Arthritis Care & Research
Vol. 75, No. 9, September 2023, pp 1986–1995
DOI 10.1002/acr.25092

American College of Rheumatology
Empowering Rheumatology Professionals

Effectiveness and Persistence in SB4- and Reference Etanercept–Treated Rheumatoid Arthritis Patients in Ordinary Clinical Practice in Norway

Glenn Haugeberg,[1] Gunnstein Bakland,[2] Erik Rødevand,[3] Inger Johanne Widding Hansen,[4] Andreas Diamantopoulos,[5] and Are Hugo Pripp[6]

Objective. Biosimilars represent cost-effective alternatives to reference biologic disease-modifying antirheumatic drugs. Our objective was to compare drug effectiveness and drug persistence in the treatment of rheumatoid arthritis (RA), assessing the etanercept biosimilar SB4 in efficacy and safety compared with reference etanercept in a Phase III, randomized controlled trial. We applied EULAR Points to Consider for Comparative Effectiveness Research in a retrospective database study of etanercept and SB4 in patients treated in clinical practice in Norway.

Methods. Patients with RA (n = 1,455) treated with etanercept or SB4 between 2010 and 2018 at 5 centers in Norway with ≥1 year of follow-up were included. Disease outcomes (Disease Activity Score in 28 joints [DAS28] at week 52) and drug persistence were compared between unmatched etanercept (n = 575) and SB4 (n = 299) cohorts and matched analyses (n = 172, both cohorts) using propensity score (PS) matching to adjust for confounders.

Results. In unmatched analyses, the difference in change from baseline between etanercept (n = 221) and SB4 (n = 106) for DAS28 at week 52 was mean –0.02 (95% confidence interval [95% CI] –0.32, 0.27), demonstrating equivalence by the predetermined equivalence margin (±0.6). In PS-matched analyses, the difference between etanercept (n = 49) and SB4 (n = 49) was 0.03 (95% CI –0.46, 0.52), within the predefined equivalence margin. Persistence using the drug at week 52 was similar between etanercept (0.62 [95% CI 0.57, 0.65]) and SB4 (0.66 [95% CI 0.60, 0.71]) cohorts in the unmatched analysis; in PS-matched cohorts, persistence at week 52 was 0.52 (95% CI 0.44, 0.59) for etanercept and 0.68 (95% CI 0.61, 0.75) for SB4.

Conclusion. Outcomes for disease status/drug persistence at week 52 were similar between patients with RA treated with etanercept or SB4.

INTRODUCTION

The introduction of biologic disease-modifying antirheumatic drugs (bDMARDs) in 1999 led to a paradigm shift in the treatment of chronic inflammatory arthritis disorders. The tumor necrosis factor inhibitors infliximab and etanercept (ETN), licensed for the treatment of rheumatoid arthritis (RA), were the first to reach the market, receiving approvals for use in the European Union in 1999 and 2000, respectively. The costs of biologic drugs present challenges, causing restrictions to the prescribing of these drugs in several countries and subsequently contributing to inequalities of care (1–3). However, the expiration of patents for bDMARDs allowed the manufacture of biosimilars, which can be sold at lower prices. Since the first biosimilar tumor necrosis factor inhibitor, infliximab CT-P13, was approved in the European Union in 2013, additional biosimilars have become available. The ETN biosimilar SB4 was approved by the European Medicines Agency in January 2016. SB4 demonstrated similarity to reference ETN in a comprehensive biosimilarity exercise, which included a

Previously presented in part as a poster at the European Congress of Rheumatology, June 2–5, 2021, Paris. Haugeberg G, Bakland G, Rødevand E, et al. POS0612 exploring equivalence between biosimilar SB4 and reference etanercept by assessing effectiveness in rheumatoid arthritis patients treated in ordinary clinical practice.

Supported by Biogen International GmbH.

[1]Glenn Haugeberg, MD, PhD: Sørlandet Hospital, Kristiansand, and Norwegian University of Science and Technology, Trondheim, Norway; [2]Gunnstein Bakland, MD, PhD: University Hospital of Northern Norway, Tromsø, Norway; [3]Erik Rødevand, MD: St. Olavs Hospital, Trondheim, Norway; [4]Inger Johanne Widding Hansen, MD: Sørlandet Hospital, Kristiansand, Norway; [5]Andreas Diamantopoulos, MD, PhD, MPH: Martina Hansens Hospital, Sandvika, Norway; [6]Are Hugo Pripp, PhD: Oslo University Hospital and Oslo Metropolitan University, Oslo, Norway.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25092&file=acr25092-sup-0001-Disclosureform.pdf.

Address correspondence via email to Glenn Haugeberg, MD, PhD, at glenn.haugeberg@sshf.no.

Submitted for publication March 23, 2022; accepted in revised form January 24, 2023.

SIGNIFICANCE & INNOVATIONS

- SB4 has demonstrated similarity to etanercept (ETN) in a comprehensive biosimilarity exercise, which included a 52-week, double-blind, Phase III, randomized controlled trial in patients with rheumatoid arthritis (RA). However, published real-world data on SB4 are limited.
- This study reports similar effectiveness, persistence, and safety for patients with RA who initiated treatment with ETN or SB4 for up to 2 years as part of routine clinical care at outpatient clinics in Norway. However, when accounting for differences between cohorts at baseline using propensity score matching, persistence was greater on SB4 than on ETN. Effectiveness was maintained in patients with RA who had a mandatory switch from ETN to SB4.
- These findings support outcomes from earlier biosimilarity studies and indicate that SB4 is an effective option for switching from ETN for the treatment of patients with RA.

52-week, double-blind, Phase III, randomized controlled trial (RCT) in patients with RA (4,5).

Biosimilar drugs follow a tailored approval pathway compared with reference drugs, including a Phase III RCT with high internal but low external validity. Therefore, observational studies with high external validity are important to reassure patients and physicians that no clinically meaningful differences exist between a biosimilar and its reference drug when used in routine clinical practice. Unfortunately, recent comparative effectiveness studies often do not disclose applied analytical methods in sufficient detail, with many not adjusting for confounders (6) or accounting for attrition or missing data, according to a EULAR task force systematic review (7). Compliance with these recommendations for conducting comparative effectiveness studies may contribute toward high-quality observational studies.

Although SB4 has been on the market for several years, published real-world data are limited for patients with RA (8–12). The objective of this real-world study was to compare drug effectiveness and drug persistence in ETN treatment-naive patients with RA who received treatment with ETN or the biosimilar SB4, applying the EULAR Points to Consider for Comparative Effectiveness Research. Further, we aimed to examine drug effectiveness and drug persistence in patients with RA treated with SB4 after a mandatory nonmedical switch from ETN and to explore reasons for cessation among the 3 RA treatment cohorts: ETN, SB4, and SB4 switch.

