was entered in the effects table.

A rapid depletion of circulating T and B lymphocytes, caused by the anti-CD52 mechanism of alemtuzumab action, results in nearly all patients in MS clinical trials experiencing lymphopenia following treatment. In analyses of worst post-baseline platelet count by CTC Grade, alemtuzumab-treated patients had lower frequencies of grade 1 and grade 2 changes in lymphocytes, compared to IFN- β -1a patients, because the majority of alemtuzumab patients had grade 3 or grade 4 abnormalities. A total of 479/919 (52.1%) patients in the alemtuzumab 12 mg group had a post-baseline lymphopenia of CTC Grade 4, compared to 1/492 (0.2%) patients in the IFN- β -1a group (Lemtrada FDA background package, Table 84).³⁷

In controlled trials, in the alemtuzumab 12 mg group, one patient each had SAEs of bradycardia, sinus bradycardia and sick sinus syndrome (Lemtrada FDA background package, Table 21).³⁷ Therefore, 3/919 (0.3%) of patients in the alemtuzumab 12 mg group had a bradycardia (compared to none in the IFN- β -1a group). A value of 0.3% was included in the effects table for bradycardia.

There were no AEs of elevated liver enzymes or AV block and no cases of PML in patients treated with alemtuzumab 12 mg and so values of zero were entered in the effects table for these risks.

Dimethyl fumarate

The most commonly reported side effects in clinical trials were flushing and GI side effects. The incidence of GI events was higher early in the course of treatment (primarily in the first month) and usually decreased over time in patients treated with dimethyl fumarate compared with placebo. An increased incidence of elevations of hepatic transaminases in patients treated with dimethyl fumarate was seen primarily during the first 6 months of treatment, and most patients with elevations had levels $<3 \times$ ULN. A transient increase in mean eosinophil counts was seen during the first two months of therapy. The rate of serious adverse events in the Phase III trials was lower than that of placebo, and there were no opportunistic infections or differences between treatment groups with regard to malignancies.¹⁴

In November 2014, the FDA released a safety warning in the wake of a fatal case of PML that occurred in an MS patient who was being treated with dimethyl fumarate. In December 2014, the US

and European labels for dimethyl fumarate were updated to include information about the risk of PML, outlining more frequent complete blood count (CBC) monitoring (every 6 to 12 months, as opposed to annually), and establish a lymphocyte concentration threshold at which to consider interruption of treatment with the drug ($0.5 \times 10^9/L$).^{14,39}

Apart from the CBC requirement, no additional monitoring is required or recommended in the US label, but in the European label, liver and kidney function tests are recommended prior to starting treatment, after 3 and 6 months of treatment, and annually thereafter, as well as at physicians' discretion.

The AE profile in the dataset including all placebo-controlled studies, is presented in Section 2.6.2 of the Tecfidera EPAR.¹⁶ Overall the incidence of infections reported as an AE was 60% for dimethyl fumarate 240 mg bid (463 patients) and 56% for placebo (469 patients).

The incidence of serious infections was similar (2% vs 2%) in patients treated with dimethyl fumarate 240 mg bid or placebo, respectively (Section 5.3 of the Tecfidera US PI).¹⁴

AEs in the gastrointestinal disorders SOC were reported at an increased incidence with dimethyl fumarate 120 mg bid compared with placebo in both placebo-controlled studies, Study 109MS301 (DEFINE): 44% vs 36% on placebo and Study 109MS302 (CONFIRM): 36% vs 26% on placebo.¹⁶ The higher value (44%) has been entered into the effects table.

Elevations of liver transaminases were reported as AEs at a slightly higher incidence with dimethyl fumarate 120 mg bid compared to placebo: AEs of increased alanine transaminase (ALT) were reported for 6% of patients in the dimethyl fumarate 120 mg bid group compared with 5% in the placebo group.¹⁶

Including patients with longer term treatment, the incidence of malignancies overall among dimethyl fumarate-treated patients was 0.375 per 100 patient years (95% CI 0.219, 0.601).⁴⁰

In the placebo-controlled studies, most patients (>98%) had normal lymphocyte values prior to initiating treatment. Upon treatment with dimethyl fumarate, mean lymphocyte counts decreased over the first year with a subsequent plateau. On average, lymphocyte counts decreased by approximately 30% of baseline value. Lymphocyte counts $<0.5 \times 10^9/L$

were observed in <1% of patients treated with placebo and 6% of patients treated with dimethyl fumarate. A lymphocyte count $<0.2 \times 10^9/L$ (CTC Grade 4) was observed in 1 patient treated with dimethyl fumarate and in no patients treated with placebo.³⁹ It was not clear from the SmPC if the patient was in the bid (769 patients) or tid (771 patients) group but the incidence was calculated as 0.13% and this value has been entered in the effects table.

