IQWiG Reports – Commission No. A21-159

Ponesimod (multiple sclerosis) –

Benefit assessment according to §35a Social Code Book V[1]

Extract

[1] Translation of Sections 2.1 to 2.5 of the dossier assessment *Ponesimod (multiple Sklerose) – Nutzenbewertung gemäß § 35a SGB V* (Version 1.0; Status: 24 February 2022). Please note: This document was translated by an external translator and is provided as a service by IQWiG to English-language readers. However, solely the German original text is absolutely authoritative and legally binding.

Publishing details

Publisher
Institute for Quality and Efficiency in Health Care

Topic
Ponesimod (multiple sclerosis) – Benefit assessment according to §35a Social Code Book V

Commissioning agency
Federal Joint Committee

Commission awarded on
2 December 2021

Internal Commission No.
A21-159

Address of publisher
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln
Germany

Phone: +49 221 35685-0
Fax: +49 221 35685-1
E-mail: berichte@iqwig.de
Internet: www.iqwig.de

Medical and scientific advice

No advisor on medical and scientific questions was available for the present dossier assessment.

Patient and family involvement

The questionnaire on the disease and its treatment was answered by 2 people.

IQWiG thanks the respondents for participating in the written exchange about how they experienced the disease and its treatment as well as their treatment goals. The respondents were not involved in the actual preparation of the dossier assessment.

IQWiG employees involved in the dossier assessment

- Michael Köhler
- Kirsten Janke
- Thomas Kaiser
- Matthias Maiworm
- Annika Orland
- Sabine Ostlender
- Dominik Schierbaum
- Anja Schwalm
- Sibylle Sturtz

Keywords: Ponesimod, Multiple Sclerosis, Benefit Assessment, NCT02425644

Table of contents

List of tables[2]

Page

[2] Table numbers start with “2” as numbering follows that of the full dossier assessment.

List of abbreviations

Abbreviation	Meaning
9-HPT	9-Hole Peg Test
ACT	appropriate comparator therapy
AE	adverse event
CI	confidence interval
C-SSRS	Columbia Suicide Severity Rating Scale
EDSS	Expanded Disability Status Scale
FDA	Food and Drug Administration
FSIQ-RMS	Fatigue Symptoms and Impacts Questionnaire – Relapsing Multiple Sclerosis
G-BA	Gemeinsamer Bundesausschuss (Federal Joint Committee)
Gd	gadolinium
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Institute for Quality and Efficiency in Health Care)
MCS	Mental Component Summary
MSFC	Multiple Sclerosis Functional Composite
PASAT-3	Paced Serial Addition Test-3
PCS	Physical Component Summary
PGI-S	Patient Global Impression of Severity
PT	Preferred Term
RCT	randomized controlled trial
RMS	relapsing multiple sclerosis
SAE	serious adverse event
SDMT	Symbol Digit Modalities Test
SF-36v2	Short Form-36 Health Survey Version 2
SGB	Sozialgesetzbuch (Social Code Book)
SOC	System Organ Class
SPC	Summary of Product Characteristics
T25-FW	Timed 25-Foot Walk
WPAI:MS	Work Productivity and Activity Impairment Questionnaire: Multiple Sclerosis

2 Benefit assessment

2.1 Executive summary of the benefit assessment

Background

In accordance with § 35a Social Code Book (SGB) V, the Federal Joint Committee (G-BA) commissioned the Institute for Quality and Efficiency in Health Care (IQWiG) to assess the benefit of the drug ponesimod. The assessment is based on a dossier compiled by the pharmaceutical company (hereinafter referred to as the "company"). The dossier was sent to IQWiG on 14 June 2021. A benefit assessment on this dossier has already been completed and was sent to the G-BA on 13 September 2021 (dossier assessment for commission A21-83). With its decision dated 2 December 2021, the G-BA temporarily suspended decision-making regarding the benefit assessment of ponesimod for patient group a (research question 1 of dossier assessment A21-83) and commissioned IQWiG with a reassessment of benefit on the basis of the dossier already submitted. This decision was taken due to a change in the ACT for this patient group based on information submitted in the written and oral commenting procedure.

Research question

The aim of this report is to assess the added benefit of ponesimod in comparison with the appropriate comparator therapy (ACT) in adult patients with active relapsing multiple sclerosis (RMS) without prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active.

The G-BA's specification of the ACT results in the research question presented in Table 2.

Table 2: Research question of the benefit assessment of ponesimod

Therapeutic indication	**ACT**[a]
Adult patients with active RMS without prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active	IFN-β 1a or IFN-β 1b or glatiramer acetate or dimethyl fumarate or **teriflunomide** or ocrelizumab, taking into account approval status
a. Presented is the ACT specified by the G-BA. In cases where the ACT specified by the G-BA allows the company to choose a comparator therapy from several options, the respective choice of the company is printed in bold. ACT: appropriate comparator therapy; G-BA: Federal Joint Committee; IFN: interferon; RMS: relapsing multiple sclerosis	

The company has selected teriflunomide as the ACT, thereby following the ACT specified by the G-BA for the present research question.

The assessment was conducted by means of patient-relevant outcomes on the basis of the data provided by the company in the dossier. Randomized controlled trials (RCTs) with a minimum duration of 12 months were used to derive added benefit.

Study pool and study design

The OPTIMUM study was included in the benefit assessment.

Study design

The OPTIMUM study is a randomized, double-blind study comparing ponesimod with teriflunomide. It included adult patients with active RMS and an Expanded Disability Status Scale (EDSS) score of 0 to 5.5. These patients had either no prior treatment with disease-modifying therapies or had received prior treatment with interferons, glatiramer acetate, natalizumab, or dimethyl fumarate.

The study randomized a total of 1133 patients at a 1:1 ratio to treatment with either ponesimod or teriflunomide. Ponesimod and teriflunomide treatment were each administered in line with the Summary of Product Characteristics (SPC) for a period of 108 weeks.

The primary outcome of the study was annualized relapse rate. Patient-relevant secondary outcomes were from the morbidity, health-related quality of life, and side effects categories.

Several characteristics of the OPTIMUM study are relevant for this benefit assessment, particularly the questionable suitability of the overall population, the high number of major protocol violations, and the change in the recording of the primary outcome (relapse) over the course of the study.

- Some OPTIMUM participants were not covered under the research question of this benefit assessment (highly active disease despite adequate prior treatment with a disease-modifying therapy). The company's approach for defining this subpopulation is adequate. Since this patient group was small (about 7% of the total population), however, the company used the total population. Highly active disease can be associated with an increased relapse frequency and consequently faster disability progression. For patients with highly active disease despite adequate prior treatment with a disease-modifying therapy, a deviating ACT excluding teriflunomide has been additionally defined (see dossier assessment A21-83). Overall, it remains unclear whether the results for the total population of the OPTIMUM study can be fully extrapolated to the target population of treatment-naive and pretreated patients whose disease is not highly active.
- Study documents show that all patients experienced at least 1 protocol deviation and 47% had at least 1 major protocol deviation. Some of these protocol deviations concern the recording of patient-relevant outcomes for the present benefit assessment. For the most part, protocol deviations rated both as major and as other deviations are spread evenly between study arms, and each of the individual reasons for deviation typically occurred in only a few patients. Overall, however, it remains unclear whether the deviations affect the OPTIMUM study's results. They conceivably might, particularly regarding outcomes with few total events, such as bradycardia.

- The procedure used to diagnose and confirm relapses was materially changed over the course of the OPTIMUM study through protocol version 4 dated 5 February 2016. A series of changes were introduced which rendered the processes of relapse documentation substantially more precise. In particular, these changes virtually ruled out any mutual influencing between treating neurologists and blinded outcome-recording persons who carried out the EDSS assessment. Due to these adjustments, relapse diagnosis and evaluation are expected to have been more reliable under protocol versions 4 and later.

Risk of bias and assessment of the certainty of conclusions

Due to the large number of protocol deviations, the risk of bias on the study level is high for the OPTIMUM study. The certainty of conclusions for the study results was reduced for the present research question, in part due to the inclusion of the irrelevant subpopulation of patients with highly active disease despite appropriate prior treatment. Based on the OPTIMUM study, at most hints, e.g. of an added benefit, can be derived for all outcomes presented.

Results

Mortality

Overall survival

The results on all-cause mortality are based on data on fatal adverse events (AEs). There was no statistically significant difference between treatment groups. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the outcome of all-cause mortality; an added benefit is therefore not proven.

Morbidity

Confirmed relapses (EDSS-based)

A statistically significant difference between treatment groups was found for the outcome of confirmed relapses, operationalized using the annualized relapse rate. There was an effect modification by the characteristic of baseline EDSS score. For patients with an EDSS score ≤ 3.5, this results in a hint of added benefit of ponesimod in comparison with teriflunomide. For patients with an EDSS score > 3.5, this results in no hint of added benefit of ponesimod in comparison with teriflunomide; an added benefit is therefore not proven for these patients.

Confirmed disability progression (EDSS-based)

No statistically significant difference between treatment groups was found for the outcome of confirmed disability progression. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for this outcome; an added benefit is therefore not proven.

Disability severity (MSFC)

For the outcome of disability severity, recorded using the Multiple Sclerosis Functional Composite (MSFC) z score, there is a statistically significant difference between treatment groups in favour of ponesimod. However, the 95% confidence interval (CI) of Hedges' g was

not completely above the irrelevance threshold of 0.20. Therefore, the effect cannot be inferred to be relevant. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the outcome of disability severity; an added benefit is therefore not proven.

Fatigue (PGI-S)

There was no statistically significant difference between treatment groups for the outcome of fatigue, recorded with the Patient Global Impression of Severity (PGI-S). Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the outcome of fatigue; an added benefit is therefore not proven.

Health-related quality of life recorded with the Short Form-36 Health Survey Version 2 (SF-36v2)

For the outcome of health-related quality of life, recorded with the SF-36v2, the company submitted analyses of responder analyses of both improvement and deterioration from baseline. Given this data situation, both operationalizations are taken into account, and the overall picture of results is interpreted in the assessment of added benefit.

For the Physical Component Summary (PCS) of the SF-36v2, there was no statistically significant difference between treatment groups on the basis of the analyses of improvement from baseline. For the outcome of deterioration from baseline, a statistically significant difference was found in favour of ponesimod. The researchers found an effect modification which was caused by the characteristic of baseline EDSS score and is consistent with the effect modification in the outcome of annualized relapse rate. For patients with an EDSS score ≤ 3.5, this results in a hint of added benefit of ponesimod in comparison with teriflunomide regarding the SF-36v2 PCS. For patients with an EDSS score > 3.5, this results in no hint of added benefit of ponesimod in comparison with teriflunomide; an added benefit is therefore not proven for these patients.

For the Mental Component Summary (MCS) of the SF-36v2, there was no statistically significant difference between treatment groups to show either improvement or deterioration from baseline. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the SF-36v2 MCS; an added benefit is therefore not proven.