MATERIALS AND METHODS

Study design. This was a retrospective database study of ETN-naive patients who received treatment with ETN or SB4 and ETN-treated patients who switched to SB4 and had at least 1 year of follow-up data. Data extraction from the participating centers was performed between June 26 and July 1, 2019. The study followed the recommendations outlined in the EULAR Points to Consider When Analysing and Reporting Comparative Effectiveness Research with Observational Data in Rheumatology (13), as well as the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and Good Research for Comparative Effectiveness (GRACE) guidelines (14,15).

Study population. ETN-naive patients with RA started ETN treatment between January 2010 and July 2018 at 5 centers in Norway and were followed for up to 2 years. The participating centers were University Hospital of North Norway, Tromsø; St. Olavs Hospital, Trondheim; Haukeland University Hospital, Bergen; Sørlandet Hospital, Kristiansand; and Martina Hansens Hospital, Sandvika.

Data collection. Data collection at participating centers was performed at clinical visits made as part of routine practice. Data variables collected by all centers included age, sex, body mass index (BMI), duration of disease, anti–cyclic citrullinated peptide antibodies (anti-CCP), C-reactive protein (CRP) level, erythrocyte sedimentation rate (ESR), 28 swollen and tender joint counts (SJC28 and TJC28), patient global assessment (PtGA) reported on a 0–100-mm visual analog scale, Disease Activity Score in 28 joints (DAS28), modified Health Assessment Questionnaire (MHAQ), current use of methotrexate, current conventional synthetic DMARDs (csDMARDs), including methotrexate, and order of bDMARDs. Data extraction also included reasons for drug cessation registered in the hospital clinical GoTreatIT Rheuma databases. Data were collected for ETN-naive patients who started treatment on ETN or SB4 and for patients who switched from ETN to SB4. At the time of data extraction, we examined how many patients had remained on ETN and how many patients who switched to SB4 had switched back to ETN. Data are available upon reasonable request.

Study objectives. The primary objective was to compare drug effectiveness and drug persistence of ETN and SB4 at week 52 in treatment-naive patients with RA treated in ordinary clinical practice in Norway. Secondary objectives were to further assess drug effectiveness and persistence at week 52 and week 104 in patients with RA treated with SB4 after a mandatory nonmedical switch from ETN and to explore reasons for drug cessation across the 3 RA treatment cohorts (ETN, SB4, and SB4 switch).

Study end points. Primary outcome measures were disease outcomes (DAS28 at week 52, assessed as a continuous variable) and drug persistence (measured as time to treatment discontinuation during a 52-week follow-up). The equivalence of DAS28 was determined based on a predefined equivalence

margin of ±0.6 (16). Unmatched (primary) and propensity score (PS)–matched (supportive) analyses were conducted, including a sensitivity analysis of PS-matched samples using all available data in the statistical analysis.

Secondary outcome measures included DAS28 at week 104, assessed as a continuous variable, and DAS28 at week 52 and week 104, assessed as a categorical variable based on EULAR response criteria of good, moderate, and no response to treatment (17). Other clinical outcomes assessed at week 52 and week 104 included CRP levels, ESR, SJC28, TJC28, PtGA, and MHAQ. Reasons for drug cessation were recorded. At data extraction, the number of patients who had remained on ETN without switching to SB4 was quantified, as was the number of patients who had switched back to ETN from SB4. Reasons for discontinuing treatment were also assessed; where an adverse event (AE) was given as the reason for discontinuation, specific AEs were reported if they had been recorded in the registry.

Statistical analysis. Independent samples *t*-tests and chi-square tests were used to compare baseline characteristics by treatment cohort in unmatched data. For matched data, paired samples *t*-test and McNemar's test were used to compare baseline characteristics for continuous variables and proportions, respectively.

For the primary, unmatched analysis of DAS28 outcomes between patients treated with ETN and patients treated with SB4, a conventional independent samples *t*-test was used (model 0). For the supportive, matched analyses of DAS28, PS matching was used and analyzed with a paired samples *t*-test. The primary PS model was based on clinical knowledge and adjusted for the following confounders at baseline: age, sex, DAS28, order of biologics, and concomitant csDMARDS. Additional supportive models that matched for different sets of confounders were also investigated (model 1 [M1], M2, M3, and M4): M1 adjusted for age; M2 adjusted for age and sex; M3 adjusted for age, sex, and DAS28; and M4 adjusted for age, sex, DAS28, order of biologics, and concomitant use of csDMARDs and the other clinical outcome measures (CRP level, ESR, SJC28, TJC28, PtGA, and MHAQ). The primary PS-matched model was found to be the most supportive based on clinical knowledge and data availability. A standardized difference of <0.1 indicates a good match. Drug persistence was analyzed by the Kaplan-Meier method. Kaplan-Meier estimates were calculated for week 52 and week 104 in unmatched (primary) and matched (supportive) analyses.

Secondary efficacy end points were analyzed based on the same approach as used for DAS28 and PS-matched models for supportive analyses. No imputation of missing data was performed for the yearly assessments. However, a sensitivity analysis for DAS28 that included all available data in the matched samples using regression analysis with standard errors for matched clusters was performed. For example, PS-matched pairs with only DAS28 data for 1 of the drugs at week 52 were excluded in the main matched analysis, but included in the sensitivity analysis.

Ethics approval and patient involvement. The study was approved by the regional ethical committee (Regional etisk komite Midt-Norge 2010/3078). No consent from patients was required by the committee, as all data were anonymized and collected as part of routine clinical care. Patients were not involved in the design, conduct, reporting or dissemination of this research.

RESULTS

Disposition and baseline characteristics. A total of 1,455 patients with RA from 5 participating outpatient clinics were included in this analysis (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25092), including 575 patients in the ETN cohort, 299 patients in the SB4 cohort, and 581 patients who had switched from ETN to SB4. Based on unmatched comparisons between the ETN and SB4 cohorts, there was a difference in DAS28 at baseline, with a mean ± SD of 4.3 ± 1.2 and 4.0 ± 1.3, respectively (Table 1). This results in a standardized difference (*d*) of 0.25. Differences were also observed at baseline between the 2 cohorts in age (*d* = 0.16), BMI (*d* = –0.13), SJC28 (*d* = 0.29), TJC28 (*d* = 0.19), and order of bDMARDs (*d* = 0.46).

After matching based on the primary PS model, there were 172 patients each in the ETN and SB4 cohorts; the mean ± SD DAS28 was 4.1 ± 1.3 and 4.1 ± 1.3, respectively (*d* = 0.00) (Table 1). Baseline characteristics showed a good overlap between the PS-matched cohorts based on $d \leq 0.1$, with the exceptions of BMI (–0.24), anti-CCP positivity (0.25), and CRP level (–0.17). For patients who switched from ETN to SB4, the mean ± SD DAS28 at baseline was 2.7 ± 1.2.

