Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Montalban X, Arnold DL, Weber MS, et al. Placebo-controlled trial of an oral BTK inhibitor in multiple sclerosis. N Engl J Med 2019;380:2406-17. DOI: 10.1056/NEJMoa1901981

(PDF updated June 5, 2019)

SUPPLEMENTARY MATERIAL

Table of Contents

Supplementary Appendix 1. Participating Centers, Investigators and Independent Data Monitoring Committee.

Steering Committee

- Xavier Montalban - Vall d'Hebron University Hospital, Barcelona, Spain; St Michael's Hospital, University of Toronto, Toronto, Canada
- Douglas L. Arnold - Montreal Neurological Institute, Montreal, Canada; NeuroRx Research, Montreal, QC, Canada
- Martin S. Weber - Institute of Neuropathology and Department of Neurology, University Medical Center, Göttingen, Germany
- Jerry S. Wolinsky - McGovern Medical School, University of Texas Health Science Center at Houston (UTHealth), Houston, USA

The Evobrutinib Phase 2 Study Group

Center	Principal Investigator
Bulgaria	
MHAT Tokuda Hospital Sofia, Sofia	Ivan Staikov
UMHAT Sv. Ivan Rilski EAD, Sofia	Penko Shotekov
MHAT - National Heart Hospital, Sofia	Lyubomir Haralanov
MHAT Avis - Medica OOD, Pleven	Hristo Lilovski
Military Medical Academy of Sofia, Sofia	Krasimir Genov
Medical Centre Asklepii OOD, Dupnitsa	Krasimir Kirilov
First MHAT - Sofia EAD, Sofia	Dimitar Maslarov
Military Medical Academy of Pleven, Pleven	Plamen Tzvetanov
NEOCLINIC Medical Center, Sofia	Rosen Ikonomov
MHAT Lyulin EAD, Sofia	Rosen Kalpachki
MHAT Puls AD, Blagoevgrad	Sasho Kastrev

MHAT - Ruse AD, Ruse	Neli Petrova
Fakultni nemocnice u sv. Anny, Brno	Michal Dufek
Nemocnice Jihlava, Jihlava	Radek Ampapa
Faculty hospital Pharmacy, Hradec Kralove	Martin Valis
Lekarna Alma, Hradec Kralove	Jan Mares
Fakultni nemocnice v Motole, Prague	Petra Rockova
Nemocnicni lekarna, Teplice	Radomir Talab
Poland	
Novo-Med Zielinski i wspolnicy, Katowice	Tomasz Zielinski
Neurologiczny NZOZ Centrum, Plewiska	Karolina Piasecka-Stryczynska
Centrum Medyczne NeuroProtect, Warszawa	Maciej Czarnecki
Indywidualna Praktyka Lekarska, Lublin	Konrad Rejdak
Prywatna Praktyka Lekarska, Katowice	Maciej Maciejowski
B and B Robert Bonek, Bydgoszcz	Robert Bonek
Centrum Neurologii K. Selmaj, Lodz	Krzysztof Selmaj
NZOZ NEURO-KARD Ilkowski, Poznan	Jan Ilkowski
Medicome Sp. z o.o., Oswiecim	Magdalena Boczarska-Jedynak
Centrum Medyczne Medyk, Rzeszow	Halina Bartosik-Psujek
Russian Federation	
SAIH Rep. Clin. Neurol. Center, Kazan	Farit Khabirov
Siberian Clin. Center of FMBA, Krasnoyarsk	Dmitry Pokhabov
Republican Clin. Hospital No. 4, Saransk	Nikolay Dorogov
Siberian Clin. Center of FMBA, Novosibirsk	Denis Sazonov
CCH No. 2 of JSC Russian Railway, Moscow	Olga Vorobyeva
LLC Neyrovector, Novosibirsk	Evgeny Pankratov
Perm State MU named after Acad. E. A. Wagner, Perm	Tatiana Trushnikova

Serbia	
Clinical Center of Serbia, Belgrade	Jelena Drulovic
Clinical Center Kragujevac, Kragujevac	Gordana Toncev
General Hospital Uzice, Uzice	Mira Gavric-Kezic
Clinical Center Nis, Nis	Slobodan Vojinovic
Slovakia	
MUDr. Beata Dupejova, neurologicka ambulancia, s.r.o., Banska Bystrica	Beata Dupejova
EMPATHY s.r.o., Bratislava	Sona Feketova
Medicinske centrum KONZILIUM, Dubnica nad Vahom	Magdalena Perichtova
Ukraine	
M.V. Sklifosovskyi Poltava, Poltava	Nataliya Lytvynenko
CI Zaporizhzhia Regional Clinic, Zaporizhzhia	Nataliia Buchakchyiska
CI CCH No. 6 Dept. of General Neurology, Zaporizhzhia	Olexandr Kozyolkin
SI Institute of Neurology, Kharkiv	Nataliya Voloshyna
City Clinical Hospital 4, Kyiv	Larysa Sokolova
MI of PH RCH Center of Emergency, Kharkiv	Olena Tovazhnyanska
Lviv Regional Clinical Hospital, Lviv	Yanosh Sanotskyy
Ivano-Frankivsk Regional C.H., Ivano-Frankivsk	Galyna Chmyr
SMPI Central Clin. Hosp. of Ukraine, Kharkiv	Ihor Pasiura
MMI Chernivtsi Regional P.H., Chernivtsi	Valeriy Pashkovskyy
MI of Kyiv RC Kyiv Region, Kyiv	Larysa Bosenko
Spain	
Hospital del Mar, Barcelona	Jose Enrique Martinez Rodriguez
Complejo Hospitalario Universitario de A Coruna, A Coruna	Ana Lopez Real

Independent Data Monitoring Committee (IDMC)

- David Burr Clifford, Washington University School of Medicine, St. Louis, MO (IDMC Chairperson)
- Bruce Anthony Campbell Cree, UCSF Multiple Sclerosis Center, San Francisco, CA
- Ludwig Achim Kappos, University Hospital Basel, Basel, Switzerland
- Emmanuel Lesaffre, Leuven Biostatistics and Statistical Bioinformatics Centre, Leuven, Belgium (IDMC Biostatistician)

Supplementary Appendix 2. Special Precautions.

Evobrutinib

Trial treatment should be temporarily withheld or permanently withdrawn if the following abnormalities occur or re-occur, as relevant, and re-initiation after temporarily withholding trial medication should be discussed with the Medical Monitor:

- For a neutrophil count <500/mm^3 or platelet count <25,000/mm^3 (Grade 4) or neutrophil count 500–999/mm^3 (Grade 3) with fever or platelet count 25,000–49,999/mm^3 (Grade 3) with bleeding, trial medication should be permanently withdrawn.
    - For a Grade 3 decrease in neutrophil count without fever or Grade 3 decrease in platelet count without bleeding, temporarily hold trial medication and recheck the value. If the value is still Grade 3, permanently discontinue trial medication. For a decrease to Grade 2, temporarily hold the trial medication and recheck the value. Re-initiate trial medication after discussion with the Medical Monitor if no further downward trend is observed.
- For an increase in aspartate aminotransferase (AST) or alanine aminotransferase (ALT) to >3× upper limit of normal (ULN) and increase in bilirubin to >1.5× ULN (Grade 2 or higher), trial medication should be permanently withdrawn. The patient should be followed with additional testing as needed until a return to ULN. Consultations with specialists, such as a hepatologist, can be considered at the discretion of the Investigator and in conjunction with the Medical Monitor.
    - For an increase in AST or ALT to >5× ULN without bilirubin elevation, trial medication should be permanently withdrawn. The patient should be followed with additional testing as needed until a return to ULN. Consultations with specialists, such as a hepatologist, can be considered at the discretion of the Investigator and in conjunction with the Medical Monitor.

