EXTENDED REPORT

A phase III randomised, double-blind, parallel-group study comparing SB4 with etanercept reference product in patients with active rheumatoid arthritis despite methotrexate therapy

Paul Emery,1,2 Jiří Vencovsky,3 Anna Sylwestrzak,4 Piotr Leszczyński,5 Wiesława Porawska,6 Asta Baranauskaitė,7 Vira Tseluyko,8 Vyacheslav M Zhdan,9 Barbara Stasiuk,10 Roma Milasiene,11 Aaron Alejandro Barrera Rodriguez,12 Soo Yeon Cheong,13 Jeehoon Ghil13

ABSTRACT

Objectives To compare the efficacy and safety of SB4 (an etanercept biosimilar) with reference product etanercept (ETN) in patients with moderate to severe rheumatoid arthritis (RA) despite methotrexate (MTX) therapy.

Methods This is a phase III, randomised, double-blind, parallel-group, multicentre study with a 24-week primary endpoint. Patients with moderate to severe RA despite MTX treatment were randomised to receive weekly dose of 50 mg of subcutaneous SB4 or ETN. The primary endpoint was the American College of Rheumatology 20% (ACR20) response at week 24. Other efficacy endpoints as well as safety, immunogenicity and pharmacokinetic parameters were also measured.

Results 596 patients were randomised to either SB4 (N=299) or ETN (N=297). The ACR20 response rate at week 24 in the per-protocol set was 78.1% for SB4 and 80.3% for ETN. The 95% CI of the adjusted treatment difference was −9.41% to 4.98%, which is completely contained within the predefined equivalence margin of −15% to 15%, indicating therapeutic equivalence between SB4 and ETN. Other efficacy endpoints and pharmacokinetic endpoints were comparable. The incidence of treatment-emergent adverse events was comparable (55.2% vs 58.2%), and the incidence of antidrug antibody development up to week 24 was lower in SB4 compared with ETN (0.7% vs 13.1%).

Conclusions SB4 was shown to be equivalent with ETN in terms of efficacy at week 24. SB4 was well tolerated with a lower immunogenicity profile. The safety profile of SB4 was comparable with that of ETN.

Trial registration numbers NCT01895309, EudraCT 2012-005026-30.

INTRODUCTION

Etanercept is a recombinant human tumour necrosis factor (TNF) receptor p75Fc fusion protein. Etanercept is well established and has been widely used in clinical practice for about 15 years, with a well-characterised pharmacological, efficacy and safety profile.1,3 Originally licensed for use in moderate to severe rheumatoid arthritis (RA), the therapeutic indications have been stepwise extended and comprise treatment of patients with polyarticular juvenile idiopathic arthritis, psoriatic arthritis, ankylosing spondylitis, psoriasis and also paediatric psoriasis. Recently, etanercept has been also approved for use in non-radiographic axial spondyloarthritis by the European Medicines Agency (EMA).6

A biosimilar is a biological medicinal product that contains a version of the active substance of an already authorised original biological medicinal product (reference medicinal product). A biosimilar demonstrates similarity to the reference medicinal product in terms of quality characteristics, biological activity, safety and efficacy based on a comprehensive comparability exercise.7–9

SB4 has been developed as a biosimilar to the reference product etanercept (ETN). SB4 is produced by recombinant DNA technology in Chinese hamster ovary mammalian cell expression system. Similar structural, physicochemical and biological activities of SB4 and ETN have been shown using state-of-the-art analytical methods including peptide mapping, TNF-α binding assay and TNF-α neutralisation cell-based assay. Equivalence in the pharmacokinetics (PK) between SB4 and ETN was demonstrated in a phase I study conducted in healthy male subjects.10 The objective of this study was to compare the efficacy, safety, PK and immunogenicity of SB4 and ETN in patients with RA.

METHODS

Patients

Patients aged 18–75 years who have been diagnosed with RA according to the revised 1987 American College of Rheumatology (ACR) criteria for ≥6 months and ≤15 years prior to screening were eligible for the study. Patients had to have active disease defined as ≥6 swollen and ≥6 tender joints and either erythrocyte sedimentation rate (ESR) ≥28 mm/h or serum C reactive protein (CRP) ≥1.0 mg/dl. Despite methotrexate (MTX) treatment for ≥6 months (stable dose of 10–25 mg/week for ≥4 weeks prior to screening). Non-steroidal anti-inflammatory drugs and oral glucocorticoids (equivalent to ≤10 mg prednisolone) were permitted if received at a stable dose for ≥4 weeks prior to randomisation.