Side effects

Serious adverse events (SAEs), discontinuation due to AEs

There was no statistically significant difference between treatment groups for either of the outcomes of SAEs or discontinuation due to AEs. For these outcomes, there was therefore no hint of greater or lesser harm from ponesimod versus teriflunomide; consequently, there is no proof of greater or lesser harm.

Specific AEs

Bradycardia (PT, AE)

A statistically significant difference between treatment groups to the disadvantage of ponesimod was found for the outcome of bradycardia. For this outcome, there is therefore a hint of greater harm from ponesimod in comparison with teriflunomide.

Infections and infestations (System Organ Class [SOC], SAE)

No statistically significant difference between treatment groups was shown for the outcome of infections and infestations. For this outcome, there was therefore no hint of greater or lesser harm from ponesimod versus teriflunomide; consequently, there is no proof of greater or lesser harm.

Alopecia (PT, AE)

A statistically significant difference between treatment groups in favour of ponesimod was shown for the outcome of alopecia. For this outcome, there is therefore a hint of lesser harm from ponesimod in comparison with teriflunomide.

Probability and extent of added benefit, patient groups with therapeutically important added benefit[3]

On the basis of the presented results, the probability and extent of added benefit of the drug ponesimod in comparison with the ACT have been assessed as follows:

Overall, both favourable and unfavourable effects of ponesimod in comparison with teriflunomide were found for adult patients with active RMS without prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active. Some of the favourable effects were found only for the subgroup with lesser disease severity (baseline EDSS score ≤ 3.5). Below, favourable and unfavourable effects are therefore weighed separately for patients with an EDSS score ≤ 3.5 versus those with a score > 3.5.

Patients with an EDSS score ≤ 3.5

For patients with an EDSS score ≤ 3.5, exclusively favourable effects of ponesimod versus teriflunomide were found regarding morbidity (confirmed relapses) and health-related quality of life (SF-36v2 PCS), each with the extent of considerable. In the side effects category, both

[3] On the basis of the scientific data analysed, IQWiG draws conclusions on the (added) benefit or harm of an intervention for each patient-relevant outcome. Depending on the number of studies analysed, the certainty of their results, and the direction and statistical significance of treatment effects, conclusions on the probability of (added) benefit or harm are graded into 4 categories: (1) "proof", (2) "indication", (3) "hint", or (4) none of the first 3 categories applies (i.e., no data available or conclusions 1 to 3 cannot be drawn from the available data). The extent of added benefit or harm is graded into 3 categories: (1) major, (2) considerable, (3) minor (in addition, 3 further categories may apply: non-quantifiable extent of added benefit, added benefit not proven, or less benefit). For further details see [1,2].

favourable effects (alopecia) and unfavourable effects (bradycardia) of ponesimod were found for individual specific AEs.

Overall, for patients with an EDSS score ≤ 3.5, this results in a hint of considerable added benefit of ponesimod in comparison with teriflunomide.

Patients with an EDSS score > 3.5

For patients with an EDSS score > 3.5, no favourable or unfavourable effects of ponesimod versus teriflunomide were found for morbidity or for health-related quality of life. In the side effects category, both favourable effects (alopecia) and unfavourable effects (bradycardia) of ponesimod were found for individual specific AEs.

In summary, there is no hint of added benefit of ponesimod in comparison with teriflunomide for patients with an EDSS score > 3.5; an added benefit is therefore not proven.

Table 3 shows a summary of the probability and extent of added benefit of ponesimod.

Table 3: Ponesimod – probability and extent of added benefit

Therapeutic indication	**ACT[a]**	**Probability and extent of added benefit**
Adult patients with active RMS without any prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active	IFN-β 1a or IFN-β 1b or glatiramer acetate or dimethyl fumarate or **teriflunomide** or ocrelizumab, taking into account approval status	▪ Patients with an EDSS ≤ 3.5: hint of considerable added benefit ▪ Patients with an EDSS > 3.5: added benefit not proven
a. Presented is the ACT specified by the G-BA. In cases where the ACT specified by the G-BA allows the company to choose a comparator therapy from several options, the respective choice of the company is printed in bold. ACT: appropriate comparator therapy; G-BA: Federal Joint Committee; IFN: interferon; RMS: relapsing multiple sclerosis		

The approach for deriving an overall conclusion on added benefit is a proposal by IQWiG. The G-BA decides on the added benefit.

2.2 Research question

The aim of this report is to assess the added benefit of ponesimod in comparison with the ACT in adult patients with active RMS without prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active. The present benefit assessment does not discuss patients with highly active disease despite treatment with disease-modifying therapy. Said patient group has already been investigated in dossier assessment A21-83 [3].

The research question presented in Table 4 results from the ACT specified by the G-BA.

Table 4: Research question of the benefit assessment of ponesimod

Therapeutic indication	**ACT[a]**
Adult patients with active RMS without any prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active	IFN-β 1a or IFN-β 1b or glatiramer acetate or dimethyl fumarate or **teriflunomide** or ocrelizumab, taking into account approval status
a. Presented is the ACT specified by the G-BA. In cases where the ACT specified by the G-BA allows the company to choose a comparator therapy from several options, the respective choice of the company is printed in bold. ACT: appropriate comparator therapy; G-BA: Federal Joint Committee; IFN: interferon; RMS: relapsing multiple sclerosis	

The company has selected teriflunomide as the ACT, thereby following the ACT specified by the G-BA for the present research question in accordance with the decision dated 2 December 2021 [4,5].

The assessment was conducted by means of patient-relevant outcomes on the basis of the data provided by the company in the dossier. RCTs with a minimum duration of 12 months were used for deriving added benefit. This concurs with the company's inclusion criteria.

2.3 Information retrieval and study pool

The study pool of the assessment was compiled on the basis of the following information:

Sources of the company in the dossier:

- study list on ponesimod (status: 13 April 2021)
- bibliographical literature search on ponesimod (last search on 22 April 2021)
- search in trial registries / study results databases on ponesimod (last search on 20 April 2021)
- search on the G-BA website on ponesimod (last search on 10 May 2021)

To check the completeness of the study pool:

- search in trial registries for ponesimod (last search on 7 July 2021); see Appendix A of the full dossier assessment for the search strategies

The check did not identify any additional relevant study.

2.3.1 Studies included

The study listed in the following table was included in the benefit assessment.

Table 5: Study pool – RCT, direct comparison: ponesimod vs. teriflunomide

Study	Study category			Available sources		
	Study for the approval of the drug to be assessed (yes/no)	Sponsored study[a] (yes/no)	Third-party study (yes/no)	Clinical study report (CSR) (yes/no [citation])	Registry entries[b] (yes/no [citation])	Publication[c] (yes/no [citation])
AC-058B301 (OPTIMUM[d])	Yes	Yes	No	Yes [6,7]	Yes [8,9]	Yes [10,11]

a. Study for which the company was sponsor.
b. Citation of the study registry entries and, if available, of the reports on study design and/or results listed in the study registries.
c. Other sources: documents from the search on the G-BA website and other publicly available sources.
d. In the following tables, the study is referred to by this acronym.

G-BA: Federal Joint Committee; RCT: randomized controlled trial

2.3.2 Study characteristics

Table 6 and Table 7 describe the study used for the benefit assessment.

Table 6: Characteristics of the studies included – RCT, direct comparison: ponesimod versus teriflunomide

Study	Study design	Population	Interventions (number of randomized patients)	Study duration	Location and period of study	Primary outcome; secondary outcomes[a]
OPTIMUM	RCT, double-blind, parallel	Adult patients (18–55 years) with active[b] RMS and a baseline EDSS score of 0–5.5, either not pretreated or pretreated with IFN β-1a, IFN β-1b, glatiramer acetate, natalizumab or dimethyl fumarate	Ponesimod (N = 567) Teriflunomide (N = 566)	Screening: up to 45 days before randomization Treatment duration: 108 weeks[c] Follow-up observation: until 37 days after the end of the treatment phase[d]	172 centres in Belarus, Bosnia and Herzegovina, Bulgaria, Canada, Croatia, Czech Republic, Finland, France, Georgia, Germany, Greece, Hungary, Israel, Italy, Latvia, Lithuania, Mexico, Poland, Portugal, Romania, Russia, Serbia, Spain, Sweden, Turkey, Ukraine, United Kingdom, United States 04/2015 – 05/2019	Primary: annualized relapse rate Secondary: morbidity, health-related quality of life, AEs

a. Primary outcomes include information without consideration of the relevance for this benefit assessment. Secondary outcomes only include information on relevant available outcomes for this benefit assessment.
b. Active disease was defined as ≥ 1 relapse within 12 to 1 months prior to the first EDSS assessment or ≥ 2 relapses within 24 to 1 months prior to the first EDSS assessment or ≥ 1 Gd-enhancing lesion within 6 months prior to the first EDSS assessment.
c. After the end of randomized treatment, patients were eligible for participation in a 1-arm extension study.
d. Patients who participated in the extension study were also followed up to 37 days after treatment end; patients who discontinued treatment early were followed up until Week 108.

AE: adverse event; EDSS: Expanded Disability Status Scale; Gd: gadolinium; IFN: interferon; N: number of randomized patients; RCT: randomized controlled trial; RMS: relapsing multiple sclerosis

Table 7: Characteristics of the intervention – RCT, direct comparison: ponesimod vs. teriflunomide

Study	Intervention	Comparison
OPTIMUM	Ponesimod, orally, once daily Titration phase: ▪ Days 1 and 2: 2 mg ▪ Days 3 and 4: 3 mg ▪ Days 5 and 6: 4 mg ▪ Day 7: 5 mg ▪ Day 8: 6 mg ▪ Day 9: 7 mg ▪ Day 10: 8 mg ▪ Day 11: 9 mg ▪ Days 12 through 14: 10 mg + placebo in each case Maintenance phase (from Day 15): 20 mg	teriflunomide, orally, once daily Days 1 through 14: 14 mg + placebo Maintenance phase (from Day 15): 14 mg
	Disallowed prior and concomitant treatment ▪ ≤ 7 days prior to study start: interferons, glatiramer acetate ▪ ≤ 15 days before study start: beta blockers, diltiazem, verapamil, digoxin, or other antiarrhythmics, systemic therapies for lowering the heart rate, colestyramine[a], or activated charcoal[a] ▪ ≤ 30 days prior to study start: adrenocorticotropic hormone, systemic corticosteroids (unless merely for short-term use to treat relapses)[b], dimethyl fumarate, live vaccines ▪ ≤ 90 days prior to study start: plasmapheresis, cytapheresis, immunoglobulins (i.v.), experimental therapies other than biologics ▪ ≤ 180 days prior to study start: azathioprine, methotrexate, cyclophosphamide, natalizumab, other systemic immunosuppressants, experimental biologics which do not have a lymphocyte-depleting effect (e.g. daclizumab) ▪ ≤ 24 months prior to study start: lymphocyte-depleting biologics (e.g. rituximab, ocrelizumab), cladribine ▪ At any time prior to study start: alemtuzumab, mitoxantrone, leflunomide, teriflunomide, fingolimod, ponesimod, other experimental S1P modulators, stem cell transplantation ▪ Other disease-modifying therapies ▪ Radiotherapy of lymphatic tissue **Permitted concomitant treatment:** ▪ Dalfampridine if ≥ 90 days prior to randomization at a constant dosage	

a. Permitted for accelerated drug elimination, where necessary.
b. Methylprednisolone 1 g/day, i.v. for 3–5days for the treatment of relapse and prednisone equivalent ≤ 10 mg for short-term treatment (≤ 2 weeks/cycle with subsequent pause for ≥ 8 weeks) permitted.

i.v.: intravenous; RCT: randomized controlled trial; S1P: sphingosine-1-phosphate

Study design

The OPTIMUM study is a randomized, double-blind study comparing ponesimod with teriflunomide. It included adult patients with active RMS and an EDSS score of 0 to 5.5. Active disease was defined as the occurrence of

- ≥ 1 relapse within 12 months to 1 month
- ≥ 1 relapse within 24 months to 1 month or
- ≥ 1 gadolinium (Gd)-enhancing lesion within 6 months

each prior to the first EDSS assessment.

