Ther Adv Neurol Disord
2021, Vol. 14: 1–14
DOI: 10.1177/17562864211021177

A propensity-matched comparison of long-term disability worsening in patients with multiple sclerosis treated with dimethyl fumarate or fingolimod

Amber Salter, Samantha Lancia, Gary Cutter, Ruth Ann Marrie, Jason P. Mendoza, James B. Lewin and Robert J. Fox Mellen

Abstract

Background: Although the aggregate of data among patients with multiple sclerosis (MS) have shown similar efficacy between dimethyl fumarate (DMF) and fingolimod (FTY), most studies have not assessed long-term worsening of disability. We compared long-term disability worsening over 5 years, as assessed by the Patient-Determined Disease Steps (PDDS), among participants with MS treated with DMF or FTY.

Methods: We identified individuals in the North American Research Committee on Multiple Sclerosis (NARCOMS) registry who had relapsing-remitting MS (RRMS), residing in the United States (Spring 2011 to Spring 2018), who initiated treatment with DMF ($n = 689$) or FTY ($n = 565$) and had ⩾1 year follow-up on index treatment. Participants receiving DMF who were previously treated with FTY and those on FTY previously treated with DMF were excluded. Propensity score matching at baseline was used to match FTY-treated to DMF-treated participants. Time to 6-month confirmed disability worsening (⩾1-point increase on PDDS, sustained for ⩾6 months) was estimated using Cox regression. A sensitivity analysis was conducted to account for differences in the duration of index exposure between DMF and FTY groups.

Results: After propensity score matching, 468 DMF-treated participants were matched with 468 FTY-treated participants. Median treatment duration was 3.0 years for DMF and 4.0 years for FTY. At 5 years, 68.3% [95% confidence interval (CI): 62.4–73.5] of DMF-treated participants and 63.3% (95% CI: 59.6–70.1) treated with FTY were free from 6-month confirmed PDDS worsening [hazard ratio (HR) 1.01 (95% CI: 0.79–1.28); $p = 0.95$]. Results were similar in the sensitivity analysis: 70.5% (95% CI: 61.8–77.6) of DMF-treated participants and 72.7% (95% CI: 65.4–78.6) of FTY-treated participants were free from 6-month confirmed PDDS worsening [HR: 1.04 (95% CI: 0.71–1.51); $p = 0.84$].

Conclusions: In participants with MS from the NARCOMS registry, there was no significant difference in confirmed disability (PDDS) worsening over 5 years between those treated with DMF *versus* FTY.

Keywords: comparative effectiveness, dimethyl fumarate, disability worsening, fingolimod, relapsing-remitting multiple sclerosis

Received: 19 January 2021; revised manuscript accepted: 8 May 2021.

Correspondence to:
James B. Lewin
Biogen, 225 Binney St. Cambridge, MA 02142, USA
jim.lewin@biogen.com

Amber Salter
Samantha Lancia
Division of Biostatistics, Washington University School in St. Louis, School of Medicine, St. Louis, MO, USA

Gary Cutter
Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA

Ruth Ann Marrie
Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

Jason P. Mendoza
Biogen, Cambridge, MA, USA

Robert J. Fox Mellen
Center for Multiple Sclerosis Treatment and Research, Cleveland Clinic, Cleveland, OH, USA

Introduction

Oral disease-modifying therapies (DMT) such as dimethyl fumarate (DMF) and fingolimod (FTY) have proven effective in the treatment of multiple sclerosis (MS).[1–4] In clinical trials, the two drugs reduced relapse rates and disease activity, and slowed disability worsening.[1–4] Several studies have demonstrated similar efficacy between DMF

and FTY on measures of disease activity.[4–10] Over 2 years, a German NeuroTransData GmbH registry study in 4172 patients showed no evidence of difference in effectiveness between DMF and FTY, but relatively greater effectiveness of DMF *versus* teriflunomide, interferons, and glatiramer acetate.[7] In a propensity score-matched analysis of claims data for 3906 patients switching from injectable to oral therapies, DMF significantly reduced the risk of relapse as compared with teriflunomide, but showed similar effectiveness to FTY.[4] Several single-center and multicenter retrospective chart review studies that used propensity score methods have also consistently shown similar effectiveness between DMF and FTY.[4–8,11,12] Many of these comparative studies have evaluated relapses,[4–8,10–14] although some have also compared magnetic resonance imaging (MRI) outcomes.[5,6,8,15–17] In these MRI comparisons, the radiologic efficacy was similar between DMF and FTY.

Only a few studies have evaluated disability worsening outcomes in patients treated with DMF or FTY. These studies have shown similar effectiveness on disability-related outcomes. However, most of these studies did not evaluate disability worsening beyond 2 years.[5,7,10,12] Therefore, in the current study, we compared long-term disability worsening over 5 years, as assessed by Patient-Determined Disease Steps (PDDS), in participants from the North American Research Committee on Multiple Sclerosis (NARCOMS) registry who were treated with DMF or FTY.

Methods

Participants

The NARCOMS registry is a voluntary self-report registry for people with MS.[18] Participants provide demographic and health-related information at enrollment and in survey updates every 6 months,[19] allowing their information to be used for research purposes. In the current analysis, we selected NARCOMS participants with relapsing-remitting multiple sclerosis (RRMS), living in the United States, who initiated index DMT (either DMF or FTY) between Spring 2011 and Spring 2018. Participants who had ⩾1 year of follow-up on index DMT and ⩾1 year of available PDDS data were included. Participants for the DMF group were excluded if they had received prior FTY treatment, while participants in the FTY group were excluded if they had prior DMF treatment.

Outcomes

The PDDS is a patient-reported measure of disability.[20] Scores range from 0 (no disability) to 8 (bedridden) in 1-point increments.[20] The PDDS was developed as a surrogate to the physician-scored Expanded Disability Status Scale (EDSS).[20] PDDS is a validated measure and correlates highly with the EDSS.[18,20,21] Confirmed disability progression was defined by identifying a worsening of ⩾1-point on the PDDS, independent of relapses, and confirmation that the worsening was sustained at the next available survey at least 6 months later. Performance Scales©, a self-report measure for MS-associated disability,[22] is used to assess eight domains: mobility, bladder/bowel, fatigue, sensory, vision, cognition, spasticity, and hand. Scores for all domains, except mobility, range from 0 (normal) to 5 (total disability); scores for the mobility domain range from 0 to 6. Both the PDDS and Performance Scales have good internal consistency and reliability and adequate test-retest reliability.[18,23] The Performance Scales' mobility, bladder/bowel, fatigue, vision, and hand subscales have each been validated against their clinical criterion measures (timed 25-foot walk, nine-hole peg test, low-contrast visual acuity, modified fatigue impact scale, and bladder control scale).[18]

In addition, the NARCOMS Depression Scale was used to assess depression.[24] These outcomes were collected every 6 months in each update survey.