Primary outcome measure: DAS28 at week 52 (continuous). Before PS matching, the mean DAS28 at week 52 was 3.2 (95% confidence interval [95% CI] 3.0, 3.3) for the ETN cohort (n = 268) and 2.9 (95% CI 2.7, 3.1) for the SB4 cohort (n = 134) (Table 2 and Figure 1). After matching based on the primary PS model, the mean DAS28 in the baseline to week 52 period was 3.0 (95% CI 2.7, 3.3) for the ETN cohort (n = 49) and 3.2 (95% CI 2.8, 3.7) for the SB4 cohort (n = 49) (Table 3). For the switch cohort, the mean DAS28 was 2.4 (95% CI 2.3, 2.5) for the same period (n = 235). Details on the availability of patient data for the primary analysis are reported in Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25092.

The disease status in patients in the ETN cohort and the SB4 cohort at week 52 is shown in Figure 2. In the primary unmatched analysis, the mean difference in change from baseline/week 0 between ETN (n = 221) and SB4 (n = 106) cohorts was –0.02 (95% CI –0.32, 0.27) at week 52, demonstrating equivalence

Table 1. Baseline clinical data for unmatched (observed) and PS-matched ETN-naive patients with RA treated with ETN or SB4 and in patients with RA switched from ETN to SB4*

	Unmatched					PS-matched			
	ETN (n = 575)	SB4 (n = 299)	*P*	*d*	Switch SB4 (n = 581)	ETN (n = 172)	SB4 (n = 172)	*P*	*d*
Age, years	60.5 ± 15.2	58.1 ± 14.9	0.033	0.16	63.0 ± 13.1	59.0 ± 14.1	60.2 ± 14.0	0.324	−0.09
No.	474	299	–	–	581	–	–	–	–
Female, no. (%)	413 (71.8)	225 (75.3)	0.280	0.08	399 (68.7)	131 (76.2)	129 (75.0)	0.803	0.03
BMI, kg/m^2	25.8 ± 4.6	26.4 ± 5.0	0.100	−0.13	25.9 ± 4.2	25.4 ± 4.6	26.5 ± 5.0	0.039	−0.24
No.	368	278	–	–	463	143	143		
Disease duration, years	8.0 ± 9.5	8.2 ± 9.3	0.670	−0.02	14.4 ± 9.1–21.2	8.3 ± 9.7	8.6 ± 10.0	0.784	−0.03
Anti-CCP positive, no./total (%)	195/258 (75.6)	86/176 (48.9)	0.700	0.04	40/49 (81.6)	33/36 (91.7)	30/36 (83.3)	0.317	0.25
C-reactive protein, mg/liter	13.7 ± 15.8	12.3 ± 16.3	0.270	0.09	5.2 ± 8.3	11.5 ± 13.6	14.4 ± 18.6	0.065	−0.17
No.	397	282	–	–	399	169	169	–	–
ESR, mm/hour	23.4 ± 16.7	21.8 ± 16.3	0.250	0.10	16.2 ± 14.9	21.8 ± 17.0	22.4 ± 16.9	0.725	−0.04
No.	349	223	–	–	364	–	–	–	–
SJC28 (range 0–28)	4.1 ± 3.7	3.1 ± 3.1	<0.001	0.29	0.9 ± 1.7	3.6 ± 3.7	3.3 ± 3.4	0.439	0.09
No.	408	281	–	–	415	–	–	–	–
TJC28 (range 0–28)	5.6 ± 4.8	4.7 ± 4.3	0.013	0.19	1.6 ± 2.9	5.1 ± 4.5	5.0 ± 4.6	0.885	0.01
No.	408	281	–	–	415	–	–	–	–
PtGA (range 0–100 mm)	50.0 ± 22.7	51.1 ± 21.9	0.540	−0.05	32.2 ± 23.2	50.6 ± 23.1	49.4 ± 21.7	0.597	0.06
No.	414	270	–	–	406	–	–	–	–
DAS28	4.3 ± 1.2	4.0 ± 1.3	0.006	0.25	2.7 ± 1.2	4.1 ± 1.3	4.1 ± 1.3	0.965	0.00
No.	327	202	–	–	331	–	–	–	–
MHAQ (range 0–3)	0.7 ± 0.5	0.7 ± 0.4	0.950	0.01	0.4 ± 0.5	0.7 ± 0.4	0.6 ± 0.4	0.333	0.11
No.	395	272	–	–	405	167	167		
Current methotrexate, no. (%)	355 (61.7)	175 (58.5)	0.360	0.07	356 (61.3)	97 (56.4)	97 (56.4)	1.000	0.00
Current csDMARDs, no. (%)	417 (72.5)	212 (70.9)	0.610	0.04	396 (68.2)	121 (70.3)	122 (70.9)	0.903	0.01
Order of bDMARDs, no. (%)									
First	394 (68.5)	143 (47.8)	<0.001	0.46	0 (0.0)	85 (49.4)	87 (50.6)	0.872	0.02
Second	127 (22.1)	91 (30.4)	–	–	455 (78.3)	54 (31.4)	53 (30.8)	–	–
≥Third	54 (9.4)	65 (21.7)	–	–	126 (21.7)	33 (19.2)	32 (18.6)	–	–
Center, no. (%)									
University Hospital of North Norway	67 (11.7)	57 (19.1)	<0.001	0.46	46 (7.9)	26 (15.1)	31 (18.0)	<0.001	1.12
St. Olavs Hospital	156 (27.1)	113 (37.8)	–	–	107 (18.4)	69 (40.1)	50 (29.1)	–	–
Haukeland University Hospital/Helse Bergen	101 (17.6)	24 (8.0)	–	–	144 (24.8)	0 (0.0)	14 (8.1)	–	–
Sørlandet Hospital	119 (20.7)	34 (11.4)	–	–	83 (14.3)	73 (42.4)	25 (14.5)	–	–
Martina Hansens Hospital	132 (23.0)	71 (23.7)	–	–	201 (34.6)	4 (2.3)	52 (30.2)	–	–

* Values are the mean ± SD unless indicated otherwise. For the propensity score (PS)–matched population, the logistic regression model used for the PS matching adjusted for age, sex, Disease Activity Score in 28 joints (DAS28), order of biologics, and conventional synthetic disease-modifying antirheumatic drugs (csDMARDs). A *d* of <0.1 indicates a good match. anti-CCP = anti–cyclic citrullinated peptide; bDMARD = biologic DMARD; BMI = body mass index; CRP = C-reactive protein; *d* = standardized difference; ESR = erythrocyte sedimentation rate; ETN = etanercept; MHAQ = modified Health Assessment Questionnaire; no. = number of patients with available data; PtGA = patient global assessment; RA = rheumatoid arthritis; SB4 = ETN biosimilar drug; SJC28 = 28 swollen joint count; TJC28 = 28 tender joint count.

