



Disease-Modifying Therapies for Relapsing- Remitting and Primary-Progressive Multiple Sclerosis: Effectiveness and Value

Response to Public Comments on Draft Evidence Report

January 26, 2017



Table of Contents

Manufacturers	3
AbbVie	3
Bayer	12
Biogen	13
EMD Serono	18
Genentech.....	27
Novartis.....	33
Sanofi-Genzyme	39
Teva	44
Patients and Advocacy Organizations.....	48
MS Coalition.....	48
National MS Society	52
Lisa Carr.....	57
Mary Holmstrand	57

#	Comment	Response
Manufacturers		
AbbVie		
1	<p>We acknowledge the difficult nature of comparing the effectiveness of the DMTs approved for the treatment of RRMS, given the limited number of head-to-head trials and lack of real-world data. We agree with the many limitations of this review, all of which add to uncertainty when comparing across trials:</p> <ul style="list-style-type: none"> • The MS trial population is somewhat different over time, as MS diagnostic criteria have evolved significantly since the early MS trials. • The definition of relapses is not consistent across trials. • EDSS was used to measure disability in most of the DMT clinical trials; EDSS is frequently criticized for being insensitive to small changes, being heavily dependent on mobility, being subjective in some assessments with high intra- and inter-rater variability, & not capturing the full range of patient disabilities. • The definition of sustained disability progression is not consistent across trials; some trials measured sustained disability progression over 12 weeks and some measured over 24 weeks. Because some patients have resolution of symptoms between 12 and 24 weeks of follow-up; sustained disability progression observed at 24 weeks is generally lower than at 12 weeks. A critical step to conducting a quality network meta-analysis is to ensure you are comparing the same endpoint. These endpoints are not the same. • MRI technology has evolved significantly since the early MS trials, leading to challenges in comparing MRI outcomes across studies. <p>Because of the limitations when comparing across trials and until more real world data is available, treatment choice for initial therapy as well as subsequent therapies following treatment failure should continue to be a shared decision between the patient and the treating physician, and payers should make every effort to keep a variety of therapeutic options available for their patients</p>	<p>Thank you for agreeing with our delineation of the limitations of the body of evidence. We agree that they introduce considerable uncertainty about the relative efficacy of the DMTs for both initial and subsequent therapy.</p>
2	<p>Page 2 – Scope of the Assessment: We recommend expanding the scope of the assessment to include the use of the DMT prescribing information when necessary to provide more comprehensive data on safety & efficacy.</p>	<p>We did include information from the DMT prescribing information when quantitative values were provided.</p>
3	<p>Page 2, 4, and 14 – Analytic Framework, Figure 1: There seems to be a disconnect between the Analytic Framework and the actual ICER Cost Effectiveness/Budget Impact analysis performed. The analytical framework includes health care utilization outcomes as well as many clinical & patient-centered outcomes measures that are important for employers and patients– permanent disability, days at work, cognitive</p>	<p>We did include a scenario analysis with these costs. Due to limited data available in published literature, we are not able to break out each of these costs separately. The model is over the patient lifetime time horizon, so covers</p>

	<p>function, quality of life, etc. These costs do not appear to be included in the cost effectiveness model. The data used for ICER's analysis had an age of MS onset of 29 years (Page 60), a population for which employers (the ultimate payer for 160MM lives) would use a 10-15 year time horizon to evaluate. We encourage ICER to take a look at cost effectiveness over a longer time horizon (minimum of 10-15 years). If the MS patient becomes permanently disabled, the model should also include lost lifetime wages and the lost taxes tied to those wages, as well as the disability payments and other costs that those patients would incur.</p>	<p>the full time from treatment initiation until death.</p>
4	<p>Page 2 – Interventions: The review should be limited to products FDA-approved for the treatment of RRMS and PPMS; Rituximab should not be included in the review and ocrelizumab should only be included if it is FDA-approved prior to the final report publication. Additionally, there was not enough evidence to assess disability progression, net health benefit, or cost effectiveness of rituximab, further supporting our position that it should not be included in the report</p>	<p>We often assess drugs that are soon to be approved by the FDA in order to provide timely context for policy decisions. We also consider drugs that are being used off-label when there is significant interest in that therapy from the patient and/or provider community. Both groups expressed interest in rituximab during our scoping process.</p>
5	<p>Page 6 – Timing: The report evaluated evidence on harms from studies of at least three months' duration. Most of the MS DMT clinical trials were relatively short duration and some new adverse events were identified through postmarketing surveillance, so the SAE rates from the prescribing information should be used, rather than clinical trial data alone, for a more accurate picture of the safety profiles of the DMTs and to better estimate costs associated with SAEs.</p>	<p>As noted above (#2) we included adverse event data from the prescribing information to inform both the evidence review and the cost model.</p>
6	<p>Page 7 - The Topic in Context: There should be more discussion about the reasons payers should offer more treatment choice and about the need for treatment switch. In addition to the patient preferences mentioned and important factors for shared decision-making about choice of DMT, please consider adding the following:</p> <ul style="list-style-type: none"> • Patients with MS present differently; you never know when a relapse will be a minor inconvenience or will result in substantial irreversible disability and significant cost; preventing relapses is critical. • Individual patient responses to DMTs and response to DMTs as the disease progresses are also components of treatment decisions. • As the disease progresses, patients may experience suboptimal response to their current therapy, necessitating a treatment switch. This is common in patients who develop neutralizing antibodies to beta interferons. • The Multiple Sclerosis Coalition suggests a therapy with a different MOA be considered in the event of signs or symptoms of suboptimal response, including continued clinical and/or MRI disease activity while on treatment. 	<p>We believe that we have adequately addressed the issues. This assessment is not intended to be a practice guideline. Existing guidelines by the AAN address these issues in more detail.</p>

7	Page 8: In Paragraph 2, please change “progressive multifocal encephalopathy” to “progressive multifocal leukoencephalopathy.”	This has been updated in the Evidence Report.
8	Page 9 – Table 1: The following corrections should be made to the daclizumab row: <ul style="list-style-type: none"> • Zinbryta is now registered; please replace the “TM” with “®”. • Zinbryta’s Class is an IL-2 Modulator (anti-CD25 monoclonal antibody) • Zinbryta’s FDA-approved dose is 150 mg once a month; not every 4 weeks 	We have made the changes in the report.
9	Page 9 – Table 1: There appears to be a calculation error for Year 1 WAC, based on the Redbook unit price extrapolated to 365 days for the following DMTs: PriceRx, Medispan, and Redbook <ul style="list-style-type: none"> • Interferon B-1a (Avonex): \$78,710 (not \$75,881) – Calculation: WAC x (365/28) • Interferon B-1b (Betaseron): \$81,065 (not \$69,220) - Calculation: 6218.71 x (365/28) • Interferon B-1b (Extavia): \$67,625 (not \$57,743) - Calculation: 5558.21 x (365/30) • Interferon B-1a (Rebif): \$81,911 (not \$77,827) – Calculation: 6283.57 x (365/28) • Peginterferon B-1a (Plegridy): \$75,881 (not \$73,017) – Calculation: 5821.00 x (365/28) 	This has been updated in the model and report. Current WACs at time of publication were used.
10	Page 9 – Table 1: The FDA-approved dose for alemtuzumab is as follows: 12 mg/day x 5 consecutive days followed by 12 mg/day x 3 consecutive days 12 months later	This has been updated in the Evidence Report.
11	Page 10 – Table 1: If ICER decides to keep rituximab in the report, under “FDA Approved Dose” there should be a note stating that rituximab is not FDA-approved for MS; otherwise the information is misleading.	This has been updated in the Evidence Report.
12	Page 13- Measures Using MRI: MRI technology has evolved significantly since the early MS trials, leading to challenges in comparing results across studies. Please add “According to the MS Coalition, evidence of new MRI activity suggests suboptimal response to DMT, & a change in DMT therapy/MOA should be considered.”	There is not consensus in the clinical community about the degree of new MRI activity that warrants a change in therapy, so we do not feel it is appropriate to highlight one organization’s recommendation. In addition, the MS Coalition guideline is summarized in Section 3.2.
13	Page 14 – Paragraph 1 states the economic hardships that are underappreciated in most economic analyses of MS. It does not appear that these costs were built into this model either, making the model not true to real life.	We included a scenario analysis with indirect costs including productivity losses. Beyond that, we are limited by a lack of data on economic hardships related to MS, and therefore are unable to include further costs in the model.
14	Page 14 - Paragraph 2 states that “For instance, Medicare patients pay an average of more than \$6,000 in out-of-pocket costs per year for Avonex, Tecfidera, or Copaxone.” It should be mentioned that year-to-date in 2016, 63% of MS claims were	We feel that the current paragraph is appropriate as is, but thank you for the additional context.

	Commercial and 26% were Medicare and Commercial patients have significantly lower out-of-pocket costs than Medicare patients. Additionally, the majority of Medicare patients taking DMTs are eligible for extra help, so their out-of-pocket cost is minimal.	
15	Page 14 - Paragraph 3 mentions that patients would like more data regarding the effect of DMTs on patient-reported outcomes. While patient-reported outcomes are not primary endpoints in clinical trials, they are often measured. When available, detailed clinical trial data on patient-reported outcomes should be included in the report, even if they can't be compared across DMTs.	We have expanded the section of the report on patient reported outcomes: Quality of Life section in report section 4.2
16	Page 16 – Paragraph 3 mentions that all payers made use of step therapy to manage therapies for MS, typically a contraindication, intolerance, or inadequate response to one or more preferred injectable therapies (not including daclizumab) or an oral agent. This is not consistent with the model assumption of all DMTs as first-line agents.	For all included DMTs, the label does not preclude use as a first-line agent, therefore all DMTs were modelled as such for completeness.
17	Pages 17-19 - Table 4 – Please update the daclizumab information as follows: <ul style="list-style-type: none"> • Humana: Daclizumab is Tier 5, ST Yes, PA Yes, Preferred Agent No • Health Net: Take N/A out of preferred agent for DAC (it is not there for any other agent) • BSCA: DAC is SP tier, ST Yes, PA Yes 	In a 1/22/2017 search of Humana's commercial formulary, daclizumab was listed as not covered. The information on BSCA's formulary placement for daclizumab has been updated.
18	Page 20: Canadian Agency for Drugs and Technology in Health & Page 21: NICE: While this information provides a perspective of how MS is treated in other countries, the AAN draft guidelines and the MS Coalition consensus guidelines are the only guidelines established by U.S. physicians	Inclusion of CADTH and NICE guidelines is consistent with ICER's approach to summarize key US and ex-US guidelines.
19	Page 21 – MS Coalition, 2016: Please add “The MS Coalition also suggests considering alternative regimens, such as a different MOA, in the event of suboptimal response suggested by additional clinical and/or MRI disease activity.”	We added the sentence “Clinicians should consider alternative regimens using a different mechanism of action when changing therapy.”
20	Pages 26-28: Quality of Individual Studies (Daclizumab) and Pages 117-119: Table C3. Quality Assessment of Included RCTs of DMTs for RRMS: SELECT and DECIDE should meet Good criteria.	See below
21	On Page 27, it states “We judged the study to be of fair quality, primarily because disability progression sustained for 24 weeks was not reported as well as the short follow-up (one year) and relatively large loss to follow-up (11%) for a one-year study. Please consider the following information: While 24-week confirmed disability progression (CDP) was not a pre-specified endpoint in the SELECT trial (FDA requirement to include 12-week CDP), a post-hoc analysis presented atECTRIMS 2012 showed a 56% relative risk reduction for daclizumab (pooled doses) vs. PBO in 24-week CDP (95% CI: 16–77) P=0.012. Havrdova E, et al. ECTRIMS 2012, P949. The majority of MS	The 24-week CDP for the FDA approved dose does not appear to be reported in the ECTRIMS abstract nor in the publication of the post-hoc analyses of Havrdova et al in the MS Journal 2013. The only results in the abstract are the combined results for the 150 and 300 mg dose. Thus the key outcome (24-week CDP is not reported). Hence the “Fair” rating. The 11% was a

	<p>clinical trials are one year or less. The stated 11% loss to follow-up for SELECT is not consistent with the publication cited or the table on Page 119; both indicate 9% loss to follow-up. Additionally, the USPSFT criteria for a Good study states it meets all criteria: Comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent) (Loss to follow up no greater than 20%) Based on this evidence, SELECT appears to meet all of the criteria of a Good study.</p>	<p>transcription error – we have corrected it.</p>
22	<p>On Page 119, for SELECT, the table states “No” under Key Outcomes Assessed column and “Fair” under the Quality column. Please update the columns to “Yes” and “Good”, respectively.</p>	<p>As noted above, this remains “No” and “Fair.”</p>
23	<p>On Pages 27-28 it states that the DECIDE trial was judged to be of Poor quality primarily because of the large loss to follow-up [23%, >20% at 144 weeks]. DECIDE was one of the largest and longest studies in MS. The DECIDE trial allowed treatment up to 144 weeks or until the last patient reached 96 weeks. The mean length of treatment in the trial is beyond 96 weeks; comparing it to standards which may have been developed for trials for 1-2 years duration may not be appropriate. The loss to follow-up rate in the DECIDE trial at 48 weeks was 12% for Avonex and 7% for Zinbryta and at 96 weeks (a duration longer than most of the other DMTs’ follow-up), 20% for Avonex and 16% for Zinbryta (18% blended). Both of these fit within the USPSFT criteria for a Good study. There is no mention on page 28 of why the table states “No” for “Key Outcomes Assessed.” All key outcomes were included in the DECIDE trial. The 24-week CDP was reported in the Kappos 2015 Supplement, Table 3, and should be added to page 28 as follows: Compared to IFNβ-1a at Week 144, Zinbryta was associated with a 27% relative risk reduction of 24-week confirmed disability progression (HR = 0.73) [95% CI: 0.55, 0.98] p=0.0332). On Page 119 for DECIDE, the table states “No-23%” in the Maintain Comparability column. Please replace this with Yes – 18% to reflect the 96 week rate. Additionally, please change the “No” to “Yes” for Key Outcomes Assessed, as indicated above. Please change “Poor” to “Good” under the Quality column. The DECIDE trial was a double-blind, double-dummy clinical trial; this stands in contrast to the rater-blinded design utilized for several other agents, including Rebif and Lemtrada. Was a sensitivity analyses conducted to evaluate the impact of the trial design?</p>	<p>We recognize that the long follow-up of the DECIDE trial is one of its strengths. However we feel that we need to follow the USPSTF guidelines explicitly.</p> <p>We have included the 24-week CDP outcome in the report and in the NMA. We used the “Assume no cases confirmed” result to be consistent with the primary analysis for the 12-week CDP and because it is consistent with the analysis approach used for other agents.</p>
24	<p>Page 127, Table C5: Please consider including the SELECT 24-week CDP reported by Havrdova E, et al. ECTRIMS 2012, P949. Additionally, please update the table with the DECIDE 24-week CDP data according to the Kappos supplement, Table S3 [18% Avonex, 13% Zinbryta; HR=0.73 (0.55 to 0.98) P=0.03]. Please consider including this information in the Base Case NMA, as it</p>	<p>The Havrdova result is for the combined data for two different doses.</p> <p>As noted above, we included the 24-week data from the supplement, but used the “Assume no cases confirmed”</p>

	is not appropriate to compare across different end points in a NMA	analysis approach for consistency with the 12-week result and other studies.
25	Page 27: The statement that the HR for CDP sustained for at least 12 weeks was 0.45 is not consistent with the table on page 131, which states the HR is 0.43. Both should state 0.43.	We corrected the typo.
26	Page 27: The report states “There were also significant improvements in quality of life as measured by the Multiple Sclerosis Impact Scale (MSIS-29) physical score, the EuroQol five dimensions (EQ-5D) summary health index, and the 12-item short form health survey (SF-12) physical and mental health components for daclizumab compared to placebo.” However, MSIS-29 PHYS was not statistically significant within the sequential closed testing procedure. We suggest adding EQ-VAS, as it is mentioned elsewhere in the report	EQ-VAS added.
27	Page 28, Paragraph 1, Line 5 states ARR for daclizumab was lower for daclizumab compared to placebo. Please replace “placebo” with “interferon beta-1a 30 mcg IM”	This has been corrected in the Evidence Report.
28	Page 28: The report states “The DECIDE trial randomized 1,841 patients to daclizumab or interferon β -1a 30 mcg IM each week for up to 144 weeks (median 108.7 weeks).” However, 108.7 weeks was the median duration in daclizumab group only. Median duration for IFN was 111.4 wks. Additionally, we suggest adding “150 mg sc” to clarify the daclizumab dose in the DECIDE trial.	We clarified the median follow-up.
29	Page 28: The report states “The primary outcome compared the relapse rate for each arm using negative binomial regression adjusted for the number of relapses in the year prior to study entry as well as baseline EDSS score and age.” We suggest changing "number of relapses in the year prior to study entry" to "baseline relapse rate," which was not determined solely by prior year but rather the number of relapses in the 3 years before study entry divided by 3. Also need to add adjusted for prior IFNbeta use.	We have clarified the report with this information.
30	Page 28: The report states “There were significant improvements in quality of life as measured by the MSIS-29 physical score and the EQ-5D summary health index for daclizumab compared to placebo. There were also statistically significant improvements on the MSFC at 96 weeks (0.091 vs. 0.055, p	
31	Pages 39-45: The NMA for disability progression should include the 24-week CDP for daclizumab.	The table has been updated with this information.
32	Page 40: The report states “...and the DECIDE study of daclizumab versus interferon beta-1a 30 mcg (RR 0.79 and 0.84).” The RR for daclizumab should be 0.73, not 0.79.	No change: The RR 0.79 and 0.84 use the same analysis approach: assume no cases confirmed.
33	Page 47: The report states “Finally, in the daclizumab trials there were significant differences between the daclizumab 150 mg group and the placebo group in the MSIS-29 physical impact score, but not the psychological impact score.” However, MSIS-	We have revised the report to indicate that the differences were not significant.

