Five-year efficacy outcomes of ocrelizumab in relapsing multiple sclerosis: A propensity-matched comparison of the OPERA studies with other disease-modifying therapies in real-world lines of treatments

Journal of Central Nervous System Disease
Volume 16: 1–17

DOI: 10.1177/11795735241260563
Sage

Erwan Muros-Le Rouzic[1], Yanic Heer[2,*], Sean Yiu[3], Viola Tozzi[2,*], Stefan Braune[4], Philip van Hövell[2,*], Arnfin Bergmann[4], Corrado Bernasconi[1], Fabian Model[1,**], Licinio Craveiro[1] and on behalf of the NTD study group

[1]F. Hoffmann-La Roche Ltd., Basel, Switzerland. [2]PricewaterhouseCoopers AG, Zürich, Switzerland. [3]Roche Products Limited, Welwyn Garden City, UK. [4]NeuroTransData GmbH, Neuburg an der Donau, Germany.

ABSTRACT

BACKGROUND: Clinical trials comparing the efficacy of ocrelizumab (OCR) with other disease-modifying therapies (DMTs) other than interferon (IFN) β-1a in relapsing multiple sclerosis (RMS) are lacking.

OBJECTIVES: To compare the treatment effect of OCR vs six DMTs' (IFN β-1a, glatiramer acetate, fingolimod, dimethyl fumarate, teriflunomide, natalizumab) treatment pathways used in clinical practice by combining clinical trial and real-world data.

METHODS: Patient-level data from OPERA trials and open-label extension phase, and from the German NeuroTransData (NTD) MS registry, were used to build 1:1 propensity score-matched (PSM) cohorts controlling for seven baseline covariates, including brain imaging activity. Efficacy outcomes were time to first relapse and time to 24-week confirmed disability progression over 5.5 years of follow-up. Intention-to-treat analysis using all outcome data irrespective of treatment switch was applied.

RESULTS: The analyses included 611 OPERA patients and 7141 NTD patients. We built 12 paired-matched cohorts (six for each outcome, two for each DMT) to compare efficacy of OCR in OPERA with each DMT treatment pathway in NTD. Post-matching, baseline covariates and PS were well balanced (standardized mean difference <.2 for all cohorts). Over 5.5 years, patients treated with OCR showed a statistically significant reduction in the risk of relapse (hazard ratios [HRs] .30 to .54) and disability progression (HRs .51 to .67) compared with all index therapies and their treatment switching pathways in NTD. Treatment switch and/or discontinuation occurred frequently in NTD cohorts.

CONCLUSION: OCR demonstrates superiority in controlling relapses and disability progression in RMS compared with real-world treatment pathways over a 5.5-year period. These analyses suggest that high-efficacy DMTs and high treatment persistence are critical to achieve greatest clinical benefit in RMS.

REGISTRATION: OPERA I (NCT01247324), OPERA II (NCT01412333)

KEYWORDS: Multiple sclerosis, ocrelizumab, real-world data, comparative effectiveness research, treatment pathways

RECEIVED: December 5, 2023. **ACCEPTED:** May 17, 2024.

TYPE: Original Research Article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: F. Hoffmann-La Roche Ltd., Basel, Switzerland, provided financial support for the study and publication of this manuscript. Writing and editorial support was provided by Articulate Science, UK. **E Muros-Le Rouzic** is an employee of and shareholder in F. Hoffmann-La Roche Ltd. **Y Heer** was an employee of PricewaterhouseCoopers and contracted to perform statistical projects for NeuroTransData at the time of this analysis. **S Yiu** is an employee of Roche Products Ltd. **V Tozzi** is an employee of PricewaterhouseCoopers and contracted to perform statistical projects for NeuroTransData at the time of this analysis. **S Braune** has received honoraria from Kassenärztliche Vereinigung Bayerns and health maintenance organizations for patient care; and from Biogen, Merck, NeuroTransData, Novartis, and Roche for consulting, project management, clinical studies, and lectures; he also received honoraria and expense compensation as a board member of NeuroTransData. **P van Hövell** was an employee of PricewaterhouseCoopers and contracted to perform statistical projects for NeuroTransData at the time of this analysis. **A Bergmann** has received consultancy fees from advisory board, speaker, and other activities for NeuroTransData; project management and clinical studies for and travel expenses from Novartis and Servier. **C Bernasconi** is a consultant for F. Hoffmann-La Roche Ltd. **F Model** was an employee of and shareholder in F. Hoffmann-La Roche Ltd. at the time of this analysis and is now an employee of Denali Therapeutics. **L Craveiro** is an employee of and shareholder in F. Hoffmann-La Roche Ltd.

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by F. Hoffmann-La Roche Ltd., Basel, Switzerland.

NTD REGISTRY: The data acquisition protocol was approved by the ethical committee of the Bavarian Medical Board (Bayerische Landesärztekammer; June 14, 2012; No. 11144) and reapproved by the ethical committee of the Medical Board North-Rhine (Ärztekammer Nordrhein; April 25, 2017; ID 2017071). Patients included in this analysis provided their informed consent (via tablets in NTD practices, electronic questionnaires, or via a patient portal) to the NTD registry. Patients explicitly agreed to any secondary use of their data.

SUPPLEMENTAL MATERIAL: Supplemental material for this article is available online.

CORRESPONDING AUTHOR: Erwan Muros-Le Rouzic, F. Hoffmann-La Roche Ltd., Basel 4070, Switzerland.
Email: erwan.muros@roche.com

Introduction

Ocrelizumab (OCR) is a humanized anti-CD20^{+} monoclonal antibody approved for the treatment of relapsing forms of multiple sclerosis (MS) and primary progressive MS (PPMS).[1,2] In patients with relapsing MS (RMS), the clinical benefit of OCR on disease activity and disability progression was demonstrated vs interferon (IFN) β-1a at a dose of 44 µg (Rebif, EMD Serono) in a 2-year phase 3 randomized controlled trial (RCT).[3] While comparison between different disease-modifying therapies (DMTs), including OCR, have been performed using meta-analytical approaches,[4-8] no head-to-head RCTs have directly compared the efficacy of OCR with other DMTs.

High-quality real-world observational studies are increasingly used for comparisons of effectiveness and safety among the growing number of DMTs, providing valuable insights that can inform treatment choices for patients with MS.[9-12] Despite the many potential benefits of such studies, they are subject to important biases, in particular selection bias and confounding by indication, which can result in an unbalanced between-groups distribution of variables that may influence the outcome.[13] These limitations can be mitigated by using statistical techniques such as propensity score (PS) methods, which allow for adjustment of baseline characteristics of important measured confounders.[14-16]

Several studies comparing DMT effectiveness in real-world MS datasets using PS techniques have been published,[9-12,17-19] including a recent MSBase study comparing switch to OCR, cladribine, or natalizumab after fingolimod treatment cessation in patients with relapsing-remitting MS.[20] However, at present, datasets capturing OCR use in real-world settings remain limited in terms of size and duration of follow-up. Merging randomized trials with real-world data is an alternative approach, which offers the advantage of allowing comparisons between newly approved and multiple established DMTs, and to rapidly expand to comparisons of long-term data from open-label extensions (OLE) and real-world datasets. To our knowledge, this approach was previously followed in only one study that used data from the pivotal trial assessing cladribine tablets vs placebo (CLARITY), and PS-matched (PSM) data from the Italian multicenter database, i-MuST.[21] However in this study, patients were censored at their treatment switch, an approach that does not mimic the high diversity of treatment pathways observed in clinical practice characterized by the typical sequencing of DMTs.[20,22]

In the present study, we applied PSM methods to compare time to onset of disability progression and time to first clinical relapse between patients with RMS treated with OCR in two large clinical trials,[3,23] and those on multiple treatment pathways with six different index DMTs in the real-world German NeuroTransData (NTD) MS registry, over a period of up to 5.5 years.

Methods

Data sources

OPERA I and II randomized clinical trials. Pooled data from the double-blind, phase 3 OPERA I and II RCTs were selected for this analysis. Study design and key eligibility criteria in OPERA I/II have been described previously[3] and, briefly, included: age 18-55 years; MS diagnosis according to the 2010 revised McDonald criteria;[24] Expanded Disability Status Scale (EDSS) score ≤5.5 at screening; at least two documented clinical relapses within the previous 2 years or one clinical relapse within the year before screening; and no neurologic worsening for at least 30 days before screening.[3]

German NeuroTransData (NTD) registry. The NTD network is a Germany-wide physicians network, which includes currently about 25,000 patients with MS and captures demographics, clinical history, patient-related outcomes, and clinical variables in real time during clinical visits (average of 3.5 visits per patient each year [2010-2018]) using the DESTINY platform and ensuring confidentiality, integrity, and high data quality.[9,25,26]

Study population

OPERA trials. All patients randomized to the OCR arms (n = 611) were included in the study population, with the exception of patients from the US. Previous subgroup analyses showed a difference on 12-week confirmed disability progression (CDP) outcomes between patients enrolled from the US and from the other countries, with a significant treatment-by-subgroup interaction observed for body mass index (BMI).[27] Given that information on BMI is not systematically collected in the NTD registry, adjustment for this interaction was not possible, thus justifying the exclusion of US patients. Finally, the non-US population in OPERA is 64% European and is therefore a more comparable cohort to the German NTD population.

Patients randomized to the IFN β-1a arms (excluding US patients) were also included in the analysis as a feasibility step for the use of clinical trial data for group matching.

German NeuroTransData (NTD) registry. From the NTD MS registry database, patients satisfying the following inclusion/exclusion criteria were extracted: aged ≥18 years with RMS diagnosis, initiated treatment (index therapy) after January 1, 2009 with either IFN β-1a, natalizumab (NTZ), glatiramer acetate (GA), fingolimod (FTY), dimethyl fumarate (DMF), or teriflunomide (TERI). We did not include IFN β-1b therapy assuming identical pharmacodynamic effects and similar clinical efficacy with IFN β-1a in RRMS and rather decided to examine

real-world lines of treatments including DMTs with different mechanisms of action.[28,29] To ensure neurologic stability at index therapy, patients must have experienced no relapse nor undergone treatment with steroids 8 weeks prior to index therapy. Other exclusion criteria included diagnosis of progressive MS (either secondary progressive MS or PPMS), and prior treatment with anti-CD20 therapies such as OCR or rituximab (RTX; NB: ofatumumab was not available in the NTD at time of these analyses). The NTD registry data extraction date was July 1, 2020.