Table 2. Disease status prior to start of treatment, at baseline, and up to 104 weeks follow-up in ETN-naive patients with RA treated with ETN or SB4 in unmatched patient cohorts and in patients switched from ETN to SB4*

	52 weeks before baseline		Baseline		Baseline to 52 weeks		52–104 weeks	
Variable/treatment	No.	Mean (95% CI)	No.	Mean (95% CI)	No.	Mean (95% CI)	No.	Mean (95% CI)
DAS28								
ETN	174	3.8 (3.6, 4.0)	327	4.3 (4.2, 4.5)	268	3.2 (3.0, 3.3)	178	3.0 (2.8, 3.1)
SB4	125	3.5 (3.3, 3.8)	202	4.0 (3.8, 4.2)	134	2.9 (2.7, 3.1)	46	2.5 (2.2, 2.9)
Switch SB4	334	2.6 (2.5, 2.7)	331	2.7 (2.5, 2.8)	235	2.4 (2.3, 2.5)	200	2.4 (2.3, 2.5)
CRP, mg/liter								
ETN	216	13.2 (10.8, 15.6)	397	13.7 (12.1, 15.2)	336	7.6 (6.4, 8.9)	228	6.7 (4.7, 8.7)
SB4	189	8.4 (7.0, 9.8)	282	12.3 (10.4, 14.2)	227	6.9 (5.2, 8.5)	85	5.2 (3.2, 7.2)
Switch SB4	413	4.7 (3.9, 5.4)	399	5.2 (4.4, 6.0)	318	5.2 (4.2, 6.1)	259	4.6 (3.6, 5.7)
ESR, mm/hour								
ETN	194	21.7 (19.5, 23.9)	349	23.4 (21.6, 25.2)	286	17.1 (15.5, 18.7)	185	15.7 (13.8, 17.5)
SB4	140	20.4 (17.8, 22.9)	223	21.8 (19.6, 23.9)	155	15.7 (13.2, 18.2)	57	12.8 (8.5, 17.0)
Switch SB4	367	15.9 (14.5, 17.4)	364	16.2 (14.7, 17.7)	271	15.3 (13.4, 17.1)	231	15.6 (13.4, 17.7)
SJC28 (range 0–28)								
ETN	233	2.7 (2.3, 3.1)	408	4.1 (3.7, 4.5)	350	2.0 (1.7, 2.3)	233	1.5 (1.2, 1.9)
SB4	193	1.9 (1.5, 2.3)	281	3.1 (2.7, 3.5)	225	1.4 (1.1, 1.8)	85	0.9 (0.6, 1.3)
Switch SB4	418	0.9 (0.8, 1.1)	415	0.9 (0.7, 1.1)	316	0.6 (0.5, 0.7)	263	0.6 (0.5, 0.8)
TJC28 (range 0–28)								
ETN	233	4.1 (3.6, 4.7)	408	5.6 (5.2, 6.1)	350	3.0 (2.6, 3.4)	233	2.2 (1.8, 2.6)
SB4	193	3.5 (2.9, 4.1)	281	4.7 (4.2, 5.2)	225	2.8 (2.2, 3.3)	85	2.4 (1.6, 3.1)
Switch SB4	418	1.3 (1.1, 1.6)	415	1.6 (1.3, 1.9)	316	1.2 (0.9, 1.5)	263	0.9 (0.7, 1.2)
PtGA (0–100 mm)								
ETN	226	43.7 (40.7, 46.6)	414	50.0 (47.9, 52.2)	355	34.9 (32.4, 37.4)	237	31.6 (28.6, 34.6)
SB4	191	42.3 (38.9, 45.7)	270	51.1 (48.5, 53.7)	223	37.2 (34.0, 40.4)	90	39.1 (33.4, 44.7)
Switch SB4	420	30.0 (27.7, 32.2)	406	32.2 (29.9, 34.4)	310	30.4 (27.7, 33.1)	260	30.3 (27.5, 33.2)
MHAQ (range 0–3)								
ETN	213	0.6 (0.5, 0.6)	395	0.7 (0.6, 0.7)	351	0.5 (0.4, 0.5)	236	0.4 (0.4, 0.5)
SB4	191	0.5 (0.5, 0.6)	272	0.7 (0.6, 0.7)	224	0.5 (0.4, 0.5)	90	0.5 (0.4, 0.6)
Switch SB4	416	0.4 (0.4, 0.4)	405	0.4 (0.4, 0.5)	309	0.4 (0.4, 0.5)	258	0.4 (0.4, 0.5)

* 95% CI = 95% confidence interval; CRP = C-reactive protein; DAS28 = Disease Activity Score in 28 joints; ESR = erythrocyte sedimentation rate; ETN = etanercept; MHAQ = modified Health Assessment Questionnaire; PtGA = patient global assessment; RA = rheumatoid arthritis; SB4 = ETN biosimilar drug; SJC28 = 28 swollen joint count; TJC28 = 28 tender joint count.

based on an independent samples *t*-test. In the PS-matched analysis, the mean difference between ETN (n = 49) and SB4 (n = 49) for the primary outcome DAS28 at week 52 was 0.03 (95% CI –0.46, 0.52) using a paired samples *t*-test of matched pairs with complete data. Outcomes were consistent between the unmatched and PS-matched analyses for disease status based on DAS28 at week 52 (Figure 2). In the primary and supportive PS models for the matched analyses, 95% CIs included zero, but equivalence between the ETN and SB4 cohorts could not be determined in all models, based on 95% CIs not being entirely confined within the predefined equivalence margin of ±0.6. Equivalence was shown in all PS models, with the sensitivity analysis using all available data for PS-matched pairs.

DAS28 at week 104 (continuous). Before PS matching, the mean DAS28 in the week 52–104 period was 3.0 (95% CI 2.8, 3.1) for the ETN cohort (n = 178) and 2.5 (95% CI 2.2, 2.9) for the SB4 cohort (n = 46) (Table 2 and Figure 1). After PS matching, the mean DAS28 in the week 52–104 period was 3.4 (95% CI 2.7, 4.2) for the ETN cohort (n = 11) and 2.4 (95% CI 1.7, 3.0) for the SB4 (n = 11) cohort. In both the unmatched and PS-matched analyses, there was a reduction in disease activity as assessed by DAS28 at week 104 compared with baseline, in both the ETN- and SB4-treated patients (Figure 1).