Propensity score matching

We used 1:1 propensity score-matching to match DMF to FTY participants. The propensity score was estimated using logistic regression with treatment (DMF *versus* FTY) as the outcome and the following baseline predictors as independent variables: age, sex, disease duration, number of prior DMTs, education level, PDDS, cognition score, depression score, relapses in last 6 months, and cardiovascular comorbidities. The pre-baseline relapses are based on patient-reported relapses, and not confirmed by a neurologist. Each semi-annual update used a single question to assess whether the participant had a relapse in the past 6 months (yes/no/unsure). On the questionnaire, relapses or exacerbations of MS were defined as

the development of new symptoms or worsening of old symptoms that last longer than 48 hours and when it has been at least 30 days since the last relapse; in a relapse or exacerbation, MS symptoms generally worsen over a period of days to weeks; they then improve partially or completely over several weeks or months. A relapse can be associated with several different symptoms getting worse at the same time. Cardiovascular comorbidities were included because of their potential effect on disability worsening and because the US product insert for FTY contains warnings and precautions for some cardiovascular conditions (i.e. bradyarrhythmia and atrioventricular blocks, as well as hypertension).[25]

The predicted probability of treatment derived from the fitted logistic regression model was used to create a matched sample of participants. A caliper of width equal to 0.2 of the standard deviation of the logit of the propensity scores and a nearest neighbor matching without replacement algorithm was employed to form pairs of DMF and FTY participants.[26] In the propensity score model, we assumed a linear relationship between continuous covariates, and we did not include any interactions. To determine the balance between matched cohorts, we calculated the standardized difference for each baseline variable that was used in the propensity score before matching and after matching, with a target of a standardized differences <0.2 after matching. The C-statistic is a measure of balance in matched data and ranges from 0.5 to 1.0; the lowest value indicating a perfectly balanced propensity score model.

Statistical analysis. Group differences between DMF and FTY in the matched sample were evaluated using paired *t*-tests for continuous outcomes and McNemar tests for categorical outcomes. Time to 6-month confirmed disability worsening was estimated using the Kaplan–Meier method over a 5-year period. Comparison in the time to 6-month confirmed disability worsening between groups used a Cox proportional hazards regression model with robust sandwich estimators. The index survey was the survey where the participant initiated DMF or FTY. The event time for participants with confirmed PDDS worsening used the time of initial PDDS worsening (the initial event where the increase from baseline PDDS by 1 occurred) for all participants. Confirmation of PDDS worsening used their next survey from the initial worsening to assess if the increase was sustained. Participants were censored at their last available follow-up on the index DMT or at the time of DMT discontinuation, whichever came first, in those without an event.

Results report a hazard ratio (HR) and 95% confidence interval (CI). DMF was the reference group; therefore, a statistically significant result with $HR > 1.0$ would indicate a higher risk of disability progression with DMF treatment; a statistically significant result with $HR < 1.0$ would suggest lower risk of disability progression with DMF treatment. The significance level was set at 0.05.

Sensitivity analysis. FTY has been available for treatment of RRMS longer than DMF; therefore, we conducted a sensitivity analysis to account for differences in the duration of index DMT exposure between FTY and DMF. We restricted the index period from Fall 2013 through Spring 2018 (period in which both DMTs were available in the United States) and included the index survey (the survey at which FTY or DMF treatment was first reported) as a variable in the propensity score-matching analysis. Statistical analyses used SAS® v9.4 (SAS Institute, Cary, NC, USA).

Results

Participants

Overall, we included 689 participants who were prescribed DMF and 565 participants prescribed FTY (Figure 1). The survey completion response was high in both groups, with >93% of participants completing ⩾50% of surveys while on treatment. Baseline characteristics of the unmatched population are provided in Table 1.

Propensity score matching

Standardized differences were <0.1 across all baseline factors included in the propensity score matching, which is shown in Figure 2(a). In addition, we observed large overlap in propensity scores in the matched cohort [Figure 2(b)]. Together, these results indicated that propensity score matching achieved well-balanced groups. The C-statistic was 0.60 for the logistic model used for propensity matching.

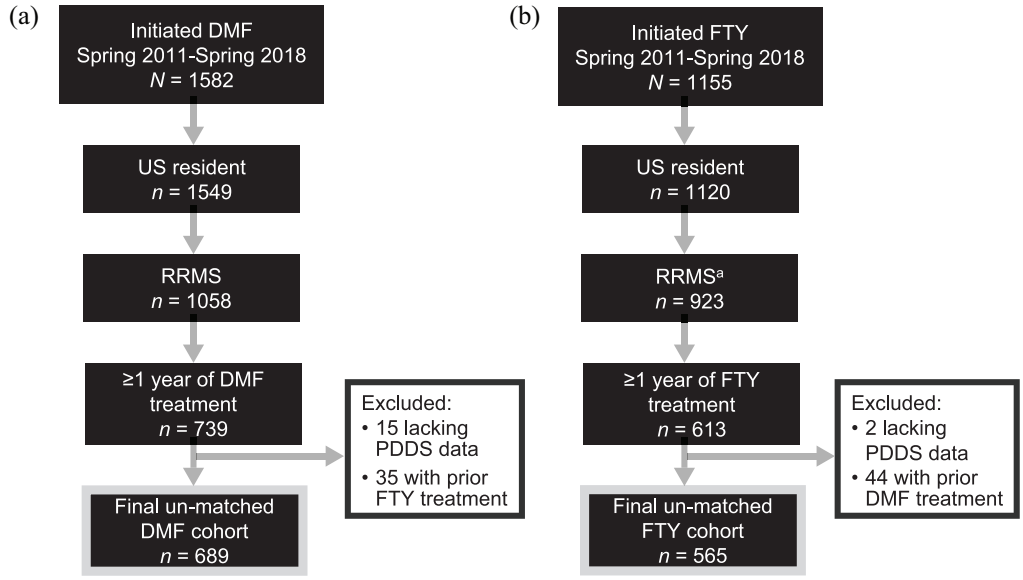

Figure 1. Disposition for (a) DMF and (b) FTY-treated participants.
[a]MS clinical course was not included as a question on the NARCOMS surveys until Spring 2012. Of the 923 patients with RRMS identified in the FTY cohort, 349 reported FTY initiation on the Spring 2011 or Fall 2011 survey; therefore, the MS subtype indicating 'RRMS' for these 349 patients was obtained from the first post-index survey that collected MS subtype information.
DMF, dimethyl fumarate; FTY, fingolimod; MS, multiple sclerosis; NARCOMS, North American Research Committee on Multiple Sclerosis; PDDS, Patient-Determined Disease Steps; RRMS, relapsing-remitting multiple sclerosis.