To November 2015 there have been four reported cases of PML according to the Biogen website. An incidence of 0.029 cases per 1000 patients was calculated using the denominator of 135,000 patients treated to March 2013 with dimethyl fumarate (again taken from the Biogen website). This value is therefore an overestimate as more patients will have received dimethyl fumarate since 2013.

For the remaining unfavourable effects in the analysis, (autoimmune disease, AV block, bradycardia, and herpetic infections) a value of zero was entered for dimethyl fumarate as they are not an issue for this drug.

Teriflunomide 14 mg

In the clinical studies the most common adverse reactions were headache, an increase in ALT, diarrhea, alopecia, and nausea. There were no deaths or serious opportunistic infections in the Phase III trials, and no data have confirmed an increased risk of malignancy or infections.¹⁸

Teriflunomide's US label carries a black-box warning of hepatotoxicity based on preclinical studies, and according to teriflunomide's prescribing information, initial liver function and hematological monitoring is required. Transaminase and bilirubin levels must be obtained prior to initiation and ALT levels must be monitored at least monthly for six months; recent blood count and a tuberculin skin test should be obtained before initiation and patients should be monitored for signs of infection. Teriflunomide has been shown to be teratogenic in animal tests.

Treatment during pregnancy is contraindicated and treatment is not recommended if a patient is planning a pregnancy. Teriflunomide has a long elimination half-life and, in the event of pregnancy or toxicity, an accelerated elimination procedure using cholestyramine or activated charcoal for 11 days (as

described in the prescribing information) may be required.¹⁸

The AE profile in the dataset including all placebo-controlled studies, is presented in Table 33 of the Aubagio EPAR.²¹ Overall the incidence of infections reported as an AE was 256/415 (61.7%) for teriflunomide 14 mg and 242/421 (57.5%) for placebo. SAEs in the Infections and infestations SOC were reported by 2.2% of patients in the teriflunomide 14 mg group and 2.1% of patients in the placebo group.²¹

AEs in the gastrointestinal disorders SOC were reported by 45.3% of patients in the teriflunomide 14 mg group and 34.4% of patients in the placebo group.²¹

AEs of increased alanine transaminase (ALT) were reported for 150/1002 (15.0%) patients in the teriflunomide 14 mg group compared with 89/997 (8.9%) in the placebo group (Hepatic Effects Table 1c in the Aubagio RMP, Module SVII).⁴¹

The number of patients with malignant and benign tumor AEs were 0.3% on teriflunomide 14 mg and 0.5% on placebo (Table 11 in the Aubagio RMP).⁴¹ Exposure-adjusted incidence rate (EAIR) per 100 patient-years was 0.2 in the teriflunomide 14 mg group compared with 0.3 in the placebo group (Table 11 in the Aubagio RMP, Module SVII).⁴¹

The number of patients with serious opportunistic infections, including PML, is summarized in Table 9a in the Aubagio RMP, Module SVII.⁴¹ No case of PML was reported in the whole teriflunomide program (Aubagio RMP, Module SVII).⁴¹ The number of patients who had an AE of the PT "Herpes virus infection" was 5/1002 (0.5%) in the teriflunomide 14 mg group (Table 9a in the Aubagio RMP, Module SVII)⁴¹ and this value (0.5%) was entered into the effects table. This is an underestimate of the incidence of herpetic infections as other PTs in the table were not taken into account (eg, oral herpes 22/1002 [2.2%], Herpes simplex 9/1002 [0.9%] and Herpes zoster 9/1002 [0.9%]). However, it was not possible to determine which of these events occurred in the same patients. The value entered in the effects table favours teriflunomide.