These patients had either no prior treatment with disease-modifying therapies or had received prior treatment with interferons, glatiramer acetate, natalizumab, or dimethyl fumarate.

The study randomized a total of 1133 patients at a 1:1 ratio to treatment with either ponesimod (N = 567) or teriflunomide (N = 566).

Ponesimod or teriflunomide treatment was administered in accordance with the SPC [12,13] for a period of 108 weeks. After the end of the blinded treatment phase, patients were eligible for inclusion in a 1-arm extension study on ponesimod treatment.

The primary outcome of the study was annualized relapse rate. Patient-relevant secondary outcomes were from the morbidity, health-related quality of life, and side effects categories.

Several characteristics of the OPTIMUM study are relevant for this benefit assessment, particularly the questionable suitability of the overall population, the high number of major protocol violations, and the change in the recording of the primary outcome (relapse) over the course of the study. These topics are discussed in detail below.

Suitability of the OPTIMUM study's total population for this benefit assessment

Some of the patients included in the OPTIMUM study had not received any prior treatment, while others had been pretreated with interferons, glatiramer acetate, or dimethyl fumarate. The research question of the present benefit assessment covers patients without prior disease-modifying therapy as well as pretreated adults whose disease is not highly active (see Section 2.2). However, pretreated OPTIMUM participants additionally included patients whose disease was highly active despite disease-modifying therapy. These patients are irrelevant for this benefit assessment.

The study protocol provided for subgroup analyses for the OPTIMUM study's outcome of annualized relapse rate based on the presence of highly active disease (yes/no).

However, the definition used for this purpose is unsuitable for differentiating the population to be investigated under the present research question; this is because any prior treatment had to

have been insufficient and inappropriate for assuming highly active disease. The definition of highly active disease to be used for the subgroup analysis was therefore too broad.

Therefore, it is appropriate for the company to use a different definition of highly active disease in the dossier and, on this basis, to describe a subpopulation of the OPTIMUM study to distinguish patients with highly active disease despite treatment with disease-modifying therapy. For this population, the company assumes prior treatment to have been adequate if patients had received disease-modifying therapy for at least 6 months of the year prior to enrolment. The company defined high disease activity in this population as follows:

- ≥ 1 relapse in the year prior to enrolment during or immediately following adequate prior treatment (maximum time after end of prior treatment: 2 months) and
- ≥ 1 Gd-enhancing T1 lesion in the baseline magnetic resonance imaging despite adequate prior treatment.

According to the company's Module 4 A, 33 patients (6%) in the ponesimod arm and 45 patients (8%) in the teriflunomide arm met these criteria. The company did not submit any analyses of the present research question's target population excluding these patients. Because the percentage of patients with highly active disease despite treatment with disease-modifying therapy is small, the company's dossier nevertheless used the OPTIMUM study's total population.

The company's approach for differentiating the subpopulation with highly active disease despite adequate prior treatment with disease-modifying therapy from the patients in the present research question is generally plausible. However, highly active disease can be associated with an increased relapse frequency and, consequently, faster disability progression. In addition, a deviating ACT excluding teriflunomide was defined for patients with highly active disease despite adequate prior treatment with a disease-modifying therapy (see dossier assessment A21-83 [3]). Although at 7%, this patient group makes up only a small percentage of the study population, whether the results for the OPTIMUM study's total population can be fully extrapolated to the target population of treatment-naive patients and pretreated patients whose disease is not highly active remains unclear. This issue has been taken into account in the assessment of the certainty of conclusions (see Section 2.4.2).

Study conduct

Protocol violations

Module 4 A of the company's dossier states that all patients included in the study had at least 1 protocol deviation. Study documents show that 47% of patients had at least 1 major protocol deviation. Module 4 A of the company's dossier presents a list of major deviations and their frequencies. They include deviations related to blinding or outcome recording, e.g. missed safety assessments or EDSS assessments of relapse events which departed from the protocol. Study documents further show that the group of all protocol deviations, irrespective of their

classification as major, includes deviations relating to the survey of patient-relevant outcomes for the present benefit assessment which were not described by the company in Module 4 A. For instance, about 16% of patients received no first-dose monitoring after administration of the first dose of the study drug or after reinitiation of the study drug despite the fact that monitoring was necessary. According to the ponesimod SPC, cardiovascular monitoring may be necessary though, e.g. for bradycardia [12].

In Module 4 A, the company reports that sensitivity analyses were carried out regarding major protocol deviations related to benefit outcomes and that these analyses produced results consistent with those of the primary analyses. However, the company failed to submit an evaluation of these analyses' results assessing the effects of each of these deviations on the results of patient-relevant outcomes. For the most part, both the protocol deviations rated as major and other deviations are balanced equally between study arms, and each of the individual reasons for deviation typically occurred in only a few patients. Overall, however, it remains unclear whether the deviations affect the OPTIMUM study's results. They conceivably might, particularly regarding outcomes with few total events, such as bradycardia (see Table 12 in Section 2.4.3). This issue has been taken into account in the assessment of the risk of bias of results (see Section 2.4.2).

Survey of relapses

Upon suggestion by the United States regulatory authority (Food and Drug Administration, FDA), the procedure for diagnosing and confirming relapses was substantially changed in the course of the OPTIMUM study through protocol version 4 dated 5 February 2016. Protocol version 4 left in place the general approach for relapse diagnosis and evaluation, with a patient history being taken by the treating neurologist and the EDSS assessment being performed by a blinded outcome-recording person. However, it introduced a series of changes which markedly increased the precision of the relapse documentation processes. For instance, between regular visits, patients were surveyed in structured phone interviews to determine whether they had developed symptoms suggesting relapse. Most importantly, responsibilities and communication channels were more precisely defined and standardized when compared to earlier protocol versions. In particular, these changes virtually ruled out any mutual influencing between treating neurologists and blinded outcome recorders who carried out the EDSS assessment. Due to these adjustments, relapse diagnosis and evaluation are expected to have been more reliable under protocol version 4 and later.

Module 4 A of the company's dossier presents subgroup analyses broken down by protocol version, with patients randomized under versions 1 through 3 being analysed separately from those randomized under version 4 and later. For the outcome of confirmed relapses, an effect modification by protocol version was found. However, the observed effect for the study's total population is dominated by the effect found in patients who were randomized after protocol version 4 entered into force (see Table 13 in Section 2.4.3 for the results on the total population and Table 25 in Appendix B of the full dossier assessment for results on the subgroup analyses). For the present benefit assessment, it was therefore assumed that conclusions on the added

benefit of ponesimod in comparison with teriflunomide regarding the outcome of confirmed relapses can be drawn on the basis of the overall population.

Patient characteristics

Table 8 characterizes the patients in the included study.

Table 8: Characterization of the study population – RCT, direct comparison: ponesimod vs. teriflunomide (multipage table)

Study **Characteristic** **Category**	**Ponesimod** **N[a] = 567**	**Teriflunomide** **N[a] = 566**
OPTIMUM		
Age [years], mean (SD)	37 (9)	37 (9)
Sex [f/m], %	64/36	66/35
Ancestry, n (%)		
White	551 (97)	553 (98)
Black	3 (1)	2 (< 1)
Native American or Alaska Native	0 (0)	1 (< 1)
Other	5 (1)	2 (< 1)
Unknown	8 (1)	8 (1)
Region, n (%)		
EU + UK	289 (51)	284 (50)
Non-EU Europe + Russia	233 (41)	239 (42)
North America	32 (6)	24 (4)
Other	13 (2)	19 (3)
EDSS at baseline, median [Q1; Q3]	2.5 [1.5; 3.5]	2.5 [1.5; 3.5]
Gd-enhancing T1-lesions, n (%)		
Yes	226 (40)	256 (45)
No	341 (60)	308 (55)
Number of relapses in the year prior to enrolment, n (%)		
0	20 (4)	28 (5)
1	416 (73)	390 (69)
2	105 (19)	123 (22)
3	22 (4)	19 (3)
≥ 3	4 (1)	5 (1)
Number of relapses in the 2 years prior to enrolment, n (%)		
0	6 (1)	9 (2)
1	277 (49)	270 (48)
2	205 (36)	197 (35)
3	57 (10)	61 (11)
≥ 3	22 (4)	28 (5)
Time from first MS symptoms to randomization [years], mean (SD)	7.6 (6.8)	7.7 (6.8)
Time from initial diagnosis to randomization [years], mean (SD)	4.3 (5.2)	4.8 (5.6)
MS subtype, n (%)		
RRMS	552 (97)	552 (98)
SPMS	15 (3)	14 (2)

Table 8: Characterization of the study population – RCT, direct comparison: ponesimod vs. teriflunomide (multipage table)

Study Characteristic Category	Ponesimod N[a] = 567	Teriflunomide N[a] = 566
Prior treatment with disease-modifying therapy[b], n (%)		
Yes	243 (43)	245 (43)
No	324 (57)	321 (57)
Treatment discontinuation, n (%)	94 (17)	93 (16)
Study discontinuation, n (%)	77 (14)	71 (13)

a. Number of randomized patients. Values which are based on different patient numbers are marked in the corresponding line, provided the deviation is relevant.
b. Any prior treatment before randomization; in the 2 years prior to randomization, 213 patients (38%) in the ponesimod arm versus 211 patients (37%) in the teriflunomide arm were treated with disease-modifying therapies.