Matched population

After propensity score matching, 468 participants in the DMF group were matched with 468 participants in the FTY group. Baseline characteristics were well-balanced after matching (Table 2).

DMF versus *FTY*

The median (25%, 75%) treatment duration was 3.0 (1.5, 4.5) years for DMF and 4.0 (2.0, 6.0) years for FTY. In the unmatched population, 69.7% (95% CI: 65.0–73.8) of participants prescribed DMF and 64.1% (95% CI: 58.9–68.9) of those receiving FTY were free from 6-month confirmed PDDS worsening at 5 years. In the propensity score-matched population, 68.3% (95% CI: 62.4–73.5) of participants in the DMF group and 63.3% (95% CI: 59.6–70.1) of those in the FTY group were free from 6-month confirmed PDDS worsening at 5 years (HR 1.01; 95% CI: 0.79–1.28; $p = 0.95$; Figure 3).

DMF versus *FTY sensitivity analysis*

Baseline characteristics of the unmatched population and the matched population for the sensitivity analysis are presented in Table 3. Standardized differences were <0.1 across most of the baseline factors that were included in the propensity score matching [Figure 4(a)]. In addition, we observed large overlap in propensity scores in the matched cohort [Figure 4(b)]. The results of this sensitivity analysis were consistent with the primary analysis: 70.5% (95% CI: 61.8–77.6) of DMF-treated participants and 72.7% (95% CI: 65.4–78.6) of FTY-treated participants were free from 6-month confirmed PDDS worsening at 5 years (Figures 5 and 6).

Discussion

In this large, real-world sample of participants with MS from NARCOMS, the rate of long-term disability worsening was compared for DMF *versus* FTY. Results showed that ~65% of participants prescribed either therapy were free of 6-month confirmed disability worsening at 5 years, with no significant differences between the DMF and FTY cohorts. Similar findings were obtained for DMF *versus* FTY after accounting for differences in exposure time between the two DMTs. These results are consistent with previous studies that evaluated comparative efficacy of DMF *versus* FTY on measures of active inflammation.

Table 1. Baseline characteristics (unmatched population).

Characteristic	DMF *n* = 689	FTY *n* = 565	*p*-value
Age, median (IQR), y[a]	53 (46–59)	50 (43–57)	<0.001[b]
Disease duration since symptoms onset, median (IQR), y[a]	20 (14–28)	18 (12–26)	0.01[b]
Sex			0.31[c]
Male	99 (14)	93 (16)	
Female	590 (86)	472 (84)	
Number of patients with 0 prior DMTs	121 (18)	87 (15)	
Number of prior DMTs			0.12[c]
1 or fewer	418 (61)	318 (56)	
2 or more	271 (39)	247 (44)	
Education level[a]			0.38[c]
High School diploma, Associates or Technical degree	264 (40)	231 (42)	
Bachelor's degree or higher education	402 (60)	317 (58)	
PDDS at index, median (IQR)[a]	2 (1–4)	3 (1–4)	0.04[b]
Cognition at index, median (IQR)[a]	1 (1–3)	1 (1–2)	0.98[b]
Depression at index, median (IQR)[a]	1 (0–2)	1 (0–2)	0.10[b]
Relapse in last 6 months			0.26[c]
No	538 (78)	426 (75)	
Yes	151 (22)	139 (25)	
Reported cardiovascular risk factors prior to index			0.48[c]
No	459 (67)	387 (68)	
Yes	230 (33)	178 (32)	

Values presented as *n* (%) unless otherwise indicated.
[a]Data not available for all participants. Missing values, *n*, age = 1; disease duration since symptoms onset, education level = 40; PDDS at index = 2; cognition at index = 2; depression at index = 89.
[b]Wilcoxon rank sum test.
[c]Pearson's chi–square test.
DMF, dimethyl fumarate; DMT, disease-modifying therapy; FTY, fingolimod; IQR, interquartile range; PDDS, Patient-Determined Disease Steps.

Most studies have shown that efficacy of DMF is similar to that of FTY. In 2016, Fox et al.[5] compared pooled data from three DMF ($n = 769$) and three FTY ($n = 783$) clinical trials over 2 years of treatment. Their analyses indicated that both DMF and FTY showed a similar reduction in risk of 12-week confirmed disability (EDSS) worsening (risk ratio 0.92; 95% CI: 0.51–1.64; $p = 0.77$). This is similar to the HR (1.01 95% CI: 0.79–1.28; $p = 0.95$; Figure 3) for 6-month disability worsening in the present study. In their analyses of retrospective clinician-reported data from a single academic center in 2017, Vollmer et al.[6] found there was no significant difference between treatments with DMF ($n = 271$) and FTY ($n = 342$) on effectiveness over 2 years. Their outcome measures included clinical relapse, MRI contrast enhancement, or new T2 lesions, but did