Major exclusion criteria consisted of previous treatment with any biological agents, history of

lymphoproliferative disease, congestive heart failure (New York Heart Association Class III/IV) or demyelinating disorders, diagnosis of active tuberculosis (TB) and pregnancy or breast feeding at screening.

Additional eligibility criteria are listed in online supplementary appendix 1.

Study design
This phase III, randomised, double-blind, parallel-group study was conducted at 73 centres across 10 countries in Europe, Latin America and Asia. Patients were randomised in a 1:1 ratio to receive 50 mg of either SB4 or ETN (see online supplementary appendix 2). Patients self-administered SB4 or ETN once weekly for up to 52 weeks via subcutaneous injection. All patients had to take MTX (10–25 mg/week) and folic acid (5–10 mg/week) during the study. This study is currently ongoing, and this report represents efficacy data up to 24 weeks of treatment and safety data up to the 24-week interim report data cut-off point (21 July 2014).

Study endpoints
The primary endpoint was the ACR20 response rate at week 24. Other efficacy endpoints were the ACR50 and ACR70 responses, the numeric index of the ACR response (ACR-N), change in the disease activity score in 28 joints (DAS28) based on ESR, the area under the curve (AUC) of the ACR-N, AUC of the change in DAS28 and the European League Against Rheumatism (EULAR) response. Safety endpoints included incidence of adverse events (AEs) and serious adverse events (SAEs).

PK analyses were performed in the PK population, which included a subset of patients from pre-designated study sites. Key PK endpoints included serum trough concentration (C\text{trough}) and area under the concentration–time curve during the dosing interval (AUC\text{C}) at steady state. Serum concentrations were determined using a validated ELISA, and PK parameters were calculated by non-compartmental analyses (WinNonlin V5.2 or higher, Pharsight, Mountain View, California, USA).

Immunogenicity was measured in all patients. The immunogenicity endpoints were incidence of antidrug antibodies (ADAs) and neutralising antibodies (NAbs). A single-assay approach with SB4 tag was used to assess immunogenicity. ADAs were measured using validated electrochemiluminescence immunoassays, and NAbs were measured using a competitive ligand-binding assay.

Details on the serum measurement and ADA detection assay can be found in online supplementary appendix 3.

Statistical analyses
Sample size was determined using the historical data for the equivalence test. The expected ACR20 response rate at week 24 for both SB4 and ETN was expected to be 60% from the previous ETN pivotal studies.\textsuperscript{11-13} Based on the expected response rate, the equivalence margin of −15% to 15% at week 24 for ACR20 response rate was calculated in line with the US Food and Drug Administration Guidance for Industry Non-Inferiority Clinical Trials and the Committee for Medicinal Products for Human Use Guideline on the Choice of the Non-inferiority Margin and was also agreed with the regulatory agencies.\textsuperscript{14 16}

Given a two-sided α level of 0.05 and 80% power, the two-sided 15% equivalence margin required 438 patients for the per-protocol set (PPS). Assuming 20% loss of patients from the PPS, the study required a minimum of 548 randomised patients.

The primary efficacy analysis for ACR20 response at week 24 was performed in the PPS in which patients completed week 24 visit, received 80–120% of both the expected number of study drug administrations and the expected sum of MTX doses, and did not have any major protocol deviations affecting the efficacy assessment. To declare the equivalence between the two treatment groups, the 95% CI of the adjusted treatment difference had to be entirely contained within the equivalence margin of −15% to 15%. The 95% CI of the difference of ACR20 response rates was estimated non-parametrically using the Mantel–Haenszel weights for region while adjusting for the baseline CRP. As a sensitivity analysis, the same analysis was repeated for the full analysis set (FAS) with missing data at week 24 considered as non-responses to explore the robustness of the results. Similar analyses were performed for ACR50 and ACR70 responses at week 24. Other secondary endpoints are summarised descriptively.

In addition, the exponential time–response model for ACR20 response rate was used to investigate the treatment difference during the time course of the study up to week 24.\textsuperscript{16} Details on the time–response model are provided in online supplementary appendix 4.

Safety and immunogenicity endpoints were analysed descriptively on the safety set that included all patients who received at least one dose of study drug. PK endpoints were summarised descriptively on the PK population who had at least one PK sample collected.

The analyses were performed using SAS V9.2 software (SAS Institute, Cary, North Carolina, USA).