	29 PHYS was not statistically significant within the sequential closed testing procedure.	
34	Page 47: Table 11. Harms of DMTs: It is unclear how these percentages were derived. Hepatic injury data is from SELECT, autoimmune hepatitis data is from integrated analysis, and not sure about immune-reactions (PI has 4% for DECIDE and 0.5% for SELECT). Would suggest using the pooled data included in the prescribing information for each DMT. The D/C rate for daclizumab is listed as 15% . What is the source for this? Is this supposed to be D/C rate due to AEs? SAEs? What is the source for the 22% SAE?	Table 11 includes discontinuation rates due to AEs and the SAEs for the FDA approved dose as reported in the clinical trials. Neither the discontinuation rates nor the SAEs are consistently reported in the PI.
35	Page 48 Table: For fair balance, a disclaimer statement should be added such as “Since ocrelizumab is not yet FDA approved, it is not known if it will have a black box warning.” Ocrelizumab clinical trial data also shows higher death rates and higher malignancy rate vs. placebo. These should be included in the table	We have added this disclaimer.
36	Page 50: In paragraph 3, the report states that only one of the 39 reviewed RCTs studied a population that had received at least one prior treatment with a DMT. Both daclizumab pivotal trials included patients with previous DMT experience. Are these really the only two studies that included MS patients who received at least one prior treatment with a DMT?	We have clarified the report to state that only one trial enrolled 100% patients with prior treatment.
37	Pages 51-52, 54: There needs to be more transparency around the objective criteria that differentiate a moderate to large net health benefit and incremental or better net health benefit and the ICER rating on the comparative net health benefit of newer DMTs for RRMS compared to the interferons and glatiramer acetate.	This is a judgement and will be discussed in detail at the public meeting.
38	Pages 58-59: Cost-Effectiveness Model: Methods: The RRMS model is based on treatment-naïve RRMS patients starting on any of the included DMTs as first line, then switching to a second line agent, then transitioning to best supportive care. This is not consistent with current clinical practice, payer coverage policies, or the labeled indication for some of the DMTs. The model does not seem to take into consideration the impact of DMT sequencing (the use of a higher efficacy agent after failure of interferon beta or glatiramer acetate). Both Zinbryta and Lemtrada are generally reserved for use after failure of 2 or more DMTs and payer coverage policies generally restrict use accordingly, so it may not be appropriate for Zinbryta and Lemtrada to be included in the model as first line agents. Additionally, in clinical practice, patients who fail a second DMT move to subsequent therapies; they do not move to supportive care. In a 2009 retrospective analysis of 606 patients with relapsing forms of MS taking a DMT, the average time for those patients who switched DMTs was 3.4 years after DMT initiation. Teter B, Agashivala N, Kavak K, Chouhfeh L, Hashmonay R, Weinstock-Guttman B. Characteristics	Though some DMT labels suggest use later in treatment sequences, no label precludes use as a first-line agent. Therefore, all DMTs were modeled as such for completeness. In the case of MS, there is no standard practice recommended, either in DMT labels, published literature, or clinical guidelines. It is not feasible to model every potential combination of DMTs over time, therefore we chose a more parsimonious model structure. Although many patients do not move to supportive care after only 2 DMTs, there is limited data regarding efficacy of DMTs in this population. Our approach aggregates future treatments to apply averages to all patients. This approach may decrease the resulting relative differences between DMTs, but would

	<p>influencing therapy switch behavior after suboptimal response to first-line treatment in patients with multiple sclerosis. <i>Mult Scler.</i> 2014;20(7):830-836. According to a market research study of 2,374 patients, an MS patient is on initial DMT therapy an average of 4.1 years before transitioning to a second line DMT and the time of second switch is on average 3.4 years later. Data on file H16.DoF.033.</p>	<p>not substantially bias any particular DMT, or the class overall, to look more or less cost-effective. Text has been added to the report accordingly.</p>
39	<p>Page 61: Table 15. Key Model Assumptions: The model assumes that the DMT discontinuation rate was constant for all DMTs and EDSS levels. While clinical trial data may not represent real world discontinuation rates, it is not appropriate to assume all DMTs have the same discontinuation rate. The model also assumes that second-line treatment was evenly distributed across natalizumab, fingolimod, and alemtuzumab. Both Zinbryta and Lemtrada are generally reserved for use after failure of 2 or more DMTs and payer coverage policies generally restrict use accordingly. What is the rationale for including Lemtrada but not Zinbryta as a second line agent in the model? The model also assumes that patients who discontinued on second-line treatment were assumed to follow the natural history progression of disease. Again, in clinical practice, patients who fail a second DMT move to subsequent therapies; they do not move to supportive care. This assumption results in the natural progression of disability and the associated costs being built into the model for all patients after failure of 2 DMTs, which is not seen in the real world.</p>	<p>Discontinuation rates: We have changed our strategy to use differential discontinuation rates by DMT based on trial data. This is described in the report.</p> <p>Second-line: We have decided to use multiple commonly used agents for second-line according to expert clinical option, and have included natalizumab, fingolimod, alemtuzumab, daclizumab, and dimethyl fumarate.</p> <p>Supportive care: see p. 9 comment 38.</p>
40	<p>Page 64 & Page 154: Table E4: Again, most of the MS DMT clinical trials were relatively short duration and some new adverse events were identified through post-marketing surveillance, so the SAE rates from the prescribing information should be used, rather than clinical trial data alone, for a more accurate picture of the safety profiles of the DMTs and to better estimate costs associated with SAEs.</p>	<p>We used SAE rates from prescribing information when quantitative rates were available in the label. Aside from that, we were unable to identify consistent sources of SAE data for all DMTs from observational datasets.</p>
41	<p>Page 65: Drug Acquisition Costs – More transparency about the methodology used by SSR Health to calculate (discounted) drug acquisition costs is necessary; a cost-effectiveness model should be able to be replicated.</p>	<p>In general, SSR’s net price reflects total discounts and rebates. Companies retain discretion over which price concessions are included in reported net sales, but in financial filings typically describe them as encompassing “all usual and customary items.” This information has been added to the report. The WAC and discount rates applied are provided in the report.</p>
42	<p>Page 65: “For alemtuzumab, costs were applied as calculated for year 1 and year 2. For years 3-5, the year 2 cost was applied to 19%, 13%, 16%, and 9% of patients who received an additional course in that year”. Why are there four percentage numbers for the 3rd, 4th and 5th year? How is alemtuzumab modeled after year 5? Do they go to natural progression or</p>	<p>This was correct to year 3-6 in the report. After year 6, patients on ALE do not incur further DMT costs but maintain the health outcomes for ALE until they discontinue.</p>

	second line DMT? This has a significant impact on cost-effectiveness.	
43	<p>Page 66: Table 19. DMT Acquisition Costs – More transparency is needed about the DMT Acquisition Cost for Year 1 and subsequent years. Are drug administration and monitoring costs built in? Are SAE costs built in? Please update the costs for Avonex, Betaseron, Extavia, Rebif, and Plegridy using the annualized methodology suggested for Page 9, Table 1.</p> <ul style="list-style-type: none"> • Interferon B-1a (Avonex): \$62,968 (not \$60,705) – Calculation: $[WAC \times (365/28)] \times (1-20\%)$ • Interferon B-1b (Betaseron): \$52,692 (not \$44,993) - Calculation: $[6218.71 \times (365/28)] \times (1-35\%)$ • Interferon B-1b (Extavia): \$43,956 (not \$37,533)- Calculation: $[5558.21 \times (365/30)] \times (1-35\%)$ • Interferon B-1a (Rebif): \$69,624 (not \$66,153) – Calculation: $[6283.57 \times (365/28)] \times (1-15\%)$ • Peginterferon B-1a (Plegridy): \$68,292 (not \$65,715) – Calculation: $[5821.00 \times (365/28)] \times (1-10\%)$ 	Acquisition costs include only WAC and discounts, as described. These values have been updated in the report. Current WACs at time of publication were used.
44	Page 67: The report states that all drug monitoring costs for alemtuzumab are directly billed to the manufacturer by the laboratory. However, the manufacturer is not permitted to pay drug monitoring costs for Medicare patients. Therefore, some monitoring costs should be built into the model.	We have confirmed with the manufacturer of alemtuzumab that monitoring costs for alemtuzumab are covered for Medicare patients as well, therefore have left the model as is.
45	Page 67: Annual Costs by EDSS State - How were the indirect and direct costs for EDSS state extrapolated from the equations: Direct cost= $4427.7 \times EDSS + 27443$; Indirect cost = $1594.1 \times EDSS + 2,217.5$ to get the direct cost for EDSS=0 \$2825 and indirect cost \$10,711?	Updated in report
46	Page 69-75: Cost-Effectiveness Model: Results - The model should be run again, using the accurate annualized pricing for the beta interferons.	Updated in report
47	Pages 77-80: Potential Budget Impact Model: Methods - The report states that they assumed a lower uptake for daclizumab based on its relatively modest effectiveness, its likely use mainly in JC virus-positive patients, and its potential displacement of only one other drug (natalizumab). This is not consistent with the comparative effectiveness results from the NMA, which shows Zinbryta as the most efficacious self-administered DMT. There needs to be greater transparency around the assumption that daclizumab is only displacing natalizumab in JC viruspositive patients. This is not consistent with clinical practice or physician market research that has been conducted. Additionally, the uptake (10% market share) may be overstated if the assumption is that daclizumab would only displace natalizumab in JC virus-positive patients.	The Evidence Report uses a different approach to estimating potential budget impact, details of which can be found in the full report and in the comment below. We no longer assume that daclizumab is solely replacing natalizumab in JC virus positive patients.
48	Page 80: Table 25: The eligible population is overstated. The number in the table reflects the number of diagnosed MS patients in the U.S., not the number of treated MS patients. DMT treatment rate in the first 2 years after diagnosis of MS is	Under our new approach to budget impact, the eligible population is meant to represent the upper bound of number of patients potentially treated.

	only about 67%. Milliman, April 2016 Multiple Sclerosis: New Perspectives on the Patient Journey. Additionally, the market share of natalizumab patients who are JC virus-positive does not likely equate to 10% of the MS market.	We did not attempt to estimate market share for particular drugs in the revised report.
49	<p>Comments on the Voting Questions:</p> <p>Questions 4, 5 etc.: There is no head-to-head trial between these products and there are many limitations to the NMA & comparative effectiveness model. Therefore, there is not enough evidence to make “Yes or No” determinations. We request the question be restated as “Better than, Similar to, or Worse than”. Currently, if the answer is “No”, it could mean both “similar to” or “worse than”.</p> <p>Question 4. Zinbryta should be both Biogen Inc. and AbbVie Inc. Question 4 compares daclizumab against fingolimod and dimethyl fumarate in terms of clinical effectiveness, but for long term value for money, Question 8 compares daclizumab Glatopa. Shouldn't the same comparator be used for both questions?</p>	We frame our questions to address relevant policy issues and, as such, these questions are appropriate. It is fair to ask whether evidence is adequate as a yes/no question. Whenever we ask about long-term value for money, uncertainty in the evidence comes into play. It is also reasonable to choose a specific comparator when evaluating incremental value, even if other comparators could also be evaluated.
Bayer		
1	<p>The Topic in Context</p> <p>In the discussion of the “Topic in Context” the ICER draft noted that it did not review studies in patients with clinically isolated syndrome (CIS). However, the report goes on to say, “...many patients with CIS never go on to MS, so the results are not directly applicable to the role of DMTs in RRMS.”</p> <p>We believe it is important that ICER mention that there are several publications that look at predictors of the progression from CIS to RRMS which may provide interesting context to this issue. Multiple studies have also been undertaken to assess the impact of early DMT treatment (pre-diagnosis of clinically definite MS) on long-term health outcomes and progression of disease. Thus, the statement from ICER is somewhat misleading as it seems to imply that there is no evidence to support the use of DMTs in CIS patients (despite the fact that the FDA has approved CIS indications for several DMTs) or which CIS patients are more likely to progress and therefore appropriate candidates for early treatment.</p>	We have consistently stated from the draft scope on that we are not considering CIS in our report. Studies of predictors for progression to MS and treatment of CIS are beyond the scope of this report.
2	<p>Table 4: Representative Private Payer Policies for MS DMTs</p> <p>Currently, Table 4 combines Betaseron and Extavia as part of the coverage policy description for Interferon β-1b 250 mcg. However, there are differences in the way these products are covered. For example, in the reference cited by ICER for Anthem BCBS, Betaseron is listed as tier 4 with a prior authorization (PA) but Extavia is listed as Tier 5 with a PA (see the notes column page 72 and 73). We feel it would be more appropriate to list these separately rather than grouping these together as specialty. Betaseron and Extavia have different injectors, syringe sizes (27G vs 30G), and support programs and,</p>	We have updated the representative coverage policy table in the revised report to include separate information for Betaseron and Extavia.

	furthermore, are covered differentially by insurance companies and thus should be considered separately in this table.	
3	<p>Table 11: Harms of DMTs</p> <p>In Table 11, ICER notes that “flu-like symptoms are common.” However, it is recommended that the descriptor be revised to “flu-like symptoms are common (57%) but the incidence decreased over time with 10% of patients reporting flu-like symptom complex at the end of studies” to reflect the language in the product label.</p>	The PI reports 49% and does not discuss the decreasing incidence over time.
4	<p>Limitations</p> <p>In the scoping document, many clinical and quality of life related outcomes were listed as outcomes of interest. While we understand that these cannot be included in the network metaanalysis, the scoping document promised a descriptive evaluation of these outcomes. However, there is little to no discussion of these additional outcomes for each product under review despite the existence of supportive evidence. This should both be added as a limitation of the review (as the benefit of products assesses only relative risk of relapse and progression of disease) and addressed descriptively for each product under review that has supportive evidence on the outcomes of interest reported in the scoping document. This is particularly critical in our opinion for those long-term outcomes such as survival, productivity, and progression to SPMS.</p>	The section on quality of life has been expanded.
Biogen		
1	<p>ICER’s Approach Does Not Fully Capture Important Elements of Value (Perspective)</p> <p>ICER’s report on multiple sclerosis therapies is limited in its perspective. Because MS is complex, we believe the conversation should be about a broader impact and include more real world elements of value. The base scenario’s inclusion of direct medical costs only does not include many aspects of value that are important to society. Societal benefits include the value to caregivers, the value to the healthy from having products available should they become sick, and the reduction/delay in co-morbid conditions due to delay in disease progression. Inclusion of these elements of value has shown to have a significant impact on the value of MS treatments .</p> <p>Caregiver burden has been defined as the type of stress or strain that caregivers experience related to the problems and challenges they face as a result of the status of the care recipient. As multiple sclerosis is a chronic, degenerative, neurological disease, the stress on caregivers will only increase as the disease progresses. There is a significant body of evidence on the impact of multiple sclerosis on patients’ caregivers including impact on utility and quality of life. We believe that any assessment of value of MS treatments must</p>	<p>Section 5: Other benefits or Disadvantages has been expanded to reflect these elements.</p> <p>We have also expanded the section on quality of life measures.</p> <p>In addition, the utilities used in the cost-model incorporate the decrements in quality of life from fatigue, depression, and other mood disorders at each disability state.</p>

	<p>include utility decrements and quality of life changes of caregivers.</p> <p>Recent research has also demonstrated that pharmaceutical interventions provide value to the population who are at risk of contracting a disease. Moreover, among persons at risk, this value rises with increasing impact of the disease. The most current analysis demonstrates that for multiple sclerosis, a significant portion of the value to society was attributed to those who were at risk of multiple sclerosis due to risk aversion. The benefits to the healthy of MS treatment availability should be included in any value assessment.</p> <p>Finally, research has demonstrated that multiple sclerosis is linked to many co-morbid conditions including depression, fatigue, mood swings, and irritability. There is also recent evidence that demonstrates that increasing severity of MS is associated with increases in co-morbid conditions. Therapies that delay progression could also delay the acquisition of co-morbid conditions. The current assessment of value does not incorporate these costs of co-morbid conditions. Incorporating the value of delaying co-morbid conditions should be included in any assessment of value.</p>																																									
2	<p>ICER Excludes Important Published Data on Patient-Relevant Outcomes</p> <p>ICER’s selection of studies, which only included publications that either reported relapse rates or sustained disability progression, failed to recognize many important publications on quality of life, as these outcomes are typically reported in other publications than the main clinical trial publication. Our research indicates that the following outcomes have been measured in the listed clinical trials which meet the inclusion criteria.</p> <table border="1" data-bbox="180 1360 972 1703"> <thead> <tr> <th>SF-36</th> <th>EQ5D Index</th> <th>SF12</th> <th>EQ5D VAS</th> <th>MSIS-29</th> </tr> </thead> <tbody> <tr> <td>AFFIRM</td> <td>ADVANCE</td> <td>ADVANCE</td> <td>CARE MS I</td> <td>ADVANCE</td> </tr> <tr> <td>CARE MS I</td> <td>CONFIRM</td> <td>SELECT</td> <td>DECIDE</td> <td>DECIDE</td> </tr> <tr> <td>CONFIRM</td> <td>DECIDE</td> <td></td> <td>FREEDOMS II</td> <td>SELECT</td> </tr> <tr> <td>DEFINE</td> <td>DEFINE</td> <td></td> <td>SELECT</td> <td></td> </tr> <tr> <td>TOWER</td> <td>FREEDOMS II</td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>SELECT</td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>TEMSO</td> <td></td> <td></td> <td></td> </tr> </tbody> </table> <p>We believe that patient reported outcomes are important in treatment decisions and should be incorporated in an assessment of comparative effectiveness.</p>	SF-36	EQ5D Index	SF12	EQ5D VAS	MSIS-29	AFFIRM	ADVANCE	ADVANCE	CARE MS I	ADVANCE	CARE MS I	CONFIRM	SELECT	DECIDE	DECIDE	CONFIRM	DECIDE		FREEDOMS II	SELECT	DEFINE	DEFINE		SELECT		TOWER	FREEDOMS II					SELECT					TEMSO				<p>We have expanded the section on quality of life outcomes to include all of these results.</p>
SF-36	EQ5D Index	SF12	EQ5D VAS	MSIS-29																																						
AFFIRM	ADVANCE	ADVANCE	CARE MS I	ADVANCE																																						
CARE MS I	CONFIRM	SELECT	DECIDE	DECIDE																																						
CONFIRM	DECIDE		FREEDOMS II	SELECT																																						
DEFINE	DEFINE		SELECT																																							
TOWER	FREEDOMS II																																									
	SELECT																																									
	TEMSO																																									

3	<p>Multiple Inaccuracies and Inconsistencies in ICER’s Analysis and Recommendations</p> <p><i>Population</i></p> <p>ICER’s final scoping document stated that this analysis was done for treatment naïve patients. There are inconsistencies between the current analysis and the stated scope.</p> <ul style="list-style-type: none"> • According to the indication statement for both Daclizumab¹⁰ and Alemtuzumab¹¹, due to their safety profile, the usage of these products should generally be reserved for patients who have had inadequate response to two or more drugs indicated for the treatment of MS. As this review is for treatment naïve patients, these two products should not be included in the analysis. • Irrespective of indication statement, the CARE-MS II¹² trial consisted solely of patients who were previously treated with either interferon or glatiramer acetate. Because there were no treatment naïve patients in this trial, it should be excluded from the analysis of effectiveness. 	<p>The scoping document does not state that the population is treatment naïve patients.</p> <p>The trial populations vary in composition as noted in Appendix Table C1. The trials that looked for effect modification by prior treatment did not find significant differences.</p>
4	<p>While we agree with ICER’s use of three distinct comparative frameworks, inclusion of Rituximab among the “other injectable or infused MS agents” is not acceptable. The draft scoping document states that patient associations recommended the inclusion of nearly all DMTs with current or projected FDA labeled indications for RRMS. Rituximab is neither approved nor projected to be approved by the Federal Drug Administration (FDA) or any regulatory agency for the use in patients with RRMS. Furthermore, there have been no phase III trials in RRMS which tested Rituximab efficacy and safety and only one phase II trial has been completed³. Moreover, a recent Cochrane Review has concluded that there is not sufficient evidence to support the use of Rituximab as a DMT in patients with RRMS⁴. Thus, Rituximab’s inclusion in the framework is unlikely to lend itself to rigorous indirect comparisons and will introduce bias into the overall decision framework. We strongly object to the inclusion of Rituximab in any indirect treatment comparison for RRMS.</p>	<p>Please see response to comment number 4 on page 4.</p>
5	<p>In addition, we caution against ICER’s decision to focus on the direct comparison of intramuscular interferon β1a (Avonex[®]) and subcutaneous interferon β1a (Rebif[®]). While both products are considered injectable platform agents, along with both versions of interferon β1b (Betaseron[®] and Extavia[®]), pegylated interferon β1a (Plegridy[®]), and Glatiramer Acetate (Copaxone[®]), they all differ with respect to posology, dosing frequency, incidence of neutralizing anti-bodies^{5,6} and the degree of use in clinical practice. These are significant differences that warrant assessment of all interferons and Glatiramer Acetate as part of the broader “platform” framework rather than arbitrarily comparing these two interferons directly.</p>	<p>We did not limit the MA to the comparison between Avonex and Rebif. Our analysis compares each of the interferons and glatiramer acetate to each other and to each of the newer DMTs.</p> <p>We were asked by stakeholders to specifically comment on the comparison between Avonex and Rebif – hence the extra focus on that one comparison.</p>