EDSS: Expanded Disability Status Scale; f: female; Gd: gadolinium; m: male; MS: multiple sclerosis; n: number of patients in the category; N: number of randomized patients; Q1: 1st quartile; Q3: 3rd quartile; RCT: randomized controlled trial; RRMS: relapsing-remitting multiple sclerosis; SD: standard deviation; SPMS: secondary progressive multiple sclerosis

The distribution of patient characteristics is balanced between groups. Patients were on average 37 years of age, almost exclusively of white ancestry and from Europe, with half of them being from EU countries. There were nearly twice as many women as men.

At a median EDSS score of 2.5, most patients exhibited no severe physical impairment at baseline. Slightly more than 2 thirds of the population had 1 relapse in the year prior to enrolment, about 20% had 2 relapses, and about 4% had 3 relapses. Within 2 years prior to study start, in contrast, nearly 50% of patients had 1 relapse, slightly more than 1 third had 2 relapses, and about 10% had 3 relapses. On average, patients had been diagnosed with multiple sclerosis about 4.6 years prior to randomization, and the first symptoms of disease had developed more than 7 years prior. The course of disease was almost exclusively relapsing-remitting MS (RRMS), while secondary progressive MS (SPMS) was found only in isolated cases. Prior to study inclusion, 43% of patients had received prior disease-modifying therapies.

About 17% of patients discontinued the study drug before the end of the planned treatment duration, and 14% discontinued participation in the study altogether.

Transferability to the German health care context

The company explains that it compared the OPTIMUM study's patient characteristics of sex, age, and disease severity with MS registry data available for the German healthcare system [14]. In the company's view, the comparison of study and registry data shows that in terms of these demographic characteristics, results from the study population at baseline can be extrapolated. The average age in the OPTIMUM study is reportedly slightly lower than the age found in the cited registry, albeit with the OPTIMUM study limiting participant age to a

maximum of 55 years. In addition, the majority of study participants was reportedly recruited in European centres. Therefore, the company does not see any evidence of significant deviations of the patient populations from the German healthcare context.

The company has not provided any further information on the transferability of the study results to the German health care context.

Risk of bias across outcomes (study level)

Table 9 shows the risk of bias across outcomes (risk of bias at study level).

Table 9: Risk of bias across outcomes (study level) – RCT, direct comparison: ponesimod vs. teriflunomide

Study	Adequate random sequence generation	Group allocation concealment	Blinding		Nonselective reporting	Absence of additional aspects	Risk of bias at study level
			Patients	Treatment providers			
OPTIMUM	Yes	Yes	Yes	Yes	Yes	No[a]	High

a. High number of protocol violations which affect, among others, the survey of patient-relevant outcomes (e.g. deviations in EDSS evaluations, failure to conduct necessary monitoring).

EDSS: Expanded Disability Status Scale; RCT: randomized controlled trial

The risk of bias on the study level was rated as high for the OPTIMUM study. This is due to the study's high number of protocol violations, which also particularly concern the survey of patient-relevant outcomes (for a detailed explanation, see “Study conduct” in the section above).

2.4 Results on added benefit

2.4.1 Outcomes included

The following patient-relevant outcomes were to be considered in the assessment:

- Mortality
  - all-cause mortality
- Morbidity
  - confirmed relapses (EDSS-based, operationalized through the annualized relapse rate)
  - confirmed disability progression (EDSS-based, confirmed over a 24-month period)
  - disability severity (surveyed using the MSFC)

  - fatigue (recorded using the Fatigue Symptoms and Impacts Questionnaire – Relapsing Multiple Sclerosis [FSIQ-RMS] or PGI-S)
- Health-related quality of life
  - measured using the SF-36v2
- Side effects
  - SAEs
  - discontinuation due to AEs
  - bradycardia (Preferred Term [PT], AE)
  - infections and infestations (SOC, SAE)
  - further specific AEs, if any

The choice of patient-relevant outcomes deviates from that made by the company, which used further outcomes in the dossier (Module 4 A).

Table 10 shows the outcomes for which data were available in the included study.

Table 10: Matrix of outcomes – RCT, direct comparison: ponesimod vs. teriflunomide

Study	Outcomes											
					Fatigue							
	All-cause mortality[a]	Confirmed relapses (EDSS-based)[b]	Confirmed disability progression[c] (EDSS)	Disability severity (MSFC)[d]	FSIQ-RMS	PGI-S	Health-related quality of life (SF-36v2)	SAEs	Discontinuation due to AEs	Bradycardia (PT, AE)	Infections and infestations (SOC, SAE)	Alopecia (PT, AE)
OPTIMUM	Yes	Yes	Yes	Yes	No[e]	Yes	Yes	Yes	Yes	Yes	Yes	Yes

a. The results on all-cause mortality are based on the information on fatal AEs.
b. Operationalized through the annualized relapse rate; defined as an increase by ≥ 0.5 points (or ≥ 1.0 points if prior EDSS = 0), or an increase by ≥ 1.0 points in at least 2 functional systems, or an increase by ≥ 2.0 points in at least 1 functional system (excluding bowel/bladder and cerebral nervous system) after prior evaluation as being clinically stable and provided that the increase is consistent with the patient's symptoms.
c. Defined as an increase by at least 1.5 points in EDSS score in patients with an EDSS score of 0.0 at baseline; an increase by at least 1.0 point in patients with an EDSS score of 1.0 to 5.0 at baseline; or an increase of at least 0.5 points in patients with an EDSS score ≥ 5.5 at baseline; confirmed over a 24-week period.
d. The validated version of the instrument comprises T25-FW (walking ability), 9-HPT (coordination), and PASAT-3 (cognition).
e. The results from the FSIQ-RMS questionnaire are unusable because the study suffered from problems with the survey of the questionnaire, leading to a high number of missing values (see Section 2.4.1 for a discussion).

9-HPT: 9-Hole Peg Test; AE: adverse event; EDSS: Expanded Disability Status Scale; FSIQ-RMS: Fatigue Symptoms and Impacts Questionnaire – Relapsing Multiple Sclerosis; MSFC: Multiple Sclerosis Functional Composite; PASAT-3: Paced Auditory Serial Addition Test-3; PGI-S: Patient Global Impression of Severity; PT: Preferred Term; RCT: randomized controlled trial; SF-36v2: Short Form-36 Health Survey Version 2; SAE: serious adverse event; SOC: system organ class; T25-FW: Timed 25-Foot Walk

Disability severity (surveyed using MSFC): Symbol Digit Modalities Test

According to the manual [15], the MSFC z score is calculated from the results of the Timed 25-Foot Walk (T25-FW) test for walking ability, the 9-Hole Peg Test (9-HPT) for coordination, and the Paced Serial Addition Test-3 (PASAT-3) for cognition. In addition to the MSFC, the company presents results on the Symbol Digit Modalities Test (SDMT). The SDMT is a test for measuring attention and cognitive processing speed in patients with MS; it is occasionally used to replace PASAT-3 as part of the MSFC. The OPTIMUM study surveyed both SDMT and PASAT-3, with PASAT-3 also being included as a component in the MSFC analyses. The present benefit assessment uses these MSFC analyses to depict disability severity. Therefore, in this benefit assessment, cognitive impairment has already been taken into account via the inclusion of PASAT-3 as a component of the MSFC. Consequently, the SDMT results were disregarded in the present benefit assessment.

Fatigue (surveyed using FSIQ-RMS or PGI-S)

FSIQ-RMS is a questionnaire developed to measure fatigue-related symptoms as well as their effects on the daily lives of patients with RMS. The instrument consists of 4 subscales, the results of which the company presented in Module 4 A of its dossier and used for deriving added benefit. The symptom scale records the severity of fatigue-related symptoms, while the other 3 subscales (referred to as impact subdomains in Module 4 A of the company's dossier) reflect different aspects of the impairment of activities of daily living [16]. FSIQ-RMS analyses are generally relevant for the present benefit assessment. However, the FSIQ-RMS results presented by the company for the present benefit assessment are unusable due to problems which occurred in the OPTIMUM study regarding the surveying of the questionnaire. This is further explained below.

In the study, patients had to complete the questionnaire for a 7-day period per survey time point. In the first 6 days of this survey, only the questions on the symptom scale were asked. On the 7th day of the survey, the other 3 subscales on impairment by fatigue-related symptoms were additionally recorded. For the symptom scale, analyses in the form of a weekly symptom score require data to be available for at least 4 out of 7 days. For the impairment subscales, in contrast, only 1 data point is available.

Already at baseline, the results for all scales exhibit a high number of missing values (symptom scale: about 17% of patients; impairment subscales: about 36% of patients each). The number of missing values further increased over the course of the study to about 32% of patients for the symptom scale and 43% for the other impairment subscales by Week 108. The markedly differing return rates for the various questionnaire subscales suggest a problem in the study with regard to the surveying of the questionnaire over a time period of 7 days because, unlike the symptoms scale, the impairment subscales were surveyed only on the 7th day and return rates were already substantially lower for those scales at baseline. This idea is supported by information from the FDA's assessment report indicating that already at baseline, return rates decreased markedly across the 7 days of the symptom scales survey, from about 96% on Day 1 to 83% on Day 4 and 58% on Day 7 [11]. This suggests that the high number of missing values for the FSIQ-RMS is due to some patients not fully completing the questionnaire survey for 7 consecutive days. It is conceivable that this might particularly affect patients with marked fatigue. Consequently, this patient group would need to be disregarded in analyses involving comparisons with baseline. For the present benefit assessment, the company's FSIQ-RMS analyses are therefore unsuitable. As an alternative, analyses on the outcome of fatigue surveyed with PGI-S were used.

Activity impairment due to MS (Work Productivity and Activity Impairment Questionnaire: Multiple Sclerosis [WPAI:MS])

The WPAI:MS measures the MS-related impairment of work productivity and activities outside of work. Module 4 A of the company's dossier presents analyses on a single question of the WPAI:MS to survey impairment of daily activity due to MS. However, this solitary question

does not reflect the concept of impairment of activities of daily life any better than the survey of health-related quality of life based on SF-35v2, which comprises a physical and a mental component. Therefore, the company's analysis of impairment of daily activity on the basis of the WPAI:MS was disregarded.

Suicidal ideation and behaviour based on the Columbia Suicide Severity Rating Scale

For its benefit assessment, the company used results from the Columbia Suicide Severity Rating Scale (C-SSRS) as an outcome in the side effects category, arguing that patients with MS frequently suffer from depression as a comorbidity of MS. However, the company did not submit any sources showing the C-SSRS to be validated for use in patients with MS. Therefore, the C-SSRS was disregarded in the present benefit assessment.

2.4.2 Risk of bias

Table 11 describes the risk of bias for the results of the relevant outcomes.