RESULTS
Patient disposition and baseline characteristics
Patient screening began in June 2013, and the 24-week evaluation of the last patient occurred in April 2014. Overall, 777 patients were screened, of whom 596 patients were randomised. A total of 551 patients completed 24 weeks of treatment and 481 (80.7%) patients were included in the PPS (75 patients were excluded from the PPS due to protocol deviations, see online supplementary table S1). Patients withdrew before week 24 mainly due to AEs (3.7%) and withdrawal of consent (2.7%) (figure 1). The demographic and baseline disease characteristics were comparable between treatment groups (table 1).

Efficacy
The ACR20 response rate at week 24 in the PPS was 78.1% for SB4 and 80.3% for ETN. The 95% CI of the adjusted difference (SB4—ETN) in ACR20 response rate was within the predefined equivalence margin of −15% to 15% in both the PPS (95% CI −9.41% to 9.98%) and FAS (95% CI −5.24% to 9.07%), indicating therapeutic equivalence between SB4 and ETN (figure 2). The time–response models of SB4 and ETN up to week 24 in the PPS were estimated to be equivalent since the treatment difference in terms of the two-norm difference was 12.7 and the 95% CI was −4.6 to 30.0, where the upper limit 30.0 was less than the pre-specified equivalence margin of 83.28 (figure 3).

The ACR50 and ACR70 response rates at week 24 in the PPS and FAS were equivalent between SB4 and ETN. The ACR50 response rate was 46.6% vs 42.3%, and the ACR70 response rate was 25.5% vs 22.6% in the SB4 and ETN, respectively, as shown in figure 2.

Subgroup analyses on the ACR response rates in PPS showed comparable results regardless of ADA status. The proportion of patients who achieved ACR20 response rate in patients with ADA-negative results was 78.0% in SB4 and 81.5% in ETN (95% CI −11.12% to 3.99%) (see online supplementary table S2).
Figure 1  Summary of patient disposition. A total of 777 patients were screened and 181 patients were excluded mainly due to meeting the exclusion criteria. Multiple screening failure reasons were possible. All patients randomised were included in the full analysis set and the safety set. Of the 551 patients who completed 24 weeks of treatment, 481 patients were included in the per-protocol set. ETN, reference product etanercept.