6	<p>Homogeneity of outcomes is another key consideration in indirect treatment comparisons. Since the earliest trials in RRMS, there have been advancements in MRI technology and variations in endpoint definition across clinical trial programs (variations in definition of relapses, no evidence of disease activity (NEDA), among others). In light of this heterogeneity in how outcomes have been defined, as well as the evolution of MRI technology, careful consideration must be taken when selecting endpoints for indirect treatment comparisons. For example, when comparing magnetic resonance imaging (MRI) outcomes, given that the technology utilized in assessing MRI outcomes has advanced substantially⁷ since the introduction of the first DMT, it would not be prudent to compare MRI results from 20 years ago with more recent ones.</p>	<p>We agree and we have not included MRI outcomes in the NMA.</p>
7	<p>There have been multiple attempts to define NEDA resulting in multiple iterations of the outcome (NEDA-3, NEDA-4). These are, in essence, different outcomes that should be treated as such and should not be compared unless there is similarity of underlying components. In the context of indirect treatment comparisons, we recommend against using outcomes where significant heterogeneity of definition exists.</p>	<p>We agree and have not considered NEDA outcomes in the NMA.</p>
8	<p><i>Annualized Relapse Rate (Table C4)</i> ICER’s final scoping document states that “Evidence on intervention effectiveness was derived from studies of at least one year’s duration.” The following trials do not meet the minimum 1- year duration criteria, which is inconsistent with the scope and should be excluded from the final analysis</p> <ul style="list-style-type: none"> • Comi, 2001 comparing glatiramer acetate 20mg daily to placebo is of 9 months duration • Panitch, 2002, the EVIDENCE trial comparing interferon beta 1-a 44mcg TIW to interferon beta 1-a 30mcg once weekly is of 6 months duration • Saida 2012 comparing fingolimod 0.5 mg once daily to placebo is of 6 months duration • O’Connor 2006 comparing Teriflunomide 7 mg once daily, Teriflunomide 14 mg once daily and placebo is only of 8 months duration • Kappos 2011 comparing Ocrelizumab 600 mg IV every 6 months, interferon beta 1-a 30 mcg IM once weekly and placebo is of 8 months duration <p>The ICER report states that it “limited the review to the doses that match the FDA-approved indication except for drugs that do not have a current FDA indication for MS.” The following trials/products do not meet this criteria and should be excluded from an analysis of effectiveness</p> <ul style="list-style-type: none"> • Rituximab 1000 mg IV is not approved for use in patients with multiple sclerosis and has no on-going trials in multiple sclerosis seeking approval. 	<p>Thank you for pointing this out. We have included the trials of 48 weeks duration as they are essentially one year in duration. The shorter trials have been excluded from the NMA.</p> <p>We have excluded the OWIMS trial from the network as well.</p> <p>The results in the Etamadifar study are not consistent. If the results in Table 2 represent the average number of relapses over 2 years after therapy as indicated in the table, then there was less than one relapse per participant (0.7) in the betaseron group. Since there are 30 participants in the betaseron group, the number of relapses must be less than 30. However, in the text of the paper, the investigators report 65 relapses in the betaseron group. If we assume that the 0.7 represents the annualized relapse rate during the full period of follow-up, then the follow-up would be $65/(30 \times 0.7) = 3.1$ years of follow-up. Using this logic, there would be 3.1 year FU for the betaseron group, 1.8 for the Avonex group and 3.2 for the</p>

	<ul style="list-style-type: none"> • The OWIMS 1999 trial was a comparison of interferon beta 1-a 22 mcg once weekly, interferon beta 1-a 44 mcg once weekly and placebo. <p>The following list of technical inaccuracies/inconsistencies in the draft evidence review should be corrected for the final review</p> <ul style="list-style-type: none"> • The actual number of relapses reported in the Etemadifar 2006 trial is as follows: <ul style="list-style-type: none"> o 65 relapses for the Betaferon group o 66 relapses for the Rebif group o 57 relapses for the Avonex group <p>These values have been misreported in Table C4. Given that these values and the reported number of patient years are in the same publication, the annualized relapse rates should be changed.</p> <ul style="list-style-type: none"> • There are additional inconsistencies with reporting of annualized relapse rate in Table C4 with data reported in the literature <ul style="list-style-type: none"> o The annualized relapse rate in the IFNB MS study²⁰ as reported in the original publication was interferon beta 1b 0.84 vs placebo 1.27 after two years. Please correct in the final report. o The annualized relapse rate in the interferon beta 1-a arm of the DECIDE²¹ trials should be 0.4 at 2 years. o The annualized relapse rates for the OPERA I²² trial are 0.156 for the Ocrelizumab arm and 0.292 for the interferon beta 1-a (Rebif 44mcg) arm. 	<p>Rebif group – a difference in length of follow-up that is at odds with the follow-up described in the paper (“All 90 patients...completed their treatment without interruption.” We chose our interpretation of the conflicting data in the paper to be consistent with the CADTH NMA.</p> <p>We have corrected the results for the IFNB MS Study.</p> <p>Table 2 of the DECIDE trial reports the ARR as 0.39. We have not changed this.</p> <p>We have updated the numbers with those in the NEJM publication</p>
9	<p><i>Disability Progression (Tables C5 & C6)</i></p> <p>For the analysis of disability, allowing inclusion of dichotomous data from different time points introduces bias. By not incorporating time through hazard ratios, the analysis is inherently biased by allowing certain trials to only account for one year of disability progression whereas other trials allow for double the amount of time at two years. This could essentially double the number of patients who progress in certain trials. Furthermore, as stated in our response to the draft scoping document, the combination of disability progression at 12 weeks and 24 weeks is essentially combining two separate endpoints. Given that disability progression can be temporary and caused by the residual effect of a relapse, a more robust measure is confirmed disability progression (CDP) over a six month (24 week) interval rather than a 3 month (12 week) interval²³. We again recommend that only confirmed disability progression at 24 weeks measured at 2 years be the sole endpoint used to measure disability progression.</p> <p>Consistent with the inaccuracy of including trials of less than one year in length in the analysis of annualized relapse rate, the</p>	<p>We agree that the ideal would be to have all of the trials analyze their data using a Cox proportional hazards model and use the hazard ratios as inputs to the model. However, fewer trials reported hazard ratios. We did a sensitivity analysis using this approach: see Appendix Table D6, results using continuous measures.</p> <p>We also agree that the preferred endpoint is CDP over 24 weeks/6 months. We present these results (Appendix Table D4), but only 8/15 drugs are represented in this analysis. Several trials reported both 12 and 24-week CDP. In all cases the 12-week outcome was similar, but closer to the null, so it represents a conservative estimate of the 24-week CDP. We felt that this was the best, albeit imperfect,</p>

	<p>following trials should be excluded from an analysis of effectiveness:</p> <ul style="list-style-type: none"> • Panitch 2002, the EVIDENCE14 trial comparing interferon beta 1-a 44 mcg TIW to interferon beta 1-a 30 mcg once weekly is of 6 months duration <p>Allowing for the inclusion of the EVIDENCE14 trial also introduces an inconsistent network. Using the point estimates of the CombiRx24 trial demonstrates that interferon beta 1-a 30 mcg has less disability progression than glatiramer acetate 20 mg. The REGARD25 trial demonstrates that glatiramer acetate 20 mg has less disability progression than interferon beta 1-a 44mcg. Including the EVIDENCE trial would support that interferon beta 1-a 44 mcg has less disability progression than interferon beta 1-a 30 mcg creating inconsistency within the network. Given that it doesn't meet the inclusion criteria and creates an inconsistent network, the EVIDENCE trial should be excluded from the analysis.</p> <p>The following are additional technical inaccuracies/inconsistencies with Tables C5 & C6 which should be corrected in the final report.</p> <ul style="list-style-type: none"> • There are no values for Calabresi 201426 comparing peg-interferon beta 1-a 125 mcg to placebo, yet this product was included in the analysis. For transparency, these values should be included in Table C5. • As stated in our response to ICER's data request, the hazard ratio in the AFFIRM27 trial for CDP24W comparing natalizumab 300 mg to placebo is 0.46 and not 0.58. For an accurate representation of Natalizumab, this should be corrected in the final report. • The CARE MS II12 study comparing Alemtuzumab 12 mg to interferon beta 1-a 44 mcg is solely previously treated patients and should be excluded from an analysis of treatment naïve patients. • The hazard ratio for CDP24W in the DECIDE21 trial comparing Daclizumab 150 mg to interferon beta 1-a 30 mcg is 0.73. For accurate representation of Daclizumab, this should be corrected in the final report. 	<p>approach to a comprehensive comparison across DMTs.</p> <p>We have excluded all trials with <48 weeks of follow-up from the NMA. The Evidence trial reports 48 week outcomes.</p> <p>The number of outcomes in the EVIDENCE, CombiRx, and REGARD trials is low and none of the differences in the trials are statistically significant. The relative ordering of the findings in each of these individual trials is not significant and cannot be used as evidence for inconsistency in the network. The value of the MA is to leverage the data from small, underpowered trials.</p> <p>Calabresi did not report CDP 24 week results.</p> <p>We have corrected the table.</p> <p>Other trials included a mix of patients previously treated and treatment naïve patients. There is no evidence supporting effect modification by prior treatment, so we have included the trial.</p> <p>The DECIDE trial reports 3 different HRs for CDP24W. To be consistent with their primary outcome (CDP12W) we used the result for CDP24W that used the same analytic approach (0.79 – see Supplement Table 3 DECIDE trial).</p>
EMD Serono		
1	<p>Lack of Patient-Centric Focus and Measures</p> <p>ICER acknowledges that there is a “mismatch between concepts and terms used to describe value across patients, clinicians, innovators, and payers” and rightly states that “patients should be at the center of the discussion.” However, the current draft report focuses on cost-effectiveness and budget impact - it is</p>	<p>The patient input is highlighted so that the voting members understand the values of patients and can use that information to appropriately weight the potential benefits and harms for each DMT.</p>

	<p>clear that the intent of the report is solely from the perspective of the payer. While we applaud the steps ICER has taken to solicit patient input in the generation of this report (ICER Report, p.13), it is unclear how this input has been incorporated into the value assessment, and the lack of patient-centricity is starkly apparent in the heavy focus on system costs and short-term budget impacts as outputs of the value framework. This represents a movement away from the recognizable shift towards patient-centric and personalized medicine among policy-makers and regulators.</p>	<p>In addition, patient values drive the quality of life inputs to the cost model. It is patient input that drive the quality of life adjustments.</p> <p>Unfortunately, there have been no consistently used set of patient reported outcomes in the pivotal clinical trials. We have expanded our section on quality of life outcomes in the final report, but the lack of a consistent measure precludes comprehensive comparative effectiveness evaluation.</p> <p>Finally, we have tried to highlight additional Benefits and Disadvantages that are not reported in the trials or fully captured in the model in Section 5 of the report.</p>
2	<p>In MS, where patient experience varies greatly and the condition can have severe impacts on day-to-day living, an aggregate evaluation approach is inappropriate, whereas a more nuanced approach is far more reflective of the needs of MS patients and the full value that innovative medicines offer in improving patients' ability to remain actively engaged in all aspects of their lives. EMD Serono encourages ICER to expand the range of benefits that are considered in the cost-effectiveness analysis to include more patient- and disease-specific measures. For example, recent literature affirms that patients value the ability to choose among different therapeutic options. The current report fails to formally capture patient-centric metrics such as ease of access, administration, choice, and patient peace of mind. Within MS, studies have shown that early treatment following a diagnosis of RRMS can make a significant and positive impact on long-term outcomes. Disease modifying therapies (DMTs) have been shown to reduce the rate of relapse, and with some therapies, slow disability progression. The median age of onset for patients with MS is 30 years, striking patients as they are starting their careers and families. The upfront costs associated with treatment can provide long-term benefits including increased productivity, improved quality of life, slowing of disability progression, and reduction in relapses. The heavy focus on system costs, short-term budgets, and value to the "health system" rather than the benefits and costs actually borne by patients limits the utility of the draft report by de-prioritizing the interests of patients and elevating those of a diffuse "system" community. While a short-term impact analysis may be appropriate for a limited number</p>	<p>We agree that MS is a heterogeneous condition. In order to complete comparative assessments between DMTs, we are limited to including outcomes that have been consistently reported across trials. Because the same patient reported outcomes are rarely used in multiple trials, we are limited to use of only progression and relapse rates for outcomes. We did use a lifetime time horizon for this model, so both short term and long term costs are emphasized. We did complete a scenario analysis including indirect costs.</p>

	of acute conditions, patients with RRMS are suffering from a long-term chronic disease, whose experiences may not be adequately appreciated when only short-term costs are considered as in the 5- year budget impact analysis. As such, the five-year budget impact analysis reinforces the focus that many payers have on short-term costs at the expense of long-term patient health and well-being.	
3	We therefore encourage ICER to more fully incorporate the feedback they have received from patient advocacy groups into its cost-effectiveness analysis. These groups can, and have, offered patient views on value and explorations of value beyond what is calculated into the quality-adjusted life year (QALY), especially as the use of QALY-based thresholds has been explicitly prohibited from being used by United States public payers in the Affordable Care Act (ACA). This more holistic view offered by patient advocacy groups would foster a more informed discussion on value than the current payer focus. An assessment that can be used collaboratively by patients, clinicians, and payers would contribute to “the conversation” more effectively than an assessment targeted solely at one stakeholder group.	We agree that the increased engagement with and feedback from patient groups has greatly enhanced our report content and process. We will also be releasing our proposed approach for more explicit integration of patient and contextual factors into cost-effectiveness considerations as part of our revised value framework release in early February.
4	Lack of Transparency and Methodological Limitations While EMD Serono appreciates the efforts that ICER is making to improve transparency and inclusiveness in the value assessment process, there are several components that could be improved to ensure adequate input from all stakeholders.	
5	Best practices have found that engaging key stakeholders continuously throughout the process ensures that the resulting scope and analysis are relevant to a broad range of users, while limiting investigator bias ¹⁰ . Policies made using ICER’s evaluations now risk a narrow focus given all relevant stakeholders are currently not meaningfully included in the process. EMD Serono recommends that ICER allow for more input and stakeholder participation throughout the entire evidence report development process, and acknowledge how the feedback that ICER receives is incorporated into their analysis. As such, we suggest that ICER consider the following: <ul style="list-style-type: none"> • provide opportunities for full participation during the entirety of the appraisal process; <ul style="list-style-type: none"> o EMD Serono appreciated the opportunity to review and comment upon the cost-effectiveness model analysis plan; however, the detailed network meta-analysis (NMA), budget impact, and systematic literature review analysis plans were not made available until the release of this draft report, thus limiting the opportunity to provide review and comment of the analysis plans. • make public the feedback received from all stakeholders, at all stages of the evidence report development process; 	While we recognize the natural tension that exists between the milestones necessary for a timely report process and the possibility of multiple stakeholder touchpoints, we feel that we have created multiple opportunities for stakeholder feedback and engagement beyond those discussed here, including: <ul style="list-style-type: none"> • Publication of our protocol for the evidence review and quantitative synthesis on the Open Science Framework website; • Provision of draft data tables for our network meta-analysis for review by all manufacturers • Discussion of both our network meta-analysis approach and findings as well as the cost-effectiveness model during preliminary results discussions

• take a sufficient amount of time to review these comments, re-evaluate relevant evidence, and incorporate necessary changes into the final report, including a point-by-point commentary of how each comment was considered and addressed. This type of approach typically occurs at other health technology assessment (HTA) agencies;

o EMD Serono has appreciated the opportunity to participate throughout this process and has provided extensive commentary and suggestions at each opportunity for interaction; however, based on a comparison of the initial model analysis plan with that which was included in the draft report, few of our suggestions have been incorporated, without any rationale as to why (i.e. our recommendation against conflating 3- and 6-month confirmed disability progression data in the NMA).

o To cite an additional example in more detail, EMD Serono reviewed the initial list of studies identified by ICER for the NMA and strongly recommended against the inclusion of data from the OWIMS study as this study evaluated Rebif once-weekly (qw) dosing administration, which is not an approved FDA dosing regimen for the RRMS indication. ICER is evaluating the Rebif thrice-weekly (tiw) dosing administrations and thus, the OWIMS study should be excluded to be consistent with the dosing regimen being evaluated; however, this study remains in the NMA (and is erroneously labeled as tiw dosing in the ICER Report, Tables C1-C6) and the OWIMS data continue to inform the network in the draft version of this assessment, leading to a bias of the analysis against Rebif.

• ensure a robust methodology review by (1) publishing the over-arching framework methodology in a respectable peer-reviewed journal; (2) publishing each framework per disease area - prior to preparing its scoping document – in a relevant peer-reviewed specialty journal to ensure the methodology is appropriate for the particular disease state; and (3) publishing the final report in an appropriate peer reviewed journal.

o For example, the ICER Ratings of Comparative Net Health Benefit (ICER Report, p.52) framework, which aims to summarize the relative efficacy and safety of each therapy, is qualitative and non-transparent. Surface under the cumulative ranking curve (SUCRA), a well-accepted and quantitative analytical framework for Bayesian NMAs, could have strengthened support for this analysis. EMD Serono is a business of Merck KGaA, Darmstadt, Germany.

o EMD Serono further believes that the methods in the draft report insufficiently includes the potential risks of

The OWIMS study has been excluded.