Disease response: DAS28 at week 52 and week 104 (categorical). Disease response based on DAS28 at week 52 and week 104 was also assessed as a categorical variable defined by the EULAR response criteria. Before matching, similar proportions of patients in the ETN (n = 221) and SB4 (n = 106) cohorts achieved a good response (40.3% versus 39.6%), a moderate response (28.1% versus 24.5%), or no response (31.7% versus 35.8%) at week 52 (see Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25092). After matching based on the primary PS model, 49 patients were in each of the ETN and SB4 cohorts, of which 49.0% and 30.6%, respectively, achieved a good response at week 52 (see Supplementary Figure 3). Moderate responses at week 52 were observed in 14.3% and 34.7%, respectively, of patients in the ETN and SB4 cohorts, and no response in 36.7% and 34.7%. At week 52, the proportions of patients in the switch cohort (n = 173) achieving

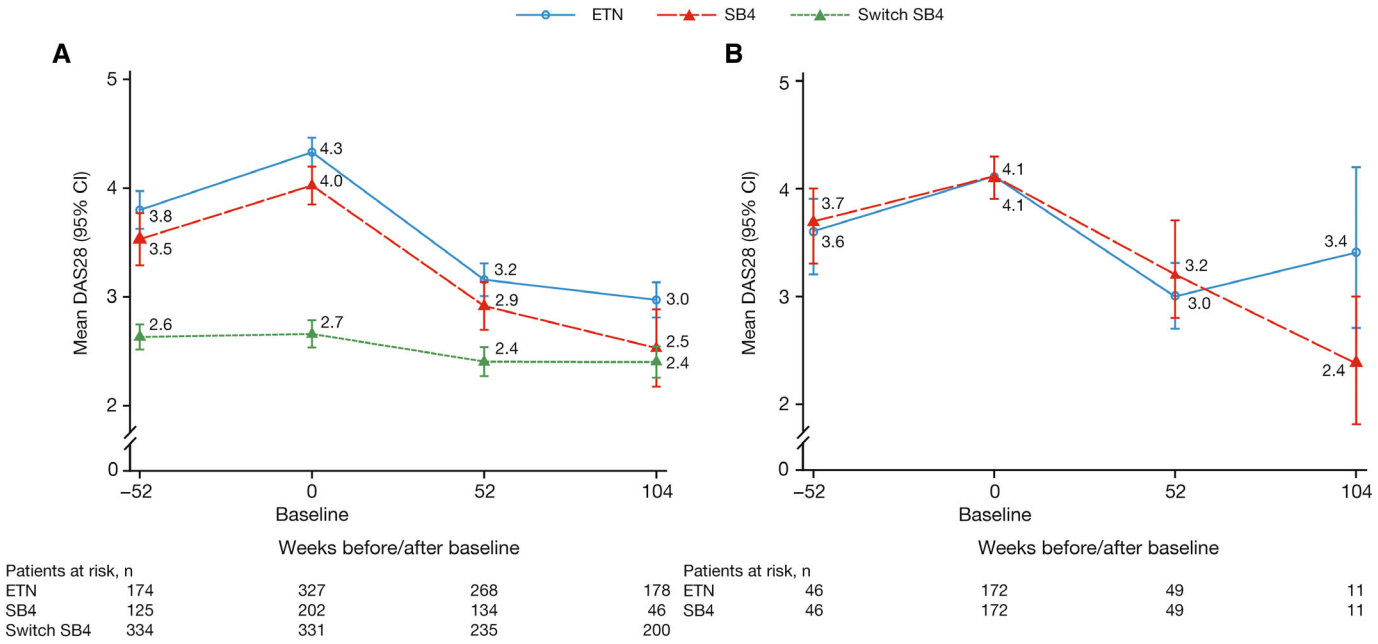

Figure 1. Disease activity expressed as Disease Activity Score in 28 joints (DAS28) over 2 years of treatment in **A**, unmatched, and **B**, propensity score–matched patients with rheumatoid arthritis. Numbers represent the numbers of patients in the unmatched and matched (primary PS model) populations. Data are shown as the mean with 95% confidence interval (95% CI). No imputation of missing data was performed. ETN = etanercept.

good or moderate responses were 9.8% and 14.5%, respectively, whereas 75.7% achieved no response.

Before matching, 37.0% versus 44.1% of patients in the ETN (n = 146) and SB4 (n = 34) cohorts, respectively, achieved a good response, 30.1% versus 14.7% achieved a moderate response, and 32.9% versus 41.2% achieved no response at week 104 (see Supplementary Figure 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25092). Based on the primary PS-matched analysis, 27.3% versus 54.6% of patients in the ETN (n = 11) and SB4 (n = 11) cohorts, respectively,

Table 3. Disease status before start of treatment, at baseline, and up to 104 weeks follow-up in ETN-naive patients with RA treated with ETN or SB4 in PS-matched cohorts in the primary PS model*

	52 weeks before baseline		Baseline		Baseline to 52 weeks		52–104 weeks	
Variable/treatment	No.	Mean (95% CI)	No.	Mean (95% CI)	No.	Mean (95% CI)	No.	Mean (95% CI)
DAS28								
ETN	46	3.6 (3.2, 3.9)	172	4.1 (3.9, 4.3)	49	3.0 (2.7, 3.3)	11	3.4 (2.7, 4.2)
SB4	46	3.7 (3.3, 4.0)	172	4.1 (3.9, 4.3)	49	3.2 (2.8, 3.7)	11	2.4 (1.7, 3.0)
CRP, mg/liter								
ETN	64	11.3 (7.8, 14.8)	169	11.5 (9.5, 13.6)	105	7.5 (5.0, 10.0)	18	7.6 (4.4, 10.9)
SB4	64	7.8 (5.6, 10.0)	169	14.4 (11.6, 17.2)	105	7.5 (4.7, 10.3)	18	4.2 (2.1, 6.3)
ESR, mm/hour								
ETN	51	19.6 (15.9, 23.2)	172	21.8 (19.2, 24.3)	63	15.4 (12.4, 18.4)	13	17.1 (10.0, 24.2)
SB4	51	20.9 (17.0, 24.8)	172	22.4 (19.8, 24.9)	63	18.0 (13.6, 22.4)	13	11.0 (6.8, 15.3)
SJC28 (range 0–28)								
ETN	76	2.7 (2.1, 3.4)	172	3.6 (3.1, 4.2)	108	2.0 (1.4, 2.5)	21	2.0 (0.7, 3.3)
SB4	76	1.6 (1.1, 2.1)	172	3.3 (2.8, 3.9)	108	1.7 (1.0, 2.3)	21	0.7 (0.1, 1.3)
TJC28 (range 0–28)								
ETN	76	4.4 (3.3, 5.5)	172	5.1 (4.4, 5.7)	108	3.2 (2.4, 4.0)	21	2.2 (0.7, 3.8)
SB4	76	3.4 (2.5, 4.2)	172	5.0 (4.3, 5.7)	108	2.9 (2.1, 3.7)	21	2.0 (0.5, 3.6)
PtGA (0–100 mm)								
ETN	76	47.2 (41.6, 52.9)	172	50.6 (47.2, 54.1)	111	35.7 (31.2, 40.1)	21	25.3 (14.0, 36.7)
SB4	76	40.7 (35.3, 46.2)	172	49.4 (46.1, 52.6)	111	35.6 (31.3, 39.8)	21	32.6 (22.5, 42.8)
MHAQ (range 0–3)								
ETN	75	0.6 (0.5, 0.7)	167	0.7 (0.6, 0.7)	112	0.5 (0.4, 0.6)	21	0.4 (0.2, 0.6)
SB4	75	0.5 (0.4, 0.6)	167	0.6 (0.6, 0.7)	112	0.5 (0.4, 0.5)	21	0.3 (0.2, 0.5)

* CRP = C-reactive protein; DAS28 = Disease Activity Score in 28 joints; ESR = erythrocyte sedimentation rate; ETN = etanercept; MHAQ = modified Health Assessment Questionnaire; PtGA = patient global assessment; PS = propensity score; RA = rheumatoid arthritis; SB4 = ETN biosimilar drug; SJC28 = 28 swollen joint count; TJC28 = 28 tender joint count.