Table 1  Baseline demographics and disease characteristics

<table>
<thead>
<tr>
<th></th>
<th>SB4 50 mg</th>
<th>ETN 50 mg</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N=299</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N=297</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N=596</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>52.1 (11.72)</td>
<td>51.6 (11.63)</td>
<td>51.8 (11.67)</td>
</tr>
<tr>
<td>Age group, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65 years</td>
<td>253 (84.6)</td>
<td>262 (88.2)</td>
<td>515 (86.4)</td>
</tr>
<tr>
<td>≥65 years</td>
<td>46 (15.4)</td>
<td>35 (11.8)</td>
<td>81 (13.6)</td>
</tr>
<tr>
<td>Gender n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>50 (16.7)</td>
<td>44 (14.8)</td>
<td>94 (15.8)</td>
</tr>
<tr>
<td>Female</td>
<td>249 (83.3)</td>
<td>253 (85.2)</td>
<td>502 (84.2)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>279 (93.3)</td>
<td>273 (91.9)</td>
<td>552 (92.6)</td>
</tr>
<tr>
<td>American Indian or Alaskan Native</td>
<td>5 (1.7)</td>
<td>7 (2.4)</td>
<td>12 (2.0)</td>
</tr>
<tr>
<td>Asian</td>
<td>11 (3.7)</td>
<td>13 (4.4)</td>
<td>24 (4.0)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (1.3)</td>
<td>4 (1.3)</td>
<td>8 (1.3)</td>
</tr>
<tr>
<td>Weight (kg), mean (SD)</td>
<td>72.5 (15.93)</td>
<td>71.0 (14.63)</td>
<td>71.8 (15.30)</td>
</tr>
<tr>
<td>Height (cm), mean (SD)</td>
<td>164.4 (8.78)</td>
<td>164.4 (8.55)</td>
<td>164.4 (8.66)</td>
</tr>
<tr>
<td>BMI (kg/m²), mean (SD)</td>
<td>26.8 (5.51)</td>
<td>26.3 (5.30)</td>
<td>26.6 (5.41)</td>
</tr>
<tr>
<td>Disease duration (years), mean (SD)</td>
<td>6.0 (4.20)</td>
<td>6.2 (4.41)</td>
<td>6.1 (4.30)</td>
</tr>
<tr>
<td>Duration of MTX use (months), mean (SD)</td>
<td>48.2 (39.87)</td>
<td>47.1 (40.77)</td>
<td>47.7 (40.29)</td>
</tr>
<tr>
<td>MTX dose (mg/week), mean (SD)</td>
<td>15.6 (4.52)</td>
<td>15.5 (4.60)</td>
<td>15.5 (4.56)</td>
</tr>
<tr>
<td>Swollen joint count (0–66), mean (SD)</td>
<td>15.4 (7.48)</td>
<td>15.0 (7.30)</td>
<td>15.2 (7.39)</td>
</tr>
<tr>
<td>Tender joint count (0–68), mean (SD)</td>
<td>23.5 (11.90)</td>
<td>23.6 (12.64)</td>
<td>23.5 (12.26)</td>
</tr>
<tr>
<td>HAQ-DI (0–3), mean (SD)</td>
<td>1.49 (0.553)</td>
<td>1.50 (0.560)</td>
<td>1.50 (0.556)</td>
</tr>
<tr>
<td>Physician global assessment VAS (0–100), mean (SD)</td>
<td>62.2 (15.09)</td>
<td>63.2 (14.76)</td>
<td>62.7 (14.92)</td>
</tr>
<tr>
<td>Subject global assessment VAS (0–100), mean (SD)</td>
<td>61.7 (18.97)</td>
<td>63.0 (17.70)</td>
<td>62.4 (18.35)</td>
</tr>
<tr>
<td>Subject pain assessment VAS (0–100), mean (SD)</td>
<td>61.8 (20.22)</td>
<td>62.3 (19.22)</td>
<td>62.1 (19.71)</td>
</tr>
<tr>
<td>DAS28 (ESR), mean (SD)</td>
<td>6.5 (0.91)</td>
<td>6.5 (0.88)</td>
<td>6.5 (0.89)</td>
</tr>
<tr>
<td>C reactive protein (mg/dL), mean (SD)</td>
<td>1.5 (2.00)</td>
<td>1.3 (1.50)</td>
<td>1.4 (1.81)</td>
</tr>
<tr>
<td>Erythrocyte sedimentation rate (mm/h), mean (SD)</td>
<td>46.5 (22.10)</td>
<td>46.4 (22.62)</td>
<td>46.5 (22.24)</td>
</tr>
<tr>
<td>Rheumatoid factor positive, n (%)</td>
<td>237 (79.3)</td>
<td>231 (77.8)</td>
<td>468 (78.5)</td>
</tr>
</tbody>
</table>

BMI, body mass index; DAS28, disease activity score in 28 joints; ETN, reference product etanercept; ESR, erythrocyte sedimentation rate; HAQ-DI, Health Assessment Questionnaire-Disability Index; MTX, methotrexate; VAS, visual analogue scale.
The mean improvement in DAS28 from baseline was 2.6 and 2.5 at week 24 in SB4 and ETN, respectively (95% CI −0.14 to 0.28) (figure 4A). The proportion of patients achieving good or moderate EULAR response (figure 4B), low-disease activity score or remission (figure 4C) at week 24 according to DAS28 were similar between SB4 and ETN. The ACR-N at week 24 was 45.0% in SB4 and 43.7% in ETN. The AUC of ACR-N up to week 24 (5822.2 vs 5525.0) and the AUC of change in DAS28 from baseline up to week 24 (358.3 vs 343.5) were comparable between SB4 and ETN.

Safety
Overall, 165 (55.2%) patients in SB4 and 173 (58.2%) patients in ETN reported at least one treatment-emergent adverse event (TEAE). Frequently occurring TEAEs by preferred term are shown in table 2, and the most frequently reported TEAE were upper respiratory tract infection (7.0%) and alanine aminotransferase increased (5.0%) in the SB4 and injection site erythema (11.1%), upper respiratory tract infection (5.1%) and nasopharyngitis (5.1%) in ETN. Most of the TEAEs were mild to moderate in severity, and TEAEs considered related to the study drug were reported in 83 (27.8%) and 106 (35.7%) patients for SB4 and ETN, respectively. Serious TEAEs were reported in 13 patients each in SB4 and ETN and 34 patients discontinued treatment due to TEAE (15 (5.0%) patients vs 19 (6.4%) patients).