	<p>each respective therapy, in part because safety outcomes have been poorly analyzed and reported by ICER. A mixed treatment comparison or an indirect treatment comparison could have been performed to more robustly evaluate the comparative safety risks associated with each product.</p>	<p>SUCRA has significant limitations as highlighted recently (see Trinquart et al, Ann Int Med 2016 164:666-673). We felt that presenting SUCRA results would lead the reader to place too much confidence in the SUCRA findings. Given the many uncertainties in the NMA assumptions and the limits of the data, we felt it better to present the results and allow the reader to draw his/her own conclusions about the data.</p>
6	<p>Further, the lack of supporting detail and transparency related to the assumptions made by ICER in the methods section of the draft report is a source of vulnerability for entry of bias into the analyses. As a specific example, ICER's use of SSR Health data as its basis for drug discounts is likely to result in an inaccurate measure of budget impact or cost-effectiveness for two reasons. First, the specific methods are not described (ICER Report, p.65) and the report lacks transparency into the calculations performed to arrive at the discount rate for each intervention. We cannot know, given the information provided, whether this approach incorporates numerous discounts, including mandatory discounts to Medicaid or 340B sales channels, or discounts to the DOD or VA. Secondly, drug discounts vary significantly from payer to payer; by failing to highlight this point in its results, ICER risks limiting the transferability of its findings to key payer audiences. For example, in the report's base case, despite comparable wholesale acquisition costs (WAC), Betaseron's net acquisition cost is 32% lower than Rebif's net acquisition cost (ICER Report, Table 19), which has a significant impact on the overall cost-effectiveness results for this pairwise comparison. EMD Serono recommends instead that WAC be used in the model's base case, with differing levels of discount applied to each intervention in a sensitivity analysis.</p>	<p>We have heard strong and consistent recommendations from various stakeholders that WAC prices do not reflect real world payments, and that most purchasers receive some level of discount from WAC. In general, SSR's net price reflects total discounts and rebates. Companies retain discretion over which price concessions are included in reported net sales, but in financial filings typically describe them as encompassing "all usual and customary items." This information has been added to the report.</p>
7	<p>Additionally, although off-label use of therapies may be utilized to treat conditions, EMD Serono is concerned that ICER was</p>	<p>All knowledge about therapies evolves over time as new evidence becomes</p>

	<p>unable to find extensive published literature and evidence for therapies that have yet to receive FDA approval for certain indications, as evidenced by the inclusion of Rituxan in the Scoping Document and NMA but its exclusion from the cost-utility model due to lack of usable disability progression data. As the body of evidence on both the costs and benefits of newer medicines increases over time, ICER’s evaluation of emerging medicines may not adequately capture the true value of these interventions. EMD Serono recommends that ICER categorize any assessments of emerging therapies as preliminary, and schedule these medicines for reassessment as more evidence becomes available.</p>	<p>available. However, all stakeholders need to make decisions based on currently available evidence.</p>
8	<p>Along these lines, despite ICER’s reported intention to focus on assessing effectiveness, the employed approach represents a missed opportunity to highlight the real-world value of therapies that have delivered a meaningful benefit to patients. Long-term outcomes data, such as that from PRISMS-15, has demonstrated that after 15 years, cumulative exposure to and long-term treatment with Rebif is associated with sustained efficacy on key measures of clinical disease activity. Yet, these long-term studies are not considered in ICER’s calculation of net health benefits. While the design of long-term extension studies may be associated with certain limitations, in this report ICER nonetheless highlights the benefits of other products that have different types of limitations in their evidence bases, given their lack of FDA approval, as highlighted above.</p>	<p>PRISMS-15 includes data on only 52% of the original sample and thus suffers from selection bias which makes any conclusions about effectiveness suspect.</p>
9	<p>Finally, after discontinuation of an initial treatment, there is limited guidance on the sequence of therapies that may be used in the treatment of MS. The choice of subsequent therapy may depend on the reasons for stopping prior therapy. Thus, the decision to switch may depend on the history of the patient up to the point of discontinuation. In all cases, the impact of treatment switching on health outcomes and costs is complicated by the lack of trial evidence on the effect and safety of these therapies in a sub-optimally treated patient group. We therefore reiterate our concern with ICER’s choice of approach in modelling treatment sequencing in MS, where all patients discontinuing first-line treatment will receive a weighted average of alemtuzumab / natalizumab / fingolimod in second-line. We do not believe that this assumption is supported by current DMT treatment patterns, as recent work has shown a large proportion of patients currently taking self-injectable DMTs (42.2%) had received one or more prior DMT treatment (ie., self-injectable or other). We also believe that it is inappropriate to consider alemtuzumab or daclizumab as first- or second-line treatment options for treatment-naïve patients, as the FDA recommends that both DMTs “should generally be reserved for patients who have had an inadequate response to two or more drugs indicated for the treatment of</p>	<p>See p.9 question 38.</p>

	MS". We therefore believe it is not appropriate to exclude a switch to other treatment options (i.e., self-injectable or oral DMTs) in second-line.	
10	<p>Consistency and interpretation of results</p> <p>Each of the beta-interferons is unique, in terms of their approved doses, frequency/mode of administration, efficacy, and safety. A well-controlled large, randomized, single blinded head-to-head trial of Rebif 44 mcg tiw and Avonex 30 mcg qw interferon formulations in RRMS (EVIDENCE) has demonstrated the superiority of Rebif 44 mcg tiw over Avonex 30 mcg qw on clinical efficacy, including proportion of patients without relapse, and MRI outcomes. A statistically significant relapse benefit for Rebif 44 mcg qiw-treated patients has been demonstrated out to 16 months. The strength of this evidence enabled the marketing of Rebif prior to the expiration of the orphan status exclusivity period of Avonex¹⁹. Based on the results of this study, other health services researchers, including the Drug Effectiveness Review Project, have concluded that "there is fair evidence that Avonex is less effective than Rebif for preventing relapse in patients with relapsing-remitting multiple sclerosis". ICER itself highlights in its report that "we judge there to be moderate certainty of a small-to-substantial net health benefit for Rebif compared to Avonex, with high certainty of at least a small net health benefit" (ICER Report, p.54-55), and indeed, the 95% credible interval for Rebif 44 mcg tiw versus Avonex 30 mcg qw from the NMA on annualized relapse rate (ARR) does not contain 1 (ICER Report, Table 6). It is therefore curious that, elsewhere in the report, the interferons are either grouped together indiscriminately on measures of efficacy (i.e., Figure 5), or no difference is found in their comparative net health benefit (ICER Report, Table 12 where both Rebif 44 mcg tiw and Avonex 30 mcg qw are rated by ICER as a "B"). EMD Serono recommends adjusting the report accordingly to correct the misrepresentation of these DMTs as identical to each other.</p>	<p>Thank you for supporting our detailed analysis of the evidence comparing Rebif and Avonex.</p> <p>That said, when placed in the context of the newer DMTs, the effect sizes for the interferons group together even though there are small differences between them. Others in analyses using a variety of data sources and methodologies published in 2016 have reached similar conclusions (Tolley et al, PLOS ONE 2016:10(06); Fogarty et al, MS and Related Disorders, 2016; Einarson et al, Current Medical Research and Opinion 2016.</p>
11	<p>Second, researchers have suggested that, given that outcomes in MS are more based on disability than on life expectancy, a QALY-based incremental cost-effectiveness ratio (ICER) may not be a viable metric in this disease area. As relapses are a key contributing factor to sustained disability progression²², it has been suggested that the cost per relapse avoided should instead be considered as a more relevant measure of cost-effectiveness. On this endpoint, a wealth of prior research has shown the superiority of Rebif 44 mcg tiw over Avonex 30 mcg qw, when both short- and long-term horizons are considered. Cost per relapse avoided was considered in ICER's analysis, but its results were not highlighted in the report's conclusion. When considering ICER's analysis, Rebif 44 mcg tiw's ICER vs. Avonex 30 mcg qw is cost-effective with an ICER of \$98,684 per</p>	<p>Cost per relapse avoided has been emphasized in the report and added to the executive summary.</p>

	relapse avoided (calculated from data in ICER Report, Table 21), a value that is in line with prior analyses	
12	Third, EMD Serono reiterates our concern with including OWIMS data in the NMA. OWIMS was a trial evaluating the efficacy of Rebif dosed qw versus placebo. No statistically significant impact on relapse or disability was demonstrated, and EMD Serono did not apply for FDA approval of this Rebif dosing regimen. The inclusion of this data in an analysis evaluating the comparative efficacy of Rebif dosed tiw is inappropriate and likely to significantly bias the results of the analysis against Rebif. We recommend that the analysis be adjusted to exclude data from the OWIMS study.	We excluded the data from OWIMS.
13	Fourth, EMD Serono recognizes the difficulties in conducting NMAs in MS, especially the likelihood and impact of heterogeneity. In part this may arise due to the period of time over which evidence has been generated in pivotal and other studies. In the case of the longer acting interferon (Plegridy), the least is known from only one placebo controlled study, with a study duration of 48 weeks. Inclusion of this one study with only 48-week data in the network for Plegridy, alongside other interferons with predominantly 96-week data from multiple studies, is likely to bias the results of the analysis in favor of Plegridy. There is also evidence comparing the interferons with glatiramer acetate through head-to-head studies (BEYOND, REGARD and COMBIRx), but no such comparison is available involving Plegridy that can inform the network, also likely biasing the results. In summarizing the cost-effectiveness results versus best supportive care (ICER Report, Table 23), ICER reports that Plegridy is a more cost-effective option than Rebif. Given the points we highlight concerning the evidence base for Plegridy, this conclusion cannot be qualified from the data presented.	Thank you for highlighting the uncertainty due to the paucity of data on Plegridy.
14	Fifth, as we have highlighted in our prior feedback responses, we remain unclear on the reason why Rebif 22 mcg is the only interferon in the model for which serious AEs are accounted for (ICER Report, Table E4), given that Rebif 22 mcg has a similar adverse event profile to other interferon therapies (including the higher dose Rebif 44 mcg preparation, for which, counterintuitively, no serious adverse events are accounted for in the model). Along these lines, the risk of progressive multifocal leukoencephalopathy (PML) appears to only be accounted for with natalizumab, while post-marketing experience has also demonstrated a risk of PML with other treatments. More generally, it is unclear why ICER limits themselves to sourcing adverse events for the DMTs from clinical trial data, when many DMTs have a wealth of real-world safety data available, which could be used in addition to clinical trial data to inform the risk profile of DMTs in real clinical practice.	<p>The report focused on uniformly available and quantifiable adverse event rates, therefore we focused on clinical trials and packaged inserts. We are including risks for Rebif 44 mcg equivalent to 22 mcg, although risks were not identified from trial data. We have updated the report to clarify this point. PML was included only for natalizumab because it was the only DMT to report a quantitative risk in the label.</p> <p>The risk for PML is noted in the Harms section of the report for other DMTs, but there are no measures of their incidence because they are so</p>

		uncommon. Since the incidence is very low, including some estimate would have little impact on the overall model results.
15	Sixth, ICER presents the results of their cost-effectiveness analysis versus two reference treatments: best-supportive care, and Glatopa. Payer decision-making relies on assessing the incremental benefits and costs of an intervention versus a current standard of care. We do not believe the current standard of care for treatment-naïve MS patients is adequately represented by ICER’s choices of referent therapies, given that (1) recent evidence suggests that the majority of newly diagnosed patients in the United States are treated with a DMT29 and (2) Glatopa is not the most prescribed DMT in the United States. We suggest revising the reference treatment in the cost-effectiveness analysis accordingly	We selected this universal comparator because glatiramer acetate is the most commonly prescribed DMT, and Glatopa is the least costly option. There is no evidence of differences in efficacy between branded and generic 20mg glatiramer acetate.
16	Finally, the NMA validation relative to existing literature, as well as the model’s cost-effectiveness results, are presented in the main text tables and the conclusions section without any accompanying assessment of uncertainty, limiting their transparent interpretation. Early in the report, ICER itself highlights the limitations of indirect evidence: “The credible intervals for most of the drugs are quite wide, highlighting the limitations of indirect evidence to distinguish one drug or set of drugs from the others” (ICER Report, p.40). Indeed, when considering ICER’s NMA results for disability progression (ICER Report, Table 9), the majority of self-injectable pairwise comparisons have 95% credible intervals that contain 1. ICER’s use of Forest plots are open to misinterpretation: by arbitrarily placing one intervention higher than another in these figures, ICER implies rank-ordered meaningful differences in efficacy, when in fact the credible intervals are so wide that this conclusion would be unwarranted. Given that disability progression is a key model driver, and the point estimates from the NMA were used as model inputs in the base case, it is inappropriate to draw conclusions on cost-effectiveness based solely on the results of this deterministic analysis. This is underlined by the huge variability in the analysis results when deterministic and probabilistic sensitivity analyses are undertaken (ICER Report, Tables E11-16). EMD Serono suggests that (1) measures of uncertainty be presented in the main text tables as well as the report conclusions, and (2) language pertaining to ‘superiority’ be removed from the report, in order to better inform decision-makers with a fair view of the uncertainty of the evidence.	<p>We have included 95% confidence intervals in all presentations of the estimates from our NMA including the league tables, the Forest plots, and all of the sensitivity analyses (Appendix Tables D1-2, D4-5) We did not feel that they were helpful or necessary in the comparisons with prior NMAs.</p> <p>We specifically elected not to include SUCRAs in the report, because we do believe that they give unrealistic weight to the true probability that one or a group of therapies is superior.</p> <p>We agree that there is considerable uncertainty – that is why we have included the sensitivity and probabilistic analyses and tried to highlight the sources of heterogeneity throughout the text.</p>
17	EMD Serono would like to emphasize that findings from this assessment could potentially – and in our opinion, mistakenly - be used to influence individual patient care. In the spirit of transparency, ICER should share use cases for how they see this	We believe that stakeholders can make reasonable judgments about how to use evidence such as that contained in ICER’s report.

	<p>information being utilized, as it is unrealistic to assume that this information will be used only at a theoretical population level and not trickle down to clinical decision-making. ICER should clarify that any assessments based on this Value Framework should only be used as an aid to stimulate discussion on the idea of value in healthcare from a broader population-based perspective. As the data used to develop the assessments are primarily drawn from population-based studies, ICER should explicitly clarify that value assessments based on this framework should not be used to guide clinical decisions or individual patient treatment plans. The American Academy of Neurology has “urge[d] access to all DMT for treating MS individuals when they have the potential to provide clinical benefit”. Given the results from ICER’s own NMA clearly show that all FDA-approved therapies reduce ARR relative to placebo (ICER Report, Table 6), EMD Serono strongly believes that under no circumstance should any result from this report be used to compromise patient access to treatment.</p>	
Genentech		
1	<p>We are concerned about errors and the lack of transparency of the methodology in different areas including the network meta-analysis (NMA) and economic models, for which Genentech was unable to replicate the results. Key areas that Genentech would like to address are correcting errors in the NMA of the annualized relapse rate (ARR), upgrading the ratings of ocrelizumab based on the publication of the manuscript, and deferring the economic evaluation of ocrelizumab until a price is available.</p>	<p>Thank you for identifying the error in our NMA – there was a transcription error that has been corrected.</p>
2	<p>Correct ocrelizumab’s point estimate and 95% credible interval for the rate ratio of ARR in the NMA and re-check numbers in all NMAs, including subgroup analyses and metaregressions. The corrected numbers should be reflected in the evaluation of the rating for ocrelizumab and the cost-effective analysis.</p> <ul style="list-style-type: none"> ● There was a discrepancy in the rate ratios for ARR between the OPERA trials and the ICER conducted NMA, which is not typically observed between direct and indirect evidence. <ul style="list-style-type: none"> ○ The OPERA I and OPERA II trials compared ocrelizumab to IFNβ-1a 44mcg and a rate ratio of 0.54 and 0.53, respectively, was reported in the trials; a much higher ratio of 0.66 was estimated in the NMA in the draft report (Page 37, Table 6), drawing questions on the face validity of the NMA. ○ In Genentech’s attempt to replicate the NMA of ARR using ICER’s methodology and the inputs listed on Table C4, our result for the rate ratio for ocrelizumab vs. placebo was 0.34 with a 95% credible interval of [0.27, 0.42], which is different than ICER’s result of 0.43 [0.34, 0.54]. 	<p>As noted above, we identified the data input error in our NMA and have corrected it in the final report.</p>

	<ul style="list-style-type: none"> ○ We were able to replicate all other point estimates to +/- 0.02 for rate ratios of DMTs vs. placebo, except for ocrelizumab, in Figure 3. ● Given the short public comment period, we did not have sufficient time to replicate all the analyses, and request that the numbers for all NMAs should be verified, including the subgroup analyses and meta-regressions. 	
3	Conduct and add separate analyses for confirmed disability progression (CDP) at 12- and 24- week confirmation, as the two endpoints are not directly comparable.	See Appendix Tables D3 and D4.
4	<p>Upgrade the quality of OPERA I and OPERA II trials from “fair” to “good” because all quality criteria have been met as described in ICER’s quality criteria.</p> <ul style="list-style-type: none"> ● In Table C3, the OPERA trials received an “Unclear” for “Comparable Groups.” Please refer to Table 1 of the recently published manuscript for the baseline demographics and disease characteristics, which were similar between treatment groups in the two trials. ● On page 28, The OPERA trials were judged to be of fair quality because they were “presented in abstract form” and “due to relatively high loss to follow-up (14% and 18% respectively).” The OPERA manuscript is now published and is enclosed for your review. In addition, the percentage of patients lost to follow-up in the OPERA trials is comparable to other trials which meet the criteria for “Maintain Comparability”, such as the FREEDOMS trial, which had a dropout rate of 19% and was judged to be of “good” quality (Table C3, Page 118). ● The quality and scientific rigor of study design should be considered in rating the quality of trials. OPERA I and OPERA II were the first double-blind, double-dummy active comparator trials with duration of two-years completed in MS. The trials had separate treating and examining investigators and central MRI readings, all blinded throughout the study. 	The quality assessment has been updated to reflect the published data and is now “Good”
5	<p>Upgrade the ocrelizumab rating from “B+” to “A” in both PPMS and RRMS (compared to best supportive care and compared to interferons and glatiramer acetate) based on ocrelizumab’s benefit-risk profile compared to other DMTs and the availability of the manuscripts.</p> <ul style="list-style-type: none"> ● In Genentech’s attempt to replicate the NMA of ARR using ICER’s methodology and inputs, our result for the rate ratio for ocrelizumab vs. placebo was 0.34 with a 95% credible interval of [0.27, 0.42], which is different than ICER’s result of 0.43 [0.34, 0.54]. The corrected numbers should be reflected in the rating for ocrelizumab. ● Despite ICER’s emphasis on benefit-risk of DMTs, ratings seem biased toward efficacy and it is unclear how potential harms were included. <ul style="list-style-type: none"> ○ Per the draft report, the evidence rating reflects a joint judgment on two components (Page 24): “Net 	Thank you for forwarding the published randomized trials. We have updated the report with those results.