    - For any other increase in AST, ALT or bilirubin to Grade 2, temporarily hold trial medication and recheck the value. Re-initiate trial medication after discussion with the Medical Monitor if no further upward trend is observed.
- For an increase in amylase or lipase to >5× ULN (Grade 4), trial medication should be permanently withdrawn.
    - For an increase in amylase or lipase to >2 to 5× ULN (Grade 3), temporarily hold trial medication and recheck the value within 24 hours of receipt. If the value is still Grade 3, permanently discontinue trial medication. For an increase to Grade 2, temporarily hold trial medication and recheck the value within 24 hours of receipt. Discontinue trial medication if the value does not decrease, or re-initiate trial medication after discussion with the Medical Monitor if a downward trend is observed.
- For an increase in serum creatinine to >3× from baseline (Grade 3 or higher), trial medication should be permanently withdrawn.
    - For any other increase in serum creatinine >1.5× from baseline, temporarily hold trial medication and recheck the value within 24 hours of receipt. Discontinue trial medication if the value does not decrease, or re-initiate trial medication after discussion with the Medical Monitor if a downward trend is observed.
- For any other laboratory abnormality of Grade 4 severity, trial medication should be permanently withdrawn.
    - For any other laboratory increase/decrease (as relevant) from baseline to a clinically significant higher severity grade, temporarily hold trial medication and recheck the value within 24 hours of receipt. Discuss restarting trial medication with the Medical Monitor if an improving trend is observed.
    - For an absolute lymphocyte count <200/mm^3 (Grade 4), trial medication should be temporarily withdrawn and follow-up testing should be conducted. When the absolute lymphocyte count returns to 800/mm^3 (i.e. returns to Grade 2), trial medication can be resumed.

Dimethyl fumarate

- For a lymphocyte count <500/mm^3 for >24 weeks, dimethyl fumarate should be temporarily withheld and the patient monitored until lymphocyte counts are back to the lower limit of normal (LLN). Once lymphocyte counts are back to LLN, trial medication can be restarted with additional follow-up monitoring of lymphocyte counts.
    - For an absolute lymphocyte count <200/mm^3 (Grade 4), dimethyl fumarate should be permanently withdrawn and the lymphocyte count of the patient monitored.
- For a serious infection, after discussion with the Medical Monitor, consideration should be given to temporarily withholding dimethyl fumarate until resolution of the infection.
- At the first sign or symptom suggestive of progressive multifocal leukoencephalopathy (PML), dimethyl fumarate should be withheld and an appropriate diagnostic evaluation conducted. Magnetic resonance imaging findings may be apparent before clinical signs or symptoms. Typical symptoms associated with PML are diverse, progress over days to weeks, and include progressive weakness on one side of the body or clumsiness of limbs, disturbance of vision, and changes in thinking, memory, and orientation leading to confusion and personality changes.
- For a flushing reaction (e.g. warmth, redness, itching, and/or burning sensation), dimethyl fumarate should be temporarily withheld until symptoms have resolved. After the flushing reaction has resolved, dimethyl fumarate should be restarted at a reduced dose.
    - Should a flushing reaction occur again, dimethyl fumarate should be permanently discontinued.

Supplementary Appendix 3. Inclusion and Exclusion Criteria.

Inclusion Criteria:

- Subjects with a diagnosis of relapsing multiple sclerosis (MS) - may include subjects with Secondary Progressive Multiple Sclerosis (SPMS) with superimposed relapses provided they meet the other criteria - in accordance with revised McDonald (Polman CH, et al. *Ann Neurol* 2011) criteria for MS and Lublin and Reingold (Lublin FD, et al. *Neurology* 2014).
- Male or female aged 18 to 65 years.
- One or more documented relapses within the 2 years before Screening with either: a) one relapse which occurred within the last year prior to randomization or b) the presence of at least one T1 gadolinium-enhancing lesion within 6 months prior to randomization.
- Expanded Disability Status Scale (EDSS) score of 0 to 6 at baseline.
- Women of childbearing potential had to use a supplementary barrier method together with a highly effective method of contraception (according to International Council for Harmonisation guidance M3[R2]) for 4 weeks prior to randomization, throughout the trial, and for 90 days after the last dose of the investigational medicinal product (IMP).
- Signed and dated informed consent (subject had to be able to understand the informed consent) indicating that the subject was informed of all the pertinent aspects of the trial prior to enrollment and would comply with the requirements of the protocol.

Exclusion Criteria:

- Primary progressive MS and secondary progressive MS without relapses.
- Disease duration >15 years in subjects with EDSS of 2 or less.
- Treatment with rituximab, ocrelizumab, mitoxantrone, or lymphocyte-depleting therapies (e.g. alemtuzumab, anti-CD4, cladribine, cyclophosphamide, total body irradiation, bone marrow transplantation) within 48 weeks prior to randomization.
- Use of lymphocyte trafficking blockers (e.g., natalizumab, fingolimod) within 24 weeks prior to randomization.

- Use of intravenous (IV) immunoglobulins (Ig), plasmapheresis, and immunosuppressive treatments within 4 weeks prior to randomization.
- Treatment with β-interferons or glatiramer acetate within 4 weeks prior to randomization.
- Systemic glucocorticoids within 4 weeks prior to randomization.
- Treatment with teriflunomide within 12 weeks prior to randomization.
- Treatment with daclizumab within 12 weeks prior to randomization.
- Exposure to dimethyl fumarate within 6 months prior to randomization.
- Any allergy, contraindication, or inability to tolerate dimethyl fumarate.
- Treatment with dalfampridine or fampridine unless on a stable dose for ≥30 days prior to randomization.
- Inability to comply with magnetic resonance imaging scanning.
- Immunologic disorder other than MS, with the exception of secondary well-controlled diabetes or thyroid disorder, or any other condition requiring oral, IV, intramuscular, or intra-articular corticosteroid therapy.
- Vaccination with live or live-attenuated virus vaccine within one month prior to Screening.
- Severe drug allergy or history of anaphylaxis, or allergy to the IMP or any of its incipients.
- Active, clinically significant viral, bacterial, or fungal infection, or any major episode of infection requiring hospitalization or treatment with parenteral anti-infectives within 4 weeks of Screening, or completion of oral anti-infectives within 2 weeks before or during Screening, or a history of recurrent infections (i.e., three or more of the same type of infection in a 12-month rolling period). Vaginal candidiasis, onychomycosis, and genital or oral herpes simplex virus considered by the Investigator to be sufficiently controlled would not be exclusionary.
- History of, or positive testing for, human immunodeficiency virus, hepatitis C antibody and/or polymerase chain reaction, hepatitis B surface antigen (+) and/or hepatitis B core total, and/or IgM antibody (+) at Screening.
- The subject:

  - Has a history of or current diagnosis of active tuberculosis (TB) or
  - Is currently undergoing treatment for latent TB infection (LTBI) or
  - Has an untreated LTBI or
  - Has a positive QuantiFERON®-TB test at Screening.
- Indeterminate QuantiFERON®-TB tests may be repeated once and will be considered positive if retest results are positive or indeterminate.
- Subjects with current household contacts with active TB will also be excluded.
- History of splenectomy or any major surgery within 2 months prior to Screening.
- History of myocardial infarction or cerebrovascular event as per the protocol.
- History of attempted suicide within 6 months prior to Screening or a positive response to items 4 or 5 of Columbia-Suicide Severity Rating Scale.
- An episode of major depression within the last 6 months prior to Screening.
- On anticoagulation, fish oil supplements, or antiplatelet therapy other than daily aspirin for cardio-protection and treatment of dimethyl fumarate induced flushing.
- History of cancer, except adequately treated basal cell or squamous cell carcinoma of the skin.
- Breastfeeding/lactating or pregnant women.
- Participation in any investigational drug trial within one month or five half-lives of the investigational drug, whichever is longest, prior to Screening.
- Subjects currently receiving (or unable to stop using prior to receiving the first dose of IMP) medications or herbal supplements known to be potent inhibitors of cytochrome P450 3A (CYP3A).
- History of or current alcohol or substance abuse.
- Clinically significant abnormality on electrocardiogram or screening chest X-ray.
- Clinically significant laboratory abnormality.