Additionally, at a minimum, data on sex, age, date of first MS symptom, dates of clinical relapses, and EDSS score at index therapy, recorded within 3 months before the start of therapy had to be available for each patient.

Cohort matching

Methods are described following recommendations for the use of PS methods in MS research from Simoneau et al.[16] A PSM algorithm was used to derive a matched sample of comparable patients with RMS between the OCR cohort in the OPERA OLE study and from the NTD registry.[16]

Selection of matching covariates. The baseline covariates described in Table 1 were included in the PS as they were considered to be confounders for disability and relapse outcomes based on empirical clinical knowledge and published evidence.[11] Different covariates were used for each outcome in line with Laplaud et al and previous subgroup analyses from the OPERA trials.[11,27] Missing values were not replaced, with the exception of partial dates for date of birth. A sensitivity analysis that excluded "existence of gadolinium-enhancing (Gd+) lesions" as a matching factor for the relapse outcome was conducted with the MRI measured ±3 months within the therapy start.

Propensity score method. A 1:1 ratio pair-matching was applied to each OCR comparison cohort (eg, OCR–NTZ, OCR–GA) using a 5-to-1-digit-greedy-nearest-neighbor-matching algorithm, in which a NTD participant was selected at random and then matched to the OPERA participant whose PS was closest to that of the NTD participant.[15] The PS was estimated using logistic regression with the treatment cohort as dependent variable and covariates at index-therapy initiation as independent variables; the covariates selected were considered important potential confounders.[14] No caliper restriction was applied.

The same method was followed to match an OPERA IFN β-1a cohort to an NTD IFN β-1a cohort. This pair-matching was performed as part of a feasibility assessment, under the assumption that no significant differences in outcomes should be observed between patients treated with similar IFN β-1a formulations in the two cohorts (null hypothesis). Full comparative analyses were performed only on the condition that this feasibility assessment was positive.

Balance assessment. Quality assessment for PSM included comparison of summary statistics for the baseline covariates between the OPERA and NTD registry pre- and post-matching for each DMT and each outcome, and assessment of standardized mean difference (SMD) of the confounder distributions. Additionally, density functions of PS distribution were created for each DMT and each outcome to show the distribution balance of covariates for each comparison cohort before and after PSM.

Outcomes

The primary outcomes were time to first relapse and time to onset of 24-week (24W-) CDP in post-matching cohorts. For every NTD-matched cohort, frequencies of DMT switches, reasons for index-therapy discontinuation, and time to index-therapy discontinuation were also reported (ie, secondary outcomes).

The baseline for each NTD-matched cohort was defined as the index-therapy initiation date. While there was no specific and predefined clinical visit schedule in the NTD registry, most patients however had clinical visits approximately every 3 months, a frequency that was similar to the OPERA trial. Baseline EDSS for the NTD cohorts was defined as the closest EDSS value to the index-therapy start within a window of ±3 months.

In OPERA, clinical relapses were reported in monthly patient phone calls, and confirmed as protocol-defined relapses confirmed via systematic neurologic examinations. In the NTD registry, only medically confirmed relapses were considered for the analysis. In NTD, a relapse can be reported by the patient during a scheduled visit, or at any time, to the NTD neurologist, which triggers a visit to assess and document the relapse. Thereafter, it will be recorded if medically confirmed in the NTD MS registry database. Additionally, relapses documented in letters from hospital inpatient treatments are captured by the NTD treating physicians. Overall, this aligns with the OPERA clinically confirmed definition of relapses and limits the risk of introducing a differential bias in the outcome ascertainment between the two studies, which could arise from a systematic difference in soliciting and recording information on relapses (eg, patient-reported relapses only).

Statistical analysis

Feasibility analyses comparing IFN β-1a cohorts were conducted over a period of 96 weeks (2 years) as this was limited to the double-blind period of OPERA. The main analyses were conducted using the data of patients randomized to OCR during the phase 3 trial and their 3-year follow-up in the OLE phase,[23] leading to 288 weeks (5.5 years) of total follow-up.

Primary outcomes. CDP events were defined as ≥24 weeks confirmed EDSS increases of ≥.5 points for patients with a

Table 1. Propensity score-matching factors.

OUTCOME / CATEGORY OF VARIABLE	AGE	SEX	TIME SINCE SYMPTOM ONSET	EDSS (BASELINE)[A]	RELAPSES IN PREVIOUS YEAR	PREVIOUS TREATMENT	GD+ LESIONS
	CONTINUOUS	DICHOTOMOUS(1)	ORDINAL(2)	CONTINUOUS	CONTINUOUS	DICHOTOMOUS(3)	DICHOTOMOUS(4)
Time to first relapse	X	X	X	X	X	X	X
Time to first relapse (sensitivity analysis)	X	X	X	X	X	X	
Time to 24W-CDP	X	X		X	X	X	

Based on Laplaud D-A *et al. Neurology* 2019[11].
24W-CDP, 24-week confirmed disability progression; EDSS, Expanded Disability Status Scale; Gd+, gadolinium-enhancing.
[a]EDSS or MRI measurements collected within a window of ±3 months relative to index-therapy initiation.
(1) Sex: Male vs Female; (2) Time since symptom onset: ≤3 years, >3 to ≤5 years, >5 to ≤10 years, >10 years; (3) Previous treatment: Yes vs No; (4) Gd+ lesions: Present vs Not present.

baseline EDSS score >5.5, and ≥1.0 points for those with a baseline EDSS score between .0 and 5.5, inclusive. Progression could be confirmed at any EDSS assessment, including during the 30 days from a protocol-defined relapse. Patients who discontinued from the OPERA study with reasons "lack of efficacy" or "death" during the double-blind period or the OLE period were imputed to have a CDP or a relapse event at the time of treatment discontinuation. All follow-up data were used for every NTD cohort irrespective of treatment. NTD patients that discontinued from the registry were censored at time of discontinuation. They were also censored when switching to OCR/RTX. Patients without a relapse or CDP event were administratively censored at 288 weeks, or at the end of follow-up, whichever came first. For patients in NTD and OPERA with an initial CDP event recorded close to the end of study, the next EDSS assessment outside of the 288-week window was considered for confirmation of the progression event (CDP).

This study compared patients receiving OCR treatment in a clinical trial setting with patients following real-world treatment paths defined by an index therapy (eg, IFN β-1a, GA, etc.) from the NTD registry. For OCR cohorts, data until treatment discontinuation were used, with outcome imputation based on the reason for discontinuation, approximating the hypothetical outcome if all patients would have completed the 288-week OCR treatment period. For NTD cohorts, an intention-to-treat (ITT) approach was applied, considering all available information collected following the start of index therapy, including data related to the outcomes of interest following treatment switch. Thus, NTD cohorts were created for each index therapy allowing that the subsequent evolution of DMT treatment pathway, including any DMT switch, was unique to each patient.

Time to first relapse and time to onset of 24W-CDP between treatment cohorts were evaluated using a Cox proportional hazard model with cohort status as the only covariate, and taking the clusters induced by matching of one OCR patient to one NTD patient into consideration via the application of the robust sandwich variance estimator when drawing inferences. The assumption of proportional hazards was verified using tests for Schoenfeld residuals and using graphical methods. Treatment effects were reported as hazard ratios (HRs) derived from Cox models, along with 95% confidence intervals (CIs) for two-sided tests. Kaplan–Meier plots were provided for each endpoint. Analyses were exploratory with no adjustment for multiple comparisons applied.

Secondary outcomes. The number/proportion of therapy switches and reasons for treatment discontinuation/switches in the NTD registry were reported in a descriptive manner. Sankey diagrams were used to visualize treatment pathways starting from the NTD index therapy, and all subsequent switches to different treatment lines are shown. Four different states are captured in the diagram: (1) "Therapy is ongoing" if the patient remained on the index therapy at the time of the cut-off date (July 1, 2020); (2) "Switched to another DMT" if the patients had documented switches to a subsequent DMT, with multiple switches being allowed; (3) "No further treatment line" if patient treatment data could not be further described because patients were lost to follow-up, were known to have moved to a clinic outside of the NTD network, or had no recorded clinical visit for ≥12 months at the time of the cut-off date (ie, patients with documented ongoing therapies without any recent NTD interactions who likely left the NTD network but this was not properly documented); (4) "Supportive care" if the patient had a period of ≥6 months of known absence of DMT treatment. Supportive care treatments shorter than 6 months may represent a wash-out period between therapies and therefore are not shown in the Sankey diagrams.

The time to index-therapy discontinuation was also reported for every NTD cohort (post-matching, for all outcomes). A survival analysis for therapy discontinuations was used with censoring at the cut-off date (July 1, 2020) if patients remained on the index therapy.

Data availability

Roche trials. For eligible studies qualified researchers may request access to individual patient-level clinical data through a data request platform. At the time of writing, this request platform is Vivli. https://vivli.org/ourmember/roche/. For up-to-date details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see here: https://go.roche.com/data_sharing. Anonymized records for individual patients across more than one data source external to Roche cannot, and should not, be linked due to a potential increase in risk of patient re-identification.

NTD registry. The data used in this study are owned by the NeuroTransData registry and sharing of the data is subject to their policies. Any reasonable requests for data access can be directed to the NeuroTransData registry by email to info@neurotransdata.com.

Patient consent. All patients in the OPERA trials provided written informed consent. Patients from NTD included in this analysis provided their informed consent (via paper forms, tablets in NTD practices, electronic questionnaires, or via a patient portal) to the NTD registry.