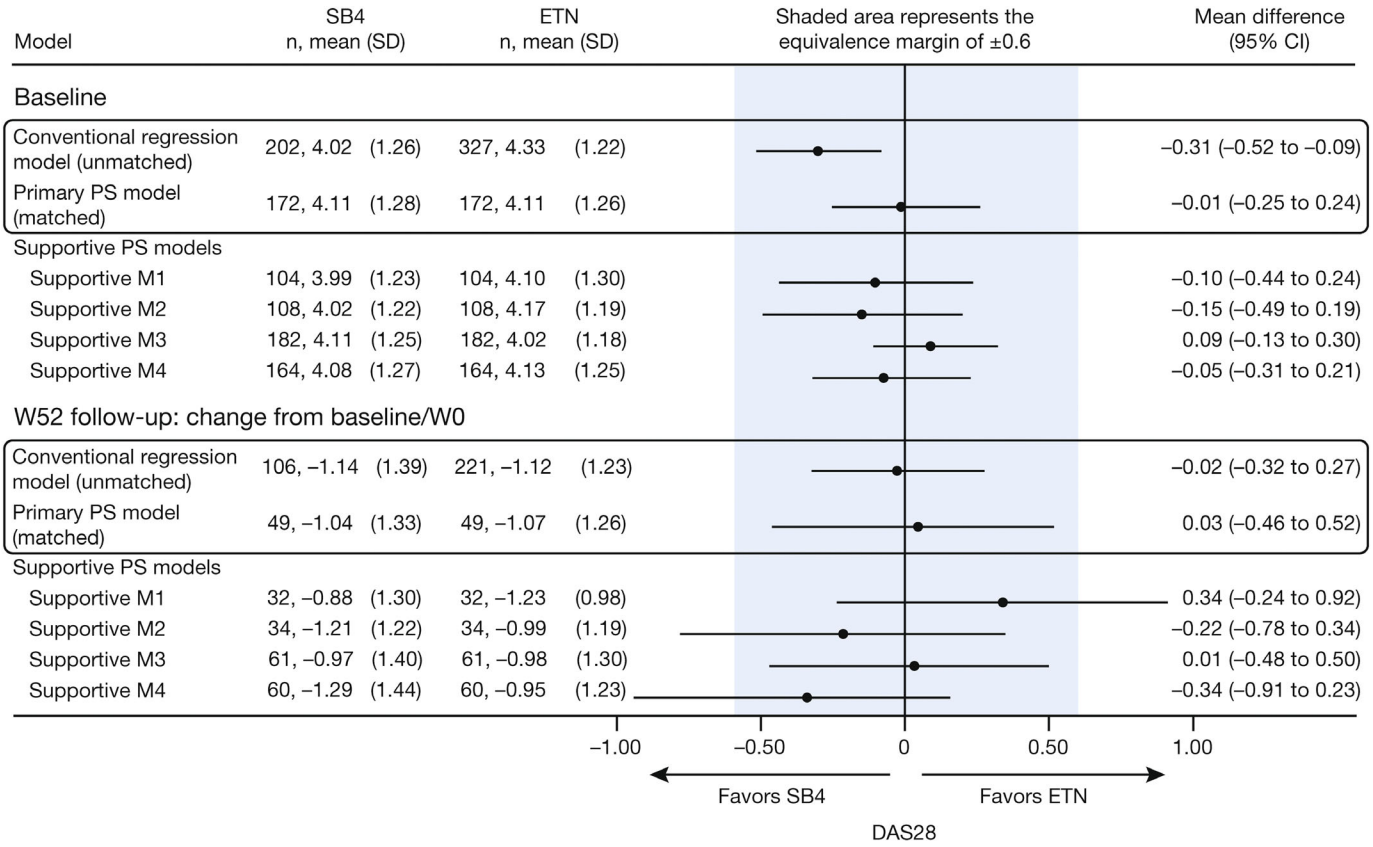

Figure 2. Comparison of effectiveness of etanercept (ETN) and the biosimilar SB4 in patients with rheumatoid arthritis on disease activity (Disease Activity Score in 28 joints [DAS28]) at baseline and week 52 (W52) follow-up for unmatched and propensity score (PS)–matched populations. Baseline shows absolute values and the differences between cohorts at baseline. Week 52 follow-up shows the change from baseline for each cohort and the differences between the cohorts at week 52. Mean differences are shown for baseline values; 1-year follow-up shows the mean difference for change from baseline. Unmatched and primary model models are highlighted. Secondary models are M1, M2, M3, and M4. 95% CI = 95% confidence interval; W0 = week 0.

achieved a good response, 27.3% versus 27.3% achieved a moderate response, and 45.5% versus 18.2% achieved no response at week 104 (see Supplementary Figure 3). At the same time point, the proportions of patients in the switch cohort (n = 148) who achieved good, moderate, or no responses were 11.5%, 15.5%, and 73.0%, respectively.

Drug persistence. In the unmatched sample, the estimated persistence at week 52 was 0.62 (95% CI 0.57, 0.65) for ETN and 0.66 (95% CI 0.60, 0.71) for SB4 (Figure 3A). The overlapping of 95% CIs for the unmatched population indicates similar persistence between ETN and SB4 cohorts. In the matched sample using the primary PS model, the estimated persistence at week 52 was 0.52 (95% CI 0.44, 0.59) for ETN and 0.68 (95% CI 0.61, 0.75) for SB4 (Figure 3B). For switched patients, the estimated persistence at week 52 was 0.80 (95% CI 0.76, 0.83).

In the unmatched sample, the estimated persistence at week 104 was 0.47 (95% CI 0.43, 0.51) for ETN and 0.56 (95% CI 0.49, 0.61) for SB4. In the matched sample using the primary PS model, the estimated persistence at week 104 was 0.37 (95% CI 0.29, 0.44) for ETN (n = 63) and 0.60 (95% CI 0.51, 0.67) for SB4 (n = 24). For nonmedical switch patients, the estimated persistence at week 104 was 0.73 (95% CI 0.69, 0.77).