	<p>health benefit” - the balance between clinical benefits and risks and/or adverse events, and level of certainty in the best point estimate of net health benefit.</p> <ul style="list-style-type: none"> o Consider the data from Table 11 titled “Harms of DMTs” when incorporating risks into ratings for DMTs. <ul style="list-style-type: none"> ● Please update the text for RRMS regarding the rating of ocrelizumab on pages 51 and 53. <ul style="list-style-type: none"> o For RRMS, ocrelizumab was given a B+ rating because results have “not yet been published” and “there is no real-world evidence supporting its efficacy”. (Page 51, similar text on Page 53) The OPERA manuscript is now published and is enclosed for your review. o The extent to which real-world evidence (RWE) informs ratings is inconsistent across DMTs and is not cited as a reason for a lower rating for other DMTs without RWE supporting efficacy. ● Please update the text for PPMS regarding the rating of ocrelizumab on page 55. <ul style="list-style-type: none"> o For PPMS, ocrelizumab was given a B+ “due to the preliminary nature of the data”. (Page 55) The ORATORIO manuscript is now published and is enclosed for your review. ● The quality of clinical trials as measured by ICER using US Preventive Services Task Force (USPSTF) criteria should be considered in the ratings. It appears that “good” quality trials were not required for an “A” rating, for example, alemtuzumab received “A” rating with clinical trials that were rated as “poor” and “fair” in Table C3. 	<p>We have updated the text about ocrelizumab, but feel that we must reflect the uncertainty from the lack of an FDA full review of the drug and the lack of real world experience with the drug in our assessment, and have left the rating at B+. As we recently saw with dimethyl fumarate (change in label due to liver toxicity), real world use of novel therapies identifies significant harms that were not seen during the clinical trials.</p> <p>We do not require a good quality study for a particular rating if the other dimensions of the strength of evidence are present.</p>
6	<p>Defer the cost-effectiveness and budget impact modeling of ocrelizumab until ocrelizumab is FDA-approved and the price is available. Including a speculative price of ocrelizumab based on an arbitrary formula into a cost-effectiveness analysis or budget impact model is inappropriate and may be expected to result in misleading conclusions.</p>	<p>We try to time our reports so that the WAC price at least will be available just prior to the public meeting, but until then we use a placeholder price. If there is no price available in time before the meeting we do not take votes on value; we only do threshold pricing at 50, 100, and 150K/QALY.</p>
7	<p>Provide clarity on the data source and methodology for the sales and utilization data used in the calculation of drug prices used in the draft report. Because ICER is using a third-party to obtain drug price, it is unclear how prices were derived and what they represent. We caution that the calculated drug price should not be stated as a definitive fact and that limitations to the analysis should be disclosed.</p>	<p>We have added some additional details on this data source. We have also noted the limitation that net prices used in our analysis may reflect an average discount from WAC, but that discounts vary widely across payers and specific discount information is usually not publicly available.</p>
8	<p>When ocrelizumab price is available, revise the PPMS budget impact model to assume a significant proportion of market uptake is replacement of off-label DMTs.</p> <ul style="list-style-type: none"> ● In the draft report, potential budget impact of ocrelizumab was estimated based on incremental costs compared to best 	<p>Ocrelizumab’s price will not be available in time to include in this report, and we did not include off-label use of DMTs because of lack of specific data on</p>

	<p>supportive care, but this does not reflect the current treatment patterns in PPMS.</p> <ul style="list-style-type: none"> ● Based on data collected from 115 physicians, currently 53% of PPMS patients (N=215) are treated with DMT (see table 1 below) (Adelphi Multiple Sclerosis DSP V (Q1 2016), Data on file, 2016.) ● Additional utilization data from a survey conducted in Spring 2015 among over 7,000 participants in the North American Research Committee on Multiple Sclerosis (NARCOMS) indicated that 33% of responders with PPMS were using an off-label DMT ● Therefore, assuming ocrelizumab will be replacing best supportive care will grossly overestimate its budget impact and therefore is misleading for payers. The revised report should instead assume a significant proportion of market uptake for ocrelizumab in PPMS is replacement of off-label DMTs. 	<p>utilization and effectiveness of off-label use in PPMS patients.</p>
9	<p>Provide full disclosure and share the cost-effectiveness and budget impact models.</p> <ul style="list-style-type: none"> ● We ask that the cost-effectiveness and budget impact models should be provided to stakeholders in a format facilitating feedback to increase confidence and credibility of ICER's evaluations. Genentech was not able to fully replicate the numbers in the cost-effectiveness analysis and budget impact analysis based on the current information in the report. ● The cost-effectiveness (CE) ratio of ocrelizumab vs. supportive care in PPMS from the probabilistic sensitivity analysis is \$702,243 per QALY gained (Table E15, Page 172), whereas the deterministic CE ratio for ocrelizumab vs. supportive care in PPMS is \$854,020 per QALY gained (Table 22, Page 72). The difference between these two results is fairly large and an explanation of this difference should be provided in the revised report. ● Please provide clarity in the final report on the budget impact analysis: <ul style="list-style-type: none"> ○ Does the calculation assume that all patients would initiate the therapy at the beginning of the year or initiate new therapy gradually over a 12 month period (e.g., patients initiating in December would only incur one month treatment during that year)? ○ How were costs estimated for patients when they enter the budget impact model in the first year vs subsequent years, for example, how were the 1st, 2nd, 3rd, 4th and 5th year costs estimated for those who initiated treatment in first year following entry? How were the 2 nd, 3rd 4 th and 5th year costs estimated for those who initiated treatment in the second year following entry, and so on? 	<p>We have attempted to provide sufficient detail on the models and their inputs in the report and technical appendices to allow for replication of results. Our potential budget impact analysis assumes that patients initiate therapy at the beginning of each year. Costs for each subsequent year were derived from the year 2 through 5 results from our cost-effectiveness model.</p>
	<p>Remove the arbitrary budget impact cap. The arbitrary budget cap (\$904 million) for societal expenditures on medical</p>	<p>As we have described repeatedly, our budget impact analysis is intended to</p>

	<p>innovation ignores the value of treatments and benefit to patients, caregivers, and society. Imposing an arbitrary budget cap for all new products can inadvertently stifle innovation particularly in areas of high unmet need, such as MS.</p>	<p>illustrate situations in which the expected impact of a new intervention might yield a signal that additional steps would need to be taken to ensure affordability and wide access to interventions that may in fact have high value at the individual patient level. It is not in any way a cap, but represents the tension that health systems face in providing access to valuable innovations while remaining responsible budgetary stewards.</p>
10	<p>Include the following limitations of the NMA in the results section of the report:</p> <ul style="list-style-type: none"> ● Different assessment time points across trials: Efficacy endpoints across various DMT trials occur at many different time points and may not be comparable across trials. If trial endpoints using different time points are analyzed, the assumptions of proportional hazards should be tested. For example, trials with a duration of 24 weeks will not provide meaningful CDP results due to the short duration of follow up. ● Different definition of outcomes across trials: The definition of key outcomes such as relapse rates and disease progression are heterogeneous across trials, which weakens the validity of indirect comparisons across trials. ● Variable quality of clinical trials: There is no consideration or weighting of the quality of the clinical trials, e.g. double-blinded vs unblinded, placebo-controlled vs open-label, which can introduce inconsistency across the network for comparisons. ● Changes in natural history of disease over time: The natural history of relapse rate has changed throughout the era of MS clinical trials, therefore comparing contemporary trials to those trials conducted a decade or more ago could result in differential background (and placebo) rates of relapse. For example, a much higher ARR was reported for IFNβ-1a 44mcg in the PRISMS 1998 trial (ARR=0.87) and the OWINS 1999 trial (ARR=0.94) than that from other trials (ranging from 0.22 to 0.52). Inclusion of such trials would impact the estimates for both IFNβ-1a 44mcg and other treatments linked to IFNβ-1a 44mcg in the NMA. ● Differences in baseline characteristics across trials: The differences in baseline characteristics between different trials such as differences in age, baseline EDSS scores, number of prior therapies, disease duration, etc. may introduce biases into the results of the NMA. Meta-regression examines the impact of each characteristic separately but the collective impact of all these variables on the results of the NMA is difficult to assess. 	<p>A brief section has been added after the comparison of our NMA results for disability progression to the other published NMAs that emphasizes many of the points that you cite.</p>
11	<p>In Table 1, add qualifier that there is no FDA-approved dose for rituximab in MS and correct the rituximab dose and</p>	<p>Please see response to comment 11 on page 5.</p>

	<p>associated whole-sale acquisition cost (WAC) to be consistent with MS clinical trials.</p> <ul style="list-style-type: none"> ● In Table 1 under the “FDA-approved Dose” column, the stated dose for rituximab is 1000 mg every 6 months. ● In the Phase 1 open-label trial in RRMS, 26 patients received rituximab 1000 mg on days 1 and 15 totaling 2000 mg every 24 weeks. ● In the Phase 2 single-dose proof-of-concept 48-week trial in RRMS, 69 patients received rituximab 1000 mg on days 1 and 15 totaling 2000 mg for one dose. ● In the Phase 3 trial in PPMS, which failed to meet its primary endpoint, 292 patients received rituximab 1000 mg on days 1 and 15 totaling 2,000 mg every 24 weeks. 	
12	<p>In Table 11, correct safety information for rituximab to include safety data from clinical trials of rituximab in MS.</p> <ul style="list-style-type: none"> ● Table 11 includes rituximab’s major safety concerns from non-MS indications listed in the prescribing information for rituximab. ● Safety profiles will differ significantly based on disease state and background therapy and thus the draft report should list published safety information for rituximab from MS clinical trials. 	<p>Given the lack of an FDA indication for MS, we felt it appropriate to include the Boxed Warning for rituximab and the other safety concerns. The discontinuation rates due to AEs and the SAE rates come from the MS Clinical Trial (HERMES).</p>
13	<p>Conduct sensitivity analysis of budget impact of ocrelizumab by varying prevalence of PPMS.</p> <ul style="list-style-type: none"> ● The prevalence of PPMS is reported in literature ranging from 10-15%. In the budget impact analysis, ICER used the higher end of the range (15%) and therefore a sensitivity analysis should be performed using lower end of the range to assess the impact on the overall budget impact. 	<p>The budget impact of ocrelizumab did not reach our budget impact threshold when using the 15% estimate, so we did not explore the impact of using lower estimates.</p>
14	<p>In Table B2 titled “Ongoing Trials of Infused DMTs for MS”, correct the comparator doses in the ocrelizumab trials and include additional ongoing trials</p> <ul style="list-style-type: none"> ● In trials NCT01412333, NCT01247324 and NCT01194570, the doses of comparator arms are listed incorrectly. <ul style="list-style-type: none"> ○ For the OPERA I and II trials (NCT01412333, NCT01247324), the comparators are ocrelizumab 600 mg and IFN beta-1a (Rebif) 44 mcg. ○ For the ORATORIO trial (NCT01194570), the comparators are ocrelizumab 600 mg and placebo. ● Add the following ongoing Genentech-sponsored trials for ocrelizumab: <ul style="list-style-type: none"> ○ CHORDS ○ CASTING ○ VELOCE ○ OBOE Biomarker Study ○ Open-label extension of the Phase 2 study in patients with RRMS. 	<p>This information has been corrected in the revised report.</p> <p>We have added the CHORDS and CASTING trials to the list of ongoing trials. VELOCE and OBOE have not been added because their outcomes are not included in this report. The open-label Phase II extension trial has not been added because its primary completion date on the ClinicalTrials website is in the past.</p>

15	<p>Include health-related quality of life data for ocrelizumab in relapsing forms of MS and PPMS.</p> <ul style="list-style-type: none"> • In OPERA II, the difference in adjusted mean change in Short Form-36 (SF-36) physical component score was 0.69 in OPERA I (p=0.22) and 1.16 in OPERA II (p=0.04). • In ORATORIO, while ocrelizumab did not demonstrate a significant change in SF-36 physical component score compared with placebo (p=0.60), post-hoc exploratory analyses, showed improvement on the SF-36 mental component score (p=0.0006) as well as reductions in fatigue as measured by the modified fatigue impact score (p=0.009). 	The results from the published studies have been added to the report.
16	<p>Include the OPERA I and OPERA II trials for the subgroup analysis that excludes trials with duration of <18 months for both ARR and CDP in Tables D1 and D5 respectively.</p> <ul style="list-style-type: none"> • For the subgroup analysis excluding trials with duration <18 months, the result is listed as “N/A” for ocrelizumab. • The OPERA I and II trials had a duration of 22 months (96 weeks) and thus would qualify for this subgroup analysis. 	Corrected.
17	<p>Correct additional errors listed in Appendix 1, organized by section of the report and page number. (see Genentech Appendix 1, included at end of this comment grid)**</p>	Thank you for the additional recommendations.
Novartis		
1	<p>The MS patient experience should be the focus and driving force of this report: The efforts with the patient survey are a step in the right direction; however, these findings were not incorporated into main components of the report. ICER should incorporate the full burden of MS from the patient perspective directly into the “Evidence Rating Matrix” and cost-effectiveness model methodology.</p>	The patient experience is central to the ICER value framework. The Evidence Rating Matrix and chosen outcomes focus on patient-important outcomes.
2	<p><i>The patient perspective should directly inform the main evaluations of the DMTs.</i> The patient survey results did not directly inform the decision-making around the draft evidence report’s main sections on comparative clinical effectiveness, cost-effectiveness, or budget impact. We recommend that ICER include all patient-centric value components in the relevant models; failing to do so may underestimate DMTs’ treatment benefit. More broadly, information on patient-reported outcomes (PRO)—available from well-controlled Phase IV data among other sources—should be considered in the report and model. Specifically, in the PREFERMS study, fingolimod patients had higher retention rates (81.3%) compared to injectable DMTs (29.2%) as well as improved treatment satisfaction. In the EPOC study, patients who were on injectable DMTs switching to fingolimod versus staying on an injectable DMT had larger improvements in patient satisfaction as measured by the Treatment Satisfaction Questionnaire for Medication (TSQM) as well as PROs on fatigue and depression.</p>	<p>We have expanded the section on quality of life / patient reported outcomes as well as the discussion of other benefits and harms to more fully represent the patient perspective.</p> <p>We have highlighted that many patients prefer oral therapies like fingolimod.</p> <p>Thank you for pointing us to the PREFER MS study.</p>
3	<p><i>ICER’s care value model should apply a societal perspective to measure DMTs full value.</i> ICER’s scoping document included a</p>	We respectfully disagree. As the 2 nd US Panel on Cost-Effectiveness has

	<p>broad list of clinical outcomes to be considered, but the draft evidence report included few outcomes in its cost-effectiveness (“care value”) model. Although incorporating work productivity and other indirect costs in scenario analyses is a step in the right direction, failing to measure these components in the baseline model biases the baseline estimates towards lower cost-effectiveness. Furthermore, ICER should include caregiver time cost and disutility into the cost-effectiveness model, as has previously been done by National Institute for Health and Care Excellence (NICE). The burden of RMS on patients and caregivers is large; in one study, 47.1% of caregivers provided more than 20 hours of care per week. Further, ICER’s approach fails to account for other key components of value including, but not limited to: (1) patient satisfaction, (2) functional status, and (3) societal value of innovation (e.g. “insurance value” ascribed to treatment from the perspective of individuals without MS).</p>	<p>appropriately noted, the original recommendation to use a societal perspective has not been feasible in an overwhelming majority of modeling situations due to a lack of data with which to compare individual drugs’ impact on indirect costs. In addition, the fragmented US health system involves a multiplicity of decision-makers, none of whom serves as the ultimate societal arbiter.</p>
4	<p>Incorporation of fingolimod’s robust long-term data: The robustness of the evidence on fingolimod, which meets ICER’s criteria on “certainty” and “magnitude,” warrants an improved “Evidence Rating Matrix” grade of an “A” versus best supportive care and a “B” or higher versus interferons or glatiramer acetate. Fingolimod has:</p> <ul style="list-style-type: none"> a) Higher quality and more “certain” data as the only oral DMT with a positive head-to-head data against an active comparator (TRANSFORMS trial) and one of only three trials receiving the highest quality evaluation by ICER (FREEDOMS trial) b) Long-term and consistent data on efficacy over 4.5 years and safety over 7 years c) Extensive real-world and Phase IV evidence on comparative effectiveness, adherence, and patient satisfaction. 	<p>Thank you for the insights into the body of randomized and observational data.</p>
5	<p><i>Fingolimod has higher quality data, as the only oral DMT with a positive head-to-head trial (TRANSFORMS) against an active comparator. In addition, the pivotal FREEDOMS trial was one of only three trials receiving the highest quality evaluation by ICER.</i> Fingolimod has proven to provide early and sustained high efficacy across a variety of disease activity measures versus standard-of-care (i.e. only oral DMT with positive results in a head-to-head trial) and placebo. Fingolimod has consistently demonstrated superior efficacy in terms of relapses vs. standard-of-care and placebo (TRANSFORMS, FREEDOMS, FREEDOMS II) and disability compared with placebo (FREEDOMS). The TRANSFORMS pivotal and extension trial demonstrated fingolimod’s superior efficacy versus intramuscular interferon beta-1a and consistent safety profile. For patients taking 0.5 mg of fingolimod, annualized relapse rates (ARR) were 0.16 (0.12-0.21), compared to 0.33 (0.26-0.42) for interferon beta-1a. The consistency of fingolimod’s clinical benefit compared to intramuscular interferon beta-1a was also</p>	

	<p>seen across several subgroups, including MS patients with highly active disease. Based on ICER’s criteria, the “certainty” of fingolimod’s evidence should be “higher” than dimethyl fumarate and natalizumab, which do not have head-to-head trials showing superiority compared to interferons or glatiramer acetate. Furthermore, despite the fingolimod trial, FREEDOMS, receiving one of only three highest quality clinical trial assessment ratings (“good”) out of 39 RRMS trials, fingolimod received a similar “Evidence Rating Matrix” grade as dimethyl fumarate, an oral DMT with only “poor” quality trials. Thus, based on the available evidence, fingolimod should receive a rating of “B” or higher grade versus interferons/glatiramer acetate and a rating of “A” versus best supportive care.</p> <p>Indirect comparisons demonstrate that fingolimod patients had a higher likelihood of achieving the “No Evidence of Disease Activity” composite endpoints (NEDA-3, NEDA-4), compared to other oral DMTs. Also, fingolimod has demonstrated a significant reduction in brain volume loss (BVL) by 32% to 35% in pivotal trials compared to intramuscular interferon beta-1a and placebo, respectively. Although ICER acknowledges the importance of these outcomes to patients, they are not reflected in the net health benefit ratings.</p> <p>Increased transparency on the value ICER’s methodology places on clinical benefits versus risks in their Evidence Rating Matrix” is needed. Despite alemtuzumab and natalizumab’s black box warnings, both treatments receive an “A” rating compared to best supportive care. Given these significant risks, their importance to patients, and their respective U.S. labels’ third line indication, alemtuzumab and natalizumab’s net health benefit grade should be lowered.</p>	
6	<p><i>Fingolimod has demonstrated a sustained long-term efficacy and safety profile.</i></p> <p>The FREEDOMS extension study (4 years) and TRANSFORMS extension study (4.5 years) confirmed the effect of fingolimod therapy in maintaining a low rate of disease activity and sustained improved efficacy, with no new safety concerns. In the TRANSFORMS extension study, patients in the continuous-fingolimod cohort demonstrated significantly lower (35%) ARR compared with those in the IFN β-1a switch cohort (0.17 vs. 0.27, respectively) after 4.5 years, suggesting improved outcomes for earlier, continuous use of fingolimod. In the FREEDOMS extension study, patients who were continuously on fingolimod had a 48% reduction in ARR, 27% to 31% reduced risk of disability progression measured, and significantly lower BVL versus placebo patients who switched to fingolimod. The LONGTERMS study, the pooled extension study of fingolimod’s Phase II/III programs over 7 years, found that fingolimod</p>	<p>Thank you. We believe that the LONGTERMS results are only available in abstract form.</p>