For the remaining unfavourable effects in the analysis, (autoimmune disease, lymphopenia, AV block, and bradycardia) a value of zero was entered

for teriflunomide 14 mg as they are not an issue for this drug.

Section 3

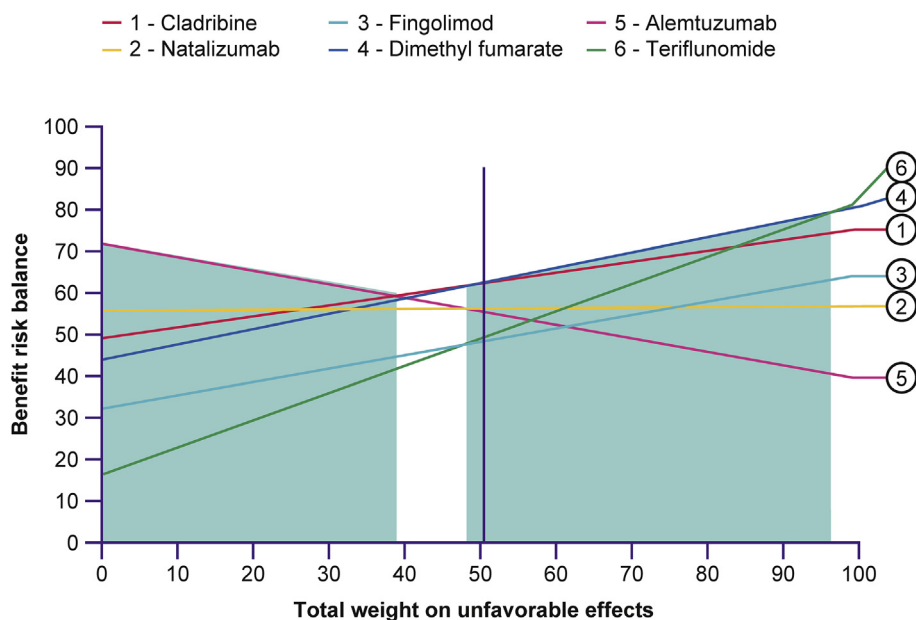
Sensitivity analyses

Two approaches to sensitivity analysis of the model have been considered. The first looks at varying the weighting applied to the individual favourable and unfavourable effects to see if applying different weights would substantially change the overall result. The second changes input data for some drugs on key criteria to see the effects on the overall result.

The experts were most unsure of the weights they had assessed when comparing the biggest swing weights for each of the favourable and unfavourable effects, EDSS 6 months compared to PML. The relative weights between those two criteria was judged to be about equal, so given weights of 100-100. As lower-level weights had already been compared, Hiview3 then normalised all weights so their sum equalled 100, preserving the ratios of all the weights. Any sum of normalised weights is termed a cumulative weight.

Sensitivity analyses then proceeded by changing the sum of the normalised weights under the unfavourable effects node for both models. For the RRMS model, shown in Supplementary Figure 1, the intersection of the vertical line (located at the cumulative weight of the unfavourable effects, 50.7) with the top-most sloping line defines the most preferred option. In the RRMS model, changing the weight on the unfavourable effects node from its value of 50.7 to any value between 0 and 100 showed that dimethyl fumarate was the most preferred drug over a range from about 49 to 96, with changes in the most preferred drug indicated by transitions to green shading. Less than 39 and alemtuzumab is most preferred; more than 96 and teriflunomide is most preferred. Interestingly, at the precise weight of 40, dimethyl fumarate, cladribine and alemtuzumab are nearly equal in preference. Over a range from 40 to 49 cladribine is most preferred, while beyond 49 cladribine is in second position compared to dimethyl fumarate.