Supplementary Appendix 4. Definition of Serious Adverse Events (SAEs).

An adverse event (AE) is any untoward medical occurrence in a subject or clinical investigation subject administered a pharmaceutical product, regardless of causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the medicinal product.

An SAE is any untoward medical occurrence that at any dose:

- Results in death.
- Is life-threatening (Note: This refers to an event in which the subject is at risk of death at the time of the event, not an event that might have caused death if it was more severe).
- Requires inpatient hospitalization or prolongs an existing hospitalization, except in the case of hospitalizations due to protocol-defined relapses.
- Results in persistent or significant disability or incapacity.
- Is a congenital anomaly or birth defect.
- Is otherwise considered to be medically important (Note: Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered as SAEs when, based upon appropriate medical judgment, they may jeopardize the subject or may require medical or surgical intervention to prevent one of the outcomes listed above. Examples of such events include allergic bronchospasm requiring intensive treatment in an emergency room or at home, blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse).

Any suspected transmission of an infectious agent via an investigational product is also considered an SAE.

Any clinical AE with severity of Grade 4 or 5 must be reported as an SAE. However, a laboratory abnormality of Grade 4, such as anemia or neutropenia, is considered serious only if the condition meets one of the criteria above.

Supplementary Appendix 5. Missing Data Handling.

The primary analysis of total gadolinium-enhancing T1 lesions at Weeks 12, 16, 20, 24 was a negative binomial (NB) model, with offset equal to the log number of available scans (from Weeks 12, 16, 20, 24) that gave rise to the total lesion count for a patient, and adjustment for baseline lesion activity. This model made use of all available data, including patients with partial data (i.e., missing data). The assumption underlying this model is that missingness is non-informative for the endpoint of interest. This assumption is partially justified by noting that missing lesion data due to treatment discontinuation during the first 24 weeks occurred at a similar frequency across treatment groups (9.3% for placebo, 9.6% for evobrutinib 25 mg QD, 9.4% for 75 mg QD, 11.1% for 75 mg BID) and was for reasons other than lack of efficacy or progressive disease.

As in other phase 2 studies in relapsing MS (e.g. Miller DH, et al. NEJM 2003; Kappos L, et al. NEJM 2006), in the primary analysis, lesion data were treated as missing if the scan giving rise to the data was collected within 3 weeks of high dose corticosteroid (HDC) use, due to the short-term effects of HDC on gadolinium-enhancing T1 lesions. The frequency of this type of missing data varied across treatment groups: one scan in one patient in the placebo group, one scan each for five patients in the 25mg QD group, one scan in one patient in the 75mg BID group, and all four scans in one patient in the 75mg BID group. However, a sensitivity analysis using all available scans, pre-specified prior to unblinding, demonstrated results nearly identical to those of the primary analysis in which HDC-affected scans were treated as missing, further justifying the assumption that missingness was non-informative for this endpoint.

In general, no imputation was needed for the NB modeling. However, for the one patient in the 75mg BID group, where all four of the scans were affected by HDC use, and treated as missing in the primary analysis, it was necessary to impute the post baseline lesion count. The pre-specified method imputed the post baseline lesion count using median values for scan count and lesions per scan among patients with available post-baseline scans, who were in the same treatment group, and who had the same value for the categorical baseline covariate. The single imputation approach employed affected the data for only one patient and did not lead to an appreciable

underestimate of variability, as evidenced by the nearly identical results for the primary analysis and the sensitivity analysis based on all available scans.

The analysis of ARR at 24 weeks was a NB model, with offset equal to the log years on study that gave rise to the relapse count for a patient, and adjustment for baseline relapse activity. The assumption underlying this model is that missingness is non-informative for the endpoint of interest. This assumption is justified based on study discontinuation during the first 24 weeks occurring at a similar frequency across treatment groups for reasons other than lack of efficacy or progressive disease.

Qualified relapse-free status at 24 weeks was analysed via a logistic model that adjusted for baseline relapse activity. A patient discontinuing study prior to 24 weeks without having a qualified relapse was treated as not being qualified relapse-free at 24 weeks: missingness was assumed to be informative for the endpoint of interest. This assumption was pre-specified in the protocol to support a conservative analysis. Study discontinuation prior to 24 weeks without qualified relapse occurred at a similar frequency across treatment groups (7.5% for placebo, 6.0% for 25mg QD, 5.9% for 75mg QD, 9.4% for 75mg BID).

The analysis of EDSS score change from baseline at 24 weeks was a nonparametric comparison of each evobrutinib dose with placebo, based on Wilcoxon rank sum test and Hodges-Lehman estimate of the shift in distribution of the change from baseline value due to evobrutinib, stratified by baseline EDSS. If the change from baseline value at Week 24 was missing for a patient, that value was imputed by the median value among patients in the same treatment group with the same value for the categorical baseline covariate (i.e., single imputation). This approach assumes missingness is non-informative for the endpoint of interest. Single imputation may lead to an underestimate of variability, but given that no treatment effect was observed, further analysis using multiple imputation was expected to yield similar results.

The analysis of 24 week change from baseline in SF-36 score was based on a Mixed-Effect Model Repeated Measure (MMRM) approach with adjustment for baseline score. The assumption

underlying this model is missing at random (MAR). This assumption may not be justified due to treatment discontinuation for adverse events during the first 24 weeks occurring at different frequencies across treatment groups (7.4% for placebo, 1.9% for evobrutinib 25 mg QD, 3.8% for 75 mg QD, 11.1% for 75 mg BID). Sensitivity analyses of the exploratory SF-36 endpoints exploring the validity of the MAR assumption have not yet been performed.

Supplementary Appendix 6. Additional Secondary Endpoints.

To evaluate the efficacy of evobrutinib compared with placebo:

- Total number of new gadolinium-enhancing T1 lesions at Week 12, 16, 20, and 24
- Mean per-scan number of gadolinium-enhancing T1 lesions at Week 12, 16, 20, and 24
- Change from baseline in the volume of gadolinium-enhancing T1 lesions at Week 24
- Total number of new or enlarging T2 lesions at Week 12, 16, 20, and 24
- Change from baseline in the volume of T2 lesions at Week 24

To evaluate efficacy within evobrutinib dose groups:

- Number of gadolinium-enhancing T1 lesions at Week 48
- Number of new gadolinium-enhancing T1 lesions at Week 48
- Change from baseline in the volume of gadolinium-enhancing T1 lesions at Week 48
- Number of new or enlarging T2 lesions at Week 48
- Change from baseline in the volume of T2 lesions at Week 48
- Change from baseline in Expanded Disability Status Scale (EDSS) at Week 48
- Annualized relapse rate (ARR), based on protocol-defined relapses, at Week 48
- Relapse-free status at Week 48

To evaluate the efficacy and safety of dimethyl fumarate:

- Total number of gadolinium-enhancing T1 lesions at Week 12, 16, 20, and 24
- ARR, based on protocol-defined relapses at Week 24
- Relapse-free status at Week 24
- Change from baseline in EDSS at Week 24
- Safety as assessed by the nature, severity, and occurrence of AEs; vital signs; ECGs; absolute concentrations and change from Baseline in Ig levels; absolute numbers and change from Baseline in B cells; and clinical laboratory safety parameters
- Total number of new gadolinium-enhancing T1 lesions at Week 12, 16, 20, 24

- Mean per-scan number of gadolinium-enhancing T1 lesions at Week 12, 16, 20, and 24
- Change from baseline in the volume of gadolinium-enhancing T1 lesions at Week 24
- Total number of new or enlarging T2 lesions at Week 12, 16, 20, and 24
- Change from baseline in the volume of T2 lesions at Week 24
- Number of gadolinium-enhancing T1 lesions at Week 48
- Number of new gadolinium-enhancing T1 lesions at Week 48
- Change from baseline in the volume of gadolinium-enhancing T1 lesions at Week 48
- Number of new or enlarging T2 lesions at Week 48
- Change from baseline in the volume of T2 lesions at Week 48
- Change from baseline in EDSS at Week 48
- ARR, based on protocol-defined relapses, at Week 48
- Relapse-free status at Week 48

Supplementary Appendix 7. Health-Related Quality of Life Analysis.