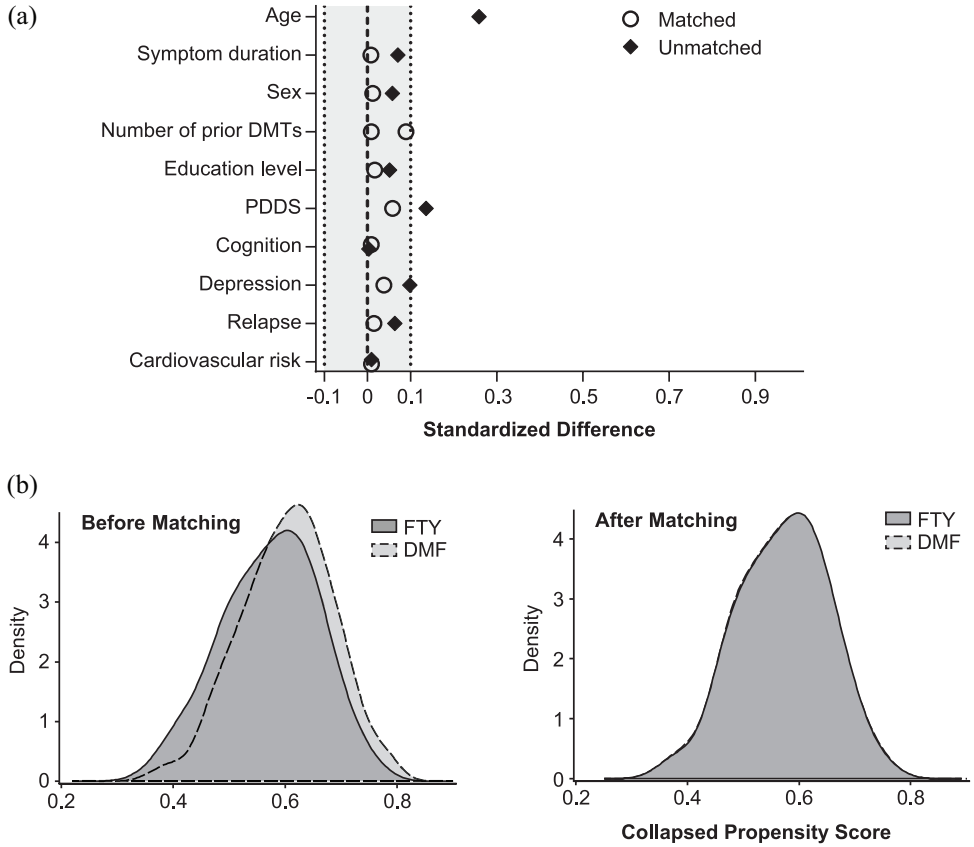

Figure 2. (a) Unmatched *versus* matched cohort standardized differences. The shaded area (standardized difference of -0.1 to 0.1) indicates well-balanced groups. (b) Density plots of propensity scores (top) before matching and (bottom) after matching.
DMF, dimethyl fumarate; FTY, fingolimod; PDDS, Patient-Determined Disease Steps.

not include disability worsening. In a subsequent (2019) analysis that pooled data from two MS centers, Vollmer et al.[8] observed no significant differences between DMF ($n=737$) and FTY ($n=535$) in clinical or MRI effectiveness outcomes, including relapses [odds ratio (OR) 1.27; 95% CI: 0.90–1.79)], gadolinium-enhancing lesions (OR 1.27; 95% CI: 0.90–1.79), new T2-hyperintense lesions (OR 0.99; 95% CI: 0.74–1.32), and brain MRI activity (OR 0.99; 95% CI: 0.74–1.31) after 36 months. In 2018, Braune et al.[7] analyzed data from the German NeuroTransData GmbH registry. They reported that there were no significant differences between patients receiving DMF ($n=457$) *versus* FTY ($n=457$) in the time to 3- and 6-month confirmed disability (EDSS) worsening events and proportion of worsening-free patients at 12 months. In a 2019 retrospective health claims analysis, Ontaneda et al.[4] compared annual relapse rates and time to relapse for patients receiving DMF ($n=1602$) or FTY ($n=534$). They observed no significant differences in either measure between DMF and FTY over 2.5 years. Kalincik et al.,[9] in their 2019 analyses of patients from the international MSBase cohort, reported a lower annual relapse rate with FTY ($n=2332$) compared with DMF ($n=782$) (0.20; 95% CI: 0.19–0.22 *versus* 0.26; 95% CI: 0.24–0.28; $p=0.01$) over 2.5 years, but found similar disability accumulation and improvement. The NARCOMS registry only captures patient-reported relapses; there is limited data to correlate patient-reported relapses with neurologist-confirmed relapses. As detailed above, there are many studies that have already compared DMF *versus* FTY on relapse activity; most of those studies have shown that there is no significant difference between DMF and FTY on relapse rates or risk of experiencing a relapse.[4–7,11–13,27–29]

Importantly, none of the preceding studies had compared the effectiveness of DMF *versus* FTY

Table 2. Baseline characteristics (matched population).

Characteristic	DMF *n* = 468	FTY *n* = 468
Age, median (IQR), y[a]	50 (43–57)	50 (43–57)
Disease duration since symptoms onset, median (IQR), y[a]	19 (12–25)	18 (12–26)
Sex		
Male	66 (14)	68 (15)
Female	402 (86)	400 (85)
Number of prior DMTs		
1 or fewer	272 (58)	270 (58)
2 or more	196 (42)	198 (42)
Education level[a]		
High School diploma, Associates or Technical degree	177 (40)	188 (41)
Bachelor's degree or higher education	268 (60)	275 (59)
PDDS at index, median (IQR)[a]	2 (1–4)	2 (1–4)
Cognition at index, median (IQR)[a]	1 (1–3)	1 (1–2)
Depression at index, median (IQR)[a]	1 (0–2)	1 (0–2)
Relapse in last 6 months		
No	358 (76)	355 (76)
Yes	110 (24)	113 (24)
Reported cardiovascular risk factors prior to index		
No	323 (69)	321 (69)
Yes	145 (31)	147 (31)

Values presented as *n* (%) unless otherwise indicated.
[a]Data not available for all participants. Missing values, *n*; age = 1; disease duration since symptoms onset, education level = 28, PDDS at index = 1; cognition at index = 2; depression at index = 25.
DMF, dimethyl fumarate; DMT, disease modifying therapy; FTY, fingolimod; IQR, interquartile range; PDDS, Patient-Determined Disease Steps.