Results

Unadjusted (before PSM) patient characteristics of the study cohorts are shown in Supplemental Table S1. A total of 611 OCR patients from OPERA (mean age: 36.7 years; 64.6% female) and 7141 patients from NTD were included in the analysis. In the NTD registry, DMF represented the largest cohort (n = 1735; mean age: 40.7 years; 74.4% female), followed

by FTY (n = 1577; mean age: 40.6 years; 72.7% female), GA (n = 1265; mean age: 39.3 years; 77.4% female), TERI (n = 1197; mean age: 45.3 years; 70.2% female), IFN β-1a (n = 792; mean age: 37.8 years; 75.5% female), and NTZ (n = 575; mean age: 36.9 years; 74.3% female). Time since symptom onset, the proportions of patients with previous treatment with DMTs, relapses in year prior to index therapy, and EDSS scores at baseline differed significantly between the OCR OPERA and NTD unmatched cohorts.

After PSM, different comparative cohorts were built for each DMT pair and each outcome respectively (a total of 12 matching groups). The matching procedure resulted in cohorts that were overall well balanced as indicated by the small SMDs between the matching factors after PSM (all SMDs <.2). Love plot for covariate balance measure by SMD after PSM are shown in Figure 1. The density functions of PS before and after matching are shown in Supplemental Figures S1 and S2. Patient demographic and clinical characteristics at baseline for each post-matching comparative cohort are shown in Table 2.

Baseline demographics and clinical characteristics of matched cohorts

Relapse outcome analysis. The number of patients included in the OCR OPERA and DMT NTD pair-matched cohort ranged from 111:111 (OCR OPERA vs TERI NTD) to 185:185 (OCR OPERA vs FTY NTD). Patients included in the OCR OPERA vs IFN β-1a, GA, and NTZ NTD-matched cohorts were younger (mean age: ≤38 years) and ≥70% were female. Patients in the OCR OPERA vs TERI NTD-matched cohort were slightly older (mean age: ~41 years) than patients in other matched cohorts. Patients included in the OCR OPERA vs IFN β-1a and GA NTD cohorts were mostly treatment naive (≥65%) and had the shortest disease duration. Patients included in the OCR OPERA vs NTZ and FTY NTD-matched cohorts were mostly previously treated (≥70%) and exhibited higher EDSS scores compared with all other cohorts. The OCR OPERA vs DMF and TERI NTD-matched cohorts had a more balanced distribution of treatment-naive and previously treated patients. Patients included in the NTZ NTD-matched cohort had higher proportions of Gd+ lesions than patients in other matched cohorts.

Disability progression outcome analysis. The number of patients included in the OCR OPERA and DMT NTD pair-matched cohort ranged from 200:200 (OCR OPERA vs NTZ NTD) to 331:331 (OCR OPERA vs GA NTD). Patients included in the OCR OPERA vs GA and NTZ NTD-matched cohorts were younger (mean age: ≤38 years) and ≥70% were female while patients in the OCR OPERA vs TERI NTD-matched cohort were slightly older (mean age: ~40 years). Patients included in the OCR OPERA vs IFN β-1a NTD cohorts were mostly treatment naive while those in the OCR OPERA vs NTZ NTD-matched cohort were mostly previously treated. The other OCR OPERA vs NTD-matched cohorts had a more balanced distribution of treatment-naive and previously treated patients. Patients included in the OCR OPERA vs GA NTD cohort had the shortest disease duration. Finally, patients in the OCR OPERA vs NTZ and FTY NTD-matched cohorts had higher EDSS scores compared with all other cohorts.

Validation of PSM approach and exchangeability of clinical trial and real-world datasets

As described, to assess the validity of the PSM approach used in this study and the exchangeability of the two datasets, primary outcomes were compared between the matched IFN β-1a cohorts from OPERA and NTD. For both outcome comparisons, cohorts were well balanced after matching (Supplemental Figure S3). Over a 96-week period, no significant difference was observed for time to first relapse (HR = .76 [95% CI .52-1.13], P = .174) or time to 24W-CDP (HR = .80

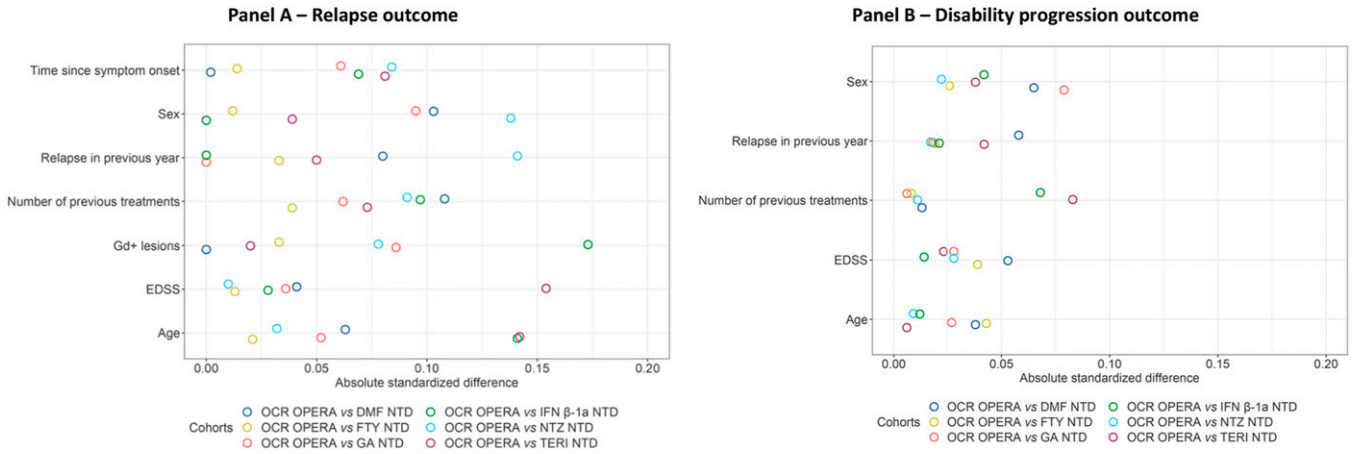

Figure 1. Love plots for absolute standardized differences after PSM for matching factors defined for relapse and 24W-CDP outcomes. 24W-CDP, 24-week confirmed disability progression; DMF, dimethyl fumarate; EDSS, Expanded Disability Status Scale; FTY, fingolimod; GA, glatiramer acetate; Gd+, gadolinium-enhancing; IFN β-1a, interferon β-1a; NTD, NeuroTransData registry; NTZ, natalizumab; OCR, ocrelizumab; PSM, propensity score-matched; TERI, teriflunomide.

Table 2. Baseline demographics and clinical characteristics of matched study cohorts.

PANEL A – RELAPSE OUTCOME													
CHARACTERISTIC	N	DMF (NTD)	OCR (OPERA)	FTY (NTD)	OCR (OPERA)	GA (NTD)	OCR (OPERA)	NTZ (NTD)	OCR (OPERA)	IFN β-1A (NTD)	OCR (OPERA)	TERI (NTD)	OCR (OPERA)
		167	167	185	185	169	169	130	130	118	118	111	111
Age (years)	Mean (SD)	39.6 (10.1)	39.1 (8.5)	38.3 (10.0)	38.5 (9.3)	37.4 (11.0)	37.9 (9.1)	37.0 (9.9)	37.3 (9.1)	38.1 (10.6)	36.8 (8.3)	42.5 (10.5)	41.1 (8.8)
EDSS	Mean (SD)	2.4 (1.6)	2.4 (1.2)	2.9 (1.6)	2.9 (1.3)	2.2 (1.6)	2.2 (1.1)	2.8 (1.7)	2.8 (1.3)	2.2 (1.5)	2.2 (1.2)	2.5 (1.8)	2.3 (1.1)
Previous treatment, *n* (%)	No	93 (55.7)	84 (50.3)	41 (22.2)	44 (23.8)	113 (66.9)	108 (63.9)	33 (25.4)	28 (21.5)	85 (72.0)	90 (76.3)	66 (59.5)	62 (55.9)
	Yes	74 (44.3)	83 (49.7)	144 (77.8)	141 (76.2)	56 (33.1)	61 (36.1)	97 (74.6)	102 (78.5)	33 (28.0)	28 (23.7)	45 (40.5)	49 (44.1)
Relapses in year prior to therapy	Mean (SD)	.9 (.6)	.9 (.4)	1.1 (.7)	1.1 (.6)	1.0 (.6)	1.0 (.4)	1.0 (.7)	1.1 (.6)	.9 (.6)	.9 (.5)	.8 (.6)	.9 (.5)
Sex, *n* (%)	Male	49 (29.3)	57 (34.1)	55 (29.7)	54 (29.2)	40 (23.7)	47 (27.8)	40 (30.8)	32 (24.6)	32 (27.1)	32 (27.1)	34 (30.6)	32 (28.8)
	Female	118 (70.7)	110 (65.9)	130 (70.3)	131 (70.8)	129 (76.3)	122 (72.2)	90 (69.2)	98 (75.4)	86 (72.9)	86 (72.9)	77 (69.4)	79 (71.2)
Time since symptom onset (years)	Mean (SD)	8.9 (8.4)	7.9 (6.0)	8.4 (6.8)	8.1 (5.5)	5.9 (7.2)	6.5 (6.1)	8.6 (6.9)	8.7 (5.5)	6.9 (8.4)	6.3 (5.9)	8.8 (7.9)	8.1 (6.0)
Gd+ lesions, *n* (%)	Present	58 (34.7)	58 (34.7)	69 (37.3)	72 (38.9)	59 (34.9)	66 (39.1)	58 (44.6)	53 (40.8)	43 (36.4)	53 (44.9)	29 (26.1)	30 (27.0)
	Not present	109 (65.3)	109 (65.3)	116 (62.7)	113 (61.1)	110 (65.1)	103 (60.9)	72 (55.4)	77 (59.2)	75 (63.6)	65 (55.1)	82 (73.9)	81 (73.0)