Health-related quality of life was measured using the Short Form 36-item Health Status Survey (SF-36v2). Each of the following eight areas of health were rated by the patient from 0–100, giving an overall maximum total score of 800:

- Physical function
- Role limitations due to health problems
- Bodily pain
- Social functioning
- General mental health
- Role limitations due to emotional problems
- Energy/fatigue
- General health perceptions.

Two summary scores, the Physical Component Summary score and Mental Component Summary score were obtained through a linear combination of weighted transformed scores from the eight subscales.

Completion of the SF-36v2 instrument took place at screening, each on-treatment visit (every 4 weeks) and at end of treatment (Week 48). This was done at the start of a patient visit, prior to any other procedure or interaction with the Investigator.

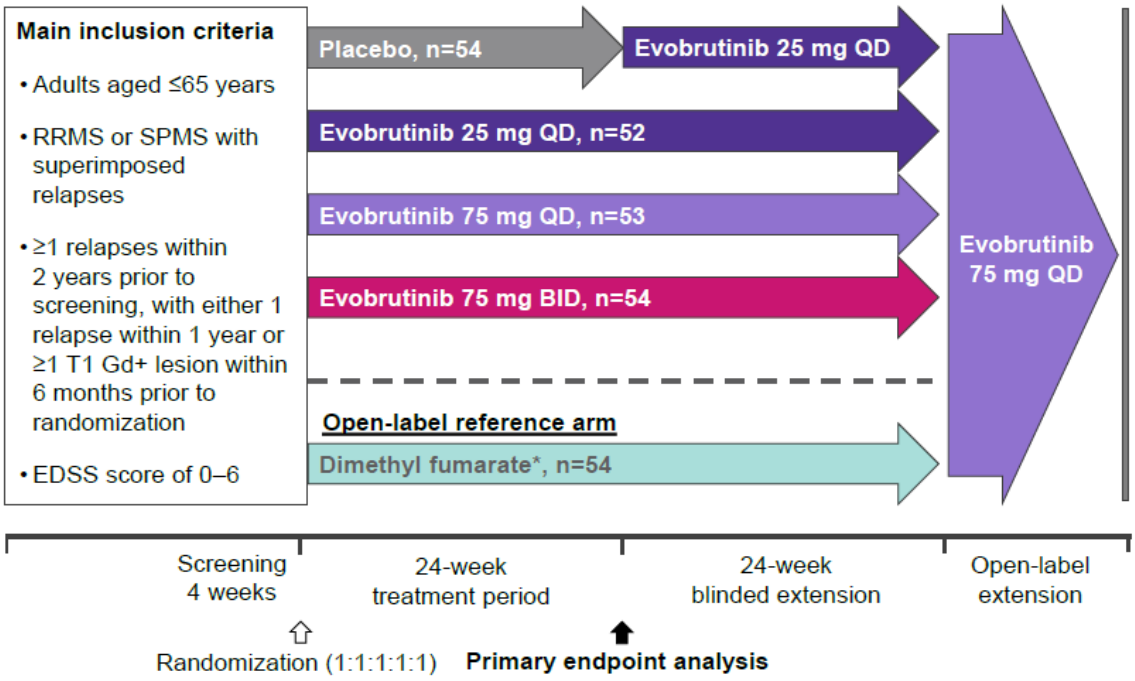
For each evobrutinib group, the difference in least squares means (LS-Mean) of SF-36 score change from baseline at Week 24, relative to placebo, was estimated, with 95% confidence interval (CI), and p-value, based on a mixed-effect model for repeated measures (MMRM) that adjusted for baseline score and included evobrutinib dose group and placebo group as factors. The LS-Mean estimate of the SF-36 score change from baseline at Week 24 (adjusted for baseline), with 95% CI, was reported for each group. An MMRM model was also used to estimate

the adjusted LS-Mean of the SF-36 score change from baseline at Week 48, with 95% CI, for each group, including the group that switched from placebo to 25 mg once daily at Week 24.

Supplementary Appendix 8. Influential Observation.

A single influential observation was identified in the DMF arm with respect to the primary endpoint, total number of gadolinium-enhancing T1 lesions at Week 12, 16, 20, and 24, based on consideration of the maximum lesions at baseline in each of the treatment groups.

Figure S1. Trial Design.

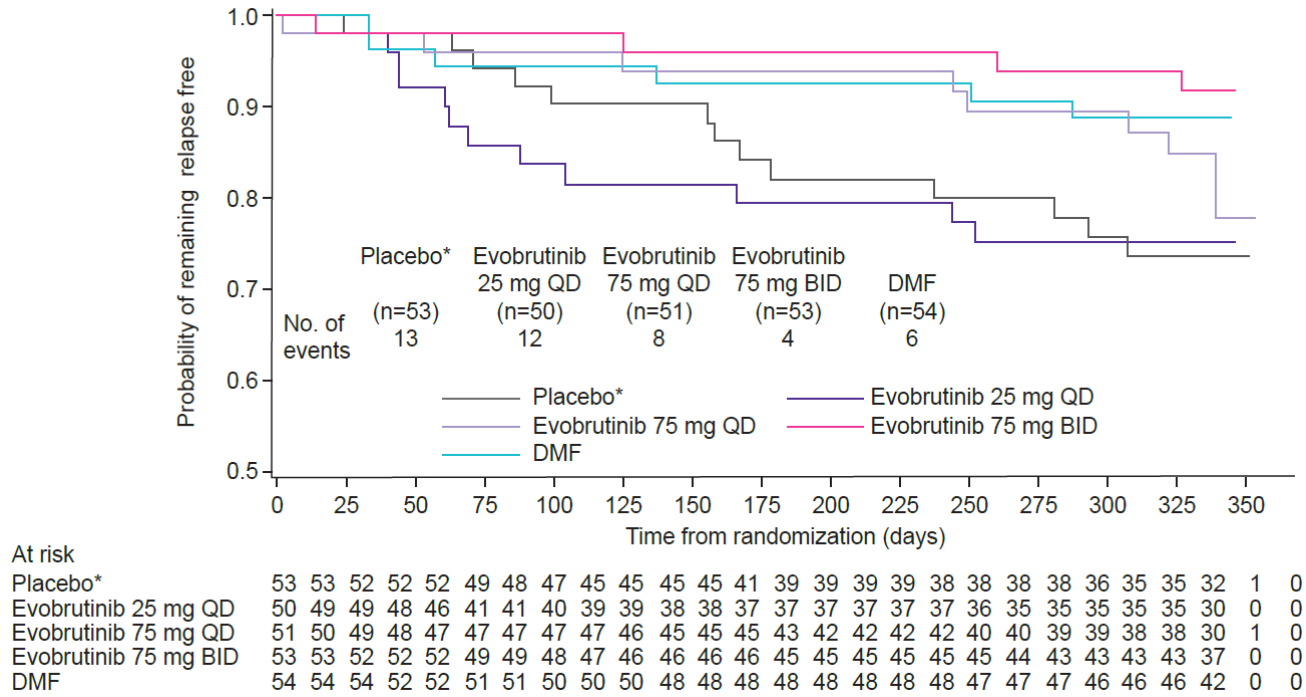


To maintain blinding, evobrutinib- and placebo-treated patients received three oral tablets twice daily (evobrutinib tablet dose: 25 mg; placebo tablets matched for color and size). There was a 4-week safety follow-up period after end of treatment. Patients switching from the reference group (dimethyl fumarate) to evobrutinib 75 mg QD at Week 48 undergo a 4–8-week washout.