on long-term disability worsening. Only one study evaluated longer-term disability outcomes. Guger et al.,[10] in their retrospective analysis of data from the Austrian MS Treatment Registry, reported similar relapse rates for DMF (0.09) and FTY (0.13) but less sustained disability worsening in 12 weeks with DMF- *versus* FTY-treated patients (HR 2.23; 95% CI: 1.14–4.38; $p = 0.02$). However, both DMF and FTY showed similar rates of disability improvement. It should be noted that the efficacy outcomes in the Guger study were only evaluated in a cohort of patients who remained on treatment for at least 2 years, and thus reflect a population of patients who are likely experiencing a positive treatment response (e.g. a responder analysis). In addition, different methodology could have contributed to the differences in the results as Guger et al.[10] use inverse probability weighting with a multinomial model in their analysis. This type of analysis is valuable, but there is still a data gap for long-term disability worsening comparisons between DMF and FTY in a broader patient population with MS. Different propensity score methods can be used to attempt

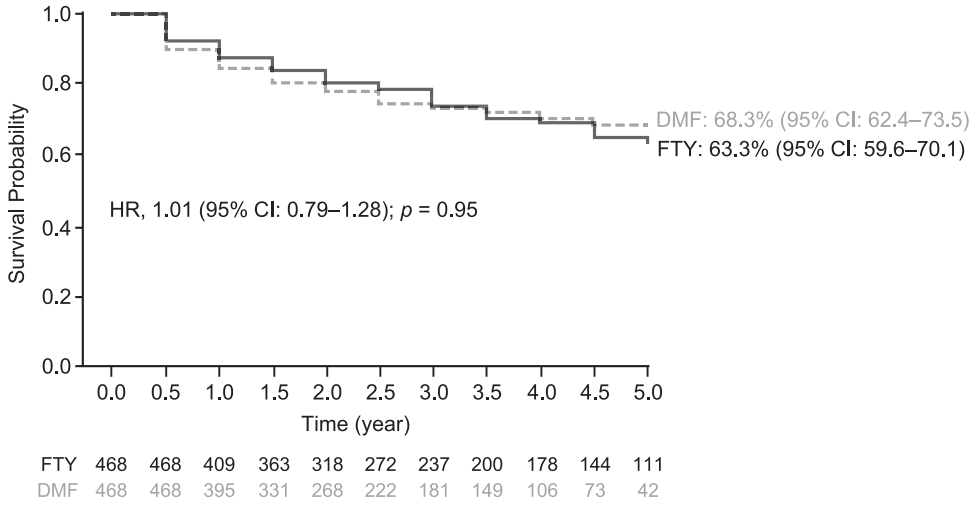

Figure 3. Freedom from 6-month confirmed PDDS worsening over 5 years (matched population).
CI, confidence interval; DMF, dimethyl fumarate; FTY, fingolimod; HR, hazard ratio; PDDS, Patient-Determined Disease Steps.

to mimic randomization when comparing outcomes based on real-world data sources. Two of the most common methods are propensity score matching and inverse probability of treatment weighting (IPTW). For our study, we chose to use propensity score matching because this method is more robust against outliers (e.g. patients with extreme propensity scores) than IPTW.

The current study represents a large real-world sample assessing long-term effectiveness of DMF *versus* FTY. It is, however, subject to limitations that arise from common issues with analyses of registry data. Propensity adjustment only adjusts for known, measured factors, whereas other factors, whether unmeasured or unknown, may affect clinicians' prescribing patterns and that, in turn, may affect what type of patient is treated with each therapy. There may still be a hidden bias from these unknown or unmeasured factors. NARCOMS registry participants are volunteers and may not be completely representative of all patients with MS. However, the NARCOMS population includes participants treated in community and tertiary care settings across the United States. Importantly, the mean age (51 years) of the NARCOMS study sample appears to be representative of US MS patients. Specifically, in the Truven US claims dataset, the mean age for DMF patients was 47 years, teriflunomide 50 years, and fingolimod 44 years with up to 2.5 years of follow-up.[4] The mean age was 48 years with 1 year follow-up in the DMF US phase IV RESPOND study,[30] mean age was 47 years in the DMF US phase IV STRATEGY study,[31] and the median age of the matched population was 51 with up to 5 years of follow-up in this NARCOMS study. In addition, Wallin et al.[32] recently reported that the prevalence of MS in 2010 was highest in the 55–64-year age group.

The smaller sample sizes at the later time points reduce our ability to detect differences in outcomes between the groups. Although the number of patients decreases over time, this is driven largely by the fact that some patients only initiated therapy 1–2 years before we conducted the analysis. However, we do not believe that this influences the primary outcome and results are consistent at earlier time points. When evaluating the differences at 2 years, we see no significant difference between DMF and FTY. This provides more confidence in the result that we see at 5 years, which also shows no significant difference between DMF and FTY. In addition, we conducted a sensitivity analysis to account for differences in the duration of index DMT exposure between DMF and FTY. We restricted the index period from Fall 2013 through Spring 2018 (period in which both DMTs were available in the United States) and included the index survey (the survey at which FTY or DMF treatment was first reported) as a variable in the propensity score-matching analysis. The results of this sensitivity analysis were consistent with the primary analysis.

Table 3. Baseline characteristics in the sensitivity analysis (unmatched population).

Characteristic	Unmatched population			Matched population	
	DMF *n* = 689	FTY *n* = 243	*p*-value	DMF *n* = 236	FTY *n* = 236
Age, median (IQR), y[a]	53 (46–59)	50 (42–58)	0.003[b]	51 (44–57)	51 (42–58)
Disease duration since symptoms onset, median (IQR), y[a]	20 (14–28)	19 (12–27)	0.006[b]	19 (13–26)	19 (12–27)
Sex			0.76[c]		
Male	99 (14)	33 (14)		25 (11)	33 (14)
Female	590 (86)	210 (86)		211 (89)	203 (86)
Number of prior DMTs			0.60[c]		
1 or fewer	418 (61)	152 (63)		141 (60)	146 (62)
2 or more	271 (39)	91 (37)		95 (40)	90 (38)
Education level[a]			0.49[c]		
High School diploma, Associates or Technical degree	264 (40)	98 (42)		92 (41)	92 (41)
Bachelor's degree or higher education	402 (60)	134 (58)		132 (59)	133 (59)
PDDS at index, median (IQR)[a]	2 (1–4)	2 (0–4)	0.33[b]	2 (1–4)	2 (0–4)
Cognition at index, median (IQR)[a]	1 (1–3)	1 (1–3)	0.48[b]	1 (1–3)	1 (1–3)
Depression at index, median (IQR)	1 (0–2)	1 (0–2)	0.031[b]	1 (0–2)	1 (0–2)
Relapse in last 6 months			0.82[c]		
No	538 (78)	188 (77)		180 (76)	181 (77)
Yes	151 (22)	55 (23)		56 (24)	55 (23)
Reported cardiovascular risk factors prior to index			0.023[c]		
No	459 (67)	181 (75)		170 (72)	175 (74)
Yes	230 (33)	62 (25)		66 (28)	61 (26)
Index survey, median (IQR)[d]	Fall 2014 (Spring 2014, Fall 2015)	Spring 2015 (Spring 2014, Fall 2016)	<0.001[c]	Spring 2015 (Spring 2014, Fall 2015)	Spring 2015 (Spring 2014, Fall 2015)