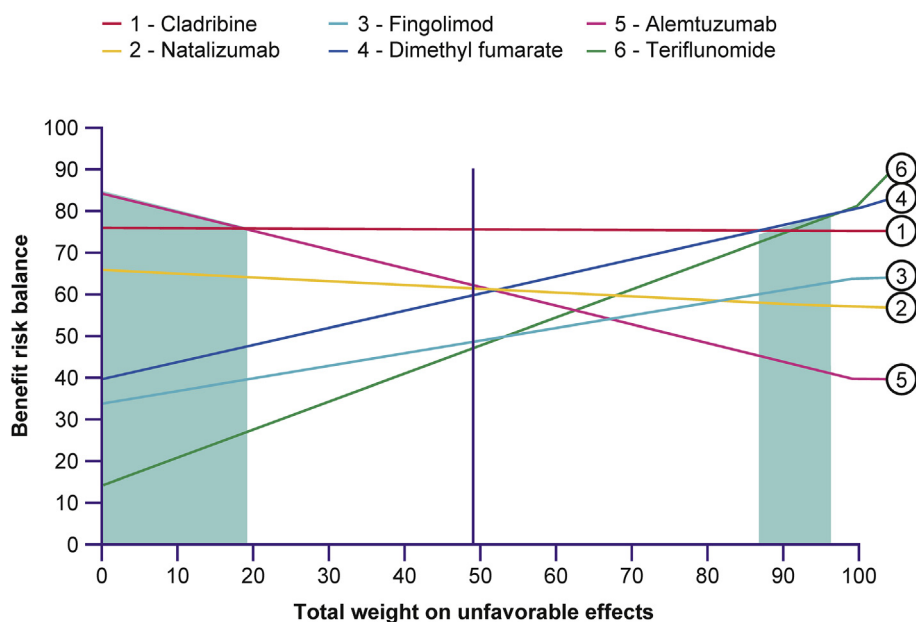
As for changing weights on individual criteria, cladribine becomes most preferred only with increases of the weight on any of Ease of Use, Durability or Infections (not shown).



Supplementary Figure 1. Sensitivity analysis for the unfavourable effects in the RRMS model

The sensitivity analysis for the HDA model showed that cladribine is most preferred over the 19 to 87 range of weights for the unfavourable

effects (Supplementary Figure 2). At a weight less than 19, alemtuzumab is most preferred; above 87, dimethyl fumarate is most preferred until the weight reaches about 96, when teriflunomide becomes best.



Changing weights for individual criteria showed that cladribine remained most preferred for any weight from 0 to 100 on 6 months confirmed EDSS progression, Ease of Use, and Any Infections. For all other criteria, an increase of more than 15 points is required for another drug to become most preferred.

Supplementary Figure 2. Sensitivity analysis for the unfavourable effects in HDA model.

Sensitivity analysis of input data for the favourable effects

The second approach changes input data in the Effects Table. For the favourable effects, more optimistic effects for the other treatment options were considered (except for alemtuzumab). For the unfavourable effects, more pessimistic effects for cladribine were assumed.

Sensitivity analysis for the favourable effects focusses on reduction in the 3-months and 6-months EDSS progression. The changes in input data of EDSS progression assigned to DMDs in this sensitivity analysis (see below) are considered to be within the realm of clinical possibilities.

For alemtuzumab, in an active comparator (IFN- β -1a) Phase III trial in treatment-experienced patients with relapsing-remitting multiple sclerosis (RRMS), there was a 42% reduction in 6-month confirmed EDSS progression, relative to IFN- β -1a.⁹ There are no placebo-controlled studies of alemtuzumab and so the 42% value was arbitrarily increased by 20%, on the assumption that the effect would have been greater relative to placebo, to give a value of 62% which was entered in the effects table. The same value was used in the effects table for the input data of the 3-month confirmed EDSS progression, because this analysis was not reported. This was a conservative approach. For the sensitivity analysis an arbitrary value of 52% was used for both the 3-month and 6-month confirmed EDSS progression in the RRMS population.

For teriflunomide, a data value of 31% was used for 3-month confirmed EDSS progression in the benefit-risk analysis in the RRMS population, based on the result in one of the placebo controlled pivotal studies (TOWER) (the corresponding result in the other pivotal study [TEMPO] was 29.8%). The input data

value used for the 6-month confirmed EDSS progression was 24%, based on the integrated analysis of the TOWER and TEMPO studies. For the sensitivity analysis, an arbitrary value of 62% was used for both the 3-month and 6-month confirmed EDSS progression.