*120 mg BID for the first 7 days, followed by 240 mg BID for the duration of treatment.

BID, twice daily; **EDSS,** Expanded Disability Status Scale; **RRMS,** relapsing-remitting multiple sclerosis; **SPMS**, secondary-progressive multiple sclerosis; **Gd+**, gadolinium-enhancing; **QD,** once daily.

Figure S2. Time to First Relapse (mITT).



At risk

Placebo*	53	53	52	52	52	49	48	47	45	45	45	45	41	39	39	39	39	38	38	38	38	36	35	35	32	1	0
Evobrutinib 25 mg QD	50	49	49	48	46	41	41	40	39	39	38	38	37	37	37	37	37	37	36	35	35	35	35	35	30	0	0
Evobrutinib 75 mg QD	51	50	49	48	47	47	47	47	47	46	45	45	45	43	42	42	42	42	40	40	39	39	38	38	30	1	0
Evobrutinib 75 mg BID	53	53	52	52	52	49	49	48	47	46	46	46	46	45	45	45	45	45	45	44	43	43	43	43	37	0	0
DMF	54	54	54	52	52	51	51	50	50	50	48	48	48	48	48	48	48	48	47	47	47	46	46	46	42	0	0

*Patients were switched to evobrutinib 25 mg QD for the second 24-week treatment period.

BID, twice daily; **DMF**, dimethyl fumarate; **mITT,** modified intention-to-treat population; **QD,** once daily

Table S1. Summary of Multiple Comparison Procedure for Primary and Key Secondary Endpoints (mITT).

Endpoint	Analysis Type	Evobrutinib 25 mg QD vs. placebo, P Value	Evobrutinib 75 mg QD vs. placebo, P Value	Evobrutinib 75 mg BID vs. placebo, P Value
Primary				
Total number of gadolinium-enhancing T1 lesions at Weeks 12, 16, 20, 24	Raw	0.29	0.002 (a)	0.03 (a)
	Adjusted	0.32	0.005 (a)	0.06
Key secondary				
ARR at Week 24	Raw	0.27	0.09	0.06
	Adjusted	0.32	0.32	0.32
Relapse-free status at Week 24	Raw	0.56	0.07	0.18
	Adjusted	0.60	0.32	0.37
EDSS score change from baseline at Week 24	Raw	0.41	0.58	0.27
	Adjusted	0.60	0.60	0.60

(a) Indicates that null hypothesis (adjusted lesion rate ratio, comparing evobrutinib to placebo, is equal to 1) was rejected. Truncated Hochberg procedure (truncation fraction = 0.9) used for hypothesis families corresponding to primary and first two key secondary endpoints. Conventional (truncation fraction = 1.0, not truncated) Hochberg procedure used for hypothesis family corresponding to third key secondary endpoint.

ARR, annualized relapse rate; **BID,** twice daily; **EDSS,** Expanded Disability Status Scale; **mITT,** modified intention-to-treat population; **QD,** once daily

Table S2. Analysis of Additional Secondary Efficacy Endpoints.

	mITT Population (N=261) (a)				
	Placebo	**Evobrutinib 25 mg QD**	**Evobrutinib 75 mg QD**	**Evobrutinib 75 mg BID**	**Dimethyl fumarate**
	n=53	**n=50**	**n=51**	**n=53**	**n=54**
Total number of new gadolinium-enhancing T1 lesions at Weeks 12, 16, 20, 24					
Mean ±SD	3.08 ±4.37	3.44 ±6.85	1.20 ±3.50	0.98 ±3.27	3.24 ±15.32
Median (range) [IQR]	1 (0–18) [5]	1 (0–35) [2]	0 (0–21) [1]	0 (0–22) [1]	0 (0–112) [1]
Lesion rate ratio (b) [95% CI]	-	1.36 [0.70–2.65]	0.27 [0.13–0.57]	0.41 [0.20–0.85]	-
Mean per-scan number of gadolinium-enhancing T1 lesions at Weeks 12, 16, 20, 24 (c)					
Mean ±SD	1.02 ±1.44	1.31 ±3.13	0.42 ±1.17	0.34 ±0.96	1.45 ±7.29
Median (range) [IQR]	0.50 (0.00–6.00) [1.50]	0.25 (0.00–19.00) [0.67]	0.00 (0.00–6.75) [0.25]	0.00 (0.00–6.25) [0.25]	0.00 (0.00–53.33) [0.25]
Location shift [95% CI]	-	0.00 [–0.25, 0.25]	–0.25 [–0.50, 0.00]	–0.50 [–0.75, –0.25]	-
Change from baseline in the volume of gadolinium-enhancing T1 lesions at Week 24 (d)					

Mean ±SD (cm^3)	–0.02 ±0.22	0.06 ±0.35	–0.11 ±0.54	–0.051 ±0.10	–0.05 ±0.48
Median (range) [IQR]	0.00 (–1.27–0.59) [0.02]	0.00 (–1.26–1.19) [0.02]	0.00 (–3.74–0.66) [0.07]	0.00 (–0.51–0.09) [0.08]	0.00 (–2.15–2.19) [0.01]
Location shift [95% CI]	-	0.00 [–0.004, 0.009]	–0.014 [–0.050, 0.000]	–0.018 [–0.042, 0.000]	-
Total number of new or enlarging T2 lesions at Weeks 12, 16, 20, 24					
Mean ±SD	5.96 ±6.99	6.52 ±11.57	3.41 ±10.75	2.19 ±4.72	5.35 ±16.67
Median (range) [IQR]	3 (0–25) [10]	1 (0–44) [7]	0 (0–73) [3]	1 (0–30) [2]	1 (0–119) [5]
Lesion rate ratio (e) [95% CI]	-	1.29 [0.63–2.65]	0.50 [0.24–1.04]	0.42 [0.20–0.87]	-
Change from baseline in the volume of T2 lesions at Week 24 (f)					
Mean ±SD (cm^3)	0.42 ±1.01	0.93 ±1.85	-0.01 ±0.56	0.06 ±0.50	0.47 ±2.96
Median (range) [IQR]	0.18 (–3.20–2.52) [1.06]	0.05 (–0.42–7.36) [0.99]	–0.01 (–2.15–1.90) [0.13]	–0.02 (–1.31–2.51) [0.25]	–0.01 (–6.35–18.23) [0.25]
Difference in LS-Mean of change from baseline in cube root of volume (cm) [95% CI]	-	0.02 [–0.24, 0.28]	–0.41 [–0.66, –0.15]	–0.36 [–0.62, –0.10]	
			mITT-BE Population (N=229) (g)		