Values presented as *n* (%) unless otherwise indicated.
[a]Data not available for all participants. Missing values, *n*: age = 1; disease duration since symptoms onset = 2; education level = 34; PDDS at index = 2; cognition at index = 2.
[b]Wilcoxon rank sum test.
[c]Pearson's chi–square test.
[d]For the purpose of propensity score matching, each index survey was assigned a number and index survey was treated as an ordinal variable.
DMF, dimethyl fumarate; DMT, disease-modifying therapy; FTY, fingolimod; IQR, interquartile range; PDDS, Patient-Determined Disease Steps.

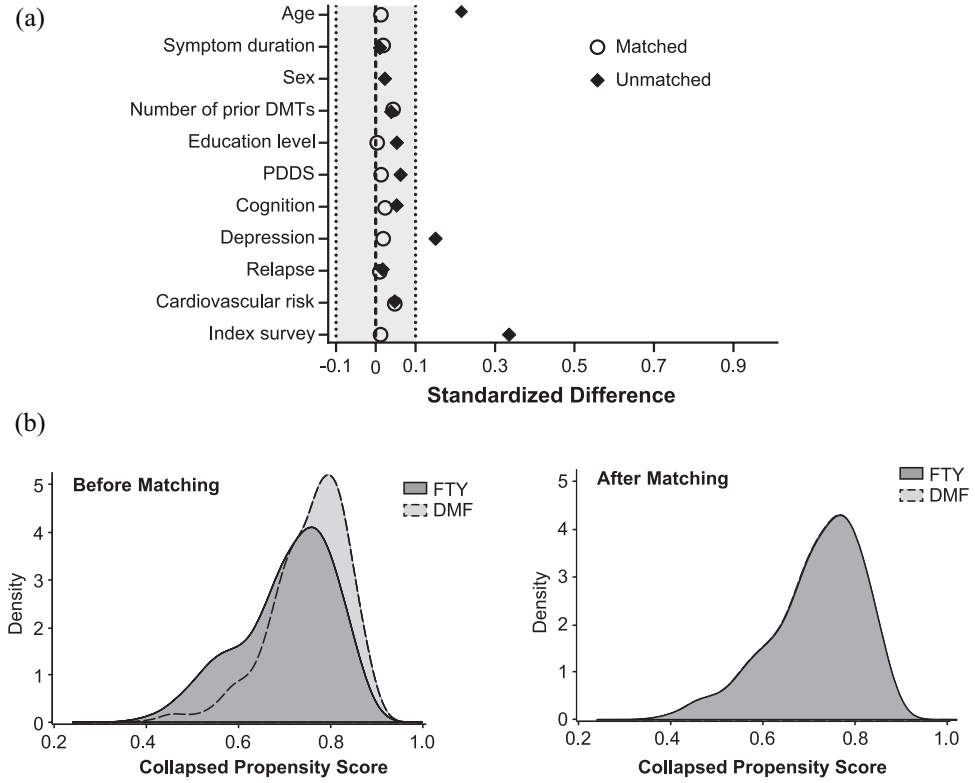

Figure 4. (a) Unmatched *versus* matched cohort standardized differences. The shaded area (standardized difference of -0.1 to 0.1) indicates well-balanced groups. (b) Propensity score density plots (top) before matching and (bottom) after matching (sensitivity analysis).
DMF, dimethyl fumarate; FTY, fingolimod; PDDS, Patient-Determined Disease Steps.

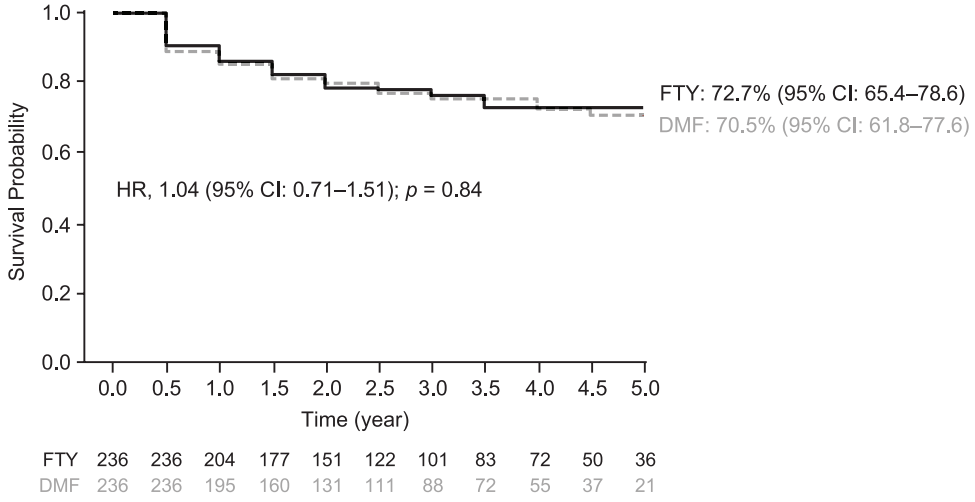

Figure 5. Freedom from 6-month confirmed PDDS worsening over 5 years (matched population sensitivity analysis).
CI, confidence interval; DMF, dimethyl fumarate; FTY, fingolimod; HR, hazard ratio; PDDS, Patient-Determined Disease Steps.