A total of 25 patients (13 patients for SB4 and 12 patients for ETN) were diagnosed at screening with latent TB but entered the study after completing at least 30 days of treatment for latent TB and while receiving treatment. None of these patients or any other patients developed active TB during the study. Other serious infections were reported in one (0.3%) patient in SB4 and four (1.3%) patients in ETN. Malignancies were reported in three (1.0%) patients in SB4 (basal cell carcinoma, breast cancer and lung cancer metastatic) and in one (0.3%) patient in ETN (invasive ductal breast carcinoma).

Injection site reactions (ISRs), counted by the high-level term of administration site reactions, occurred in fewer patients in SB4 compared with ETN. There were 22 ISRs reported in 11 (3.7%) patients vs 156 ISRs reported in 51 (17.2%) patients in SB4 and ETN, respectively (p<0.001). Most of the ISRs occurred early (between weeks 2 and 8) and were mild in severity. The incidence of ISR for SB4 and ETN were 3.7% vs 17.1% in ADA-negative patients and 0.0% vs 17.9% in ADA-positive patients, respectively (see online supplementary table S3).

One death was reported in the SB4 treatment group due to cardiorespiratory failure, which was not considered related to the study drug.

Pharmacokinetics
PK analyses were performed on 79 patients (41 patients in SB4 and 38 patients in ETN).

$C_{\text{tough}}$ were comparable at each time point between SB4 (ranging from 2.419 to 2.886 μg/mL in weeks 2–24) and ETN (ranging from 2.066 to 2.635 μg/mL in weeks 2–24) (see online supplementary figure S2). The AUC, at week 8 was 676.4 vs 520.9 μg h/mL and the inter-subject variability (CV%) was 37.7% vs 50.1% in SB4 and ETN, respectively (see online supplementary figure S3).

Immunogenicity
The incidence of ADA was significantly lower in SB4 compared with ETN. Two (0.7%) patients in SB4 and 39 (13.1%) patients in ETN tested positive at least once up to week 24 (p<0.001), and only one sample from the ETN group had neutralising capacity. The ADAs appeared early (between weeks 2 and 8), and most of the ADAs disappeared after week 12 (see online supplementary appendix 9).

DISCUSSION
In this randomised, double-blind, parallel-group, multicentre study, the efficacy, safety, PK and immunogenicity of SB4 were compared with those of ETN in patients with moderate to
severe RA despite MTX treatment. Equivalence of efficacy between SB4 and ETN was demonstrated and the safety of SB4 was generally comparable to ETN.

The primary endpoint at week 24 was met: the 95% CI of the adjusted treatment difference between SB4 and ETN in ACR20 response rate was within the predefined equivalence margin of −15% to 15%. The ACR20 responses observed in this study (73.8% for SB4 and 71.7% for ETN in FAS) were within the range of ACR20 response rates reported in pivotal studies with ETN (49–86%) but slightly higher than what was assumed (60%). Since active treatment is used in both groups, biosimilar studies tend to show higher ACR20 response rates compared with pivotal controlled studies.

As the primary efficacy assessment (ACR20 response at week 24) was evaluated at a time point in the therapeutic plateau, various efficacy endpoints and statistical methods were applied to detect any non-equivalence in efficacy and to support the robustness of the primary efficacy analysis. The ACR20 response rate, ACR-N and DAS28 were measured at several different time points early in the treatment period. The time–response curves of SB4 and ETN

Figure 3  Estimated time–response curves of American College of Rheumatology 20% (ACR20) response rate up to week 24 in the per-protocol set. For details of the estimation process, please refer to the main text. ETN, reference product etanercept.

Figure 4  Changes over time in the disease activity score in 28 joints (DAS28) and European League Against Rheumatism (EULAR) responses at week 24 in the full analysis set. (A) Change in DAS28 up to week 24. (B) EULAR response based on DAS28. (C) Proportion of patients achieving low-disease activity score (LDAS) defined as DAS28 ≤3.2 and remission defined as DAS28 ≤2.6. ETN, reference product etanercept.
up to week 24 showing the ACR20 response over time were estimated to be equivalent, and the AUC of ACR-N up to week 24 and AUC of the change in DAS28 (ESR) from baseline up to week 24 were comparable between SB4 and ETN, indicating that the efficacy of SB4 over time was similar to ETN.