Panel B – Disability progression outcome													
Characteristic	*N*	DMF (NTD)	OCR (OPERA)	FTY (NTD)	OCR (OPERA)	GA (NTD)	OCR (OPERA)	NTZ (NTD)	OCR (OPERA)	IFN β-1a (NTD)	OCR (OPERA)	TERI (NTD)	OCR (OPERA)
		328	328	252	252	331	331	200	200	264	264	222	222
Age (years)	Mean (SD)	38.7 (10.4)	38.3 (8.7)	38.5 (10.4)	38.9 (9.1)	37.6 (10.8)	37.9 (8.8)	37.2 (9.9)	37.1 (9.4)	38.1 (10.0)	38.2 (8.9)	40.4 (10.9)	40.3 (8.7)
EDSS	Mean (SD)	2.5 (1.7)	2.6 (1.2)	2.7 (1.6)	2.8 (1.3)	2.5 (1.8)	2.5 (1.2)	2.9 (1.6)	3.0 (1.4)	2.4 (1.8)	2.4 (1.2)	2.5 (1.8)	2.5 (1.2)
Previous treatment, *n* (%)	No	211 (64.3)	213 (64.9)	109 (43.3)	110 (43.7)	221 (66.8)	220 (66.5)	58 (29.0)	57 (28.5)	194 (73.5)	186 (70.5)	139 (62.6)	130 (58.6)
	Yes	117 (35.7)	115 (35.1)	143 (56.7)	142 (56.3)	110 (33.2)	111 (33.5)	142 (71.0)	143 (71.5)	70 (26.5)	78 (29.5)	83 (37.4)	92 (41.4)
Relapses in year prior to therapy	Mean (SD)	1.1 (.6)	1.0 (.4)	1.1 (.7)	1.1 (.6)	1.1 (.6)	1.0 (.4)	1.1 (.6)	1.1 (.6)	1.0 (.6)	1.0 (.5)	1.0 (.6)	1.0 (.4)
Sex, *n* (%)	Male	102 (31.1)	112 (34.1)	79 (31.3)	76 (30.2)	95 (28.7)	107 (32.3)	60 (30.0)	62 (31.0)	79 (29.9)	74 (28.0)	73 (32.9)	77 (34.7)
	Female	226 (68.9)	216 (65.9)	173 (68.7)	176 (69.8)	236 (71.3)	224 (67.7)	140 (70.0)	138 (69.0)	185 (70.1)	190 (72.0)	149 (67.1)	145 (65.3)
Time since symptom onset (years)	Mean (SD)	8.5 (8.5)	7.0 (5.9)	8.7 (7.5)	7.7 (5.7)	6.3 (6.9)	6.9 (5.7)	8.9 (7.2)	7.8 (5.5)	7.3 (8.4)	6.9 (6.1)	9.1 (9.0)	7.2 (6.0)

DMF, dimethyl fumarate; EDSS, Expanded Disability Status Scale; FTY, fingolimod; GA, glatiramer acetate; Gd+, gadolinium-enhancing; IFN β-1a, interferon β-1a; NTD, NeuroTransData registry; NTZ, natalizumab; OCR, ocrelizumab; SD, standard deviation; TERI, teriflunomide.

[95% CI .53-1.20], P = .279), suggesting an acceptable exchangeability of the two datasets (Supplemental Figure S4). However, a numerical trend of fewer relapses and 24W-CDP events was observed for patients treated with IFN β-1a in the OPERA cohort.

Time to first relapse

Treatment with OCR was associated with a statistically significant reduction in time to first relapse vs any of the treatment pathways regardless of the index therapy (Table 3, Figure 2). Relative to patients where injectable DMTs (IFN β-1a and GA) were the index therapy, treatment with OCR resulted in a significant reduction of 66-70% in the risk of relapses. By the end of the follow-up period (5.5 years), only 32-34% of patients were still on treatment with the index therapies IFN β-1a and GA, respectively (Figure 3). Compared with patients treated with oral index therapies (TERI, DMF, FTY), treatment with OCR led to a significant reduction of 49-55% in relapse activity (Table 3); higher persistence was observed relative to injectable DMTs with 41%, 40%, and 58% of patients remaining on the oral index therapy after 5.5 years (Figure 3). While patients starting on NTZ showed better relapse outcomes compared with other DMTs, treatment with OCR was still associated with a significant 46% reduction in risk to first relapse (Table 3). After 5.5 years only 49% of patients remained on NTZ (Figure 3), with most patients de-escalating to lower-efficacy DMTs (Supplemental Figure S5, panel D).

Treatment with OCR was still associated with a statistically significant reduction in time to first relapse vs any of the treatment pathways, in a sensitivity analysis where presence/absence of Gd+ lesions was not used as a matching factor (Supplemental Table S2, panel B).

Time to 24-week confirmed disability progression

Treatment with OCR was associated with a reduction in time to 24W-CDP vs any of the treatment pathways regardless of the index therapy (Table 3). The survival curves for probability of disability progression of each treatment pathway vs OCR are shown in Figure 4. Relative to patients where injectable DMTs (IFN β-1a and GA) were the index therapy, treatment with OCR resulted in a significant reduction of 33-49% in disability progression (Table 3). Similar to the analyses for the relapse outcome, by the end of the follow-up period (5.5 years), only 28% of patients were still on treatment with the index therapies IFN β-1a and GA, respectively (Figure 5). Compared with patients treated with oral index therapies, treatment with OCR led to a significant reduction of 34% (vs FTY), 36% (vs DMF), and 39% (vs TERI) in disability progression (Table 3); however, after 5.5 years, only 44% and 45% were still treated with the index therapies DMF and TERI, respectively, with most escalating to higher-efficacy therapies (Supplemental Figure S6, panels A and F). Higher persistence was observed with 57% of patients remaining on the FTY index therapy after 5.5 years. A marginally significant difference in favor of OCR on time to disability progression (34% risk reduction, P = .048) was also observed compared with patients on a NTZ treatment pathway (Table 3). After 5.5 years only 35% of patients remained on NTZ (Figure 5), with most patients de-escalating to lower-efficacy DMTs (Supplemental Figure S6, panel D).

Treatment pathway discontinuations and treatment switching for index-matched cohorts

After an observation period of 5.5 years (288 weeks), over 50% of patients in the NTD registry had discontinued their initial index therapy for each outcome analysis except for the FTY cohort (Figures 4 and 5). Overall, across different comparison cohorts, IFN β-1a and GA showed the lowest persistence (28-34%) whereas FTY showed the highest persistence, with approximately 58% and 57% of patients remaining on index therapy after 5.5 years for the relapse and disability cohorts, respectively (Figures 4 and 5). Moreover, the median time to treatment discontinuation (ie, time by which 50% of patients discontinued the treatment) of injectable index therapies and DMF occurred relatively early after initiation (IFN β-1a, 2.4 and 2.9 years for relapse and disability outcomes analyses, respectively; GA, 3.3 and 2.4 years, respectively; DMF, 2.5 and 3.5 years, respectively; Figures 4 and 5). Relative to injectable index therapies or DMF, the median time to treatment discontinuation was delayed in cohorts where NTZ and TERI were the index therapies, in particular in the relapse cohorts (5.2 and 3.9 years, respectively).

Injectable DMTs were more likely to be discontinued due to lack of efficacy (IFN β-1a, 30.1% and 35.2%; GA, 39.2% and 32.4% in the relapse and disability cohorts, respectively; Table 4), DMF and TERI were more likely to be discontinued due to side effects in both relapse and disability cohorts (DMF, 39.8% and 31.8%; TERI, 42.1% and 38.1%, respectively). For NTZ and FTY, lack of efficacy was less likely to lead to discontinuation, in particular in the disability cohort (NTZ, 8.2%; FTY, 19.2%). A positive JCV index (reason for discontinuation recorded as "Antibodies") was one of the main reasons for NTZ discontinuation. However, reasons for discontinuation should be interpreted with caution as discontinuation for other reasons ranged from 8.9% (TERI) to 37.4% (DMF) and 13.2% (TERI) to 39.2% (DMF) for the relapse and disability outcome analyses, respectively. For additional details refer to Table 4.

In general, in the pathways starting with a lower-efficacy DMT (IFN β-1a, GA, TERI, DMF) a switching pattern towards an escalating strategy to higher-efficacy DMTs was observed, whereas for higher-efficacy index therapies (NTZ and FTY) a de-escalating pattern was observed (Supplemental Figures S4 and S5).

Table 3. Effectiveness of OCR in OPERA studies compared with index therapies in NTD-matched cohorts over 288 weeks.

OCR VS INDEX-THERAPY LINE IN NTD REGISTRY	TIME TO FIRST RELAPSE				TIME TO 24W-CDP			
	N	HR	95% CI	*P*-VALUE	N	HR	95% CI	*P*-VALUE
DMF	167:167	.45	.31-.64	<.001	328:328	.64	.44-.92	.015
FTY	185:185	.47	.34-.66	<.001	252:252	.66	.44-.98	.039
GA	169:169	.34	.22-.51	<.001	331:331	.67	.47-.95	.026
NTZ	130:130	.54	.38-.76	<.001	200:200	.66	.43-1.00[a]	.048
IFN β-1a	118:118	.30	.19-.47	<.001	264:264	.51	.36-.72	<.001
TERI	111:111	.51	.31-.83	.007	222:222	.61	.39-.95	.029

24W-CDP, 24-week confirmed disability progression; CI, confidence interval; DMF, dimethyl fumarate; FTY, fingolimod; GA, glatiramer acetate; HR, hazard ratio; IFN β-1a, interferon β-1a; NTD, NeuroTransData registry; NTZ, natalizumab; OCR, ocrelizumab; TERI, teriflunomide.
[a]The 3-digit 95% CI was (.431-.996) and the upper limit was rounded to 1.00; however, 95% did not include the 1.0 value.

Discussion

In this study we used PS methods to compare the effectiveness of OCR administered in a phase 3 clinical trial and its OLE phase with matched real-world treatment pathways starting with IFN β-1a, GA, DMF, TERI, FTY, or NTZ. Over a period of up to 5.5 years, treatment with OCR was associated with a significant reduction in time to first relapse and time to 24W-CDP compared with any of these treatment pathways. We also confirm that treatment patterns in the real world are heterogeneous and dynamic, with many DMT switches that may lead to suboptimal clinical outcomes.