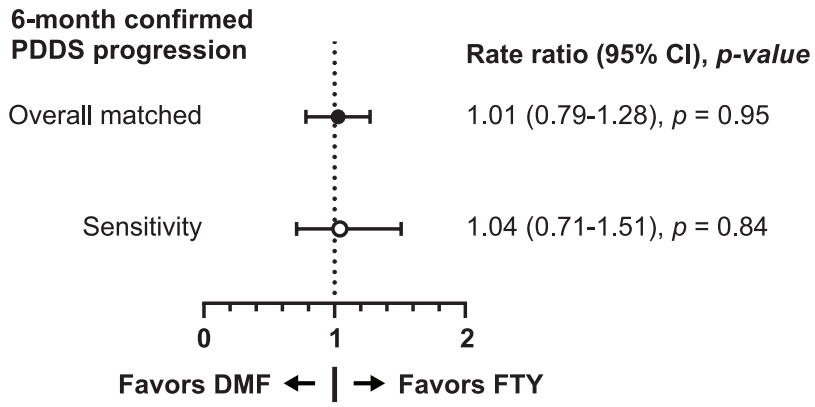

Figure 6. Hazard ratios of 6-month confirmed PDDS worsening for DMF *versus* FTY.
CI, confidence interval; DMF, dimethyl fumarate; FTY, fingolimod; PDDS, Patient-Determined Disease Steps.

This study was also limited by what data are collected by the NARCOMS registry. Although employment status (yes/no) was captured in NARCOMS, the full spectrum of work-related outcomes was not captured, such as presenteeism. Further research to compare the effect of DMTs on these additional work-related outcomes is warranted. This study was also not able to compare outcomes such as cognition and depression since there is no established definition for 'confirmed progression' on the patient-reported cognition or depression symptom scale. Self-reported change in cognition is poorly associated with performance-based changes in cognition, rather, anxiety and depression are stronger predictors of perceived change than measured changes in cognition.[33] Therefore, we are unsure if we could draw valid conclusions based on the cognition and depression symptom scale used in this study. In addition, we feel more research is needed to understand how to clearly define progression on these scales and understand the reliability of self-reported change in cognition prior to conducting a comparison between different treatments. Other limitations relate to the fact that this is patient-reported data, including relapse activity, and may be prone to recall bias.[34] However, previously published studies suggest that persons with MS can provide valid data on their clinical disease course,[35] and regarding which DMTs they use.[36] For example, baseline MRI data were not collected in this study, so we could not match the participants on this. The baseline relapse activity in this study population appeared to be low compared with the baseline relapse activity observed in other comparative effectiveness studies of DMF and FTY patients; however, these data are not directly comparable to other studies because the NARCOMS registry captures patient-reported relapses whereas other studies have primarily reported baseline relapse activity based on medical records.

Conclusions

In this propensity score-matched analysis of participants with MS, treatment with DMF and FTY demonstrated similar effectiveness on delaying confirmed disability (PDDS) worsening over 5 years. Our results expand on previous studies that showed similar effectiveness between DMF and FTY on relapse and MRI outcomes.

Acknowledgements

The authors thank the NARCOMS study participants. Writing and editorial support for the preparation of this manuscript was provided by Excel Medical Affairs (Fairfield, CT): funding was provided by Biogen (Cambridge, MA). Biogen provided funding for medical writing support in the development of this paper; Karen Spach, PhD, from Excel Medical Affairs (Fairfield, CT) wrote the first draft of the manuscript based on input from authors, and Linda Cirella from Engage Medical Affairs (Fairfield, CT) copyedited and styled the manuscript per journal requirements. Biogen reviewed and provided feedback on the paper to the authors. The authors had full editorial control of the paper and provided their final approval of all content for submission.

Performance Scales Copyright Registration Number/Date: TXu000743629/1996-04-04; assigned to DeltaQuest Foundation, Inc., effective 1 October 2005. U.S. Copyright law governs terms of use.

Conflict of interest statement

RJF: consulting fees from Actelion, Biogen, Celgene, EMD Serono, Genentech, Immunic, Novartis, Sanofi, Teva, and TG Therapeutics; advisory board for Actelion, Biogen, Immunic, and Novartis; research support from Biogen and Novartis.

SL: The author declares that there is no conflict of interest.

GC: data safety and monitoring board for AveXis, BioLineRx, BrainStorm Cell Therapeutics, CSL Behring, Galmed, Hisun, Horizon, Mapi-Pharma, Merck, Neurim, Novartis, OPKO Biologics, Orphazyme, Reata, Receptos/Celgene/BMS, Sanofi-Aventis, Teva, and Vivus; consulting fees from BioDelivery Sciences International, Biogen, Click Therapeutics, Genentech, Genzyme, GW, Klein Buendel, MedDay, MedImmune/Viela Bio, Novartis, Osmotica, Perception Neuroscience, Recursion/Cerexis, Roche, and TG Therapeutics; owner of Pythagoras Inc.

RAM: member of editorial advisory board for *Multiple Sclerosis Journal* and journal co-editor for *Multiple Sclerosis Journal–Experimental, Translational and Clinical*.

JPM and JBL: employees of and hold stock/stock options in Biogen.

AS: journal editor and member of editorial advisory board for *Circulation: Cardiovascular Imaging*.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Biogen.

Research ethics and patient consent

This NARCOMS-based research was approved by the Washington University Institutional Review Board (IRB#: 201610132) and followed the principles of the Helsinki declaration.

NARCOMS is a voluntary self-report registry for people diagnosed with MS. Registry participants give permission for their de-identified information to be used for research.

References

1. Kappos L, Radue EW, O'Connor P, *et al.* A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* 2010; 362: 387–401.
2. Gold R, Kappos L, Arnold DL, *et al.* Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N Engl J Med* 2012; 367: 1098–1107.
3. Calabresi PA, Radue EW, Goodin D, *et al.* Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol* 2014; 13: 545–556.
4. Ontaneda D, Nicholas J, Carraro M, *et al.* Comparative effectiveness of dimethyl fumarate versus fingolimod and teriflunomide among MS patients switching from first-generation platform therapies in the US. *Mult Scler Relat Disord* 2019; 27: 101–111.
5. Fox RJ, Chan A, Zhang A, *et al.* Comparative effectiveness using a matching-adjusted indirect comparison between delayed-release dimethyl fumarate and fingolimod for the treatment of multiple sclerosis. *Curr Med Res Opin* 2017; 33: 175–183.
6. Vollmer B, Nair KV, Sillau SH, *et al.* Comparison of fingolimod and dimethyl fumarate in the treatment of multiple sclerosis: two-year experience. *Mult Scler J Exp Transl Clin* 2017; 3: 2055217317725102.
7. Braune S, Grimm S, van Hovell 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: 2980–2992.
8. Vollmer B, Ontaneda D, Harris H, *et al.* Comparative discontinuation, effectiveness, and switching practices of dimethyl fumarate and fingolimod at 36-month follow-up. *J Neurol Sci* 2019; 407: 116498.
9. Kalincik T, Kubala Havrdova E, Horakova D, *et al.* Comparison of fingolimod, dimethyl fumarate and teriflunomide for multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2019; 90: 458–468.
10. Guger M, Enzinger C, Leutmezer F, *et al.* Oral therapies for treatment of relapsing-remitting multiple sclerosis in Austria: a 2-year comparison using an inverse probability weighting method. *J Neurol* 2020; 267: 2090–2100.
11. Prosperini L, Lucchini M, Haggiag S, *et al.* Fingolimod vs dimethyl fumarate in multiple