	Placebo/ Evobrutinib 25 mg QD n=44	Evobrutinib 25 mg QD n=44	Evobrutinib 75 mg QD n=46	Evobrutinib 75 mg BID n=45	Dimethyl fumarate n=50
Number of gadolinium-enhancing T1 lesions at Week 48 (h)					
Mean ±SD	1.00 ±1.65	2.00 ±4.45	0.91 ±2.96	0.56 ±1.26	0.42 ±1.44
Median (range) [IQR]	0 (0–7) [2]	0 (0–20) [2]	0 (0–18) [0]	0 (0–6) [0]	0 (0–9) [0]
Number of new gadolinium-enhancing T1 lesions at Week 48 (h)					
Mean ±SD	0.95 ±1.61	1.93 ±4.30	0.91 ±2.96	0.56 ±1.26	0.42 ±1.44
Median (range) [IQR]	0 (0–7) [2]	0 (0–20) [2]	0 (0–18) [0]	0 (0–6) [0]	0 (0–9) [0]
Change from baseline in volume of gadolinium-enhancing T1 lesions at Week 48 (h)					
Mean ±SD (cm^3)	0.10 ±0.47	0.09 ±0.47	–0.07 ±0.38	0.00 ±0.22	–0.23 ±1.01
Median (range) [IQR]	0.00 (–0.32–2.75) [0.01]	0.00 (–1.77–2.00) [0.07]	0.00 (–2.27–0.56) [0.03]	0.00 (–0.45–0.86) [0.05]	0.00 (–6.75–0.52) [0.04]
Number of new or enlarging T2 lesions at Week 48 (i)					
Mean ±SD	3.57 ±4.35	5.86 ±11.33	3.84 ±10.08	1.60 ±3.80	1.88 ±4.80
Median (range) [IQR]	2 (0–16) [6]	1 (0–57) [6]	0 (0–62) [3]	0 (0–22) [2]	0 (0–23) [1]

Change from baseline in volume of T2 lesions at Week 48 (h)					
Mean ±SD (cm^3)	0.56 ±1.39	1.46 ±3.17	0.39 ±1.09	0.22 ±0.78	–0.14 ±1.92
Median (range) [IQR]	0.05 (–0.31–6.47) [0.50]	0.08 (–1.49–16.24) [1.55]	0.00 (–0.72–4.38) [0.41]	–0.02 (–0.76–3.22) [0.39]	–0.01 (–8.79–5.15) [0.24]
Change from baseline in EDSS score at Week 48 (j)					
Median (range) [IQR]	0.0 (–1.5–0.5) [0.0]	0.0 (–2.5–1.0) [0.0]	0.0 (–2.5–0.5) [0.0]	0.0 (–1.0–1.5) [0.0]	0.0 (–1.5–1.0) [0.0]
	mITT Population (N=261) (a)				
	Placebo/ Evobrutinib 25 mg QD n=53	**Evobrutinib 25 mg QD n=50**	**Evobrutinib 75 mg QD n=51**	**Evobrutinib 75 mg BID n=53**	**Dimethyl fumarate n=54**
Relapses at Week 48					
Number of relapses, n	17	23	11	5	7
Unadjusted ARR [95% CI] (k)	0.37 [0.21–0.59]	0.52 [0.33–0.78]	0.25 [0.12–0.44]	0.11 [0.04–0.25]	0.14 [0.06–0.29]
Relapse-free status at Week 48					

Proportion relapse-free at Week 48, %	66.0	70.0	70.6	79.2	85.2

(a) All patients included in the mITT analysis set have a baseline and ≥1 post baseline MRI assessment.

(b) Lesion rate ratio estimated based on negative binomial model for lesion count (summed over scans). Model includes factors for treatment group and covariate for baseline lesions (presence/absence), with offset equal to log number of available scans. Subjects with missing scans or discontinuing early were analyzed according to the available number of scans and lesion counts.

(c) For a subject missing all post-baseline scans, endpoint imputed by the median value among subjects with available post-baseline scans, who were in the same treatment group with the same value for the baseline covariate. Hodges-Lehman estimate of location shift, strata aligned.

(d) For a subject missing volume change from baseline at Week 24, endpoint imputed by the median value among subjects in the same treatment group with the same value for the baseline covariate.

(e) Same model as in (b) with covariate corresponding to T2 lesion volume (≤ 13 cc, > 13cc) at baseline.

(f) No imputation for missing scans. MMRM for change from baseline in cube root of T2 lesion volume includes fixed effects for treatment, visit, treatment by visit interaction, and covariate for cube root of lesion volume at baseline.

(g) mITT-BE is defined as all patients in the mITT population with an MRI assessment during the second 24-week treatment period.

(h) All available scans included.

(i) New/enlarging T2 lesions defined relative to previous available scan. At Week 48, previous available scan was at or before Week 24. All available scans included.

(j) The EDSS scale ranges from 0 to 10 in 0.5-unit increments, with higher scores representing greater disability. For patients who switched from placebo to evobrutinib at Week 24, baseline is the last measurement prior to the first dose in the second 24-week treatment period. For other treatment groups, baseline is the last measurement prior to the first dose in the first 24-week treatment period.

(k) Unadjusted ARR defined as the number of relapses among patients divided by number of patient-years of follow-up. For early discontinuers from treatment, relapses and follow-up through safety follow-up visit were included.

CIs were not adjusted for multiple testing and cannot be used to infer treatment effect. Scans collected within 3 weeks of HDC use were considered missing.

ARR, annualized relapse rate; **BE**, blinded extension; **BID,** twice daily; **CFB,** change from baseline; **CI,** confidence interval; **HDC,** high-dose corticosteroids; **IQR,** interquartile range; **LS-Mean,** least squares means; **mITT,** modified intention-to-treat; **MMRM,** mixed-effect model for repeated measures; **MRI,** magnetic resonance imaging; **NB,** negative binomial; **QD,** once daily; **SD,** standard deviation.

Table S3. HRQoL Analysis: Change from Baseline in SF-36 Summary and Subdomain Scores.

	Placebo	Evobrutinib 25 mg QD	Evobrutinib 75 mg QD	Evobrutinib 75 mg BID
Change from baseline in Physical Component Summary Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	–0.18 [–1.27–0.91]	1.49 [0.37–2.60]	0.97 [–0.14–2.09]	1.33 [0.24–2.42]
Difference (evobrutinib vs. placebo) [95% CI]	-	1.59 [–0.70–3.87]	1.98 [–0.29–4.25]	1.88 [–0.39–4.14]
Week 48 LS-Mean [95% CI] (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	0.12 [–0.99–1.22]	1.51 [0.41–2.61]	0.78 [–0.30–1.87]	1.17 [0.09–2.26]
Change from baseline in Mental Component Summary Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	–0.21 [–1.64–1.22]	–1.11 [–2.58–0.37]	–0.06 [–1.52–1.41]	0.86 [–0.58–2.30]
Difference (evobrutinib vs. placebo) [95% CI]	-	–0.05 [–3.23–3.13]	–0.06 [–3.19–3.07]	1.10 [–2.05–4.24]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	–0.34 [–1.77–1.09]	–1.16 [–2.59–0.27]	–0.26 [–1.66–1.13]	1.31 [–0.10–2.73]
Change from baseline in Physical Functioning Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	0.09 [–1.05–1.24]	1.54 [0.36–2.71]	0.62 [–0.55–1.78]	0.92 [–0.22–2.07]

Difference (evobrutinib vs. placebo) [95% CI]	-	1.19 [–1.28–3.65]	0.64 [–1.79–3.08]	0.35 [–2.09–2.79]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	–0.20 [–1.40–0.99]	1.77 [0.59–2.96]	0.52 [–0.64–1.68]	1.05 [–0.12–2.22]
Change from baseline in Role-Physical Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	0.34 [–0.88–1.57]	0.66 [–0.60–1.91]	1.16 [–0.09–2.40]	1.56 [0.33–2.78]
Difference (evobrutinib vs. placebo) [95% CI]	-	–0.41 [–3.10–2.29]	0.53 [–2.14–3.19]	1.01 [–1.66–3.68]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	0.26 [–0.97–1.49]	0.74 [–0.49–1.96]	0.97 [–0.23–2.17]	1.60 [0.40–2.81]
Change from baseline in Bodily Pain Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	–0.26 [–1.74–1.22]	0.59 [–0.93–2.11]	–0.12 [–1.63–1.38]	2.49 [1.00–3.97]
Difference (evobrutinib vs. placebo) [95% CI]	-	2.08 [–1.20–5.37]	2.13 [–1.11–5.37]	3.59 [0.35–6.84]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	0.57 [–0.91–2.06]	0.60 [–0.88–2.08]	–0.57 [–2.02–0.88]	2.22 [0.76–3.68]
Change from baseline in General Health Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	–1.15 [–2.41–0.11]	0.76 [–0.54–2.05]	0.74 [–0.55–2.03]	–0.17 [–1.44–1.10]