Table 11: Risk of bias across outcomes and outcome-specific risk of bias – RCT, direct comparison: ponesimod vs. teriflunomide

Study		Outcomes												
						Fatigue								
	Study level	All-cause mortality[a]	Confirmed relapses (EDSS-based)[b]	Confirmed disability progression[c] (EDSS)	Disability severity (MSFC)[d]	FSIQ-RMS	PGI-S	Health-related quality of life (SF-36v2)	SAEs	Discontinuation due to AEs	Bradycardia (PT, AE)	Infections and infestations (SOC, SAE)	Alopecia (PT, AE)
OPTIMUM	H	H[e]	H[e]	H[e]	H[e]	–[f]	H[e]	H[e, g]	H[e]	H[e]	H[e]	H[e]	H[e]

a. The results on all-cause mortality are based on the information on fatal AEs.
b. Operationalized through annualized relapse rate; defined as an increase by ≥ 0.5 points (or ≥ 1.0 points if prior EDSS = 0), or an increase by ≥ 1.0 points in at least 2 functional systems, or an increase by ≥ 2.0 points in at least 1 functional system (excluding bowel/bladder and cerebral nervous system) after prior evaluation as clinically stable and provided that the increase is consistent with the patient's symptoms.
c. Defined as an EDSS increase by at least 1.5 points in patients with a baseline EDSS score of 0.0; an increase by at least 1.0 point in patients with a baseline EDSS score of 1.0 to 5.0; or an increase by at least 0.5 points in patients with a baseline EDSS score ≥ 5.5; confirmed over a 24-week period.
d. The validated version of the instrument comprises T25-FW (walking ability), 9-HPT (coordination), and PASAT-3 (cognition).
e. High risk of bias across outcomes due to high number of protocol violations affecting, among others, the survey of patient-relevant outcomes (e.g. deviations in EDSS evaluations, necessary monitoring which was foregone; see Section 2.3.2 for a discussion).
f. The results of the FSIQ-RMS questionnaire are unusable because the study exhibited problems with the surveying of the questionnaire which led to a high number of missing values (see Section 2.4.1 for a discussion).
g. High percentage of patients with missing values (about 10% at baseline and > 20% by study end).

9-HPT: 9-Hole Peg Test; AE: adverse event; EDSS: Expanded Disability Status Scale; FSIQ-RMS: Fatigue Symptoms and Impacts Questionnaire – Relapsing Multiple Sclerosis; H: high; L: low; PASAT-3: Paced Auditory Serial Addition Test-3; PGI-S: Patient Global Impression of Severity; PT: Preferred Term; RCT: randomized controlled trial; SF-36v2: Short Form-36 Health Survey Version 2; SAE: serious adverse event; SOC: system organ class; T25-FW: Timed 25-Foot Walk

Due to the high number of protocol deviations which also affect the survey of patient-relevant outcomes, the risk of bias across outcomes is high for the OPTIMUM study (for a detailed discussion, see Section 2.3.2). This also leads to a high risk of bias for the results of all individual outcomes surveyed in the study.

The outcome of health-related quality of life, surveyed using SF-36v2, additionally exhibited a high percentage of missing values (approx. 10% at baseline, > 20% by study end), which further contributes to the high risk of bias of results for this outcome.

Overall assessment of the certainty of conclusions

In addition to patients covered by the present research question, the OPTIMUM study includes a subpopulation of patients with highly active disease despite adequate prior treatment with disease-modifying therapy. While at 7%, this patient group represents a small percentage of the study population, it remains unclear whether the results for the total population of the OPTIMUM study can be fully extrapolated to the target population of treatment-naive or pretreated patients without highly active disease (for a detailed discussion, see Section 2.3.2). In addition, the risk of bias across outcomes for the OPTIMUM study is deemed high due to a large number of protocol deviations because it remains unclear whether the deviations affect the results of the OPTIMUM study (for a detailed discussion, see Section 2.3.2). Overall, the certainty of conclusions of the study results for the present research question is therefore reduced. Based on the OPTIMUM study, at most hints, e.g. of an added benefit, can be derived for all presented outcomes.

2.4.3 Results

Table 12, Table 13, Table 14, and Table 15 summarize the results of the comparison of ponesimod versus teriflunomide in adult patients with active RMS without any prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active. Where necessary, IQWiG calculations are provided in addition to the data from the company's dossier.

Kaplan-Meier curves on the presented time-to-event analyses can be found in Appendix C of the full dossier assessment. The results on common AEs, SAEs, and discontinuations due to AEs are presented in Appendix E of the full dossier assessment.

Table 12: Results (mortality, health-related quality of life, side effects, dichotomous) – RCT, direct comparison: ponesimod vs. teriflunomide

Study Outcome category Outcome	Ponesimod		Teriflunomide		Ponesimod vs. teriflunomide
	N	Patients with event n (%)	N	Patients with event n (%)	RR [95% CI]; p-value[a]
OPTIMUM					
Mortality					
All-cause mortality[b]	565	0 (0)	566	2 (0.4)	0.20 [0.01; 4.16]; 0.212
Health-related quality of life					
SF-36v2 PCS[c,d]					
Improvement[e]	567	58 (10.2)	566	66 (11.7)	0.88 [0.63; 1.22]; 0.533
Deterioration[f]	567	65 (11.5)	566	103 (18.2)	0.63 [0.47; 0.84]; 0.001
SF-36v2 MCS[c,d]					
Improvement[e]	567	116 (20.5)	566	122 (21.6)	0.95 [0.76; 1.19]; 0.683
Deterioration[f]	567	132 (23.3)	566	133 (23.5)	0.99 [0.80; 1.22]; 0.957
Side effects					
AEs (presented as supplementary information)	565	502 (88.8)	566	499 (88.2)	–
SAEs	565	49 (8.7)	566	46 (8.1)	1.07 [0.73; 1.57]; 0.821
Discontinuation due to AEs	565	49 (8.7)	566	34 (6.0)	1.44 [0.95; 2.20]; 0.097
Bradycardia (PT, AEs)	565	4 (0.7)	566	0 (0)	–[g, h]; 0.046
Infections and infestations (SOC, SAEs)	565	7 (1.2)	566	4 (0.7)	1.75 [0.52; 5.96][g]; 0.530
Alopecia (PT, AEs)	565	18 (3.2)	566	72 (12.7)	0.25 [0.15; 0.41]; < 0.001

a. IQWiG calculation, unconditional exact test (CSZ method according to [17]).
b. The results on all-cause mortality are based on the data on fatal AEs.
c. No data are available on the SF-36v2 subscales.
d. About 10% of values missing at baseline, likely counted as patients without event; the company's dossier does not provide any details on this topic, including on missing values over the course of the study.
e. Clinically relevant improvement is defined as an increase from baseline by ≥ 10.80 points (MCS) or ≥ 10.05 points (PCS) (scale range 2 to 74 points for MCS and 4 to 71 points for PCS; calculated using the 1998 standard sample).
f. Clinically relevant deterioration is defined as a decrease from baseline by ≥ 10.80 points (MCS) or ≥ 10.05 points (PCS) (scale range 2 to 74 points for MCS and 4 to 71 points for PCS; calculated using the 1998 standard sample).
g. IQWiG calculation of RR and CI (asymptotic); if 0 events occurred in 1 of the study arms, the calculation used the correction term of 0.5 in both study arms.
h. Discrepancy between p-value (exact) and CI (asymptotic) due to different calculation methods; effect estimation and CI not presented because not informative.

AE: adverse event; CI: confidence interval; CSZ: convexity, symmetry, z-score; MCS: Mental Component Summary; n: number of patients with (at least one) event; N: number of analysed patients; PCS: Physical Component Summary; PT: Preferred Term; RCT: randomized controlled trial; RR: relative risk; SAE: serious adverse event; SF-36v2: Short Form-36 Health Survey version 2; SOC: System Organ Class

Table 13: Results (morbidity, confirmed relapses) – RCT, direct comparison: ponesimod vs. teriflunomide

Study Outcome category Outcome	Ponesimod			Teriflunomide			Ponesimod vs. teriflunomide
	N	n_E	Annualized relapse rate [95% CI][a]	N	n_E	Annualized relapse rate [95% CI][a]	Rate ratio [95% CI]; p-value[a]
OPTIMUM							
Morbidity							
Confirmed relapses (EDSS-based)[b]							
Annualized relapse rate	567	242	0.20 [0.17; 0.23]	566	344	0.29 [0.25; 0.33]	0.69 [0.57; 0.85]; < 0.001

a. Annualized relapse rate and CI (per treatment arm) as well as rate ratio with CI and p-value (group comparison): negative binomial model adjusted for baseline EDSS (≤ 3.5; > 3.5), treatment with disease-modifying therapy within 2 years prior to randomization (yes; no), number of relapses 1 year prior to randomization (≤ 1; ≥ 2); logarithmic follow-up duration as offset variable.

b. Defined as an increase by ≥ 0.5 points (or ≥ 1.0 points if prior EDSS = 0) or an increase by ≥ 1.0 points in at least 2 functional systems, or an increase by ≥ 2.0 points in at least 1 functional system (excluding bowel/bladder and cerebral nervous system) after prior evaluation as clinically stable and provided that the increase is consistent with the patient's symptoms.

CI: confidence interval; EDSS: Expanded Disability Status Scale; N: number of analysed patients; n: number of patients with (at least 1) event; n_E: number of events; RCT: randomized controlled trial

Table 14: Results (morbidity, time to event) – RCT, direct comparison: ponesimod vs. teriflunomide

Study Outcome category Outcome	Ponesimod N	Ponesimod Median time to event in months [95% CI] Patients with event n (%)	Teriflunomide N	Teriflunomide Median time to event in months [95% CI] Patients with event n (%)	Ponesimod vs. teriflunomide HR [95% CI]; p-value[a]
OPTIMUM					
Morbidity					
Confirmed relapses (EDSS-based)[b]					
Time to first confirmed relapse until study end (supplementary information)	*567*	*NR* *166 (29.3)*	*566*	*NR* *223 (39.4)*	*0.75 [0.61; 0.92]; 0.005*
Confirmed disability progression (EDSS-based)[c]	567	NR 46 (8.1)	566	NR 56 (9.9)	0.84 [0.57; 1.24]; 0.373

a. HR, CI, and p-value: Cox proportional hazards model, likely stratified by baseline EDSS (≤ 3.5; > 3.5), treatment with disease-modifying therapy within 2 years prior to randomization (yes; no), and number of relapses 1 year prior to randomization (≤ 1; ≥ 2). According to the statistical analysis plan (SAP), the latter stratification variable was not part of the model for the outcome of confirmed disability progression. The company did not provide any reasoning for its approach deviating from the SAP. However, this deviation is not expected to relevantly influence study results.

b. Defined as an increase by ≥ 0.5 points (or ≥ 1.0 points if prior EDSS = 0) or an increase by ≥ 1.0 points in at least 2 functional systems, or an increase by ≥ 2.0 points in at least 1 functional system (excluding bowel/bladder and cerebral nervous system) after prior evaluation as clinically stable and provided that the increase is consistent with the patient's symptoms.

c. Defined as an increase by at least 1.5 points in EDSS score in patients with an EDSS score of 0.0 at baseline; an increase by at least 1.0 point in patients with an EDSS score of 1.0 to 5.0 at baseline; or an increase of at least 0.5 points in patients with an EDSS score ≥ 5.5 at baseline; confirmed over a 24-week period.