Overall, the safety profile of SB4 was comparable with that of ETN and was similar to those observed in the pivotal trials with ETN. There were no cases of active TB and only one patient in SB4 and four patients in ETN reported serious infection, which is lower than 6.3% shown in ETN product information. Malignancies were reported in three (1.0%) patients in SB4 and one (0.3%) patient in ETN. The incidence of malignancy observed in this study is similar to the previously conducted studies. Interestingly, ISRs were reported in fewer patients from SB4 compared with ETN (3.7% vs 17.2%). The proportion of patients who experienced at least one ISR from ETN group in this study (17.2%) is in line with recently conducted studies with 50 mg once weekly ETN, and most ISRs occurred in the first month, which is in accordance with the reference product label. Although it is unclear why the incidence of ISR was lower in SB4 compared with ETN, the difference in drug product formulation and container closure system may have contributed to the lower ISR. The only difference in drug composition between SB4 and ETN is the absence of L-arginine in SB4. It has not been shown that L-arginine is associated with increased risk of ISR; however, we cannot preclude the sole difference in formulation (absence of L-arginine) as the cause of ISR. In addition, natural rubber latex known to cause hypersensitivity reactions has not been used in the needle shield of SB4. There appears to be no correlation between ISR and ADA development, which is consistent with previously conducted studies.

In this study, C_{trough} and steady state PK was investigated in a subset of population to provide supporting evidence to the phase I comparative PK study in healthy subjects, which demonstrated similar PK behaviour. In the phase III study, the C_{trough} values were comparable between SB4 and ETN at each time point and AUC, at steady state was relatively higher in SB4 compared with ETN; however, the numerical difference is likely due to an inherent high inter-subject variability (37.7% vs 50.1%).

The incidence of ADA shown in the ETN group of this study (13.1%) is seemingly higher compared with what has been reported in the previous studies. In this study, the assay used to detect immunogenicity was more sensitive and immunogenicity was measured more frequently; most of the ADAs were detected at week 4 in this study while immunogenicity was not measured at these time points in the previous studies and could have resulted in higher overall incidence of ADA in this study. The characteristics of antibodies detected in this study were generally transient and non-neutralising, which is in accordance with those established with ETN in previous studies. Since SB4 tagged single-assay approach was used to detect immunogenicity, the assay method does not seem to have caused the lower incidence of ADA observed in SB4 compared with ETN (0.7% vs 13.1%). There are product-specific factors known to affect immunogenicity, such as product origin (foreign or human), product aggregates, impurities, container closure system, however, factors contributing to lower immunogenicity of SB4 are to be further investigated. Yet, according to the EMA guideline on biosimilars the lower immunogenicity of SB4 does not preclude classification as biosimilar since clinical efficacy of SB4 and ETN were equivalent in patients with ADA-negative results and no apparent correlation between ADA and clinical response or safety was observed.

To date, this is the first global, multicentre study comparing an ETN biosimilar to reference product ETN. Confirmed equivalence of SB4 and ETN in this study may provide an alternative treatment option for RA and allow better access to biologics for patients.

CONCLUSIONS

SB4 was shown to be equivalent in terms of clinical efficacy when compared with ETN. SB4 was well tolerated with a comparable safety profile to ETN.

Author affiliations

1 Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Chapel Allerton Hospital, Leeds, UK
2 NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds Teaching Hospitals NHS Trust, Leeds, UK
3 Institute of Rheumatology, Prague, Czech Republic
4 NZOZ Medica Pro Familia Sp. z o.o., Warsaw, Poland
5 Poznan University of Medical Sciences, Poznan, Poland
6 Poznanski Oskorek Medycyny NOWAMED, Pultusk, Poland
7 Lithuanian University of Health Sciences, Kaunas, Lithuania
8 Kharkiv Medical Academy of Postgraduate Education, Kharkiv, Ukraine
9 M.V. Sklofasovskyi Poltava Regional Clinical Hospital, Poltava, Ukraine
10 Medicom Sp. z o.o., Oswiecim, Poland
11 Klaipeda University Hospital, Klaipeda, Lithuania
12 UniCredito Medica e Investigación en Salud (UNAMIS), Yucatán, Mexico
13 Samsung Bioepis Co., Ltd., Incheon, Republic of Korea

Correction notice This article has been corrected since it was published Online First. Several minor textual changes have been made and the author affiliations altered.

Ukraine: Ignatenko G, Vatulin M, Gasanov Y, Grylloboyv A, Shevchuk S, Rekalov D, Povoroznyuk Y, Yatsyshyn R, Petrov A, Stanislavchuk M, UK: Haigh R. The authors also thank the study team for assistance with the logistic management and reporting of the study, specifically Isun Hong (lead clinical study manager) and Evelyn Euben Hong (scientific solutions for writing and editorial support).