A small number of patients with RA (n = 5) did not undergo the nonmedical switch from ETN to SB4, including 1 patient at Sørlandet Hospital and 4 at Martina Hansens Hospital; these patients did not contribute to this study. Similarly, a number of patients with RA (n = 48) switched back from SB4 to ETN: 3 at University Hospital of North Norway, 3 at St. Olavs Hospital, 7 at Haukeland University Hospital, 8 at Sørlandet Hospital, and 27 at Martina Hansens Hospital. Reasons for switching back to ETN were often subjective and included lack of efficacy and AEs.

Secondary effectiveness outcome measures. Secondary outcomes for the unmatched analyses at week 52 and week 104 are reported in Table 2. After PS matching based on the primary PS model, both ETN and SB4 cohorts

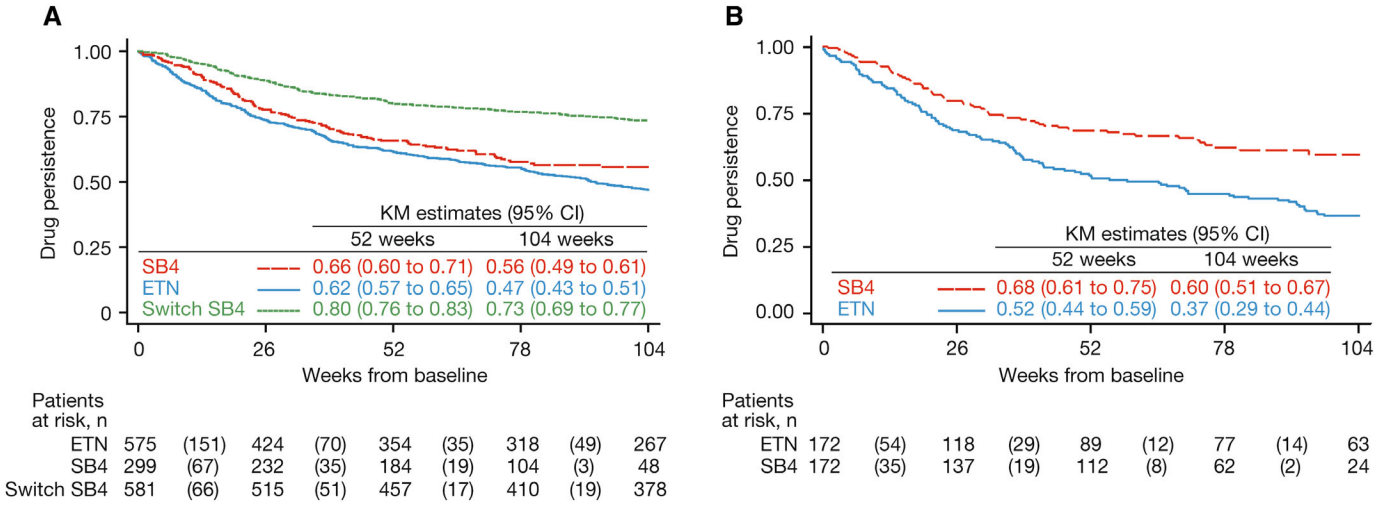

Figure 3. Kaplan-Meier (KM) plots of treatment retention rates among patients with rheumatoid arthritis, **A**, treated with etanercept (ETN) or the biosimilar SB4 or with a nonmedical switch from ETN to SB4, and **B**, treated with ETN or SB4 after propensity score (PS) matching based on the primary PS model. Listed under the graphs are the numbers of patients at risk and the numbers of patients who experienced an event and stopped treatment (shown in parentheses). 95% CI = 95% confidence interval.

experienced improvements from baseline to week 52 and week 104 in measures of disease activity (CRP level, ESR, SJC28, and TJC28) and patient-reported outcomes (PtGA and MHAQ) (Table 3).

Safety. After 104 weeks, 52.9% (n = 304) of ETN, 41.5% (n = 124) of SB4, and 26.2% (n = 152) of nonmedical switch patients had discontinued treatment (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25092). The most common reasons for drug discontinuation were AEs, occurring in 17.4%, 16.4%, and 8.1%, and lack of effect/no effect, occurring in 15.0%, 17.4%, and 9.6% of patients in the ETN, SB4, and switch cohorts, respectively. The most frequent AEs leading to discontinuation were skin involvement and infection. Reasons for stopping treatment in the PS-matched population are summarized in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25092.

DISCUSSION

This retrospective database study assessed disease activity and drug persistence in 1,455 patients with RA who were treated with ETN or the biosimilar SB4 for up to 2 years in routine clinical care at 5 outpatient clinics in Norway. Outcomes were compared between treatment cohorts using unmatched and matched analyses. For DAS28, unmatched analyses were based on independent samples *t*-tests, whereas matched analyses used PS models adjusted for confounders including age, sex, and baseline disease status. The primary outcome measure of DAS28 after 52 weeks of treatment was equivalent between cohorts of patients treated with ETN or SB4 based on independent samples *t*-tests and the applied predefined equivalence margin of ±0.6 (16). Consistent results were observed applying the primary PS model, but owing to the low number of patients with complete available disease scores at week 52 in the matched pairs, results could be uncertain. Therefore, the observed results may be limited by the nonsystematic capture of patients' disease scores.

Differences in baseline characteristics in the unmatched cohorts were observed, suggesting a selection bias in treatment initiation; hence, PS matching was investigated as a supportive analysis. The PS-matched models ensured comparability of treatment cohorts at baseline. Persistence using the drug at week 52 was similar between ETN and SB4 treatment cohorts based on the unmatched analysis, as indicated by overlapping 95% CIs. However, in PS-matched cohorts, persistence was greater for SB4 than for ETN at week 52 and week 104.

Although the frequency of drug discontinuation was higher in the ETN cohort than in the SB4 cohort (52.9% versus 41.5%), the reasons for discontinuation were consistent between cohorts and included AEs and lack of effectiveness/no effect. Further, these reasons for discontinuation occurred at similar frequencies between the 2 cohorts.

Published real-world data on SB4 are limited, particularly in patients naive to ETN. A study of the National Romanian Registry of Rheumatic Diseases followed patients with RA for 6 months and found no difference in effectiveness and safety between ETN (n = 123) and SB4 (n = 119) (8). A 2019 systematic review of SB4 real-world data found no difference in effectiveness and safety between switch or ETN-naive patients (9). Similar to

Norway, Denmark also operates a mandatory switch system. An analysis of the Danish DANBIO registry in patients with RA who switched from ETN to SB4 indicated no change in disease activity 3 months post-switch compared with the 3 months pre-switch (10). In addition, the 1-year adjusted retention rate for SB4 post-switch (0.83 [95% CI 0.79, 0.87]) was found to be somewhat lower than for a historical control group for ETN (0.90 [95% CI 0.88, 0.92]). A limitation of these analyses was that data were not reported for ETN or SB4 outcomes in treatment-naive patients who initiated treatment on ETN or SB4 (10).