CI: confidence interval; EDSS: Expanded Disability Status Scale; HR: hazard ratio; n: patients with event; N: number of analysed patients; NR: not reached; RCT: randomized controlled trial; SAP: statistical analysis plan

Table 15: Results (morbidity, continuous) – RCT, direct comparison: ponesimod vs. teriflunomide

Study Outcome category Outcome	Ponesimod			Teriflunomide			Ponesimod vs. teriflunomide
	N[a]	Values at baseline mean (SD)	Change by Week 108 mean (SE)[b]	N[a]	Values at baseline mean (SD)	Change by Week 108 mean (SE)[b]	MD [95% CI]; p-value[b]
OPTIMUM							
Morbidity							
Disability severity							
MSFC z score[c]	471	0.00 (0.72)	0.03 (0.02)	470	0.00 (0.73)	-0.04 (0.02)	0.07 [0.02; 0.12]; 0.006 Hedges' g: 0.18 [0.05; 0.31]
Cognition (PASAT-3 [correct answers])[c]	472	48.14 (10.57)	1.51 (0.27)	472	48.16 (10.83)	0.90 (0.27)	0.61 [-0.13; 1.35]
Coordination (9-HPT [seconds])[d]	474	23.59 (13.11)	-0.15 (0.14)	473	22.90 (6.60)	0.79 (0.14)	-0.94 [-1.34; -0.55]
Walking ability (T25-WT [seconds])[d]	473	5.86 (2.85)	0.35 (0.11)	471	5.87 (2.95)	0.25 (0.11)	0.10 [-0.21; 0.40]
Fatigue							
PGI-S[d, e]	520[f]	3.20 (2.38)	0.33 (0.09)	519[f]	3.25 (2.32)	0.49 (0.09)	-0.15 [-0.35; 0.05]; 0.131

a. Number of patients for whom, based on study documents, results were available for Week 108. It is unclear whether earlier measuring points were also included in the calculation of effect estimators. Baseline values may be based on different patient numbers.
b. Mean and SE (change per treatment arm) as well as MD, CI, and p-value (group differences): MMRM with treatment, visit, treatment × visit and baseline value × visit as fixed effects as well as baseline value, EDSS at study start (≤ 3.5; > 3.5), treatment with disease-modifying therapy within 2 years prior to randomization (yes; no), and number of relapses 1 year prior to randomization (≤ 1; ≥ 2) as covariates. Module 4 A of the company's dossier shows that the number of relapses in the year prior to randomization (≤ 1, ≥ 2) was included in the calculation as a covariate; according to the statistical analysis plan (SAP), it was not part of the model for the outcome of disability severity. The company did not provide any reasoning for its approach deviating from the SAP. However, this deviation is not expected to materially influence the result.
c. Higher (increasing) values indicate improved symptoms; favourable effects (intervention minus control) indicate an advantage for ponesimod.
d. Lower (decreasing) values indicate improved symptoms; unfavourable effects (intervention minus control) indicate an advantage for ponesimod.
e. Mean change over the entire course of the study.
f. Number of patients with baseline value and at least 1 subsequent value.

9-HPT: 9-Hole Peg Test; CI: confidence interval; ITT: intention to treat; MD: mean difference; MMRM: mixed effect model repeated measurement; MSFC: Multiple Sclerosis Functional Composite; N: number of analysed patients; PASAT-3: Paced Serial Addition Test; PGI-S: Patient Global Impression of Severity; RCT: randomized controlled trial; SD: standard deviation; SE: standard error; T25W: Timed 25-Foot Walk

Based on the available information, at most hints, e.g. of an added benefit, can be determined for all outcomes (see Section 2.4.2).

Mortality

All-cause mortality

The results on all-cause mortality are based on data on fatal AEs. There was no statistically significant difference between the treatment groups. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the outcome of all-cause mortality; an added benefit is therefore not proven.

Morbidity

Confirmed relapses (EDSS-based)

Operationalization

In its benefit assessment, the company used results on multiple operationalizations for the outcome of confirmed relapses, including the annualized relapse rate and time to first confirmed relapse. The present assessment operationalizes the outcome using annualized relapse rate up to study end. Time to first confirmed relapse does not allow drawing conclusions regarding the total number of relapses and additionally depends on the annualized relapse rate. In the present benefit assessment, this operationalization is therefore presented only as supplementary information.

Results

A statistically significant difference between treatment groups was found for the outcome of confirmed relapses, operationalized using the annualized relapse rate. There was an effect modification by the characteristic of baseline EDSS score. For patients with an EDSS score ≤ 3.5, this results in a hint of added benefit of ponesimod in comparison with teriflunomide. For patients with an EDSS score > 3.5, this results in no hint of added benefit of ponesimod versus teriflunomide; an added benefit is therefore not proven for these patients (see Section 2.4.4).

Confirmed disability progression (EDSS-based)

No statistically significant difference between treatment groups was found for the outcome of confirmed disability progression. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for this outcome; an added benefit is therefore not proven.

Disability severity (MSFC)

Operationalization

For the outcome of disease severity, surveyed using the MSFC z-score, the company presented analyses on the basis of mean differences over the entire course of the study as well as analyses at Week 108. This benefit assessment uses analyses at Week 108 which reflect disability severity by the end of treatment.

Results

For the outcome of disability severity, recorded using the MSFC z-score, there is a statistically significant difference between treatment groups in favour of ponesimod. However, the 95% CI of Hedges' g was not completely above the irrelevance threshold of 0.20. Therefore, the effect cannot be inferred to be relevant. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the outcome of disability severity; an added benefit is therefore not proven.

Fatigue (PGI-S)

Operationalization

For the outcome of fatigue, surveyed using the PGI-S, the company presented analyses on the basis of mean differences over the entire course of the study as well as analyses by Week 108. This benefit assessment uses analyses performed over the entire course of the study since they also reflect fluctuations over the course of the study.

Results

There was no statistically significant difference between treatment groups for the outcome of fatigue, recorded using the PGI-S. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the outcome of fatigue; an added benefit is therefore not proven.

Health-related quality of life recorded with the SF-36v2

Operationalization

In the OPTIMUM study, health-related quality of life was measured using SF-36v2. Module 4 A of the company's dossier presents responder analyses of both improvement and deterioration from baseline. For patients with RMS without prior disease-modifying therapy or patients with prior disease-modifying therapy whose disease is not highly active, both an improvement and a deterioration of health-related quality of life are generally conceivable. In the OPTIMUM study, nearly equal numbers of patients exhibited improvement versus deterioration over the course of the study. In addition, the baseline values of the majority of participants allow development in either direction (see Appendix C of the full dossier assessment). Given the available data, both operationalizations are therefore taken into account, and the results for the assessment of added benefit are interpreted using the overall picture.

Results

For the SF-36v2 PCS, there was no statistically significant difference between treatment groups on the basis of the analyses of improvement from baseline. For the outcome of deterioration from baseline, a statistically significant difference was found in favour of ponesimod. The researchers found an effect modification which was caused by the characteristic of baseline EDSS score and is consistent with the effect modification in the outcome of annualized relapse rate. For patients with an EDSS score ≤ 3.5, this results in a hint of added benefit of ponesimod in comparison with teriflunomide regarding the SF-36v2 PCS. For patients with an EDSS

score > 3.5, this results in no hint of added benefit of ponesimod versus teriflunomide; an added benefit is therefore not proven for these patients (see Section 2.4.4).

For the SF-36v2 MCS, there was no statistically significant difference between treatment groups to show either improvement or deterioration from baseline. Consequently, there is no hint of added benefit of ponesimod in comparison with teriflunomide for the SF-36v2 MCS; an added benefit is therefore not proven.

Side effects

SAEs

No statistically significant differences between treatment groups were found for the outcome of SAEs. For the outcome of SAEs, there was therefore no hint of greater or lesser harm from ponesimod versus teriflunomide; therefore, there is no proof of greater or lesser harm.

Discontinuation due to AEs

No statistically significant difference was found between treatment groups for the outcome of discontinuation due to AEs. For the outcome of discontinuation due to AEs, there was therefore no hint of greater or lesser harm from ponesimod versus teriflunomide; therefore, there is no proof of greater or lesser harm.

Specific AEs

Bradycardia (PT, AE)

A statistically significant difference between treatment groups to the disadvantage of ponesimod was shown for the outcome of bradycardia. For this outcome, there is therefore a hint of greater harm from ponesimod in comparison with teriflunomide.

Infections and infestations (SOC, SAE)

No statistically significant difference between treatment groups was found for the outcome of infections and infestations. For this outcome, there was therefore no hint of greater or lesser harm from ponesimod versus teriflunomide; consequently, there is no proof of greater or lesser harm.

Alopecia (PT, AE)

A statistically significant difference between treatment groups in favour of ponesimod was shown for the outcome of alopecia. For this outcome, there is therefore a hint of lesser harm from ponesimod in comparison with teriflunomide.

2.4.4 Subgroups and other effect modifiers

The following potential effect modifiers were considered in the present assessment:

- age (< 40 years vs. ≥ 40 years)
- sex (women vs. men)

- EDSS score at baseline (≤ 3.5, > 3.5)

The mentioned characteristics were defined a priori. Subgroup analyses were not prespecified for all outcomes in the OPTIMUM study. In the dossier, the company presented subgroup analyses on all outcomes of the present benefit assessment.

Interaction tests were performed when at least 10 patients per subgroup were included in the analysis. Moreover, for binary data, there had to be at least 10 events in at least 1 subgroup.

Only the results with an effect modification with a statistically significant interaction between treatment and subgroup characteristic (p-value < 0.05) are presented. In addition, subgroup results are only presented if there is a statistically significant and relevant effect in at least 1 subgroup.

Table 16 and Table 17 summarize the subgroup results comparing ponesimod with teriflunomide to which the mentioned criteria apply. Where necessary, IQWiG calculations are provided in addition to the data from the company's dossier.