	<p>patients in extension phases had fewer or similar serious adverse events compared to patients taking fingolimod for 1 or 2 years in the pivotal trials [incidence rate ratio = 0.73 (0.60-0.91)] ; efficacy measured by ARR and EDSS pooled of the TRANSFORMS and FREEDOMS cohorts were consistently maintained at reduced levels over 7 years.</p>	
7	<p><i>The “Harms of DMTs” table should accurately reflect the safety profile of fingolimod.</i> In the draft evidence report, ICER states that fingolimod requires a REM (Table 11 and the “Harms” section). However, the FDA recently determined that the communication plan in the Risk Evaluation and Mitigation Strategy (REMS) for Gilenya (fingolimod) is no longer needed to ensure that the benefits of the drug outweigh the risks.</p>	<p>We have updated Table 11.</p>
8	<p><i>ICER should consider fingolimod’s Phase IV comparative satisfaction data and real world effectiveness in evaluations of comparative clinical effectiveness.</i> The word “effectiveness” in “comparative clinical effectiveness” implies that data beyond clinical trials were incorporated in the report. However, the net health benefit ratings do not currently consider real-world data or Phase IV studies, which are important to all stakeholders. As mentioned previously, fingolimod’s retention, patient satisfaction, and other PROs were measured in the Phase IV studies, PREFERMS and EPOC. Evidence from the international MSBase registry indicates that patients switching to fingolimod were associated with a significant reduction (51%) in the rate of first relapse and significant slowing of disability progression, similar to natalizumab, compared to those switching to an interferon beta, glatiramer acetate, teriflunomide or dimethyl fumarate. In the prospective, non-interventional study PANGAEA, patients continuing fingolimod from trials had sustained reduction of ARR and stable EDSS over 4 years with 62.5% to 74.5% of patients with no clinical disease activity in any given year. Additional comparative effectiveness studies have shown similar results on relapse prevention. Several real-world studies based on data from national claims databases, MS centers, and international registries demonstrated that patients taking fingolimod have high adherence and persistence rates, which are consistent with findings in the Phase IV PREFERMS study. Specifically, an MSBase Registry study and MS Center study have demonstrated that patients using fingolimod were significantly less likely to discontinue treatment in addition to having a trend towards better clinical outcomes (i.e. relapses) relative to dimethyl fumarate (as well as teriflunomide and injectable DMTs). Finally, a cross-sectional PRO study found that patients taking fingolimod reported higher scores compared to dimethyl fumarate on satisfaction and tolerability.</p>	<p>While not explicitly stated, we appreciate the value of real-world data from the therapies that have FDA approval and substantial clinical experience. That will be clearer in the final report in our assessment of ocrelizumab, which now has published clinical trial results, but no FDA approval and real world data.</p> <p>That said, we felt that the selection bias inherent in continuation studies as well as the variability in the Phase IV studies of the DMTs in term of length of follow-up and loss to follow-up compared with the participants initially randomized limited the conclusions that could be drawn.</p>

9	<p>Consistency with approved indications: Alemtuzumab and daclizumab’s FDA approved indications are after “inadequate response to two or more drugs”; however, ICER incorrectly interprets line of therapy in the draft evidence report’s comparative clinical effectiveness summary and cost-effectiveness model. Misrepresentation of DMT risk profiles and approved indications may jeopardize appropriate treatment benefit/risk assessments for MS patients.</p>	<p>We have added the language from the FDA indication to the Harms section of the document.</p>
10	<p><i>ICER should apply care value modeling based on a treatment’s indicated patient population.</i></p> <p>ICER incorrectly states that there is a “lack of conclusive FDA labels” to be able to compare first-line agents to later lines. Each DMT’s FDA indication, however, is clear. Fingolimod is approved as a first-line RMS agent, while alemtuzumab and daclizumab are indicated for use after “inadequate response to two or more drugs” “because of [their] safety profile.” ICER’s methodology should reflect these indications. In the cost-effectiveness model, ICER states that “after discontinuation from second-line therapy, patients transitioned to best supportive care,” implying that alemtuzumab and daclizumab are modeled for non-indicated, second line use. Further, FDA-approved lines of therapy for all DMTs should be included in Table 1. Failing to do so may jeopardize appropriate treatment benefit/risk assessments for MS patients.</p>	<p>See p9, comment 38. Although some labels suggest use in later lines of therapy, no label explicitly precludes use as a first line agent. Because there is no formal recommendation or standard practice for order or treatment regimens, we modelled each DMT equivalently. The choice of using and aggregate second-line regimen followed by supportive care may decrease differences in results between DMTs, but would not substantially bias results in favor of any particular DMT.</p>
11	<p><i>ICER should remove voting question #4 comparing daclizumab to dimethyl fumarate or fingolimod and should revise question #1 comparing dimethyl fumarate to fingolimod.</i></p> <p>The direct comparison of two treatments with different indications is problematic. Since daclizumab is only indicated for third-line therapy, any comparisons should be made for similar, third-line therapies as per labeling. In addition, question #1 is phrased in a biased manner as it does not allow for the voters preferring fingolimod to voice their view. ICER should instead ask which DMT has the higher net health benefit, as this would allow voters to select among (i) fingolimod, (ii) dimethyl fumarate, or (iii) that the evidence is insufficient.</p>	<p>Daclizumab is not “only indicated” for third-line therapy, although there is a general recommendation to use it in this way.</p>
12	<p><i>ICER should exclude rituximab as it is not indicated for the treatment of RMS.</i></p> <p>A Cochrane Review’s NMA concluded that: “There is not sufficient evidence to support the use of rituximab as a [DMT] for RRMS.” The only comparative trial of rituximab included in the NMA was cited as high risk of bias due to high attrition. This study was not sufficiently powered to detect changes in important endpoints such as relapses, BVL, and safety. Off-label use of rituximab may have public safety consequences, particularly in light of boxed warnings and reported data on serious adverse events for on-label indications.</p>	<p>Please see response to comment 4 on page 4.</p>
13	<p>Methodological concerns and need for additional transparency:</p>	<p>LOE is discussed below. ICER reports and publications on Open Science provide</p>

	<p>ICER’s report should incorporate the timing of loss of exclusivity (LOE) and availability of generic DMTs, which are important considerations especially for patients in regards to lower out-of-pocket costs. Further, to enhance the validity of the network meta-analysis (NMA) and cost-effectiveness models, additional information on assumptions and direct access to ICER’s models are needed.</p>	<p>extensive information and transparency to allow all stakeholders to evaluate the assumptions and models.</p>
14	<p><i>The cost-effectiveness model should incorporate price decreases after loss of exclusivity.</i></p> <p>When modeling cost-effectiveness, ICER should account for future drug price adjustments to reflect the effect of loss of exclusivity (LOE). Historically, among recent generic entrants, generic oral medications had 74% lower prices than the pre-expiry brand prices within 8 months of becoming available and 93% adoption of the generic (vs. brand) within a year; these decreases are much larger than those for generic biologic (i.e. large molecules) medications, which have greater manufacturing complexities and fewer competitors. Fingolimod is an oral DMT for which the compound patent expires in the near future, which may lead to LOE during the time horizon used in ICER’s cost-effectiveness model. Incorporating the availability of lower cost generic versions within the model’s time horizon better reflects the true cost-effectiveness calculations as a drug nears its LOE and is consistent with good practice recommendations made by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR).</p>	<p>While ICER continues to examine this important question, for now we do not model price decreases after LOE. The historical trajectory of pricing in relationship to LOE is complicated, with many examples of prices, even net prices, increasing before LOE so that LOE does not result in a realized price decrease when viewed over a several-year timeframe. That, plus the unknown but often equivalent trajectory (up or down) of competitors has made the widely used standard in academic and HTA methods to not model price changes at estimated LOE.</p>
15	<p><i>ICER’s baseline analysis should separately analyze trials based on whether disability progression was measured over a 12-week or 24-week time period using high quality evidence.</i></p> <p>The evidence from trials that confirmed disability progression at 12-weeks or 24-weeks is not comparable. A 24-week disability duration leads to fewer acute incidences of improvement or worsening because patients in 24-week trials need to show change in disability for a longer duration of time compared to 12-week trials. This approach reduces the occurrence of smaller, temporary changes being reported (i.e., fewer ‘false-positives’). Thus, using 12-week and 24- week results in a single NMA results in an inappropriate comparison. Further, the disability endpoint used leads to large differences in care value estimates. In the baseline approach, fingolimod’s incremental cost-effectiveness was \$576,325 per QALY compared to glatiramer acetate, whereas in the sensitivity analysis incorporating only trials with 12-week disability progression, the cost per QALY was \$119,764 (i.e. below the ICER threshold of \$150,000 per QALY). This large difference in care value suggests that trials using 12 and 24-week disability progression measures may not be comparable.</p>	<p>We agree that the 24-week outcome is preferred, but the 12-week outcome is a reasonable surrogate and in the trials that report both outcomes, the 24-week results are consistently better than the 12-week results, so the decision to use the 12-week results when 24 week results are not available is a conservative one. The results using only 24-week outcomes are reported (Appendix Table D4).</p>
16	<p><i>NMA should measure effectiveness for the target population and test all model assumptions.</i></p>	<p>The differences in follow-up are not large: almost all trials are either 1 or 2</p>

	There has not been an evaluation whether differences in follow-up affect the ARR estimates. With the approach taken by ICER it is implicitly assumed that relapse follows an exponential distribution with a constant underlying rate. The question is whether this is true and if not, whether the within-trial rate ratios are different at different follow-up times. If this is the case, then differences in follow-up between trials might bias NMA results.	years in length. We exclude trials of less than 48 weeks duration and performed a sensitivity analysis excluding trials with less than 18 months follow-up.
17	<i>Despite adjusting for between-trial differences, it is not clear whether treatment effect estimates are relevant for the target populations of interest.</i> A meta-regression of all trials adjusting for proportion of experienced patients to obtain results for treatment-naïve as well as treatment-experienced patients would be a relevant sensitivity analysis, especially with NMA results used in the cost effectiveness model where all patients are treatment-naïve to begin with and switch to a second DMT after failure.	We are not aware of any data supporting effect modification by treatment experience. If time allows, we will conduct this additional meta-regression.
18	<i>ICER's approach for measuring drug prices and discounts should be transparent.</i> Discount and rebate information obtained from SSR Health are not shared in the current report. Details of the methodology used in these calculations should be disclosed.	The WAC and discount rates applied are provided in the report, in the table on DMT acquisition costs.
19	<i>ICER should provide clarity on whether treatment efficacy in terms of delaying disability progression has been included in the cost-effectiveness model for EDSS transitions in SPMS.</i> It is currently unclear as to whether the natural history transitions between EDSS states in SPMS are adjusted for treatment efficacy. More clarity is required to describe the modeling approach.	Clarified in text.
Sanofi-Genzyme		
1	Concerns Related to the Network Meta-Analysis (NMA) <ul style="list-style-type: none"> The network meta-analysis (NMA) as applied to MS in this ICER report is based on some questionable assumptions. Among the concerns are inclusion of inappropriate studies and exclusion of appropriate studies. We are particularly concerned by the impact that heterogeneous patient populations, changes in diagnostic criteria over time and other significant differences between the trials conducted on the multiple MS therapies approved over the last three decades has on the analysis, which ICER has not accounted for. Additional concerns are related to pooling results from a variety of studies irrespective of differences in statistical power, prespecified endpoints or length of follow up. NMA results have substantial uncertainty and limitations which should be highlighted in the report and its conclusions. The degree of consistency with other NMAs is overstated. Non-approved drugs in MS should not be included in this review as their benefit risk profile in MS has not been 	We agree. These concerns apply to all NMAs and should be borne in mind when interpreting the results. It would be helpful if you could provide data supporting effect modification by any of the sources of heterogeneity across the trials. For instance, we recognize and reported that patients in the earliest trials had higher relapse rates than the more recent trials. However, it is not clear that the relative rate of relapses or the relative hazard for disability progression differs by the underlying risk for relapses.

	<p>established and no label can therefore guide their evaluation.</p> <p>All NMAs run the risk of oversimplification because they combine information from heterogeneous studies that involve different comparators. It is important to articulate the limitations of these syntheses and recognize the potentially high degree of uncertainty in conclusions. ICER's NMA is no exception.</p> <p>An NMA tends to be valid only if the included studies are very similar to one another. If there are large variations between the included trials, broad generalizations of the effectiveness of interventions should not be made, as was done in this ICER report. Even in a conventional metaanalysis, errors or bias in an included study will result in biased conclusions. In an NMA, the problems with an errant or biased study are compounded because that study's biases will generally affect many comparisons.</p> <p>Given the above consideration for general NMAs, the application of NMA in the MS disease area needs a very careful interpretation. Clinical studies in RRMS have shown very heterogeneous outcomes, including the placebo group rates. As both the ICER clinical and cost effectiveness results are based on the NMA, limitations of these findings need to be prominently discussed and acknowledged.</p>	
2	<p><i>Study Selection for Inclusion in NMA</i></p> <ul style="list-style-type: none"> o The inclusion or exclusion of individual studies has a significant impact on the outcomes of the analysis, and therefore trial selection must be based on critical evaluation. In some cases in the MS analysis, ICER has made incorrect decisions around study selection. It is essential to include the appropriate studies to obtain the most meaningful results. o An example of a study that should not have been included is the Bornstein, et al. study from 1987 (6). This study was a pilot trial conducted three decades ago and is the oldest study included in the NMA [report p.26; Tables C1-C3 Appendix]. Furthermore, the very small treatment arm (n=25) displayed different characteristics from the placebo group that exaggerate the observed magnitude of superiority of Cop 1 (glatiramer acetate [GA]). The result from this study is a significant outlier with a relative reduction in ARR almost 5 times greater than placebo -- a result never replicated. Finally, we would note the authors cautioned against drawing conclusions as it was a pilot study. o Another study that should not be included is TENERE. Unlike other studies included in the analysis, the 	<p>We have performed extensive sensitivity analyses to examine the impact including different sets of studies in the NMAs (see Appendix Tables D1, D2, D5, and D6).</p> <p>The Bornstein study is not included in the Base Case NMA.</p> <p>It would be arbitrary to exclude the TENERE trial on the basis of its primary endpoint. Some of the trials had MRI primary endpoints, some ARR endpoints, some CDP. We included studies with at least 48 weeks follow-up that reported either relapse rates or disability progression.</p> <p>The GATE and GLACIER studies were excluded because the follow-up was less than 48 weeks.</p>

	<p>TENERE study employed a unique composite primary endpoint of time to treatment failure.</p> <ul style="list-style-type: none"> o An example of a study that was not included but should have been is the GATE study (7). GATE is a relevant Phase III study that studied three treatments: Copaxone, placebo and generic glatiramer acetate. This study meets all study selection criteria applied by ICER. It was published in the timeframe of other studies included (Dec. 2015 issue of JAMA Neurology) 	
3	<p><i>Concerns about Including Results from Non-comparable Studies</i></p> <ul style="list-style-type: none"> o Another general concern with the conduct of the ICER NMA was the decision to include results and endpoints/outcomes from studies that are not similar. NMAs lose their 3 validity if results are included from studies that are dissimilar in terms of patient populations or other study characteristics. o This is clearly evident in calculations of disability progression in the different studies, as ICER combines results from trials which included different pre-specified disability progression endpoints (e.g. disability progression at 12 or 24 weeks). For individual therapies, studies have shown that when both 12 and 24 weeks outcomes are measured, the outcomes are typically different. Combining such results in one analysis, leads to misleading and invalid conclusions. For example, teriflunomide (both 7mg and 14mg) is no longer 'dominated' when compared with generic glatiramer acetate (GA) on either the cost per additional QALY or cost per additional life year criteria when only studies that consistently measure disability progression at 12 weeks are included. 	<p>The NMA loses its validity if there is effect modification by the differences in the included populations. We found no evidence for that in our analyses.</p> <p>As noted above, the 24-week outcome is preferred for CDP, but we lose almost 50% of the DMTs if we limit the analysis to the studies reporting the 24-week outcome. The 12-week CDP is correlated with the 24-week outcome, but is a conservative estimate. Thus we feel it is an appropriate approach to the analysis.</p>
4	<p><i>Insufficient Discussion of Sensitivity and Uncertainty of Results</i></p> <p>Given the known limitations of NMAs, ICER should include in the body of the report more thorough discussion of the sensitivity analyses currently only included in the appendix. For example, both dosage forms of teriflunomide show a credible range of cost effectiveness that is not dominated by GA 20mg (Glatopa). This finding should be in the body of the report but currently does not appear until appendix table E16 on page 175.</p>	<p>We agree that probabilistic results and sensitivity analyses are of key importance. To address this, we have added a table of probability of cost-effectiveness at the \$150,000 willingness-to-pay threshold to the body of the report.</p>
5	<p><i>Misleading conclusions from ICER NMA analysis</i></p> <p>Indirect comparisons can lead to wrong conclusions about the efficacy of a given product. For example, in direct studies versus placebo, GA 20 mg showed no statistically significant difference and little numerical difference (Johnson 1995 and CONFIRM) on slowing disability progression (8, 9). Despite the lack of positive findings on slowing disability progression as described</p>	<p>One of the primary goals of MAs and NMAs is to combine data from underpowered studies to gain precision in the estimates of the effect size. It is not surprising that the combining data from multiple trials with gives a statistically significant finding, when the smaller individual studies (GA 20)</p>

	<p>above, in this NMA GA 20 mg is reported to be significantly superior to placebo on this endpoint. Therefore, GA 20mg should not be included in the relative risk reduction analysis for EDSS progression.</p>	<p>studies were underpowered to demonstrate statistical significance.</p>
6	<p><i>Consistency with Other NMAs</i></p> <p>ICER’s results for annualized relapse rate (ARR) and disability progression are not consistent with previous MS NMAs. There are clear variations in NMA results reported by Cochrane, CADTH, Tolley, and Fogarty in Tables 7 and 10 compared to ICER, and not all numbers in the table represent the same basis of measurement. The fact that each of these NMAs has different results highlights the limitations of MS NMAs.</p>	<p>We have added additional text about the differences between the NMAs just before Table 10 and an additional paragraph on the uncertainties underlying our NMA immediately after Table 10.</p>
7	<p><i>Comments on Non-Approved Products</i></p> <ul style="list-style-type: none"> ○ Rituximab should not be included in this report. There are no well-controlled Phase III studies that establish the safety and efficacy of rituximab in MS. As a result, there are no rigorous data to establish a benefit risk profile. Additionally, there is no label for its use in MS to guide data inclusion in this evidence evaluation. ○ For ocrelizumab, it must be noted the current results depend on assumptions ICER made about the clinical profile of the product, monitoring requirements, price etc., limiting the accuracy of the inputs. Therefore, some of the current results are not valid. If ocrelizumab is approved, analyses will need to be updated based on labeling and price. 	<p>Please see response to comment 4 on page 4.</p> <p>Noted.</p>
8	<p>Problems with Number Needed to Treat (NNT) Analysis</p> <ul style="list-style-type: none"> • The methodology used to compute number needed to treat (NNT) is flawed and therefore results in misleading conclusions <p>The methodology used to calculate the NNT to prevent one relapse or to prevent disability progression in the ICER analysis is flawed. ICER’s calculation of NNT to prevent one relapse is based on a “background (i.e., placebo) relapse rate of 0.5622 relapses per year” and NNT to prevent one patient from experiencing disability progression is based on “a background (i.e., placebo) risk for disability progression of 0.176”. For appropriate analysis, absolute risk reduction must be calculated based on drug and placebo (or comparator) results within a given study and its inverse (1/absolute risk reduction) must be used to calculate NNT for relapse rate. This is required because MS studies do not have a common set of baseline parameters (e.g. region, age, race, year of study conduct) or a common set of disease and prior treatment characteristics (e.g. disease duration, prebaseline relapse activity, baseline EDSS, MRI activity, prior DMTs use status), and include unique study populations. Thus the results of the placebo arm vary greatly between studies (e.g. range of placebo annualized relapse rates</p>	<p>We have removed the NNT tables, though we still think that they offer value. We considered adding additional columns with different underlying incidence data to illustrate the differences in NNTs in a high-risk population compared with a low risk population.</p> <p>While we agree that NNTs are most commonly calculated for an individual randomized trial, they are most valuable when used to compare across trials and that requires using a standard underlying incidence and length of follow-up. The assumption that there is no effect modification by degree or risk is generally true for most drug therapies.</p>