PS-based analyses are increasingly being applied in observational MS studies to compare the effectiveness and safety of DMTs.[15,16] However, merging clinical trial and real-world data remains largely unexplored in MS, with only one previous study reporting results that used data from the pivotal trial assessing cladribine tablets vs placebo (CLARITY), and PS-weighted data from the Italian multicenter database, i-MuST.[21] Similar to Signori et al[21] we demonstrated the feasibility of merging interventional and observational datasets, in our case by comparing matched IFN β-1a cohorts from the OPERA trial and the NTD registry and showing no significant difference between the two cohorts for the study outcomes. Some notable differences can be identified between the two studies. We used PSM whereas Signori et al opted for using the PS to calculate inverse probability of treatment weighting (IPTW). Whilst IPTW with stabilized trimmed weights seems to be more popular in the MS literature,[11,21,30] a recent study has shown that there are no important differences between conclusions obtained with PSM or PS weighting as long as a study is sufficiently powered, models are correctly specified, and positivity assumption is fulfilled.[31] Perhaps a more important difference is that patients in i-MuST were censored at treatment switch, whereas in our study all matched patients were retained in the groups as initial DMT allocation regardless of subsequent switches to other DMTs (ITT analysis). An ITT framework has greater advantages as it mitigates the risk of informed censoring,[31] and it allows to fully characterize long-term treatment pathways observed in clinical practice.[20]

A high diversity of treatment pathways could in fact be observed in the NTD registry, characterized by frequent switching/sequencing and discontinuation over a period of 5.5 years. By the end of the study, less than half of patients in the different NTD-matched cohorts were still being treated with the index therapy, with a higher proportion of switchers reported on platform injectable DMTs (IFN β-1a, GA), and FTY and NTZ associated with higher persistence. This is consistent with a recent study looking at real-world DMT pathways using claims administrative data, which observed that 51.3% and 26.5% of patients had evidence of a second and a third DMT over a follow-up period of 2.0-10.5 years, respectively.[20] Overall, in our study lack of efficacy was the most common reason for stopping a DMT followed by side effects and lack of tolerability, which is consistent with the results from a recent Big MS Data Network study.[32] Individual patterns reported for each DMT are also in line with previous literature (eg, [gastrointestinal] intolerability is a primary reason for DMF discontinuation, whereas a positive JCV index is the most common reason for NTZ discontinuation).[33-35]

The most important finding of our study is that patients treated with OCR achieved better clinical outcomes than matched patients in the different treatment pathways. Over a 5.5-year period, a significant risk reduction of relapse activity and disability progression was observed for OCR patients compared with all other DMTs (although the 34% risk reduction of 24W-CDP associated with OCR compared with NTZ treatment pathway was marginally significant, $P = .048$). Interestingly, the risk reduction of 24W-CDP relative to the matched cohort with NTZ as index therapy (34% risk reduction) was similar to that observed for GA (33% risk reduction). While this result may appear paradoxical, it is likely explained by the observed treatment pathways. Over a period of 5.5 years, more than 50% of patients had discontinued NTZ and GA; however, while patients discontinued NTZ mostly due to safety

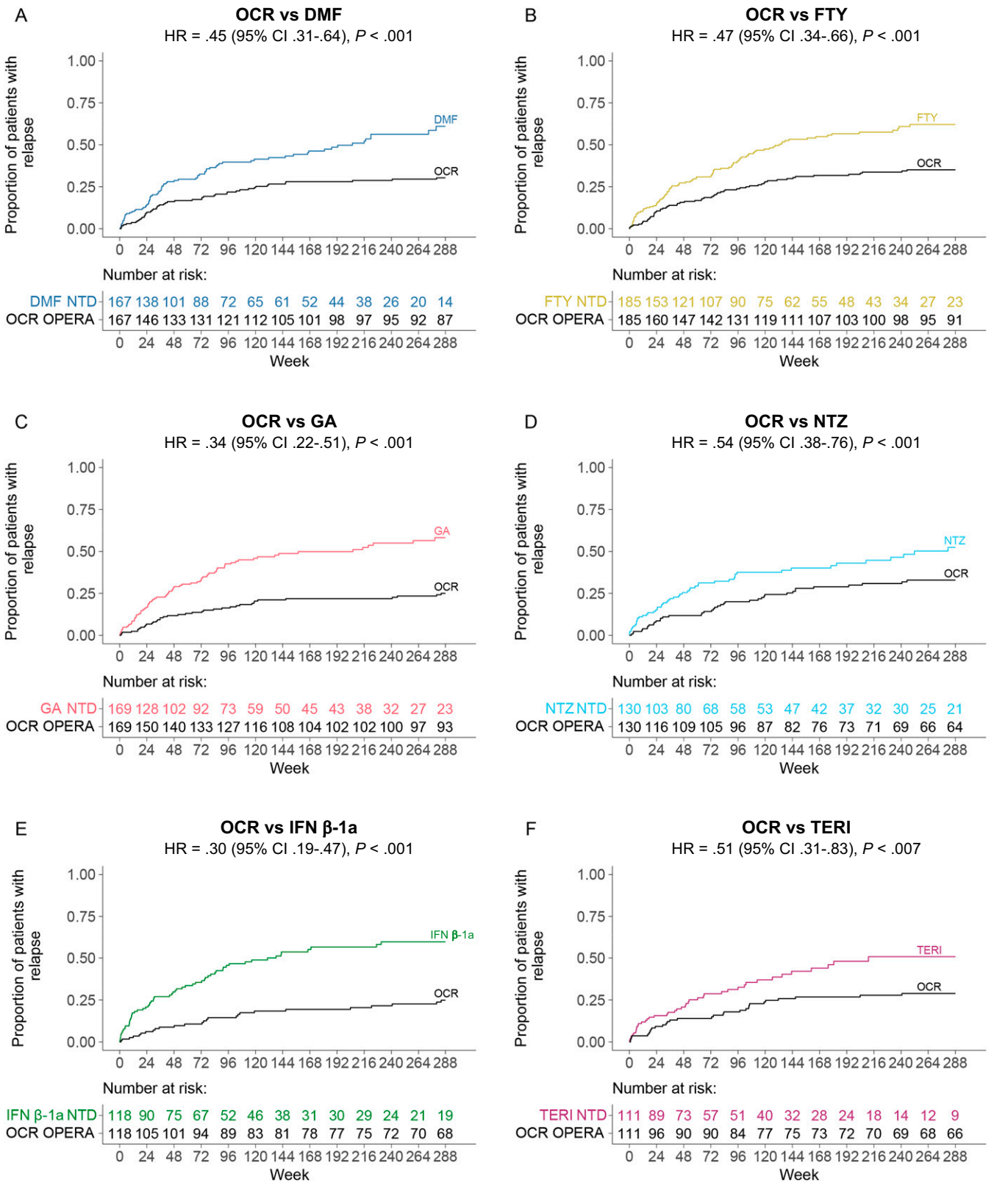

Figure 2. Time to first relapse in patients receiving OCR in OPERA vs patients receiving each index therapy in the NTD-matched cohorts over 288 weeks. CI, confidence interval; DMF, dimethyl fumarate; FTY, fingolimod; GA, glatiramer acetate; HR, hazard ratio; IFN β-1a, interferon β-1a; NTD, NeuroTransData registry; NTZ, natalizumab; OCR, ocrelizumab; TERI, teriflunomide.

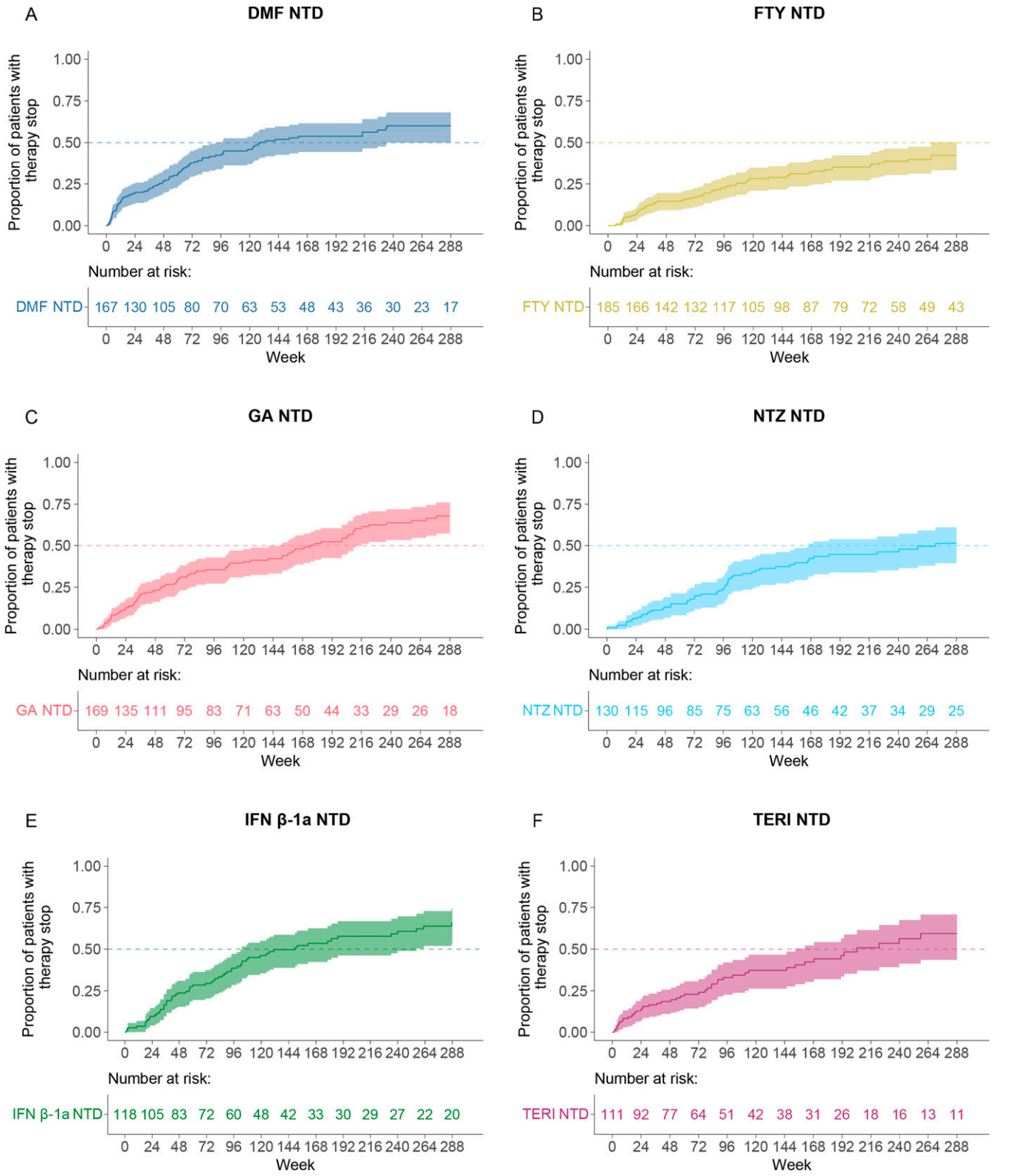

Figure 3. Time to discontinuation of each index therapy in the NTD-matched cohorts for time to first relapse outcome over 288 weeks. DMF, dimethyl fumarate; FTY, fingolimod; GA, glatiramer acetate; IFN β-1a, interferon β-1a; NTD, NeuroTransData registry; NTZ, natalizumab; TERI, teriflunomide.