Table 16: Subgroups (health-related quality of life, dichotomous) – RCT, direct comparison: ponesimod vs. teriflunomide

Study Outcome category Outcome Characteristic Subgroup	Ponesimod N	Ponesimod Patients with event n (%)	Teriflunomide N	Teriflunomide Patients with event n (%)	Ponesimod vs. teriflunomide RR [95% CI]	Ponesimod vs. teriflunomide p-value[a]
OPTIMUM						
Health-related quality of life						
SF-36v2 PCS[b]						
Improvement[c]						
Age						
< 40 years	349	43 (12.3)	342	35 (10.2)	1.20 [0.79; 1.83]	0.530
≥ 40 years	218	15 (6.9)	224	31 (13.8)	0.50 [0.28; 0.90]	0.017
Total					Interaction:	0.017
Deterioration[d]						
Baseline EDSS score[e]						
≤ 3.5	472	48 (10.2)	474	89 (18.8)	0.54 [0.39; 0.75]	< 0.001
> 3.5	95	17 (17.9)	92	14 (15.2)	1.18 [0.62; 2.25]	0.682
Total					Interaction:	0.021

a. IQWiG calculation, unconditional exact test (CSZ method according to [17]).
b. Missing values at study start are likely counted as patients without event; the company's dossier does not provide any specific details on this topic, including on missing values over the course of the study.
c. Clinically relevant improvement is defined as an increase by ≥ 10.05 points from baseline (scale range 4 to 71 points, calculated using the 1998 standard sample).
d. Clinically relevant deterioration is defined as a decrease by ≥ 10.05 points from baseline (scale range 4 to 71 points, calculated using the 1998 standard sample).
e. EDSS scores at baseline as recorded in the eCRF.

9-HPT: 9-Hole Peg Test; CI: confidence interval; CSZ: convexity, symmetry, z-score; EDSS: Expanded Disability Status Scale; MSFC: Multiple Sclerosis Functional Composite; n: number of patients with (at least 1) event; N: number of analysed patients; PASAT-3: Paced Auditory Serial Addition Test-3; PCS: Physical Component Summary; RCT: randomized controlled trial; RR: relative risk; SF-36v2: Short Form-36 Health Survey Version 2; T25-FW: Timed 25-Foot Walk

Table 17: Results (morbidity, confirmed relapses) – RCT, direct comparison: ponesimod vs. teriflunomide

Study Outcome Characteristic Subgroup	Ponesimod			Teriflunomide			Ponesimod vs. teriflunomide
	N	n_E	Annualized relapse rate [95% CI][a]	N	n_E	Annualized relapse rate [95% CI][a]	Rate ratio [95% CI]; p-value[a]
OPTIMUM							
Confirmed relapses (EDSS-based)[b]							
Annualized relapse rate							
Baseline EDSS score[c]							
≤ 3.5	472	157[d]	0.16 [0.13; 0.19]	474	268[d]	0.27 [0.23; 0.32]	0.59 [0.47; 0.74]; < 0.001
> 3.5	95	85[d]	0.47 [0.36; 0.60]	92	76[d]	0.41 [0.32; 0.54]	1.13 [0.78; 1.64]; 0.525
Total						Interaction:	0.009

a. Annualized relapse rate and CI (per treatment arm) as well as rate ratio with CI and p-value (group comparison): negative binomial model adjusted for baseline EDSS (≤ 3.5; > 3.5), treatment with disease-modifying therapy within 2 years prior to randomization (yes; no), number of relapses 1 year prior to randomization (≤ 1; ≥ 2); logarithmic follow-up duration as an offset variable.
b. Defined as an increase by ≥ 0.5 points (or ≥ 1.0 points if prior EDSS = 0) or an increase by ≥ 1.0 points in at least 2 functional systems, or an increase by ≥ 2.0 points in at least 1 functional system (excluding bowel/bladder and cerebral nervous system) after prior evaluation as clinically stable and provided that the increase is consistent with the patient's symptoms.
c. Baseline EDSS scores as recorded in the eCRF.
d. Discrepancy between information in Module 4 A and the study documents; Module 4 A shows 45 versus 84 relapses for the EDSS ≤ 3.5 subgroup and 28 versus 22 relapses for the EDSS > 3.5 subgroup.

CI: confidence interval; EDSS: Expanded Disability Status Scale; N: number of analysed patients; n: number of patients with (at least 1) event; n_E: number of events; RCT: randomized controlled trial

Morbidity

Confirmed relapses (EDSS-based)

For the outcome of confirmed relapses, operationalized through the annualized relapse rate, an effect modification by the characteristic of baseline EDSS score was found. For patients with an EDSS score ≤ 3.5, a statistically significant difference was shown in favour of ponesimod. For patients with an EDSS score ≤ 3.5, this results in a hint of added benefit of ponesimod in comparison with teriflunomide.

For patients with an EDSS score > 3.5 years, in contrast, there was no statistically significant difference between treatment groups. This results in no hint of added benefit; for patients with an EDSS score > 3.5, there is therefore no proof of added benefit of ponesimod versus teriflunomide for this outcome.

Health-related quality of life

SF-36v2 PCS

For the SF 36v2 PCS, the responder analyses on improvement or deterioration from baseline each show effect modifications for various characteristics.

An effect modification by the characteristic of age was found for the analyses of improvement from baseline. For patients ≥ 40 years of age, a statistically significant difference was shown to the disadvantage of ponesimod. For patients < 40 years, in contrast, there was no statistically significant difference between treatment groups.

An effect modification by the characteristic of baseline EDSS score was found for the analyses of deterioration from baseline. For patients with an EDSS score ≤ 3.5, a statistically significant difference was shown in favour of ponesimod. For patients with an EDSS score > 3.5, in contrast, there was no statistically significant difference between treatment groups.

The effect modification by the characteristic of EDSS is consistent with the effect modification observed for this characteristic regarding the outcome of confirmed relapses. Therefore, for the overall conclusion on added benefit, only the effect modification regarding the characteristic of EDSS is taken into account. For patients with an EDSS score ≤ 3.5, this results in a hint of added benefit of ponesimod in comparison with teriflunomide in the SF-36v2 PCS. For patients with an EDSS score > 3.5, this results in no hint of added benefit of ponesimod versus teriflunomide for this outcome; an added benefit is therefore not proven for these patients.

2.5 Probability and extent of added benefit

The probability and extent of added benefit at outcome level are derived below, taking into account the different outcome categories and effect sizes. The methods used for this purpose are explained in the IQWiG General Methods [1].

The approach for deriving an overall conclusion on the added benefit based on the aggregation of conclusions derived at outcome level is a proposal by IQWiG. The G-BA decides on the added benefit.

2.5.1 Assessment of the added benefit at outcome level

The extent of the respective added benefit at outcome level was estimated from the results presented in Section 2.4 (see Table 18).

Determination of the outcome category for outcomes on symptoms and side effects

For the symptoms outcome of confirmed relapses, it cannot be inferred from the dossier whether they are serious/severe or non-serious/non-severe. The classification of this outcome is explained below.

Only in exceptional cases did relapses lead to hospitalization. Module 4 A of the company's dossier shows an analysis of the outcome of time to first relapse leading to hospitalization. Such

relapse events occurred in 0.2% versus 0.9% of patients, compared to 29.3% versus 39.4% for relapses overall. Furthermore, the majority of patients with relapses exhibited no disability progression: disability progression by study end was found in about 9% of patients. The outcome of confirmed relapses was therefore allocated to the outcome category of non-serious/non-severe.

Table 18: Extent of added benefit at outcome level: ponesimod vs. teriflunomide (multipage table)

Outcome category **Outcome** **Effect modifier** **Subgroup**	**Ponesimod vs. teriflunomide** **Median time to event (months) or proportion of events (%) or mean change or annualized rate** **Effect estimation [95% CI];** **p-value** **Probability**[a]	**Derivation of extent**[b]
Mortality		
All-cause mortality	0% vs. 0.4% RR: 0.20 [0.01; 4.16]; p = 0.212	Lesser/added benefit not proven
Morbidity		
Confirmed relapses Baseline EDSS score ≤ 3.5	Annualized rate: 0.16 vs. 0.27 Rate ratio: 0.59 [0.47; 0.74]; p < 0.001 Probability: hint	Outcome category: non-serious/non-severe symptoms / late complications CI_u < 0.80 Added benefit; extent: considerable
> 3.5	Annualized rate: 0.47 vs. 0.41 Rate ratio: 1.13 [0.78; 1.64]; p = 0.525	Lesser/added benefit not proven
Confirmed disability progression	Median: NR vs. NR HR: 0.84 [0.57; 1.24]; p = 0.373	Lesser/added benefit not proven
Disability severity (MSFC z-score)	Change by Week 108: 0.03 vs. -0.04 MD: 0.07 [0.02; 0.12]; p = 0.006 Hedges' g: 0.18 [0.05; 0.31][c]	Lesser/added benefit not proven
Fatigue (PGI-S)	Change over the course of the study: 0.33 vs. 0.49 MD: -0.15 [-0.35; 0.05]; p = 0.131	Lesser/added benefit not proven
Health-related quality of life		
SF-36v2 PCS		
Improvement by ≥ 10.05 points	10.2% vs. 11.7% RR: 0.88 [0.63; 1.22]; p = 0.533	

Table 18: Extent of added benefit at outcome level: ponesimod vs. teriflunomide (multipage table)

Outcome category **Outcome** **Effect modifier** **Subgroup**	**Ponesimod vs. teriflunomide** **Median time to event (months) or proportion of events (%) or mean change or annualized rate** **Effect estimation [95% CI]; p-value** **Probability**[a]	**Derivation of extent**[b]
Deterioration by ≥ 10.05 points Baseline EDSS score ≤ 3.5	10.2% vs. 18.8% RR: 0.54 [0.39; 0.751]; p < 0.001 Probability: hint	▪ EDSS ≤ 3.5: Outcome category of health-related quality of life 0.75 < CI_u < 0.90 added benefit; extent: considerable ▪ EDSS > 3.5; Lesser/added benefit not proven
> 3.5	17.9% vs. 15.2% RR: 1.18 [0.62; 2.25]; p = 0.682	
SF-36v2 MCS		
Improvement by ≥ 10.8 points	20.5% vs. 21.6% RR: 0.95 [0.76; 1.19]; p = 0.683	Lesser/added benefit not proven
Deterioration by ≥ 10.8 points	23.3% vs. 23.5% RR: 0.99 [0.80; 1.22]; p = 0.957	
Side effects		
SAEs	8.7% vs. 8.1% RR: 1.07 [0.73; 1.57]; p = 0.821	Greater/lesser harm not proven
Discontinuation due to AEs	8.7% vs. 6.0% RR: 1.44 [0.95; 2.20]; p = 0.097	Greater/lesser harm not proven
Bradycardia (AEs)	0.7% vs. 0% RR: –[d] p = 0.046 Probability: hint	Outcome category of non-serious/non-severe side effects greater harm; extent: non-quantifiable[e]
Infections and infestations (SAEs)	1.2% vs. 0.7% RR: 1.75 [0.52; 5.96]; p = 0.530	Greater/lesser harm not proven
Alopecia (AEs)	3.2% vs. 12.7% RR: 0.25 [0.15; 0.41]; p < 0.001 Probability: hint	Outcome category of non-serious/non-severe side effects CI_u < 0.80 lesser harm; extent: considerable

Table 18: Extent of added benefit at outcome level: ponesimod vs. teriflunomide (multipage table)