	<p>in the ICER report range from 0.33 to 1.35 [source: Table C4 of report]. The same rationale holds for the NNTs derived for the disability progression endpoint.</p> <p>When NNTs are calculated for each individual study correctly, they can then be used to compare across trials when head to head studies are not available.</p>	
9	<p>Rating of Quality of Studies</p> <ul style="list-style-type: none"> • ICER should change the criteria used to assess the quality of studies included in the report. <p>ICER should reassess how they rated the quality of the clinical studies in the report. Specifically, classifying studies with 20% (or more) loss to follow-up as poor quality (Page 27 and Table C3) is not a sound approach. By using this criterion, several well-designed, well-conducted, high quality, long term double blinded Phase III studies have been rated as poor quality, while placebo-controlled, shorter duration Phase II studies with non-clinical primary endpoints are rated as fair-to-good quality. There is higher probability of patient drop out in longer duration studies (typically Phase III studies) compared to shorter duration, placebo-controlled Phase II studies. In addition, some Phase III studies have protocol-driven mandatory study discontinuation rules that lead to patient withdrawal, independent of patient or physician decision to do so. This results in higher frequency of drop outs. Thus the use of 20% loss to follow-up criteria results in incorrect rating of study quality. Furthermore, the Code of Federal Regulations states that an adequate and well-controlled study consists of: 1) A clear statement of the objectives of the investigation and a summary of the proposed or actual methods of analysis in the protocol for the study and in the report of its results; 2) A design that permits a valid comparison with a control to provide a quantitative assessment of drug effect; 3) The study drug being compared with an inactive preparation designed to resemble the test drug as far as possible; and 4) An analysis of the results of the study which is adequate to assess the 5 effects of the drug (21 C.F.R. § 314.126 (2016)). The fulfillment of the above criteria has led to approval of DMTs by the FDA yet ICER’s classification approach categorizes several of these studies as poor quality.</p>	<p>We recognize the issues with an arbitrary cut off, but we are using the AHRQ/USPSTF rating instrument which considers >20% lost to follow-up a fatal flaw because it introduces a significant risk for selection bias. All quality rating approaches have been criticized and have limitations. For example, see Hartling et al, Validity and Inter-Rater Reliability Testing of Quality Assessment Instruments, AHRQ, 2012.</p>
10	<p>Figure 5 is Not Data-Driven</p> <p>Figure 5, a depiction of safety and effectiveness of different products, should be removed from the report. The size and location of the elements in the graph are subjective and not based on quantitative analysis and thus this figure has no place in a data-driven report.</p>	<p>As you have noted, there is considerable uncertainty about all of the quantitative estimates made in the report. Figure 5 is roughly based on the quantitative estimates, but encourages the reader to make their own overall assessment of the safety, effectiveness, and</p>

		uncertainty of the estimates for all of the DMTs.
11	<p>Limitations Due to the Lack of Inclusion of Other Measures of Value</p> <p>These results would be more meaningful for patients and health care providers in making treatment decisions if they included attributes of a drug that have not been captured such as route of administration and aspects of tolerability and dosing. For example, ICER does not adequately address the serious challenges patients may experience with alternative routes of administration and accompanying side effects, and the resulting impacts on tolerability and adherence. Additionally, indirect costs (e.g. work absences, caregiver time, lost income, early retirement etc.) need to be considered as they often contribute significantly to total costs. For example, in the ICER analysis, indirect costs are more than double the size of direct costs in MS.</p>	Please see Section 5. Other Benefits or Disadvantages.
Teva		
1	<p>Relative Risk of Disability Progression Estimate for Copaxone® 40 mg TIW Used in the Analysis</p> <p>As ICER has noted on page 40 of the draft evidence report, “It is unlikely that glatiramer acetate 40 mg increases disability progression.” We fully concur with ICER on this item. However, we also note with concern that the ICER analyses nevertheless apply the RR estimate of 1.18 for Copaxone® 40 mg TIW (implying that the treatment increases disability progression) as the base case, which would likely lead to invalid effectiveness and value conclusions relating to Copaxone® 40 mg TIW. We suggest to ICER to use an assumption in the cost-effectiveness model that the clinical effectiveness in slowing disability progression for Copaxone® 40 mg TIW is equivalent to that estimated for Copaxone® 20 mg QD (RR=0.70; 95% CI 0.54-0.93) and provide our rationale below.</p>	We have decided to remove Copaxone 40 mg from the cost-effectiveness analysis in the report.
2	<p>The application of the statistically non-significant disability RR of 1.18 (95% CI 0.67-1.97; p=0.57) for Copaxone® 40 mg TIW in the ICER model is based upon data from the one-year randomized placebo-controlled GALA study (Khan 2013) which does not reflect the underlying true clinical benefit of Copaxone® 40 mg TIW in slowing disability progression for patients with RRMS and, moreover, lacks face validity for the following reasons:</p> <ol style="list-style-type: none"> 1. One year of observation is insufficient (too brief) to observe or infer a robust assessment of disability progression, thus this outcome must be utilized and interpreted with extreme caution. A minimum requirement to obtain two consecutive measurements of EDSS a minimum of 6 months apart to establish Confirmed Disability Progression (CDP) is suggested by the European Medicines Agency (EMA 2015). The 	We recognize these results and have addressed them in the report. We do not feel comfortable making an assumption of equal efficacy with the 20 mg dose and look forward to an adequately powered randomized trial designed to have at least 2 years of follow-up.

necessity for confirmation of the EDSS change of 1 or more points, at consecutive measurement intervals 12 or 24 weeks apart, to be considered as CDP allows only a few events of disability progression to occur and be confirmed during a one-year study, as has been observed in GALA (Table 1, Khan 2013).

Table 1. GALA (MS-GA-301) - Placebo Controlled (PC) Phase – Post Hoc Analysis. EDSS Data Distribution of the Number of Subjects with 3 Month Confirmed EDSS Progression during the PC Phase

GALA (MS-GA-301)	Placebo (n=461)		Glatiramer Acetate 40mg TIW (n=943)	
	N	%	N	%
All	461	100.0	943	100.0
No progression	444	96.3	901	95.5
Confirmed EDSS progression	17	3.7	42	4.5

International guidance suggests that longer-term assessment is required to effectively evaluate treatment effects with respect to disability owing to the slow natural progression in RRMS (EMA 2015). For example, the European Medicines Agency states that “...For a distinct claim on disability large-scale long-term parallel group trials will be required to establish clinically relevant treatment effects on disease progression. Study duration will depend on the population studied, and should be sufficient to show a reliable and relevant effect on disability. Such a study may need to last ~3 years” (EMA 2015). Based on this guidance, it can be reasonably inferred that a one-year assessment of disability progression is scientifically inadequate and lacks robustness.

It is suggested that, for a valid claim on disability, large-scale long-term parallel group trials are required for a new DMT to establish clinically relevant treatment effects on CDP (EMA 2015).

2. The development program for Copaxone® 40 mg TIW included a single one-year phase 3 placebo controlled study to assess the safety and efficacy of the new dosing regimen of the same active drug substance as Copaxone® 20 mg QD but with less frequent injections (Khan 2013). The study was designed to detect a difference between Copaxone® 40 mg TIW vs placebo in the primary endpoint, the annualized relapse rate (ARR). It was not designed to detect a significant effect on the exploratory endpoint of CDP. In fact the power to detect a significant difference in CDP in GALA would

be about 7%, compared to the usual power of 80-90% for a primary endpoint in a clinical trial. Therefore, the disability progression RR generated compared to placebo is based only upon a limited number of events occurring (Table 1), further highlighting the robustness concerns of this one-year estimate.

3. The two-year open-label extension phase of the one-year GALA controlled clinical trial provides an estimate of disability progression for Copaxone® 40 mg TIW over three years (“Early Start” group) relative to one year of placebo plus two years of Copaxone® 40 mg TIW (“Delayed Start” group). Indeed, in the GALA three-year open-label extension, there was a trend towards a reduction in disability with Copaxone® 40 mg TIW (early start vs delayed start) although this was not statistically significant [hazard ratio (HR) = 0.76, 95% CI 0.55-1.04, p=0.09] (Khan 2016). The HR and the corresponding disability RR for Copaxone® 40 mg TIW from the GALA extension study represent a conservative estimate of the actual benefits of Copaxone® 40 mg TIW in reducing disability progression as the delayed start arm data include treatment with placebo and two years of Copaxone® treatment as well. These data indicate that the one-year estimate used by ICER to estimate the Comparative Clinical Effectiveness of Copaxone® 40 mg TIW is inconsistent with the observed disability progression over longer follow-up (Khan 2016) and is therefore misleading and inappropriate.

To avoid a glaring analysis weakness related to Copaxone® 40 mg TIW and based on available data on clinical efficacy (highlighted below) we suggest ICER uses the assumption in the cost-effectiveness model that the disability progression RR for Copaxone® 40 mg TIW is equivalent to that estimated for Copaxone® 20 mg QD (RR=0.70; 95% CI 0.54- 0.93). Likewise, a prediction model and meta-analyses comparing Copaxone® 20 mg QD to Copaxone® 40 mg TIW results in similar clinical and MRI outcomes (Cutter 2014a; Cutter 2014b). The European Union (EU) regulatory agencies considered this comparison across study data, supporting the similarity of effects of Copaxone® 20 mg QD and Copaxone® 40 mg TIW on relapse rates, when approving Copaxone® 40 mg TIW in the EU, as summarized on pages 16-17 of the Copaxone® 40 mg TIW Public Assessment Report (MHRA 2015). Sormani et al. (2010) demonstrated a strong correlation of the effect of DMTs on relapses and CDP, and due to the similarity of effect on relapses between Copaxone® 20 mg QD and Copaxone® 40 mg TIW, it is a reasonable assumption, in the absence of an appropriate trial examining the effect of Copaxone® 40 mg TIW on disability, that

	<p>Copaxone® 20 mg QD and Copaxone® 40 mg TIW have similar effects on CDP.</p> <p>To provide further validity to this assumption that Copaxone® 20 mg QD and Copaxone® 40 mg TIW will have a similar effect on the CDP, using the GALA data Teva conducted additional analyses on disability progression in both arms over two and three years of duration. Using a generalized estimating equation model that adjusted for baseline EDSS and number of relapses in the two years prior to GALA, it was estimated, based on 24 months and 36 months of follow-up data for Copaxone® 40 mg TIW from the GALA extension study, that the corresponding disability RR (early start vs delayed start) estimates are: 24 months = 0.81 (95% CI 0.55-1.19); at 36 months = 0.78 (95% CI 0.59-1.05). We believe that these are conservative estimates of disability progression RR for Copaxone® 40 mg TIW compared with placebo since the delayed start (comparison) group was treated with Copaxone® 40 mg TIW for one out of the two years for the 24 month and two out of the three years for the 36 month estimate. This provides further face validity to our proposed approach of considering the disability progression RR estimate for Copaxone® 40 mg TIW to be equivalent to that estimated for Copaxone® 20 mg QD (RR=0.70; 95% CI 0.54- 0.93).</p> <p>Use of these supplemental data is consistent with the evidence used by leading health authorities. The European Union (EU) regulatory agencies considered this comparison across study data, supporting the authorities to evaluate the value and patient benefit of providing access to Copaxone® 40 mg TIW as a therapeutic option (Cutter 2014a; Cutter 2014b; Giovannoni 2015)</p>	
3	<p>Medication Discontinuation Rates</p> <p>One of the Key Model Assumptions made in the Comparative Value assessment (as noted on page 61, Table 15) is a constant medication discontinuation rate of 10% per year for the first two years of therapy across health states and medications. A few considerations related to this topic are outlined below:</p> <ul style="list-style-type: none"> a. Notable heterogeneity is reported among the DMTs for the treatment of RRMS both in medication discontinuation rates (page 47, Table 11 of the ICER draft report) and in long-term safety profiles (Mikol 2008; O'Connor 2009). b. The study cited (Tappenden 2009) in ICER's evaluation to arrive at a 20% withdrawal rate estimate over the first two years of treatment (i.e., 10% per year) for the included interventions does not reflect these observed differences. 	See p.10 question 39.

	<p>We respectfully suggest that ICER reconsiders use of treatment-specific drug discontinuation rates in place of the current assumption of same discontinuation rates across all therapies. Not accounting for drug specific discontinuation rates in the analyses could further impact validity of cost-effectiveness analysis results.</p>	
4	<p>Copaxone® and Glatopa®</p> <p>ICER has included in its cost-effectiveness analysis Glatopa®, a generic version of Copaxone® 20 mg/ml. The clinical effectiveness parameters for Glatopa® used in the model are assumed to be identical to Copaxone® clinical trial results. We suggest that ICER considers in this analysis the findings of data recently presented (Kolitz 2016) and submitted to the FDA (https://www.regulations.gov/document?D=FDA-2007-D-0369-0395), highlighting differences between Copaxone® and Glatopa®, a generic version of Copaxone® 20 mg/mL.</p> <p>Copaxone® is a synthetic complex polypeptide mixture that contains up to 1029 variants of polypeptides. Copaxone® physicochemical properties cannot be fully characterized. There is no measurable pharmacokinetic profile, and no validated pharmacodynamic markers specifically identified and robustly validated to date. The active moiety(ies) are unidentifiable within the active substance. Furthermore, the manufacturing process is extremely sensitive to minor changes in reaction conditions and specifications. Glatopa®, the first generic glatiramer acetate product, was recently approved in accordance with the criteria set forth in FDA Draft Guidance on Bioequivalence of glatiramer acetate. Results of a battery of rigorous scientific tests on multiple commercial batches of Glatopa® demonstrate that there are compositional and biological differences between Glatopa® and Copaxone® across multiple physicochemical attributes as well as inflammatory and immune-related pathways. Teva has submitted its comments to the FDA Draft Guidance, which includes a comprehensive description of the new comparative scientific data available on Glatopa® (https://www.regulations.gov/document?D=FDA-2007-D-0369-0395).</p>	<p>Thank you. The abstract highlights that gene expression studies identify many genes that are similarly modulated by Copaxone and Glatopa including the important upregulation of IL-10. However there are also important differences in gene expression in animal models that may impact safety and efficacy.</p> <p>However, the FDA approved Glatopa.</p> <p>The National MS Society in their evaluation states “If the FDA reviews and approves a generic medication, it means the medication’s maker has provided sufficient evidence that the generic will have the same therapeutic benefits as the brand-name product.”</p> <p>We agree that there is some remaining uncertainty about equivalence, but will rely on the FDA assessment at this time.</p>

Patients and Advocacy Organizations

MS Coalition

1	<p>The MS Coalition commends ICER for the comprehensive review of the current disease modifying therapies in the Draft Evidence Report. However, there was overwhelming consensus that the authors took considerable liberty comparing across</p>	<p>We have added additional language highlighting the uncertainties in the analysis that go beyond the confidence intervals for each agent in the NMA.</p>
---	---	--

	<p>trials that do not have comparable populations and extrapolating from old natural history studies. The report does identify and discuss some of the problems with mixing old and new data but in some cases, this issue is not adequately addressed. For example, (page 2) the authors make the important point that populations changed over time, making trial populations less comparable—yet they proceed with the comparisons. Specifically, the report acknowledges the substantial time span of the trials (1987-2015), the use of different definitions of MS, the use of different inclusion criteria, the use of different outcome definitions, and so on. Yet there is no indication that any adjustment or analysis using these factors was made when generating the results, and no mention of study heterogeneity is made in the presentation of the results as forest plots, league tables, etc.</p> <p>The lack of reliable estimates of disease progression in naïve patients is a major limitation for estimating the cost-effectiveness of MS treatments. These considerations make the comparative efficacy conclusions in the analysis unreliable.</p>	<p>However, the relative benefits for most drug therapies are similar across different population groups. For most drugs, they are equally effective in men and women and in Asians and those of African descent. The absolute benefits of therapy change with differences in risk (like the higher risk populations in the early studies of DMTs compared to more recent studies), but the relative benefit usually is similar. Effect modification by subgroup (different relative treatment effects in one group compared to another) is rare in medicine and usually represents something fundamentally biological. Common examples include targeted therapy in cancer: for instance, Herceptin, which targets the HER2 receptor improves outcomes in patients whose cancer expresses HER2, but does not improve outcomes in patients with tumors that do not express HER2. There is little data supporting effect modification for any of the DMTs for MS.</p> <p>The major problem with the evidence base is that the clinical trials are too short to adequately capture long term differences in disability progression.</p>
2	<p>The report does not take into account the use of more potent agents as first line therapies. In addition, the model does not account for more than one switch of therapy due to breakthrough disease. The projected number of relapses (page 69) is not realistic given current accepted medical practice. Further, extrapolation of clinical trial data which is limited to 2 years or less does not provide support or validation of assumptions made in the report.</p>	<p>The model does allow for any of the therapies as a first line agent even though the FDA indication for several of the drugs generally recommends that they be considered for use in patients who have failed two or more other DMTs. We cross validated with other models and our estimates are aligned.</p>
3	<p>It is not clear how indirect costs were included in the model. For example, on page 4, indirect costs are not included despite the high impact of these costs cited in the background. Further, the term, ‘best supportive care’ is used as a comparator yet no definition or citation is provided. This term must be clearly defined with costs allocated.</p>	<p>We completed a scenario analysis including indirect costs and have emphasized in the report. We have added further description of best supportive care in the text.</p>