Contributors PE, JG and SYC were involved in the conception or design of the work, or the acquisition, analysis or interpretation of data, drafting the work or revising it critically for important intellectual content, final approval of the version published, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. JV, AS, PL, WP, AB, VT, VZ, BS, RM and AB were involved in acquisition of data, drafting the work or revising it critically for important intellectual content, and final approval of the version published.

Funding This study was funded by Samsung Bioepis Co., Ltd.

Competing interests All authors received funding for clinical research from Samsung Bioepis: PE received consulting fees; JV, AS, PL, WP, AB, VT, VZ, BS, RM, and AA8R received research grants; SYC and JG are full-time employees of Samsung Bioepis. In addition, PE reports receiving grants/research support from AbbVie, Pfizer and consultancy fee from AbbVie, BMS, Pfizer, USB, MSD, Roche, Novartis, Takeda, Lilly; JV served on speakers bureau for USB, Pfizer, AbbVie, MSD; PL reports receiving grant/research support from Roche, MSD, Janssen, Novo-Nordisk, UCB, Pfizer, Novartis, GSK, BM; served as paid instructor for Novo-Nordisk, and served on speakers bureau for MSD, UCB, Roche, Amgen; AB reports receiving grant/research support from AbbVie.

Patient consent Obtained.

Ethics approval This study was conducted in compliance with the Declaration of Helsinki and Good Clinical Practice Guidelines established by the International Conference Harmonisation. The protocol was reviewed and approved by the institutional review board or the independent ethics committee of each investigational centre.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is appropriately cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

SUPPLEMENTARY MATERIAL

Table of Contents

Appendix 1. Eligibility Criteria
Appendix 2. Randomisation Scheme and Blinding
Appendix 3. Serum Measurement and ADA Detection Assay
Appendix 4. Time-response Model
Appendix 5. Protocol Deviations
  Table S1. Summary of major protocol deviations
Appendix 6. ACR20/50/70 Response Rates
  Figure S1. ACR20/50/70 response rates by visit in PPS
Appendix 7. Sub-group Analysis
  Table S2. Analysis of ACR20 response rate at Week 24 by overall 24-week ADA status (PPS)
  Table S3. Injection site reaction by overall 24-week ADA status
Appendix 8. Pharmacokinetic Results
  Figure S2. Mean (standard deviation) serum trough concentration ($C_{\text{trough}}$) profile
  Figure S3. Mean (standard deviation) serum concentration profiles at Week 8
Appendix 9. Immunogenicity Results
  Table S4. Incidence of ADA by Visit and Treatment Group
Appendix 1. Eligibility Criteria

Inclusion Criteria

Patients must meet all of the following inclusion criteria to be enrolled in the study:

1. Are male or female aged 18–75 years at the time of signing the consent form.

2. Have been diagnosed as having RA according to the revised 1987 ACR criteria (Appendix 1) for at least 6 months but not exceeding 15 years prior to Screening.

3. Have moderate to severe active disease despite MTX therapy defined as:
   a. More than or equal to six swollen joints and more than or equal to six tender joints (from the 66/68 joint count system) at Screening and Randomisation.
   b. Either erythrocyte sedimentation rate (ESR; Westergren) ≥ 28 mm/h or serum C-reactive protein (CRP) ≥ 1.0 mg/dL at Screening.

4. Must have been treated with MTX for at least 6 months prior to Randomisation and be on a stable dose of MTX 10–25 mg/week given orally or parenterally for at least 4 weeks prior to Screening.

5. If using NSAIDs or other analgesics for RA, must have been on a stable dose for at least 4 weeks prior to Randomisation. If taking oral glucocorticoids, must have been on a stable dose equivalent to ≤ 10 mg prednisolone for at least 4 weeks prior to Randomisation. Low potency topical, otic and ophthalmic glucocorticoid preparations are permitted.

6. Female subjects who are not pregnant or nursing at Screening and who are not planning to become pregnant from Screening until 2 months after the last dose of IP.

7. Subjects of child-bearing potential (female or male) who agree to use at least two forms of appropriate contraception (e.g., established use of oral, injected or implanted hormonal contraceptive, placement of an intrauterine device or intrauterine system, physical barrier, male sterilisation or true abstinence) from Screening until 2 months after the last dose of IP.

8. Must be able to, in the opinion of the Investigator, understand the implications of taking part in the study and be willing to follow the study requirements.