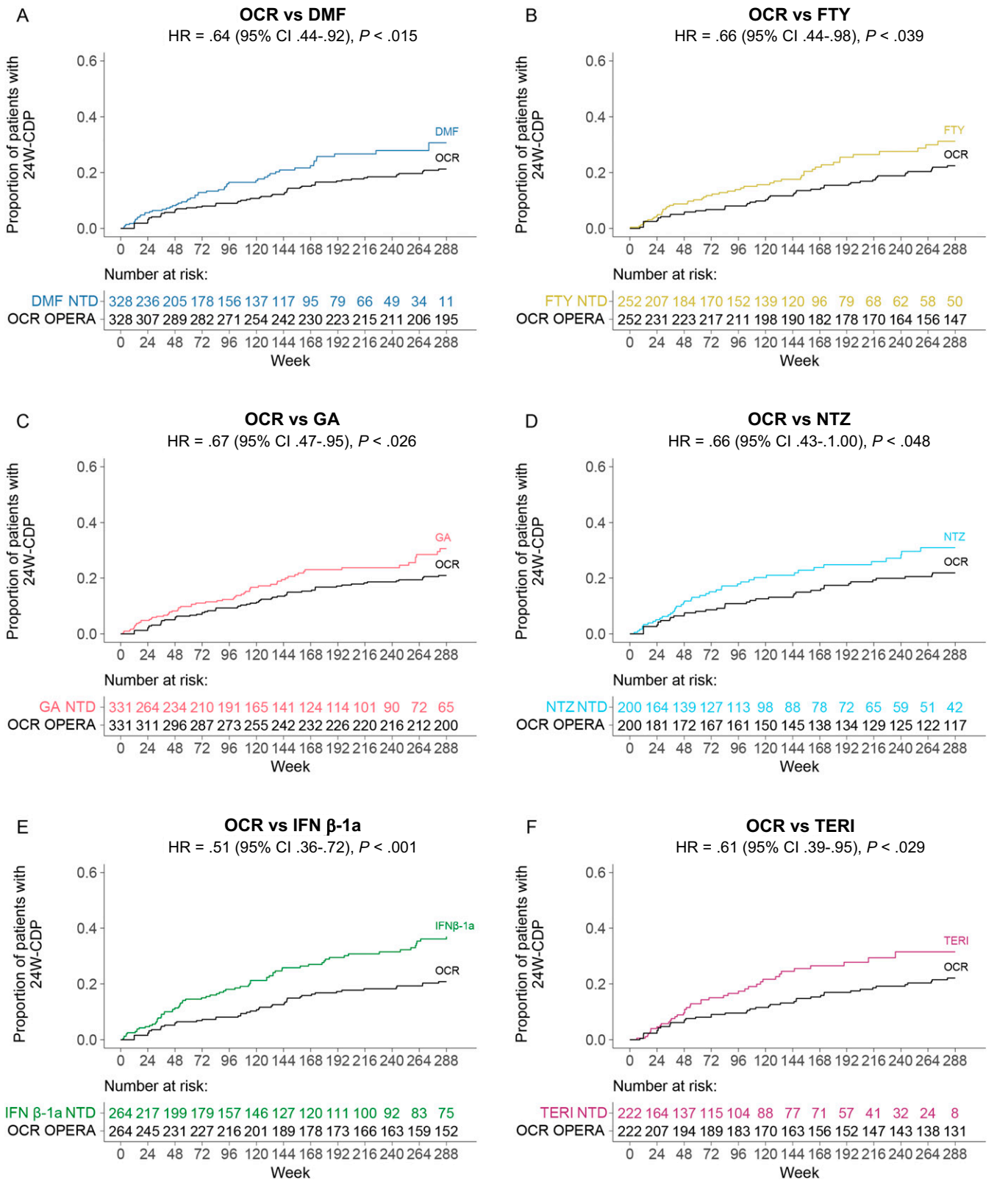

Figure 4. Time to 24W-CDP relapse in patients receiving OCR in OPERA vs patients receiving each index therapy in the NTD-matched cohorts over 288 weeks. 24W-CDP, 24-week confirmed disability progression; CI, confidence interval; DMF, dimethyl fumarate; FTY, fingolimod; GA, glatiramer acetate; HR, hazard ratio; IFN β-1a, interferon β-1a; NTD, NeuroTransData registry; NTZ, natalizumab; OCR, ocrelizumab; TERI, teriflunomide.

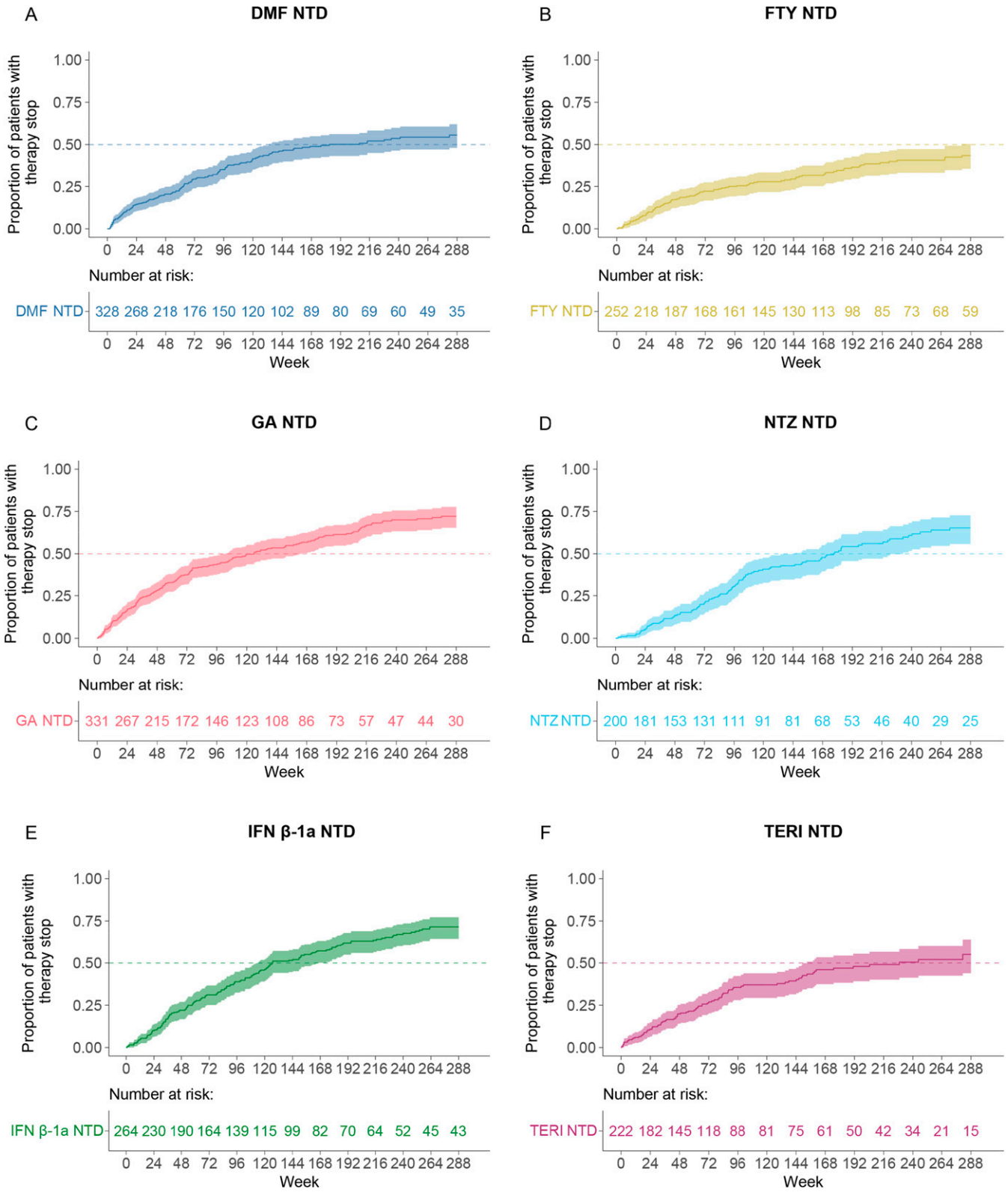

Figure 5. Time to discontinuation of each index therapy in the NTD-matched cohorts for time to 24W-CDP outcome over 288 weeks. 24W-CDP, 24-week confirmed disability progression; DMF, dimethyl fumarate; FTY, fingolimod; GA, glatiramer acetate; IFN β-1a, interferon β-1a; NTD, NeuroTransData registry; NTZ, natalizumab; TERI, teriflunomide.

Table 4. Discontinuation patterns and reasons of each index therapy in NTD-matched cohorts for time to first relapse outcome (panel A) and time to 24W-CDP (panel B).