sclerosis: a real-world propensity score-matched study. *Neurology* 2018; 91: e153–e161.

12. Lorscheider J, Benkert P, Lienert C, *et al.* Comparative analysis of dimethyl fumarate and fingolimod in relapsing-remitting multiple sclerosis. *J Neurol* 2020; 268: 941–949.
13. Boster A, Nicholas J, Wu N, *et al.* Comparative effectiveness research of disease-modifying therapies for the management of multiple sclerosis: analysis of a large health insurance claims database. *Neurol Ther* 2017; 6: 91–102.
14. Sloane J, Phillips JT, Calkwood J, *et al.* Delayed-release dimethyl fumarate demonstrated no difference in clinical outcomes versus fingolimod in patients with relapsing-remitting multiple sclerosis: results from the real-world EFFECT study. *Mult Scler J* 2017; 23: 855.
15. Feng J, Nakamura K, Hersh C, *et al.* Comparative effectiveness of dimethyl fumarate and fingolimod using fully automated algorithms to assess effects on brain atrophy and T2-lesion volume. *Neurology* 2018; 90(Suppl. 15): P1.365.
16. Moiola L, Esposito F, Di Cristinzi M, *et al.* Comparative effectiveness of dimethylfumarate and fingolimod in an Italian monocentric cohort of relapsing remitting multiple sclerosis patients. *Mult Scler J* 2018; 24: 86–87.
17. Hersh C, Harris H, Cohn S, *et al.* Comparative effectiveness and discontinuation of dimethyl fumarate and fingolimod in clinical practice at 36-month follow-up. *Mult Scler J* 2018; 24(Suppl. 2): 266.
18. Marrie RA and Goldman M. Validity of performance scales for disability assessment in multiple sclerosis. *Mult Scler* 2007; 13: 1176–1182.
19. Salter A, Thomas N, Tyry T, *et al.* Employment and absenteeism in working-age persons with multiple sclerosis. *J Med Econ* 2017; 20: 493–502.
20. Learmonth YC, Motl RW, Sandroff BM, *et al.* Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. *BMC Neurol* 2013; 13: 37.
21. Hohol MJ, Orav EJ and Weiner HL. Disease steps in multiple sclerosis: a longitudinal study comparing disease steps and EDSS to evaluate disease progression. *Mult Scler* 1999; 5: 349–354.
22. Schwartz CE, Vollmer T and Lee H. Reliability and validity of two self-report measures of impairment and disability for MS. North American Research Consortium on Multiple Sclerosis Outcomes Study Group. *Neurology* 1999; 52: 63–70.
23. Hohol MJ, Orav EJ and Weiner HL. Disease steps in multiple sclerosis: a simple approach to evaluate disease progression. *Neurology* 1995; 45: 251–255.
24. Marrie RA, Cutter G, Tyry T, *et al.* Validation of NARCOMS depression scale. *Int J MS Care* 2008; 10: 81–84.
25. Novartis. Gilenya (fingolimod) prescribing information, https://www.novartis.us/sites/www.novartis.us/files/gilenya.pdf (2019, accessed 21 December 2020).
26. Parsons LS. Reducing bias in a propensity score matched-pair sample using greedy matching techniques. In: *Proceedings of the 26th Annual SAS users group international conference*, *SAS Institute Inc*., Long Beach, CA, 2001, pp.214-223.
27. 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: 292–301.
28. Hersh CM, Love TE, Bandyopadhyay A, *et al.* Comparative efficacy and discontinuation of dimethyl fumarate and fingolimod in clinical practice at 24-month follow-up. *Mult Scler J Exp Transl Clin* 2017; 3: 2055217317715485.
29. Hersh CM, Love TE, Cohn S, *et al.* Comparative efficacy and discontinuation of dimethyl fumarate and fingolimod in clinical practice at 12-month follow-up. *Mult Scler Relat Disord* 2016; 10: 44–52.
30. Kresa-Reahl K, Repovic P, Robertson D, *et al.* Effectiveness of delayed-release dimethyl fumarate on clinical and patient-reported outcomes in patients with relapsing multiple sclerosis switching from glatiramer acetate: RESPOND, a prospective observational study. *Clin Ther* 2018; 40: 2077–2087.
31. Cohan SL, Moses H, Calkwood J, *et al.* Clinical outcomes in patients with relapsing-remitting multiple sclerosis who switch from natalizumab to delayed-release dimethyl fumarate: a multicenter retrospective observational study (STRATEGY). *Mult Scler Relat Disord* 2018; 22: 27–34.
32. Wallin MT, Culpepper WJ, Campbell JD, *et al.* The prevalence of MS in the United States: a population-based estimate using health claims data. *Neurology* 2019; 92: e1029–e1040.

33. Akbar N, Honarmand K and Feinstein A. Self-assessment of cognition in multiple sclerosis: the role of personality and anxiety. *Cogn Behav Neurol* 2011; 24: 115–121.

34. Schmier JK and Halpern MT. Patient recall and recall bias of health state and health status. *Expert Rev Pharmacoecon Outcomes Res* 2004; 4: 159–163.

35. Ingram G, Colley E, Ben-Shlomo Y, *et al.* Validity of patient-derived disability and clinical data in multiple sclerosis. *Mult Scler* 2010; 16: 472–479.

36. McKay KA, Evans C, Fisk JD, *et al.* Disease-modifying therapies and adherence in multiple sclerosis: comparing patient self-report with pharmacy records. *Neuroepidemiology* 2017; 48: 124–130.