Across different countries and regions, a nonmedical switch may follow a mandatory or nonmandatory switch model. Countries with mandatory switch models, including Denmark and Norway, have been shown to be more successful in using biosimilars than countries using nonmandatory models. In 2015, the infliximab biosimilar constituted as much as 90.6% of the total infliximab prescribed in Denmark 4 months after the patent expiration of the reference drug (18).

As for all observational studies, this study's limitations relate to measured and unmeasured confounding factors, attrition, and missing data. To counteract these limitations, we aimed to analyze the data and report the results in accordance with observational study recommendations, including GRACE and STROBE. The use of propensity statistics as supportive analyses mitigated the risk of selection bias, simulating a randomized study design. We analyzed the primary outcome measure with different propensity matching adjustments to explore the robustness of the results. In matched pairs analysis, missing data may have a substantial impact. Typically, a matched pair at baseline may only have data for 1 of them at week 52 and then be lost. Therefore, a sensitivity analysis was conducted using all available data, and produced outcomes consistent with those from the primary matched pairs analyses.

This study aimed to closely observe the recommendations of the EULAR Points to Consider When Analyzing and Reporting Comparative Effectiveness Research with Observational Data in Rheumatology (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25092) (7,13). Although most recommendations were followed closely, there were some minor deviations. First, the numbers of patients who stopped and/or changed therapies over time may not have been fully captured. Second, although other analyses that are not reported here were performed, a sensitivity analysis investigating the missing data pattern was conducted by not excluding matched pairs with partially missing data. Finally, although a full statistical analysis plan had not been prepared, an outline was developed in advance of this study.

In conclusion, after 52 weeks of treatment, disease outcomes based on DAS28 were comparable between cohorts of patients treated with ETN or SB4, and equivalence for DAS28 was demonstrated based on independent sample *t*-tests. Consistent results were observed applying the primary PS model but should be interpreted with caution owing to missing patient disease scores at week 52. Persistence was similar at week 52 between the ETN and SB4 cohorts in the unmatched populations but greater for SB4 in the PS-matched analyses.

ACKNOWLEDGMENTS

The authors thank Bjørg Tilde-Fevang for providing data from Haukeland University Hospital and Ulrich Freudensprung of Biogen International GmbH for statistical guidance. They are also grateful to all doctors, nurses, secretaries, and patients at the participating outpatient clinics who have contributed in daily clinical practice to make this real-life study possible. Medical writing support for the preparation of the manuscript was provided by David Pertab, PhD, Excel Medical Affairs.

ROLE OF THE STUDY SPONSOR

Biogen International had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Biogen International.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Haugeberg had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Haugeberg, Bakland, Rødevand, Hansen, Diamantopoulos, Pripp.

Acquisition of data. Haugeberg, Bakland, Rødevand, Hansen, Diamantopoulos.

Analysis and interpretation of data. Haugeberg, Pripp.

REFERENCES

1. Bergstra SA, Branco JC, Vega-Morales D, et al. Inequity in access to bDMARD care and how it influences disease outcomes across countries worldwide: results from the METEOR-registry. Ann Rheum Dis 2018;77:1413–20.
2. Putrik P, Ramiro S, Kvien TK, et al. Inequities in access to biologic and synthetic DMARDs across 46 European countries. Ann Rheum Dis 2014;73:198–206.
3. Putrik P, Ramiro S, Kvien TK, et al. Variations in criteria regulating treatment with reimbursed biologic DMARDs across European countries: are differences related to country's wealth? Ann Rheum Dis 2014;73:2010–21.
4. Emery P, Vencovsky J, Sylwestrzak A, et al. A phase III randomised, double-blind, parallel-group study comparing SB4 with etanercept reference product in patients with active rheumatoid arthritis despite methotrexate therapy. Ann Rheum Dis 2017;76:51–7.
5. Emery P, Vencovský J, Sylwestrzak A, et al. 52-week results of the phase 3 randomized study comparing SB4 with reference etanercept in patients with active rheumatoid arthritis. Rheumatology (Oxford) 2017;56:2093–101.
6. Kay J, Schoels MM, Dörner T, et al. Consensus-based recommendations for the use of biosimilars to treat rheumatological diseases. Ann Rheum Dis 2018;77:165–74.
7. Lauper K, Kedra K, de Wit M, et al. A systematic review to inform the EULAR points to consider when analyzing and reporting comparative

effectiveness research with observational data in rheumatology. Ann Rheum Dis 2020;79 Suppl 1:2.

8. Codreanu C, Popescu CC, Mogoşan C, et al. Efficacy and safety of original and biosimilar etanercept (SB4) in active rheumatoid arthritis: a comparison in a real-world national cohort. Biologicals 2019;62: 27–32.
9. Ebbers HC, Pieper B, Issa A, et al. Real-world evidence on etanercept biosimilar SB4 in etanercept-naïve or switching patients: a systematic review. Rheumatol Ther 2019;6:317–38.
10. Glintborg B, Loft AG, Omerovic E, et al. To switch or not to switch: results of a nationwide guideline of mandatory switching from originator to biosimilar etanercept. One-year treatment outcomes in 2061 patients with inflammatory arthritis from the DANBIO registry. Ann Rheum Dis 2019;78:192–200.
11. Madenidou AV, Jeffries A, Varughese S, et al. Switching patients with inflammatory arthritis from Etanercept (Enbrel®) to the biosimilar drug, SB4 (Benepali®): a single-centre retrospective observational study in the UK and a review of the literature. Mediterr J Rheumatol 2019;30 Suppl 1:69–75.
12. Tweehuysen L, Huiskes VJ, van den Bemt BJ, et al. Open-label, non-mandatory transitioning from originator etanercept to biosimilar SB4: six-month results from a controlled cohort study. Arthritis Rheumatol 2018;70:1408–18.
13. Courvoisier D, Lauper K, Bergstra SA, et al. Points to consider when analysing and reporting comparative effectiveness research with observational data in rheumatology. Ann Rheum Dis 2020;79 Suppl 1:2.
14. Dreyer NA, Bryant A, Velentgas P. The GRACE checklist: a validated assessment tool for high quality observational studies of comparative effectiveness. J Manag Care Spec Pharm 2016;22:1107–13.
15. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370: 1453–7.
16. Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol 2005;23 Suppl 39:S93–9.
17. Van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum 1998;41:1845–50.
18. Jensen TB, Bartels D, Saedder EA, et al. The Danish model for the quick and safe implementation of infliximab and etanercept biosimilars. Eur J Clin Pharmacol 2020;76:35–40.