PANEL A – RELAPSE OUTCOME						
DISCONTINUATION PATTERNS	DMF	FTY	GA	NTZ	IFN β-1A	TERI
No. patients with index therapy discontinued (%)	100 (60.1)	78 (42.4)	115 (68.0)	67 (51.3)	77 (65.6)	66 (59.4)
Therapy discontinuation due to[a]						
Lack of efficacy	27.8	20.5	39.2	20.0	30.2	35.5
Lack of compliance	20.5	4.7	10.7	8.1	23.8	2.2
Side effects	39.8	23.8	15.5	8.1	20.6	42.1
Pregnancy	0.0	15.8	3.6	9.9	3.2	0.0
Antibodies	1.2	0.0	0.0	16.1	0.0	0.0
Non-specified reasons	37.4	22.3	17.9	28.1	20.6	8.9
Fear of needle	na.	na.	3.6	na.	1.5	na.
Patient wish	19.3	22.3	13.1	9.9	14.2	13.3

Panel B – Disability progression outcome						
Discontinuation Patterns	DMF	FTY	GA	NTZ	IFN β-1a	TERI
No. patients with index therapy discontinued (%)	182 (55.6)	110 (43.5)	239 (72.2)	131 (65.3)	189 (71.5)	122 (55.1)
Therapy discontinuation due to[a]						
Lack of efficacy	28.3	19.2	32.4	8.2	35.2	28.6
Lack of compliance	21.6	10.3	7.8	8.2	18.9	4.8
Side effects	31.8	20.3	17.3	3.1	10.6	38.1
Pregnancy	2.9	11.5	5.7	13.3	7.0	4.8
Antibodies	0.0	0.0	0.0	18.4	0.0	0.0
Non-specified reasons	39.2	30.7	22.0	32.7	25.8	13.2
Fear of needle	na.	na.	1.0	0.0	1.8	0.0
Patient wish	17.3	16.1	15.3	16.3	13.8	13.2

24W-CDP, 24-week confirmed disability progression; DMF, dimethyl fumarate; DMT, disease-modifying therapy; FTY, fingolimod; GA, glatiramer acetate; IFN β-1a, interferon β-1a; na., not applicable; NTD, NeuroTransData registry; NTZ, natalizumab; OCR, ocrelizumab; TERI, teriflunomide.
[a]Reasons for DMT discontinuation are not mutually exclusive and more than one discontinuation reason can be recorded for a single patient.

concerns and were thus more likely to de-escalate to a less efficacious DMT (NTZ > FTY > DMF), patients discontinuing GA did so for lack of efficacy and were therefore more likely to escalate to a more efficacious DMT (FTY > DMF > NTZ). These results support the concept that the best clinical outcomes can be achieved through a combination of high efficacy[36-39] and good treatment persistence (associated with adequate tolerability).

In our study, most patients in the OCR OPERA cohort remained on continuous treatment for substantially longer than matched patients in NTD cohorts on initial comparator DMTs. While this reflects partially the interventional study design that enforces permanence in the trial in particular during the double-blind period, treatment with OCR in a real-world setting is typically characterized by high persistence rates, as seen in a number of recent large studies.[40,41] Reports from CONFIDENCE, an ongoing multicenter, non-interventional post-authorization safety study showed a 92% persistence over 2 years in patients with RMS and PPMS.[42,43] In a recent MSBase study comparing switch to OCR, cladribine, or NTZ after FTY treatment cessation in patients with relapsing-remitting MS, OCR users had an 89% lower discontinuation rate than NTZ.[30] In this study, OCR treatment was associated with a significant 33% risk reduction on time to first relapse relative to NTZ, which is consistent with our results. There was no significant difference in the cumulative hazard of disability accumulation between the OCR and NTZ users (IPTW HR, .81; 95% CI, .49 to 1.35, *P*-value, .42), while our study found marginally significant risk reduction (PSM HR, .66; 95% CI, .43 to 1.00, *P*-value, .048). This result as well as the lower risk reduction estimate on relapses (33% in MSBase vs 46% in our study) may be explained by the methodologic approach (per protocol in MSBase vs ITT in our study) and by the shorter follow-up duration in the MSBase (2.0 years in MSBase vs 5.5 years in our study).

Strengths and limitations

This study uses PSM methods that allow to control for important covariates but inevitably result in the selection of a narrower RMS population thus compromising the generalizability of results.[15,16] Overall, a relatively high number of

patients were included in each matched cohort. However, patients not included in the PSM analyses had specific but notable differences compared with those included. For example, our results may not be generalizable to patients with lower relapse activity in the year before study start, as these patients were overall not included in the matched cohorts. Results may also not be generalizable to PSM-excluded patients on injectable therapies who had in general slightly lower disability levels at baseline than the PSM-included patients. The same applies to PSM-excluded NTD patients with index therapies DMF and TERI who were in general 2-4 years older, with slightly lower disability levels at baseline, and more frequently previously treated with a DMT whereas the included patients had a more balanced combination of previously treated and treatment naive. Our study can therefore be expected to estimate the comparative effectiveness in a real-world population with characteristics similar to those of a selected interventional clinical trial, rather than overall real-world effectiveness. However, there is still a significant overlap between the OPERA trial and the overall NTD registry, and the treatment effect size observed across multiple comparative cohorts is expected to comfortably exceed the effect of any potential biases.

Another limitation of our PSM approach is that it only controls for balance in the measured confounders at baseline. Bias due to unobserved clinical factors (eg, BMI) or on-therapy confounders at baseline or any other residual confounding may still be present after matching.[44] However, we were able to adjust for the most important factors at baseline as shown by other groups,[10,11,17,21] including for MRI activity, although complete MRI information was only available for about 30% of patients. Indeed, as it was included as a matching factor in the primary analyses for effectiveness in relapse activity, the size of the comparative cohorts was considerably reduced. Nevertheless, we observed similar results in a sensitivity analysis with larger cohorts not matched for presence/absence of Gd+ lesions, which indicates limited indication bias.

A major challenge in our study was to estimate treatment effect in a clinical trial setting while accounting for expected changes in treatment pathways post-index therapy in NTD registry arising from the real-world setting. Due to the relevance for clinical decision-making, we report treatment effects of OCR vs other DMTs reflecting the treatment policy principles.[45,46] Analytically, we applied an ITT approach using all collected information, including outcome data, and essentially ignoring intercurrent events (ie, treatment switch or discontinuation) related to the complexity and diversity of patient journeys across real-world treatment pathways after initiation of the index treatment. Attrition bias may be present in our study due to patients loss to follow-up in the NTD registry, impeding adequate documentation regarding treatment discontinuation or switch. Lost-to-follow-up rates were higher in the NTD registry vs the OPERA trials, resulting in different censoring rates in time-to-event analyses, which may explain part of the treatment effects. However, our ITT analytical approach controls for some of this bias as well as informed censoring.[31]

Finally, in the NTD registry clear and unique relapse definitions are used and all information on EDSS is entered by certified raters leading to overall higher data quality and better comparability with data acquired in the context of clinical trials. The approximate 3-monthly visit schedule (including relapse and EDSS assessment) across all cohorts allows for mitigation of the risk of performance bias and led to minimal data imputation.

Conclusions

We demonstrated that, by using PSM in the context of an ITT analysis, comparison of RCT and high-quality real-world datasets was feasible. Following this approach, we showed the superiority (Class III evidence) of OCR in controlling disease activity and progression compared with complex and dynamic DMT real-world treatment pathways over a 5.5-year period. This strategy using clinical trial data with a high-quality observational cohort can therefore be used to address the knowledge gap arising from the absence of head-to-head clinical trials comparing the clinical efficacy of multiple DMTs for MS. This could prove particularly useful at the time of approval of a new DMT. Future research could explore other methods such as marginal structural model accounting for the time-dependent confounders, including treatment status and allocation at any time point, in order to further assess the effect of DMTs conditional on treatment persistence.[44,47] While the long-term benefits of early use of high-efficacy DMTs have been shown by a number of real-world studies,[36-39] our study provides novel evidence that high treatment persistence, as observed with OCR,[40-43] is also critical for realizing those benefits. However, these observations require confirmation using real-world OCR data acquired in the NTD registry. Data collection is ongoing and comparative effectiveness analyses will be conducted when longer term data become available.

Acknowledgements

We thank all patients, their families, and the investigators who participated in the OPERA trials and NTD registry.

Author contributions

Erwan Muros-Le Rouzic: Conceptualization; Methodology; Writing – original draft; Writing – review & editing; Visualization; Supervision; Project administration of the manuscript. **Yanic Heer:** Conceptualization; Methodology; Software; Validation; Resources; Data curation; Writing – reviewing & editing; Visualization; Supervision; Project administration of the manuscript. **Sean Yiu:** Conceptualization; Writing – reviewing & editing of the manuscript. **Viola Tozzi:** Conceptualization; Methodology; Software; Validation; Formal

analysis; Resources; Data curation; Writing – review & editing; Visualization of the manuscript. **Stefan Braune:** Conceptualization; Methodology; Investigation; Writing – reviewing & editing of the manuscript. **Philip van Hövell:** Conceptualization; Resources; Writing – review & editing; Supervision; Project administration of the manuscript. **Arnfin Bergmann:** Conceptualization; Investigation; Writing – review & editing of the manuscript. **Corrado Bernasconi:** Conceptualization; Methodology; Writing – review & editing of the manuscript. **Fabian Model:** Conceptualization; Methodology; Writing – review & editing of the manuscript. **Licinio Craveiro:** Conceptualization; Methodology; Writing – review & editing; Visualization; Supervision; Project administration of the manuscript.

Ethical Statement

Ethical approval

OPERA trials: The relevant institutional review boards/ethics committees approved the trial protocols (NCT01247324/ NCT01412333) [Hauser SL, Bar-Or A, Comi G, *et al.* Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis. *N Engl J Med* 2017; 376 (3): 221-234].

Informed Consent

All patients provided written informed consent. The trial was conducted in accordance with the International Conference on Harmonisation guidelines for Good Clinical Practice and the Declaration of Helsinki.