Sensitivity analysis for the high disease activity (HDA) subpopulation

For the high disease activity (HDA) subpopulation, the following amendments were made to the input data for the sensitivity analysis:

For alemtuzumab the input data for 3-month and 6-month confirmed EDSS progression was increased to 82% for both (to match cladribine). Similarly, the input data for 3-month and 6-month confirmed EDSS progression for natalizumab was increased to 82% for both (to match cladribine). For dimethyl fumarate the input data for 3-month and 6-month confirmed EDSS progression was doubled from 21% in the main analysis to 42% in the sensitivity analysis. For fingolimod, the input data for 3-month and 6-month confirmed EDSS progression was increased from 33% to 42% in the sensitivity analysis.

Sensitivity analysis of input data for the unfavourable effects

The unfavourable effects of malignancy, lymphopenia (grade 4) and PML most relevant for cladribine assessment were chosen for the sensitivity analyses. For the sensitivity analysis of these unfavourable effects, the cladribine values for lymphopenia (Grade 4) and malignancy were changed to the most negative value among the other treatment options (52.1% for lymphopenia [Grade 4] and 0.4 new cases per 100 patient years for malignancy).

There were no cases of PML reported in the cladribine studies in multiple sclerosis. The risk of PML for cladribine is currently unknown. The cladribine value for PML was changed to be the same value as fingolimod (0.104 cases per 1000 patients). A second sensitivity analysis was done with cladribine assigned a value for PML that was double that used for fingolimod in the main analysis (0.208 cases per 1000 patients).

RESULTS

The changes to the data in the effects table for input into the model are shown in Supplementary Table 1. Input data that were changed in this sensitivity analysis are shown in bold italicized font in the table.

The overall weighted preference values from the original analysis showed that dimethyl fumarate and cladribine were most preferred for the RRMS model, and cladribine was most preferred for the HDA model.

EDSS is a relevant measure included in most major DMD clinical trials conducted in the past four decades. Slowing disability progression is the ultimate goal for DMD therapies in MS and for this reason, this measure was chosen for the sensitivity analysis for the favourable effects.

The sensitivity analysis for the favourable effects included large favourable changes to values assigned to DMDs other than cladribine. Despite these favourable hypothetical values assigned to comparators, cladribine retained its place in terms of benefit-risk, as the most preferred agent for HDA and its second position (after dimethyl fumarate) for RRMS.

The unfavourable effects of malignancy, lymphopenia (Grade 4) and PML most relevant for cladribine

assessment were chosen for the sensitivity analyses. Unfavourable hypothetical values were assigned to cladribine for these effects. For the RRMS model, when the % of patients experiencing lymphopenia (Grade 4) for cladribine was changed from 0.7% to 52.1%, the overall weighting for cladribine was reduced from 62 to 56 (ie, similar to natalizumab [56] and alemtuzumab [55] and lower than dimethyl fumarate [63]. This was the only sensitivity analysis in which the relative preference for cladribine compared to another DMD was not maintained

For the HDA model, cladribine remained the most preferred option in all of the sensitivity analyses. Even when the % of patients experiencing lymphopenia (Grade 4) for cladribine was changed from 0.7% to 52.1%, the resulting overall weighting for cladribine (70) was still the highest of the six drugs being compared.

In conclusion, these sensitivity analyses indicate that cladribine has a robust benefit-risk balance compared to other marketed agents for MS (highest in HDA patients and second highest in RRMS patients) and are consistent with the original analyses performed with values derived from the clinical trials of these agents.

Supplementary Table 1. Input data for the RRMS Population and the HDA Subset (data shown in parentheses) in the sensitivity analysis (values changed shown in bold italics)

Criteria	Description	Metric	Cladribine	Natalizumab	Fingolimod	Dimethyl fumarate	Alemtuzumab	Teriflunomide
Favourable effects								
Relapse rate	Relative reduction, compared to the control, in annualized relapse rate at 2 years	%	58 (68)	68 (81)	54 (63)	53 (60)	75 (71)	36 (19)
T2 lesions	Relative reduction in mean number of active T2 lesions per patient per scan over 2 years	%	73	83	74	85	87 (92)	53
T1 Gd+ lesions	Relative reduction in mean number of T1 Gd+ lesions per patient per scan over 2 years	%	86	92	82	94	92 (97)	80
EDSS 3 months	Relative reduction in time to 3-month confirmed EDSS progression over 2 years	%	33 (72)	42 (82)	30 (42)	38 (42)	52(82)	62 (35)
EDSS 6 months	Relative reduction in the time to 6-month confirmed EDSS progression over 2 years	%	47 (82)	54 (82)	37 (42)	23 (42)	52 (82)	62 (35)
Ease of use	Ranking based on 4 sub-criteria*	1–3.5	3.5	1.5	2.0	2.5	1.0	3.0
Durability	Number of months of remaining efficacy after stopping the drug	Months	12	2	1	1	12	1
Unfavourable effects								
AR infections	Percentage of patients with any infections	%	51.8	73.7	65.1	60.0	70.9	61.7
AR GI effects	Percentage of patients with any GI disorder	%	31.6	0.0	43.0	44.0	49.0	45.3