Outcome category **Outcome** **Effect modifier** **Subgroup**	**Ponesimod vs. teriflunomide** **Median time to event (months) or proportion of events (%) or mean change or annualized rate** **Effect estimation [95% CI]; p-value** **Probability**[a]	**Derivation of extent**[b]
a. Probability provided if there is a statistically significant and relevant effect. b. Depending on the outcome category, estimations of effect size are made with different limits based on the upper limit of the confidence interval (CI_u). c. If the CI of Hedges' g is fully outside the irrelevance range [−0.2; 0.2], this is interpreted to be a relevant effect. In other cases, the presence of a relevant effect cannot be inferred. d. Discrepancy between p-value (exact) and CI (asymptotic) due to different calculation methods; effect estimation and CI not presented because not informative. e. Due to the asymptotic calculation, the confidence interval is deemed insufficiently reliable for determining extent in this case; the extent of greater harm is nonquantifiable because of the additionally reduced certainty of conclusions due to a high number of protocol violations. AE: adverse event; CI: confidence interval; CI_u: upper limit of the confidence interval; EDSS: Expanded Disability Status Scale; MCS: Mental Component Summary; MD: mean difference; PCS: Physical Component Summary; RR: relative risk; SAE: serious adverse event; SF-36v2: Short Form-36 Health Survey Version 2; SAE: serious adverse event		

2.5.2 Overall conclusion on added benefit

Overall, both favourable and unfavourable effects of ponesimod in comparison with teriflunomide were found for adult patients with active RMS without prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active. Some of the favourable effects were found only for the subgroup with lesser disease severity (baseline EDSS score ≤ 3.5). Below, favourable and unfavourable effects are therefore weighed separately for patients with an EDSS score ≤ 3.5 versus those with a score > 3.5.

Patients with an EDSS score ≤ 3.5

Table 19 summarizes the results taken into account to derive the overall conclusion on the extent of added benefit for patients with an EDSS score ≤ 3.5.

Table 19: Favourable and unfavourable effects for patients with an EDSS score ≤ 3.5 from the analysis of ponesimod in comparison with teriflunomide

Favourable effects	Unfavourable effects
Non-serious/non-severe symptoms / late complications ▪ Confirmed relapses: hint of added benefit – extent: considerable	–
Health-related quality of life ▪ SF-36v2 PCS: hint of an added benefit, extent: considerable	–
Non-serious/non-severe side effects ▪ Alopecia (AEs): hint of lesser harm – extent: considerable	Non-serious/non-severe side effects ▪ Bradycardia (AEs): hint of greater harm – extent: nonquantifiable
AE: adverse events; EDSS: Expanded Disability Status Scale; PCS: Physical Component Summary; SF-36v2: Short Form-36 Health Survey Version 2	

For patients with an EDSS score ≤ 3.5, exclusively favourable effects of ponesimod versus teriflunomide were found regarding morbidity (confirmed relapses) and health-related quality of life (SF-36v2 PCS), each with the extent of considerable. In the side effects category, both favourable effects (alopecia) and unfavourable effects (bradycardia) of ponesimod were found for individual specific AEs.

Overall, for patients with an EDSS score ≤ 3.5, this results in a hint of considerable added benefit of ponesimod in comparison with teriflunomide.

Patients with an EDSS score > 3.5

Table 20 summarizes the results taken into account to derive an overall conclusion on the extent of added benefit for patients with an EDSS score > 3.5.

Table 20: Favourable and unfavourable effects for patients with an EDSS score > 3.5 from the analysis of ponesimod in comparison with teriflunomide

Favourable effects	Unfavourable effects
Non-serious/non-severe side effects ▪ Alopecia (AEs): hint of lesser harm – extent: considerable	Non-serious/non-severe side effects ▪ Bradycardia (AEs): hint of greater harm – extent: nonquantifiable
AE: adverse events; EDSS: Expanded Disability Status Scale	

For patients with an EDSS score > 3.5, no favourable or unfavourable effects of ponesimod versus teriflunomide were found for morbidity or for health-related quality of life. In the side effects category, both favourable effects (alopecia) and unfavourable effects (bradycardia) of ponesimod were found for individual specific AEs.

In summary, there is no hint of added benefit of ponesimod in comparison with teriflunomide for patients with an EDSS score > 3.5; an added benefit is therefore not proven.

Probability and extent of added benefit – summary

Table 21 summarizes the result of the assessment of added benefit of ponesimod in comparison with the ACT.

Table 21: Ponesimod – probability and extent of added benefit

Therapeutic indication	**ACT[a]**	**Probability and extent of added benefit**
Adult patients with active RMS without prior disease-modifying therapy or adult patients with prior disease-modifying therapy whose disease is not highly active	IFN-β 1a or IFN-β 1b or glatiramer acetate or dimethyl fumarate or **teriflunomide** or ocrelizumab, taking into account approval status	▪ Patients with an EDSS ≤ 3.5: hint of considerable added benefit ▪ Patients with an EDSS > 3.5: added benefit not proven
a. Presented is the ACT specified by the G-BA. In cases where the ACT specified by the G-BA allows the company to choose a comparator therapy from several options, the respective choice of the company is printed in bold. ACT: appropriate comparator therapy; G-BA: Federal Joint Committee; IFN: interferon; RMS: relapsing multiple sclerosis		

The assessment described above deviates from the assessment by the company, which derived an indication of minor added benefit of ponesimod in comparison with teriflunomide on the basis of the results of the OPTIMUM study for all patients regardless of EDSS score.

The approach for the derivation of an overall conclusion on the added benefit is a proposal by IQWiG. The G-BA decides on the added benefit.

References for English extract

Please see full dossier assessment for full reference list.

The reference list contains citations provided by the company in which bibliographical information may be missing.

1. Institute for Quality and Efficiency in Health Care. General Methods; Version 6.0 [online]. 2020 [Accessed: 22.03.2021]. URL: https://www.iqwig.de/methoden/general-methods_version-6-0.pdf.

2. Skipka G, Wieseler B, Kaiser T et al. Methodological approach to determine minor, considerable, and major treatment effects in the early benefit assessment of new drugs. Biom J 2016; 58(1): 43-58. https://dx.doi.org/10.1002/bimj.201300274.

3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Ponesimod (multiple Sklerose) – Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung [online]. 2021 [Accessed: 15.09.2021]. URL: https://www.iqwig.de/download/a21-83_ponesimod_nutzenbewertung-35a-sgb-v_v1-0.pdf.

4. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine vorläufige Aussetzung eines Beschlusses zur Nutzenbewertung nach § 35a Absatz 3 Satz 1 SGB V zu Ponesimod (Schubförmige Multiple Sklerose) [online]. 2021 [Accessed: 03.01.2022]. URL: https://www.g-ba.de/downloads/39-261-5159/2021-12-02_AM-RL-XII_Ponesimod-Pat-a_D-702.pdf.

5. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine vorläufige Aussetzung der Beschlussfassung zur Nutzenbewertung nach § 35a Absatz 3 Satz 1 SGB V zu Ponesimod (Schubförmige Multiple Sklerose) [online]. 2021 [Accessed: 03.01.2022]. URL: https://www.g-ba.de/downloads/40-268-8078/2021-12-02_AM-RL-XII_Ponesimod-Pat-a_D-702_TrG.pdf.

6. Actelion Pharmaceuticals. OPTIMUM: Oral Ponesimod versus Teriflunomide In relapsing MUltiple sclerosis; Multicenter, Randomized, Double-Blind, Parallel-group, Active-Controlled, Superiority Study to Compare the Efficacy and Safety of Ponesimod to Teriflunomide in Subjects With Relapsing Multiple Sclerosis; study AC-058B301; Final Clinical Study Report [unpublished]. 2020.

7. Janssen Cilag. OPTIMUM: Oral Ponesimod versus Teriflunomide In relapsing MUltiple sclerosis; Multicenter, Randomized, Double-Blind, Parallel-group, Active-Controlled, Superiority Study to Compare the Efficacy and Safety of Ponesimod to Teriflunomide in Subjects With Relapsing Multiple Sclerosis; study AC-058B301; Zusatzanalysen [unpublished]. 2021.

8. Actelion. Oral Ponesimod Versus Teriflunomide In Relapsing MUltiple Sclerosis (OPTIMUM) [online]. 2021 [Accessed: 14.12.2021]. URL: https://ClinicalTrials.gov/show/NCT02425644.

9. Actelion Pharmaceuticals. Multicenter, randomized, double-blind, parallel-group, active-controlled, superiority study to compare the efficacy and safety of ponesimod to teriflunomide in subjects with relapsing multiple sclerosis [online]. [Accessed: 14.12.2021]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-000540-10.

10. Kappos L, Fox RJ, Burcklen M et al. Ponesimod Compared With Teriflunomide in Patients With Relapsing Multiple Sclerosis in the Active-Comparator Phase 3 OPTIMUM Study: A Randomized Clinical Trial. JAMA Neurol 2021; 78(5): 558-567. https://dx.doi.org/10.1001/jamaneurol.2021.0405.

11. Center for Drugs Evaluation and Research. Clinical Review(s); Application Number: 213498Orig1s000 [online]. 2021 [Accessed: 27.01.2022]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2021/213498Orig1s000MedR.pdf.

12. Janssen. Ponvory Filmtabletten [online]. 2021 [Accessed: 27.01.2022]. URL: https://www.fachinfo.de/.

13. Sanofi Genzyme. Aubagio 7 mg Filmtabletten, Aubagio 14 mg Filmtabletten [online]. 2021 [Accessed: 03.01.2021]. URL: https://www.fachinfo.de/.

14. Weih M, Rossnagel F, Dikow H et al. Daten zur Multiplen Sklerose in Deutschland und ihre Abbildung im Register des ambulanten Netzwerkes NeuroTransData (NTD). Fortschr Neurol Psychiatr 2020; 88(6): 379-385. https://dx.doi.org/10.1055/a-1130-6222.

15. Fischer JS, Amy JJ, Kniker JE et al. Multiple Sclerosis Functional Composite (MSFC); Administration and Scoring Manual [online]. 2001 [Accessed: 27.01.2022]. URL: https://nmsscdn.azureedge.net/NationalMSSociety/media/MSNationalFiles/Brochures/10-2-3-31-MSFC_Manual_and_Forms.pdf.

16. Hudgens S, Schuler R, Stokes J et al. Development and Validation of the FSIQ-RMS: A New Patient-Reported Questionnaire to Assess Symptoms and Impacts of Fatigue in Relapsing Multiple Sclerosis. Value Health 2019; 22(4): 453-466. https://dx.doi.org/10.1016/j.jval.2018.11.007.

17. Martín Andrés A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. Computat Stat Data Anal 1994; 17(5): 555-574. https://dx.doi.org/10.1016/0167-9473(94)90148-1.

The full report (German version) is published under https://www.iqwig.de/en/projects/a21-159.html.