ORCID iDs

Erwan Muros-Le Rouzic https://orcid.org/0000-0002-8925-0481
Stefan Braune https://orcid.org/0000-0002-0541-8196

REFERENCES

1. European Medicines Agency. Ocrevus [Summary of product characteristics], https://www.ema.europa.eu/en/documents/product-information/ocrevus-epar-product-information_en.pdf. (2021, accessed January 18, 2021). [Not Available in CrossRef].
2. Genentech. Ocrevus (Ocrelizumab) [Full prescribing information], https://www.gene.com/download/pdf/ocrevus_prescribing.pdf. (2020, accessed January 20, 2020).
3. Hauser SL, Bar-Or A, Comi G, et al. OPERA I and OPERA II Clinical Investigators. Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis. *N Engl J Med*. 2017;376(3):221-234.
4. Asha MZI, Al-Asaad Y, Khalil SFH. The comparative efficacy and safety of anti-CD20 monoclonal antibodies for relapsing-remitting multiple sclerosis: a network meta-analysis. *IBRO Neurosci Rep*. 2021;11:103-111
5. Bose D, Ravi R, Maurya M, Pushparajan L, Konwar M. Impact of disease-modifying therapies on MRI outcomes in patients with relapsing-remitting multiple sclerosis: a systematic review and network meta-analysis. *Mult Scler Relat Disord*. 2022;61:103760
6. Giovannoni G, Lang S, Wolff R, et al. A systematic review and mixed treatment comparison of pharmaceutical interventions for multiple sclerosis. *Neurol Ther*. 2020;9(2):359-374
7. Li H, Hu F, Zhang Y, Li K. Comparative efficacy and acceptability of disease-modifying therapies in patients with relapsing-remitting multiple sclerosis: a systematic review and network meta-analysis. *J Neurol*. 2020;267(12): 3489-3498
8. McCool R, Wilson K, Arber M, et al. Systematic review and network meta-analysis comparing ocrelizumab with other treatments for relapsing multiple sclerosis. *Mult Scler Relat Disord*. 2019;29:55-61
9. Braune S, Rossnagel F, Dikow H, Bergmann ANeuroTransData Study Group. Impact of drug diversity on treatment effectiveness in relapsing-remitting multiple sclerosis (RRMS) in Germany between 2010 and 2018: real-world data from the German NeuroTransData multiple sclerosis registry. *BMJ Open*. 2021;11(8): e042480
10. Hillert J, Tsai JA, Nouhi M, Glaser A, Spelman T. A comparative study of teriflunomide and dimethyl fumarate within the Swedish MS Registry. *Mult Scler*. 2022;28(2):237-246
11. Laplaud D-A, Casey R, Barbin L, et al. Comparative effectiveness of teriflunomide vs dimethyl fumarate in multiple sclerosis. *Neurology*. 2019;93(7):e635-e646
12. Salter A, Lancia S, Cutter G, et al. A propensity-matched comparison of long-term disability worsening in patients with multiple sclerosis treated with dimethyl fumarate or fingolimod. *Ther Adv Neurol Disord*. 2021;14:17562864211021177
13. Trojano M, Tintore M, Montalban X, et al. Treatment decisions in multiple sclerosis - insights from real-world observational studies. *Nat Rev Neurol*. 2017; 13(2):105-118
14. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399-424
15. Austin PC, Yu AYX, Vyas MV, Kapral MK. Applying propensity score methods in clinical research in neurology. *Neurology*. 2021;97(18):856-863
16. Simoneau G, Pellegrini F, Debray TP, et al. Recommendations for the use of propensity score methods in multiple sclerosis research. *Mult Scler*. 2022;28(9): 1467-1480
17. Braune S, Grimm S, van Hövell P, et al. Comparative effectiveness of delayed-release dimethyl fumarate versus interferon, glatiramer acetate, teriflunomide, or fingolimod: results from the German NeuroTransData registry. *J Neurol*. 2018; 265(12):2980-2992
18. Lorscheider J, Benkert P, Lienert C, et al. Comparative analysis of dimethyl fumarate and fingolimod in relapsing-remitting multiple sclerosis. *J Neurol*. 2021; 268(3):941-949
19. Spelman T, Herring WL, Zhang Y, et al. Comparative effectiveness and cost-effectiveness of natalizumab and fingolimod in patients with inadequate response to disease-modifying therapies in relapsing-remitting multiple sclerosis in the United Kingdom. *Pharmacoeconomics*. 2022;40(3):323-339
20. Fox RJ, Mehta R, Pham T, Park J, Wilson K, Bonafede M. Real-world disease-modifying therapy pathways from administrative claims data in patients with multiple sclerosis. *BMC Neurol*. 2022;22(1):211
21. Signori A, Saccà F, Lanzillo R, et al. Cladribine vs other drugs in MS. *Neurol Neuroimmunol Neuroinflamm*. 2020;7(6):e878
22. Hillen J, Ward M, Slee M, et al. Utilisation of disease modifying treatment and diversity of treatment pathways in relapsing remitting multiple sclerosis. *Mult Scler Relat Disord*. 2022;57:103412
23. Hauser S, Kappos L, Arnold D, et al. Five-years of ocrelizumab in relapsing multiple sclerosis: OPERA studies open-label extension. *Neurology*. 2020;95(13):e1854-e1867
24. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol*. 2011;69(2):292-302
25. Bergmann A, Stangel M, Weih M, et al. Development of registry data to create interactive doctor-patient platforms for personalized patient care, taking the example of the DESTINY system. *Front Digit Health*. 2021;3:633427
26. Wehrle K, Tozzi V, Braune S, et al. Implementation of a data control framework to ensure confidentiality, integrity, and availability of high-quality real-world data (RWD) in the NeuroTransData (NTD) registry. *JAMIA Open*. 2022;5(1):ooac017
27. Turner B, Cree BAC, Kappos L, et al. Ocrelizumab efficacy in subgroups of patients with relapsing multiple sclerosis. *J Neurol*. 2019;266(5):1182-1193
28. Bermel RA, Rudick RA. Interferon-beta treatment for multiple sclerosis. *Neurotherapeutics*. 2007;4(4):633-646
29. Plosker GL. Interferon-β-1b: a review of its use in multiple sclerosis. *CNS Drugs*. 2011;25(1):67-88
30. Zhu C, Zhou Z, Roos I, et al. Comparing switch to ocrelizumab, cladribine or natalizumab after fingolimod treatment cessation in multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2022;93(12):1330-1337
31. Lefort M, Sharmin S, Andersen JB, et al. Impact of methodological choices in comparative effectiveness studies: application in natalizumab versus fingolimod comparison among patients with multiple sclerosis. *BMC Med Res Methodol*. 2022; 22(1):155
32. Hillert J, Magyari M, Soelberg Sørensen P, et al. Treatment switching and discontinuation over 20 years in the Big Multiple Sclerosis Data Network. *Front Neurol*. 2021;12:647811.

33. Coerver EME, Wessels MHJ, van Lierop ZYG, van Kempen ZLE, Killestein J, Strijbis EMM. Natalizumab discontinuation in a Dutch real-world cohort. *Mult Scler Relat Disord*. 2021;52:102974
34. Conway DS, Hersh CM, Harris HC, Hua LH. Duration of natalizumab therapy and reasons for discontinuation in a multiple sclerosis population. *Mult Scler J Exp Transl Clin*. 2020;6(1):2055217320902488
35. Vollmer B, Ontaneda D, Bandyopadhyay A, et al. Discontinuation and comparative effectiveness of dimethyl fumarate and fingolimod in 2 centers. *Neurol Clin Pract*. 2018;8(4):292-301
36. Harding K, Williams O, Willis M, et al. Clinical outcomes of escalation vs early intensive disease-modifying therapy in patients with multiple sclerosis. *JAMA Neurol*. 2019;76(5):536-541
37. He A, Merkel B, Brown JWL, et al. Timing of high-efficacy therapy for multiple sclerosis: a retrospective observational cohort study. *Lancet Neurol*. 2020;19(4): 307-316
38. Iaffaldano P, Lucisano G, Butzkueven H, et al. Early treatment delays long-term disability accrual in RRMS: results from the BMSD Network. *Mult Scler*. 2021; 27(10):1543-1555
39. Spelman T, Magyari M, Piehl F, et al. Treatment escalation vs immediate initiation of highly effective treatment for patients with relapsing-remitting multiple sclerosis: data from 2 different national strategies. *JAMA Neurol*. 2021;78(10):1197-1204
40. Engmann NJ, Sheinson D, Bawa K, Ng CD, Pardo G. Persistence and adherence to ocrelizumab compared with other disease-modifying therapies for multiple sclerosis in U.S. commercial claims data. *J Manag Care Spec Pharm*. 2021;27(5):639-649
41. Moccia M, Affinito G, Berera G, et al. Persistence, adherence, healthcare resource utilization and costs for ocrelizumab in the real-world of the Campania region of Italy. *J Neurol*. 2022;269(12):6504-6511
42. Butzkueven H, Spelman T, Ozakbas S, et al. Real-world experience with ocrelizumab in the MSBase Registry. In: 8th Joint ACTRIMS-ECTRIMS Meeting, virtual. September 11–13, 2020, P0909. MSVirtual2020.
43. Weber MS, Buttmann M, Meuth SG, et al. Safety, adherence and persistence in a real-world cohort of German MS patients newly treated with ocrelizumab: first insights from the CONFIDENCE study. *Front Neurol*. 2022;13:863105
44. Karim ME, Tremlett H, Zhu F, et al. Dealing with treatment-confounder feedback and sparse follow-up in longitudinal studies: application of a marginal structural model in a multiple sclerosis cohort. *Am J Epidemiol*. 2021;190(5):908-917
45. Clark TP, Kahan BC, Phillips A, White I, Carpenter JR. Estimands: bringing clarity and focus to research questions in clinical trials. *BMJ Open*. 2022;12:e052953
46. Phillips A, Clark T. Estimands in practice: bridging the gap between study objectives and statistical analysis. *Pharm Stat*. 2021;20(1):68-76.
47. Kalincik T, Diouf I, Sharmin S, et al. Effect of disease-modifying therapy on disability in relapsing-remitting multiple sclerosis over 15 years. *Neurology*. 2021;96(5):e783-e797