4	The use of generic glatiramer acetate 20 mg as the universal comparator is unsupported by any human data. The product has not been studied in a clinical trial, has modest efficacy on relapses and MRI (using brand-name GA 20 mg) and lacks data demonstrating prevention of disability progression.	See p.25 comment 15.
5	Figure 4 shows GA 20 mg with better disability outcome than teriflunomide. However, two large trials of over 1,000 patients each have shown teriflunomide’s positive benefit on disability and one smaller trial of GA failed to show a significant benefit on disability. These results are inconsistent with the reports statement, “Finally, our NMA suggested that interferons, glatiramer acetate, and teriflunomide were substantially similar with respect to their effects on ARR and disability progression.”	<p>Please see our new Figure 4. Teriflunomide 14mg has a point estimate that is more favorable than GA 20. The NMA incorporates the totality of information about GA 20: for instance, in the CONFIRM study, 343 patients were randomized to GA 20 and had a 13% reduction in disability progression compared to placebo. The advantage of an NMA is that it includes all of the data for the DMTs and does not focus undue attention on one or two trials.</p> <p>As noted above, 3 publications using a variety of methods and data sources in 2016 came to this same conclusion.</p>
6	The conclusion that generic GA 20 mg was favored as a ‘good value’ gives a green light to make generic glatiramer acetate the first-line favorite by insurance companies/pharmacy benefit managers. The assumptions made on its performance (negative treatment effects) appear to be from a model-based assessment and are inconsistent with other published data	We assumed that generic glatiramer acetate has the same efficacy results as branded product, as there is no evidence to the contrary. We did not use any model to determine this efficacy results, and only assumed equivalence.
7	There is limited data available on daclizumab, particularly related to long term effects. However, the authors are less cautious in declaring daclizumab safe than for ocrelizumab.	There are limited data on daclizumab’s long term effects and no data on ocrelizumab’s long term effects as it is not yet FDA approved. There is considerable uncertainty about both, but more for ocrelizumab as we have no long term data and do not yet have the advantage of a thorough FDA review of the evidence.
8	The conclusion that generic glatiramer acetate and alemtuzumab are most cost effective has significant negative implications for the availability of MS treatment options. This conclusion is nullified if rituximab (actually the lowest cost generic) is used instead as the comparator.	We did not include rituximab in this model because it’s use for this indication in the US is currently off label and there was insufficient data on disability progression to include in the NMA. We agree that the choice of comparator is important, and chose the least costly version of the most used product with a labelled indication for this reason.
9	Another limitation of the report is the lack of strong data on patient reported outcomes which the authors acknowledge are most important to people living with MS. The importance of	We have expanded the section on quality of life / patient reported outcomes, but there remains a dearth of

	shared decision-making between people living with MS and their clinicians cannot be overstated in the report as an individualized treatment plan is the singular path to achieving both the best results of maximizing efficacy and adherence and achieving individualized goals for each person living with MS.	results to compare across DMTs and over time periods longer than 1-2 years.
10	Costs and Sensitivity Analysis There are concerns with some of the assumptions and references used on the cost side of the analysis. The costs have a source but it might be important to have multiple sources to attempt to better understand the real cost sides. The sensitivity analyses show just how imprecise the results are and thus, it leaves any interpretation pretty wide open to comment/criticism.	We agree that there is limited data on costs, however, we have incorporated the most recent data source consistent with our model structure.
11	Classification of Quality of Studies While the need for certain measurements for this purpose is important, the lack of measurement of one of the key outcomes is not necessarily a quality issue and may encourage the reader to discount a particular study when the study itself is of high quality but may not be of particular utility for this undertaking	No quality rating is perfect. We did not limit studies based on quality rating in the base case analysis, but did assess the effect of leaving out the lowest quality studies.
12	12 Week Confirmation Rates There is a fundamental flaw in using the 12-week confirmation rates without adjustment—even a ratio adjustment could be made but this scrambles the expected relationship between the relapse rates and CDP rates and falsely raises the CDP compared to the 24 week data as it has been well shown that 12 week CDP is higher than 24 weeks. The estimated increase over placebo by GA 40 mg seems inconsistent with data presented to date (page 40).	We are limited by the data. Using a ratio adjustment would give undue weight to studies that did not measure or report the most important outcome (24 week CDP). We received consistent feedback from dozens of experts and organizations that 24 week CDP is superior to 12 week CDP. As for GA 40, the only published randomized trial data reported an increase in CDP in the GA 40 group compared to the placebo group. We highlight the implausibility of the result, but that was the finding in the study. We have removed GA 40 from the cost effectiveness model because of the implausibility of the finding.
13	Bayesian Methodology The use of the credible interval is reasonable (page 41) but the assumptions of the Bayesian methodology and how MCMC borrows information in the face of the previously described time trends needs a bit more discussion. One solution to understanding would be to repeat the NMA studies after 2008 to see how many of the conclusions depend on the earlier data, where the definitions of relapse were looser, the populations potentially sicker, certainly with longer duration of disease.	We attempted to do this in our subgroup analyses by diagnostic definition. The earliest trials used the Poser criteria, while the later trials use the MacDonald criteria. Tables D1 and D5. In almost all instances, the results for the risk ratio for relapse were nearly identical. The only exception is interferon beta-1a (0.91 Poser, 0.79 Macdonald), but the confidence interval are widely overlapping and the

		<p>difference is not statistically significant. There is also no consistent direction for the differences for individual drugs comparing trials using the Poser criteria to trials using the MacDonald criteria.</p>
14	<p>Trial Discontinuation Rate The rationale for the 10% trial discontinuation rates for all drugs is not reasonable. The discontinuation rates for Tecfidera are much higher than others and this should factor into the model in some way</p>	<p>We have changed to use DMT specific rates from trial data.</p>
15	<p>Mortality Multiplier The mortality multiplier is based on 1997 data which antedates the treatment era as well as not reflecting the current diagnostic criteria. It is possible that milder cases are being included today in cohorts because there are treatments that can be offered whereas in the era before formal DMT's, little was offered and the milder cases may have been ignored or not labeled. There are at least 4 datasets that provide mortality data that might be used to inform this.</p>	<p>We were unable to find a more recent data source for mortality multipliers consisted with our model structure (i.e. stratified by EDSS states). The projected life-expectancy were in line with other epidemiology data.</p>
16	<p>Utilities The utilities come from two related studies and may not be as generalizable as might be desired. In summary, we commend the considerable effort involved in compiling the available evidence and conducting the analyses described in this report. However, as noted above, several of the choices made in the NMA analysis and cost-effectiveness model significantly undermine our confidence in the results. We also wish to emphasize that due to heterogeneity in treatment response, safety and tolerance, and individual preferences, none of the currently available MS DMTs will be beneficial and appropriate for all MS patients, but each DMT will be beneficial and appropriate for some patients. We would appreciate seeing this point reinforced in the report, with support given for access by patients and physicians to all DMTs without undue restrictions on the part of payers.</p>	<p>These utility values may not be applicable to all patients, but are useful to apply to a cohort of all patients.</p>
National MS Society		
1	<p>We commend ICER on their review of the MS medication class and for seeking to bring economic clarity to this expensive class of medications. High prices, along with increased out-of-pocket costs for people with MS, inconsistent formularies across different insurers, lack of price transparency and complex approval and appeals processes often create barriers to people with MS accessing the right treatment for them. The Society's "Make Medications Accessible" Initiative seeks to find solutions to these challenges with all stakeholders involved in the healthcare system. We hope that ICER's final evidence report can bring value to these important conversations.</p>	<p>Thank you for your comments.</p>
2	<p>We found the analysis of the clinical trial evidence to be a thorough summary of disease modifying therapies (DMTs)</p>	<p>We have expanded section 5: other benefits or disadvantage to try to</p>

	<p>approved for use in the United States (U.S) market. We were also pleased to see some incorporation of the learnings ICER accumulated from outreach to people living with MS, patient advocacy groups, healthcare providers and other stakeholders. These included recognition of the economic burdens facing people with MS, a desire for patient-reported outcomes and the critical importance of shared decision making with their healthcare provider to ensure treatment choices that meet individual needs. However, not enough attention is paid to the heterogeneity of MS and the differences in the mechanisms of action associated with the DMTs, which are of high importance when choosing treatment. The type of analysis that ICER attempts is commendable; however, it is dependent on many variables that are further complicated by the heterogeneity of MS, the variable individual response to medication, and a large number of quality of life factors. Studies show that early and ongoing treatment with a DMT effectively modifies the course of the disease, prevents the accumulation of disability and protects the brain from damage due to MS. As such, we believe that a full range of treatment options should be available to every person living with MS, so that they - in collaboration with their health care providers - can make informed treatment decisions. Further, any person who is stable on a DMT should not be forced to switch to another agent because of changes in medication coverage or cost considerations. A delay in treatment can have a negative and permanent result</p>	<p>capture more fully the less quantifiable benefits of multiple DMTs with multiple mechanisms and link that to the heterogeneity of the disease experience.</p> <p>We have also expanded the section reporting on the quality of life / patient-centered outcomes.</p>
3	<p>In our review of the draft evidence report, the Society has outlined some areas that need to be reevaluated for accuracy and to improve the usefulness of the document. Some of the below inaccuracies are regarding alemtuzumab and glatiramer acetate, both of which figure prominently in the review and conclusions as a cost effective treatment and baseline treatment respectively.</p> <ul style="list-style-type: none"> • Within Table 1, ICER has listed alemtuzumab’s dosage as 12 mg per day for 3 days every year. The label for alemtuzumab, marketed as Lemtrada, states that the drug should be administered for 5 days at baseline, and then for 3 days a year later. Additional doses are only administered after that with new disease activity. • The American Academy of Neurology Draft Guidelines do not recommend testing for antibodies to John Cunningham virus (JCV) in patients taking fingolimod or dimethyl fumarate nor avoidance of these drugs in patients with JCV antibodies. • ICER reports the CONFIRM trial of glatiramer acetate and dimethyl fumarate versus placebo as a head to head trial: however, the CONFIRM trial was not powered as a head to head assessment. • The authors state that alemtuzumab was consistently better in preventing disability progression; however, in the Care-MS1 	<ul style="list-style-type: none"> • The dosing for alemtuzumab has been clarified. • We have revised the description of the AAN guidelines to note that the JC virus testing considerations apply only to natalizumab. • CONFIRM compares DMF and GA to placebo – it may not have had adequate power, but it provides head to head evidence that is valuable, particularly when combined with other evidence in an NMA. • In CARE MS1 the HR for disability progression was 0.70, which was not statistically significant, but was consistent with the other evidence of benefit. • Thank you for pointing out the Typo for natalizumab’s rating • We have added clarification about the FDA label for alemtuzumab

	<p>trial, there was no significant difference between the alemtuzumab and IFNB-1a in preventing disability progression.</p> <ul style="list-style-type: none"> • Natalizumab, when compared to a generic glatiramer acetate, was given a B+ rating however, the accompanying table (Table 13) had its designation listed as a C+. • In the U.S., alemtuzumab has a strong recommendation from the FDA to be used as a thirdline therapy; however, within the review, the authors repeatedly refer to it as a second-line therapy 	
4	<p>Current Limitations of the Draft Evidence Report</p> <p>While the Society appreciates ICER’s thorough review, we are concerned by assumptions made within the document, the scientific validity of the comparisons used and the resulting value conclusions. Insufficient attention is paid to the heterogeneity of the MS population, quality of life factors and variable response to treatments. In the survey of people with MS, 90% rated continuing working/normal activities as important/very important- behind only delaying disability and preventing relapse (Table 3). The authors state this echoes what they heard from individual patients and patient advocacy groups, yet this doesn’t have a corresponding emphasis in the analysis. In our view, the report also draws incorrect conclusions from the widely differing opinions on treatment guidelines (American Academy of Neurology, Canadian Agency for Drugs and Technology in Health, MS Coalition and National Institute for Health and Care Excellence) and the range of coverage policies by payers. The range of these guidelines and policies indicates the need for differing options due to the heterogeneity of MS. Reviews like ICER’s that look at cost effectiveness may be used to limit access to DMTs for people living with MS. Therefore, we believe it is critical that ICER acknowledge the limitations of the review and clearly point out the many assumptions that were made that potentially undermine the validity of the cost conclusions.</p>	<p>Thank you for the input. We have expanded the section on quality of life measures. Unfortunately, none of the trial report on continuing work / maintaining normal activities as an outcome. A major weakness in the comparative trials is the short time frame: one to two years is too short to evaluate effects on work / usual home activities. Trials of 5 years duration, as are typical in cardiovascular disease, would provide a much more robust evidence base.</p> <p>We have pointed out several of the limitations of our cost-effectiveness analysis in the Summary and Comment (Section 6.7).</p>
5	<p>We remain concerned that the comparisons that ICER used to evaluate the different treatment trials are based on data that are more than two decades old. These data and the study populations for older therapies do not represent modern populations or current practice. People entering trials for relapsing remitting (RRMS) MS for the older therapies were generally in a later state of disease than those currently entering RRMS trials due to improved diagnostic tools. Further, the randomized controlled trials (RCT) only show data over a relatively short time frame (usually a maximum of 2 years). Beyond that time period, there is very limited data available to validate the assumptions that ICER makes in the document. Given these significant study population differences, the RCTs are not directly comparable, thus making the resulting comparative efficacy conclusions in the analysis unreliable.</p>	<p>As noted above, the data do not suggest any difference in the relative efficacy of drugs that were studied in the older era using the Poser criteria for trial entry and more recent trials using the MacDonald criteria. The study populations are different, but the relative rate for relapses has remained constant. This is a typical finding for drug therapy. For instance, the relative reduction in heart attacks, strokes, and death from cardiovascular disease is similar in high risk patients, intermediate risk patients, and low risk</p>

	ICER acknowledges the challenges of trying to compare therapies based on registration trials, but does not adequately account for this challenge in the result and cost-analysis. The lack of reliable estimates of MS progression in newly diagnosed patients is another major limitation for estimating cost effectiveness of MS treatment modalities; its implications on the results of any predictive modeling need more attention in the review.	patients. Effect modification by patient characteristics or risk of disease is rare.
6	Additionally, the review makes the assumption that a person with MS goes off treatment after failure with second-line therapies. This assumption is not consistent with current medical practice or payer policies. There are many reasons why someone may need to switch to another DMT after the second therapy: allergy, adverse side effect (e.g. laboratory abnormalities), new contraindication, etc. While many people with MS will take more than one DMT throughout the course of the disease, it is also common for people to take more than one medication that ICER refers to as first-line before moving to a medication that ICER refers to as second-line. Often, this is due to payer policies. People with MS may also take more than one of the “second-line” therapies. These assumptions should be changed in the final review to reflect current practices.	See p.9 comment 38.
7	The draft evidence report also lacks reliable data on patient reported outcomes, which as the authors state (Table 3) are the most important outcomes for patients. Furthermore, the utility data that the authors used in their modeling came primarily from non-U.S. studies. Utility data are known to be reflective of cultural and societal preferences, therefore it is likely that these data do not represent the true preferences of a person in the U.S. who lives with MS. Changes in relapse management, as well as other healthcare delivery changes are also likely to affect costs.	See p.18 comment 2.
8	In our review, it appears that indirect and direct health costs are missing health expenditures that are common for people with MS. For example, when a person with MS switches or begins a new DMT, this often requires additional physician visits for medication adjustments and side effect management. Regular MRIs may also be used to monitor or assess DMT effectiveness.	We added a physician visit when switching or discontinuing. We did not add periodic MRI because there are no guidelines that recommend frequency of periodic MRIs.
9	<p>Possible Areas of Improvement for Final Evidence Report and Final Voting Questions</p> <p>The Society has outlined several areas that the authors should reevaluate in the final evidence report. We believe that these revised components will improve the review for providers and people living with MS.</p> <ul style="list-style-type: none"> • The authors should reconsider the exclusion of clinically isolated syndrome studies as the implications of treatment decision on people with this early form of MS are particularly important. 	It is beyond the scope of the review to include CIS.

<ul style="list-style-type: none"> • ICER should reconsider their projected number of relapses on page 69 to better align with modern treatment guidelines. • Ocrelizumab, to date, has not received approval from the Food and Drug Administration (FDA), and is unlikely to be approved prior to the review of this report; therefore much information concerning benefit/risk and price is speculative. The authors should reevaluate the information contained in the review on ocrelizumab once FDA’s approval decision is made and more precise data and pricing information on the drug is known. • It is currently unclear how the model used calculates and allocates indirect costs. More details on components used to calculate indirect costs and how they are valued is needed to truly identify and present cost. In addition, it would enhance confidence in the model if ICER were to publish the details of the model in a peer-reviewed publication. • The economic evaluation that the authors utilize (Table 20) is based on a single NARCOMS survey from 2004. The prices of all MS therapies and reimbursement amounts for services and delivery have changed dramatically since that time. The authors should note in the final evidence review how their economic evaluation accounts for price differences since 2004 and how associated healthcare costs were estimated for other DMTs which were not available in 2004. • Real world practice and treatment should be factored in with the cost analysis. For example, alemtuzumab is FDA approved as a third-line therapy. FDA makes a strong recommendation in the labeling that this treatment is to be used only after inadequate response to two or more DMTs. Thus, even though the review rates it as cost effective, this treatment is likely not an available option for many people with MS due to the labeling information and medical practice. • ICER should add an answer choice of “insufficient data” to their draft voting questions. The limitations of the review impact the efficiency and cost conclusions drawn and currently there is a lack of scientifically validated data to answer the questions posed. 	<p>See comment regarding projected relapses above.</p> <p>As new evidence comes to light, information in ICER reports may need to be revised or updated.</p> <p>Additional detail has been added on indirect costs. We plan to publish a peer reviewed manuscript in addition to the report.</p> <p>Prices of MS therapies, monitoring, and delivery reflect current costs, as described. We were unable to identify a more recent source for cost data on other underlying healthcare costs. Increases in costs are accounted for with inflation.</p> <p>We have added additional language about the FDA labeled indication.</p> <p>The ICER report shows were there are issues of inadequate evidence. In the face of inadequate evidence, policy makers must still make decisions, and so the CTAF will be asked to make their best judgments in the face of the available evidence, however evidentiary issues will be highlighted as they arise.</p>
<p>10 As ICER moves to finalize its review and voting questions, the Society believes it is important to acknowledge the benefit of this type of analysis to inform providers and people affected by MS about the full spectrum of approved treatment options. The heterogeneity of the MS population and the clinical variability of MS between individuals make access to the full range of therapies critically important. Treatment that may be effective and well tolerated in one may fail in another person, and people with MS may utilize several treatments in their lifetime. We believe that individualized treatment plans, created by shared decision making between people with MS and their physicians will produce the best result and cost effectiveness by</p>	<p>It would be helpful if you could be more specific about how different phenotypes are defined within the heterogeneous manifestations of RRMS and how those different phenotypes should influence treatment choices.</p> <p>We recognize that patients often fail one treatment and are then tried on another, but it appears to be a process of trial and error rather than one guided</p>

	maximizing efficacy and adherence, while balancing risk tolerance.	by biology, patient characteristics, or disease manifestations.
Lisa Carr		
1	I have had RRMS since 1990. I am a RN and had to stop working in 2013. I am almost 49 years old. I read over some of your materials and have some thoughts. I'm not sure how beneficial they are, but do hope you read this and there is some consideration. I have been on numerous DMTs since 1997 and do think it is beneficial to consider them and which is more effective. However, my big question is, does decreasing relapses really effect disability? I think I'm kind of young to be out of work. Believe me, relapses can be awful, but i think we need to think of this?	Thank you for your comments.
2	Also, maybe more importantly, i have asked for antibodies at least twice, to be told insurance doesn't cover it. So stupidly, i never had it done. I thought it was not a good thought. However, i had two big relapses last year while on Tysabri for over two years and one every year before that, the same as prior to being on that med. I thought is was the strongest med. I do realize this is a progressive disease, and know it does work for some people (for relapses...I don't know statistics on disease progression) but it did not work for me. So after much research, i talked to my neurologist and discontinued it, with my last infusion March 20, 2016. I started non-traditional therapy and have not had a relapse in a year. I started the anti-inflammatory diet 4/15/16, LDN 5/20/16, and hormone therapy 8/22/16. I have not had a relapse in one year. The insurance does not cover these, so i pay cash out of pocket even though i am on a limited income now. So my insurance payed \$10,000 a month but wont cover \$100 a month. This does not seem right. And i still have not had antibody testing, but in my mind i feel those meds didn't work for me but this method is.	
3	So perhaps we need to consider voting on some additional questions? Please consider and maybe respond to me as well. I am willing to help in any way. I don't think insurance companies shouldn't cover certain meds, because I'm telling you, a person is willing to try almost anything when they are very sick, and if insurance says they won't cover something, that is not fair. Should a doctor have knowledge of drug superiority, yes, and he should share the knowledge with the patient, but drugs should all be covered. Maybe certain ones only after others have failed? But then again, are we messing with the disease process? Maybe that's why antibodies are important?	
Mary Holmstrand		
1	Where is the research for patients who have decided to NOT to DMDs because they really don't work and only have to have 30% efficacy to be approved by the FDA not to mention the side effects by introducing a "decoy" molecule to your body. A decoy may/may not be effective for a short time until the	Thank you for your comments.

<p>"decoy" does not work anymore. Oh, right, then patient goes on to yet another DMDs. Just makes no sense. What about the patients who have changed their diet, added key supplementation, increased exercise, reduced lifestyle stress, etc...and they are doing just fine without DMDs. Stop supporting Pharm and look for the root cause!</p>	
---	--