(continued on next page)

Supplementary Table 1. (Continued)

Criteria	Description	Metric	Cladribine	Natalizumab	Fingolimod	Dimethyl fumarate	Alemtuzumab	Teriflunomide
Liver functions	Percentage of patients experiencing elevated liver enzymes	%	1.5	0.0	10.1	6.0	0.0	15.0
Malignancy	Number of new cases per 100 patient-years	No/100	0.4	0.320	0.400	0.375	0.370	0.200
Autoimmune disease	Percentage of patients with any autoimmune disease	%	2.0	0.0	0.0	0.0	47.0	0.0
Lymphopenia	Percentage of patients experiencing lymphopenia Grade 4	%	52.1	0.0	18.0	0.13	52.1	0.0
AV block	Percentage of patients with first degree AV block	%	0.0	0.0	4.7	0.0	0.0	0.0
Bradycardia	Percentage of patients with bradycardia	%	0.2	0.0	4.1	0.0	0.3	0.0
Serious infections	Percentage of patients with any serious infection	%	2.5	2.4	1.6	2.0	2.7	2.2
Herpetic infections	Percentage of herpetic infections	%	7.9	8.0	9.0	0.0	15.7	0.5
PML	Number of cases of PML per 1000 patients	No/1000	0.104/0.208	2.100	0.104	0.029	0.000	0.000

Note. This table shows changes input data used in sensitivity analyses of the effects of different input data on the overall results. For the favourable effects, more optimistic effects for the other treatment options are considered. For the unfavourable effects, more pessimistic effects for cladribine were assumed. The values changed and used in the sensitivity analyses are shown here in bold italics.