Difference (evobrutinib vs. placebo) [95% CI]	-	1.87 [–0.75–4.49]	2.14 [–0.46–4.73]	1.44 [–1.16–4.03]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	–0.98 [–2.30–0.33]	0.63 [–0.67–1.94]	0.57 [–0.71–1.86]	–0.19 [–1.49–1.10]
Change from baseline in Vitality Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	–0.17 [–1.52–1.18]	–0.73 [–2.12–0.66]	0.89 [–0.49–2.27]	1.49 [0.12–2.85]
Difference (evobrutinib vs. placebo) [95% CI]	-	–0.39 [–3.36–2.57]	1.84 [–1.08–4.77]	2.22 [–0.73–5.16]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	–0.14 [–1.44–1.16]	–1.07 [–2.37–0.23]	0.83 [–0.44–2.10]	1.69 [0.40–2.98]
Change from baseline in Social Functioning Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	0.77 [–0.60–2.15]	–0.24 [–1.66–1.18]	0.44 [–0.97–1.84]	1.33 [–0.06–2.72]
Difference (evobrutinib vs. placebo) [95% CI]	-	0.85 [–2.34–4.04]	0.16 [–2.98–3.31]	1.63 [–1.52–4.79]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%)	n=41 (93.2%)	n=42 (91.3%)	n=42 (93.3%)
	1.02 [–0.29–2.33]	0.04 [–1.27–1.34]	0.23 [–1.04–1.51]	1.93 [0.64–3.22]
Change from baseline in Role-Emotional Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%)	n=42 (84.0%)	n=45 (88.2%)	n=44 (83.0%)
	0.36 [–1.25–1.97]	–0.47 [–2.13–1.19]	0.12 [–1.51–1.75]	0.51 [–1.11–2.12]

Difference (evobrutinib vs. placebo) [95% CI]	-	–0.50 [–4.26–3.26]	0.23 [–3.47–3.92]	–0.74 [–4.46–2.97]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%) –0.34 [–1.98–1.31]	n=41 (93.2%) –0.59 [–2.23–1.05]	n=42 (91.3%) –0.21 [–1.81–1.39]	n=42 (93.3%) 0.71 [–0.91–2.33]
Change from baseline in Mental Health Score				
Week 24 LS-Mean (mITT)	n=44 (83.0%) –0.81 [–2.18–0.57]	n=42 (84.0%) –0.47 [–1.89–0.95]	n=45 (88.2%) –0.32 [–1.73–1.09]	n=44 (83.0%) 1.11 [–0.28–2.49]
Difference (evobrutinib vs. placebo) [95% CI]	-	0.61 [–2.54–3.76]	–0.62 [–3.71–2.48]	1.77 [–1.34–4.89]
Week 48 LS-Mean (mITT-BE)	n=39 (88.6%) –0.64 [–1.99–0.71]	n=41 (93.2%) –0.21 [–1.56–1.14]	n=42 (91.3%) –0.55 [–1.87–0.76]	n=42 (93.3%) 1.42 [0.09–2.75]

Each of the following subscales was rated by the patient from 0–100, where higher scores represent better HRQoL: Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional, Mental Health (Ware JE Jr, et al. 2007). Physical Component Summary (PCS) and Mental Component Summary (MCS) scores were obtained through a linear combination of weighted transformed scores from the eight subscales. See Supplementary Appendix 7 for more information.

BE, blinded extension; **BID,** twice daily; **CI,** confidence interval; **HRQoL,** health-related quality of life; **LS,** least squares; **mITT**, modified intention-to-treat; **QD,** once daily; **SF-36,** Short Form 36.

Table S4. Most Common TEAEs (≥5% in any Group) from Week 0 up to 52 – Other.

Patients with TEAE, n (%)	Placebo/ Evobrutinib 25 mg QD* n=54	Evobrutinib 25 mg QD n=52	Evobrutinib 75 mg QD n=53	Evobrutinib 75 mg BID n=54	Dimethyl Fumarate n=54
GGT increased	0 (0.0)	1 (1.9)	1 (1.9)	3 (5.6)	1 (1.9)
White blood cell count decreased	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.9)	3 (5.6)
Erythema	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	7 (13.0)

*Placebo-treated patients switched to evobrutinib 25 mg QD for Weeks 25-52.

BID, twice daily; **GGT,** gamma-glutamyltransferase; **TEAE,** treatment-emergent adverse event; **QD**, once daily.

Table S5a. Shifts from Normal (Grade 0) to Highest Grade ALT* (Safety Analysis Set).

Highest NCI-CTCAE grade, n (%)	Placebo n=54	Evobrutinib 25 mg QD n=52	Evobrutinib 75 mg QD n=48	Evobrutinib 75 mg BID n=49	Dimethyl fumarate n=50
Weeks 0–24					
Grade 1	4 (7.4)	9 (17.3)	11 (22.9)	11 (22.4)	21 (42.0)
Grade 2	3 (5.6)	0 (0.0)	0 (0.0)	2 (4.1)	0 (0.0)
Grade 3	1 (1.9)	2 (3.8)	1 (2.1)	2 (4.1)	0 (0.0)
Grade 4	0 (0.0)	1 (1.9)	0 (0.0)	1 (2.0)	0 (0.0)
Weeks 0–52					
Grade 1	-	13 (25.0)	19 (39.6)	14 (28.6)	23 (46.0)
Grade 2	-	0 (0.0)	0 (0.0)	1 (2.0)	0 (0.0)
Grade 3	-	2 (3.8)	1 (2.1)	3 (6.1)	0 (0.0)
Grade 4	-	1 (1.9)	0 (0.0)	1 (2.0)	0 (0.0)

*Shifts in AST grade generally mirrored ALT grade shifts but were overall less pronounced (data not shown).

ALT, alanine aminotransferase; **AST**, aspartate aminotransferase**; BID,** twice daily; **NCI-CTCAE,** National Cancer Institute Common Terminology Criteria for Adverse Events**, QD,** once daily.

Table S5b. Shifts from Baseline to Highest Grade (>Baseline) ALT* (Safety Analysis Set).

Highest NCI-CTCAE Grade, n (%)	Placebo n=54	Evobrutinib 25 mg QD n=52	Evobrutinib 75 mg QD n=53	Evobrutinib 75 mg BID n=54	Dimethyl fumarate n=54
Weeks 0–24					
Grade 2	3 (5.6)	0 (0.0)	0 (0.0)	2 (3.7)	0 (0.0)
Grade 3	1 (1.9)	2 (3.8)	2 (3.8)	3 (5.6)	0 (0.0)
Grade 4	0 (0.0)	1 (1.9)	0 (0.0)	2 (3.7)	0 (0.0)
Weeks 0–52					
Grade 2	-	0 (0.0)	0 (0.0)	1 (2.0)	0 (0.0)
Grade 3	-	2 (3.8)	2 (3.8)	4 (7.4)	0 (0.0)
Grade 4	-	1 (1.9)	0 (0.0)	2 (3.7)	0 (0.0)

*Shifts in AST grade generally mirrored ALT grade shifts but were overall less pronounced (data not shown).

ALT, alanine aminotransferase; **AST**, aspartate aminotransferase**; BID,** twice daily; **NCI-CTCAE,** National Cancer Institute Common Terminology Criteria for Adverse Events**, QD,** once daily.

Table S5c. Shifts from Normal (Grade 0) to Highest Grade ALT* during Weeks 25–52 for Subjects Switching from Placebo to Evobrutinib 25 mg QD (SAF-BE Analysis Set).

Highest NCI-CTCAE Grade, n (%)	Placebo + Evobrutinib 25 mg QD n=47
Weeks 25–52	
Grade 1	10 (21.3)
Grade 2	0 (0.0)
Grade 3	0 (0.0)
Grade 4	0 (0.0)

*Shifts in AST grade generally mirrored ALT grade shifts but were overall less pronounced (data not shown).

Baseline is defined as the last measurement prior to the first dose in the second treatment period.

SAF-BE Analysis Set defined as the Safety Analysis Set (subjects who receive at least one dose of trial treatment) restricted to those who entered the second 24-week treatment period.

ALT, alanine aminotransferase; **AST**, aspartate aminotransferase**; NCI-CTCAE,** National Cancer Institute Common Terminology Criteria for Adverse Events**, QD,** once daily; **SAF-BE,** safety analysis set of the blinded extension.

Table S5d. Shifts from Baseline to Highest Grade (>Baseline) ALT* during Weeks 25–52 for Subjects Switching from Placebo to Evobrutinib 25 mg QD (SAF-BE Analysis Set).

Highest NCI-CTCAE Grade, n (%)	Placebo + Evobrutinib 25 mg QD n=49
Weeks 25–52	
Grade 2	0 (0.0)
Grade 3	0 (0.0)
Grade 4	0 (0.0)

*Shifts in AST grade generally mirrored ALT grade shifts but were overall less pronounced (data not shown).

Baseline is defined as the last measurement prior to the first dose in the second treatment period.

SAF-BE Analysis Set defined as the Safety Analysis Set (subjects who receive at least one dose of trial treatment) restricted to those who entered the second 24-week treatment period.

ALT, alanine aminotransferase; **AST**, aspartate aminotransferase**; NCI-CTCAE,** National Cancer Institute Common Terminology Criteria for Adverse Events**, QD,** once daily; **SAF-BE,** safety analysis set of the blinded extension.

Table S6a. Shifts from Normal (Grade 0) to Highest Grade Decrease in Lymphocytes (Safety Analysis Set).

Highest NCI-CTCAE Grade, n (%)	Placebo n=52	Evobrutinib 25 mg QD n=49	Evobrutinib 75 mg QD n=52	Evobrutinib 75 mg BID n=53	Dimethyl fumarate n=51
Weeks 0–24					
Grade 1	3 (5.8)	2 (4.1)	2 (3.8)	3 (5.7)	10 (19.6)
Grade 2	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.9)	7 (13.7)
Grade 3	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Weeks 0–52					
Grade 1	-	2 (4.1)	4 (7.7)	3 (5.7)	10 (19.6)
Grade 2	-	1 (2.0)	1 (1.9)	1 (1.9)	13 (25.5)
Grade 3	-	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.0)
Grade 4	-	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

BID, twice daily; **NCI-CTCAE,** National Cancer Institute Common Terminology Criteria for Adverse Events, **QD,** once daily.

Table S6b. Shifts from Baseline to Highest Grade (>Baseline) Decrease in Lymphocytes (Safety Analysis Set).

Highest NCI-CTCAE Grade, n (%)	Placebo n=54	Evobrutinib 25 mg QD n=52	Evobrutinib 75 mg QD n=53	Evobrutinib 75 mg BID n=54	Dimethyl fumarate n=54
Weeks 0–24					
Grade 2	1 (1.9)	1 (1.9)	0 (0.0)	1 (1.9)	7 (13.0)
Grade 3	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Grade 4	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Weeks 0–52					
Grade 2	-	2 (3.8)	1 (1.9)	1 (1.9)	15 (27.8)
Grade 3	-	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.0)
Grade 4	-	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

BID, twice daily; **NCI-CTCAE,** National Cancer Institute Common Terminology Criteria for Adverse Events, **QD,** once daily.

Table S6c. Shifts from Normal (Grade 0) to Highest Grade Decrease in Lymphocytes during Weeks 25–52 for Subjects Switching from Placebo to Evobrutinib 25 mg QD (SAF-BE Analysis Set).

Highest NCI-CTCAE Grade, n (%)	Placebo + Evobrutinib 25 mg QD n=48
Weeks 25–52	
Grade 1	2 (4.2)
Grade 2	0 (0.0)
Grade 3	0 (0.0)
Grade 4	0 (0.0)

Baseline is defined as the last measurement prior to the first dose in the second treatment period.

SAF-BE Analysis Set defined as the Safety Analysis Set (subjects who receive at least one dose of trial treatment) restricted to those who entered the second 24-week treatment period.

NCI-CTCAE, National Cancer Institute Common Terminology Criteria for Adverse Events, **QD,** once daily; SAF-BE, safety analysis set of the blinded extension.

Table S6d. Shifts from Baseline to Highest Grade Decrease in Lymphocytes during Weeks 25–52 for Subjects Switching from Placebo to Evobrutinib 25 mg QD (SAF-BE Analysis Set).

Highest NCI-CTCAE Grade, n (%)	Placebo + Evobrutinib 25 mg QD n=49
Weeks 25–52	
Grade 2	0 (0.0)
Grade 3	0 (0.0)
Grade 4	0 (0.0)

Baseline is defined as the last measurement prior to the first dose in the second treatment period.

SAF-BE Analysis Set defined as the Safety Analysis Set (subjects who receive at least one dose of trial treatment) restricted to those who entered the second 24-week treatment period.

NCI-CTCAE, National Cancer Institute Common Terminology Criteria for Adverse Events, **QD,** once daily; **SAF-BE**, safety analysis set of the blinded extension.

Table S7. Most Common TEAEs Leading to Withdrawal (≥2% in any Group) from Week 0 up to 52.

Patients with TEAE, n (%)	Placebo/ Evobrutinib 25 mg QD* n=54	Evobrutinib 25 mg QD n=52	Evobrutinib 75 mg QD n=53	Evobrutinib 75 mg BID n=54	Dimethyl Fumarate n=54
Hepatobiliary disorders	1 (1.9)	1 (1.9)	1 (1.9)	2 (3.7)	0 (0.0)
Investigations	3 (5.6)	2 (3.8)	5 (9.4)	6 (11.1)	0 (0.0)
ALT increased	1 (1.9)	2 (3.8)	1 (1.9)	4 (7.4)	0 (0.0)
AST increased	0 (0.0)	1 (1.9)	1 (1.9)	2 (3.7)	0 (0.0)
Lipase increased	1 (1.9)	0 (0.0)	2 (3.8)	1 (1.9)	0 (0.0)

*Placebo-treated patients switched to evobrutinib 25 mg QD for Weeks 25-52.

ALT, alanine aminotransferase; **AST,** aspartate aminotransferase; **BID**, twice daily; **TEAE,** treatment-emergent adverse event; **QD**, once daily.

References

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline: Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals M3(R2). Available at: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M3_R2/Step4/M3_R2__Guideline.pdf (Accessed April 2019), 2009.

Kappos L, Antel J, Comi G, et al. Oral fingolimod (FTY720) for relapsing multiple sclerosis. N Engl J Med 2006;355:1124-40.

Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology 2014;83:278-86.

Miller DH, Khan OA, Sheremata, WA, et al. A controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med 2003;348:15-23.

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

Ware JE Jr, Kosinski M, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. User's manual for the SF-36 v2 Health Survey (2nd edition). Lincoln, RI: Quality Metric Incorporated, 2007.