REFERENCES

- EMA. *Benefit-risk Methodology Project. Work Package 4 Report: Benefit-Risk Tools and Processes*; 2012. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2012/03/WC500123819.pdf.
- Giovannoni G, Comi G, Cook S, et al. A placebo-controlled trial of oral cladribine for relapsing multiple sclerosis. *N Engl J Med*. 2010;362(5):416–426.
- Tysabri US PI. Available from: <http://www.tysabri.com/prescribingInfo>. Accessed Nov 2015.
- Polman CH, O'Connor PW, Havrdova E, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med*. 2006;354(9):899–910.
- Tysabri EPAR. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/000603/WC500044688.pdf. Accessed Nov 2015.
- Kappos L, Radue EW, O'Connor P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med*. 2010;362(5):387–401.
- Gilenya EPAR. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_Public_assessment_report/human/002202/WC500104529.pdf. Accessed Nov 2015.
- Lemtrada US PI. Available from: <http://products.sanofi.us/Lemtrada/Lemtrada.pdf>. Accessed Nov 2015.
- Coles AJ, Twyman CL, Arnold DL, et al. Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. *Lancet*. 2012;380(9856):1829–1839.
- Cohen JA, Coles AJ, Arnold DL, et al. Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. *Lancet*. 2012;380(9856):1819–1828.
- Lemtrada EPAR. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_Public_assessment_report/human/003718/WC500150522.pdf. Accessed Nov 2015.
- Fox RJ, Miller DH, Phillips JT, et al. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *N Engl J Med*. 2012;367(12):1087–1097.
- Gold R, Kappos L, Arnold DL, et al. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N Engl J Med*. 2012;367(12):1098–1107.
- Tecfidera US PI. Available from: <http://www.tecfidera.com/pdfs/full-prescribing-information.pdf>. Accessed Nov 2015.
- Fox R, Miller D, Phillips JT. Clinical efficacy of BG-12 (dimethyl fumarate) in relapsing–remitting multiple sclerosis (RRMS): an integrated analysis of the Phase 3 DEFINE and CONFIRM Studies. *Neurology*. 2013. Abstract P07.097.
- Tecfidera EPAR. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002601/WC500162070.pdf. Accessed Nov 2015.
- Aubagio EU SmPC. Available from: <http://www.medicines.org.uk/emc/medicine/28533>. Accessed Nov 2015.
- Aubagio US PI.
- Confavreux C, O'Connor P, Comi G, et al. Oral teriflunomide for patients with relapsing multiple sclerosis (TOWER): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol*. 2014;13(3):247–256.
- O'Connor P, Wolinsky JS, Confavreux C, et al. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N Engl J Med*. 2011;365(14):1293–1303.
- Aubagio EPAR. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002514/WC500148684.pdf. Accessed Nov 2015.
- O'Connor PW, Li D, Freedman MS, et al. A Phase II study of the safety and efficacy of teriflunomide in multiple sclerosis with relapses. *Neurology*. 2006;66(6):894–900.
- Al-Sabbagh A, B R, Kazma C. Medication gaps in disease-modifying drug therapy for multiple sclerosis are associated with an increased risk of relapse: findings from a national managed care database. *J Neurol*. 2008;255(Suppl 2):S79.
- Steinberg SC, Faris RJ, Chang CF, Chan A, Tankersley MA. Impact of adherence to interferons in the treatment of multiple sclerosis: a non-experimental, retrospective, cohort study. *Clin Drug Investig*. 2010;30(2):89–100.
- Ivanova JI, Bergman RE, Birnbaum HG, Phillips AL, Stewart M, Meletiche DM. Impact of medication adherence to disease-modifying drugs on severe relapse, and direct and indirect costs among employees with multiple sclerosis in the US. *J Med Econ*. 2012;15(3):601–609.
- Oleen-Burkey MA, Dor A, Castelli-Haley J, Lage MJ. The relationship between alternative medication possession ratio thresholds and outcomes: evidence from the use of glatiramer acetate. *J Med Econ*. 2011;14(6):739–747.

27. Tan H, Cai Q, Agarwal S, Stephenson JJ, Kamat S. Impact of adherence to disease-modifying therapies on clinical and economic outcomes among patients with multiple sclerosis. *Adv Ther.* 2011;28(1):51–61.
28. Bloomgren G, Richman S, Hotermans C, et al. Risk of natalizumab-associated progressive multifocal leukoencephalopathy. *N Engl J Med.* 2012;366(20):1870–1880.
29. Panzara, M.A. Natalizumab safety. Available from: <http://www.fda.gov/ohrms/dockets/ac/06/slides/2006-4208S1-01-02-%20Biogen-Panzara.pdf>.
30. Tysabri FDA Assessment report. Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/125104s000_Natalizumab_Medr_P3.pdf.
31. Gilenya EU SmPC. Available from: <http://www.medicines.org.uk/emc/medicine/24443>. Accessed Nov 2015.
32. Calabresi PA, Radue EW, Goodin D, et al. Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol.* 2014;13(6):545–556.
33. Gilenya US PI. Available from: <https://www.pharma.us.novartis.com/product/pi/pdf/gilenya.pdf>. Accessed Nov 2015.
34. Gilenya CDER review. Application number 22-527. August 26, 2010.
35. Gilenya RMP. Fingolimod EU Safety Risk Management Plan (EMA/357270/2012). 18 June 2012.
36. Havrdova E, Horakova D, Kovarova I. Alemtuzumab in the treatment of multiple sclerosis: key clinical trial results and considerations for use. *Ther Adv Neurol Disord.* 2015;8(1):31–45.
37. Lemtrada FDA background package for alemtuzumab (BLA 103948\5139). November 2013.
38. Willis MD, Harding KE, Pickersgill TP, et al. Alemtuzumab for multiple sclerosis: long term follow-up in a multi-centre cohort. *Mult Scler.* 2016;22(9):1215–1223.
39. Tecfidera EU SmPC. Available from: <http://www.medicines.org.uk/emc/medicine/28593>. Accessed Nov 2015.
40. Tecfidera RMP. BG00012 EU Risk Management Plan. March 2013 (Version 5).
41. Aubagio RMP. EU Safety Risk Management Plan. March 2015.