9. Must be able to provide informed consent, which must be obtained prior to any study related procedures.
Exclusion criteria

Subjects meeting any of the following criteria must not be enrolled in the study:

1. Have been treated previously with any biological agents including any TNF-α inhibitor.

2. Have a known hypersensitivity to human immunoglobulin proteins or other components of Enbrel or SB4.

3. Have been taking any of the following concomitant medications, within the timeframe specified:
   a. Corticosteroids above levels equivalent to 10 mg prednisolone daily within 4 weeks prior to Randomisation.
   b. Any DMARDs/systemic immunosuppressive agents, other than MTX, including hydroxy-chloroquine, chloroquine, sulfasalazine, azathioprine, cyclosporine or mycophenolate mofetil within 4 weeks prior to Randomisation.
   c. Leflunomide within 12 weeks prior to randomisation or within 4 weeks prior to randomisation if the subject had washout with 8 g of cholestyramine three times daily for at least 11 days.
   d. Alkylating agents within 12 months prior to Randomisation.
   e. Live/live-attenuated vaccine within 8 weeks prior to Randomisation.
   f. Injectable corticosteroids within 4 weeks prior to Randomisation.
   g. Investigational product from another study within five half-lives of that product prior to Randomisation or use of an investigational device at Screening.

4. Have abnormal renal or hepatic function at Screening defined as the following:
   a. Serum creatinine $\geq 2 \times$ the upper limit of normal (ULN).
   b. Serum alanine transaminase or aspartate transaminase $\geq 2 \times$ ULN.

5. Have abnormal haematological parameters at Screening defined as the following:
   a. Haemoglobin < 8.0 g/dL.
   b. White blood cell count < $3.5 \times 10^3$ cells/$\mu$L ($< 3.5 \times 10^9$ cells/L).
   c. Neutrophil count < $1.5 \times 10^3$ cells/$\mu$L.
   d. Platelet count < $100 \times 10^3$ cells/$\mu$L.
   e. Lymphocyte count < $800$ cells/$\mu$L.

6. Have a positive serological test for hepatitis B or hepatitis C or have a known history of infection with human immunodeficiency virus.

7. Have a current diagnosis of active tuberculosis.

8. Have been recently exposed to a person with active tuberculosis, or are considered to have latent TB from the screening tests (QuantiFERON® Gold test and chest X-ray).
If such subjects complete at least 30 days of isoniazid prophylaxis or other anti-TB therapy according to country-specific guidelines and are willing to complete the entire course of recommended anti-TB therapy they may be enrolled into the study following re-screening.

9. Have had a serious infection (such as sepsis, abscess, opportunistic infections or invasive fungal infection including histoplasmosis) or have been treated with intravenous antibiotics for an infection within 8 weeks or oral antibiotics within 2 weeks prior to Randomisation. Non-significant infections do not need to be considered exclusionary at the discretion of the Investigator.

10. Have a history of an infected joint prosthesis which has not been removed or replaced.

11. Have any of the following conditions:
   a. Bone marrow hypoplasia which, in the opinion of the Investigator, will put the subject at risk if they are enrolled.
   b. Significant systemic RA involvement (e.g., vasculitis, pulmonary fibrosis etc) which, in the opinion of the Investigator, will put the subject at risk if they are enrolled.
   c. Other inflammatory or rheumatic diseases, including but not limited to PsA, AS, systemic lupus erythematosus, Lyme disease or fibromyalgia, which may confound the evaluation of the effect of IP.
   d. History of any malignancy within the previous 5 years prior to Screening except completely excised and cured squamous carcinoma of the uterine cervix, cutaneous basal cell carcinoma, or cutaneous squamous cell carcinoma.
   e. History of lymphoproliferative disease including lymphoma
   f. History of congestive heart failure (New York Heart Association Class III/IV) or unstable angina.
   g. Uncontrolled diabetes mellitus or uncontrolled hypertension.
   h. History of organ transplantation.
   i. Physical incapacitation (ACR functional Class IV (see Appendix 2) or wheelchair-/bed-bound).
   j. History of demyelinating disorders (such as multiple sclerosis or Guillain-Barré syndrome).
   k. Any conditions significantly affecting the nervous system (e.g., neuropathic conditions or nervous system damage) which may interfere with the Investigator’s assessment on disease activity scores including joint counts.
   l. Any other disease or disorder which, in the opinion of the Investigator, will put the subject at risk if they are enrolled.

12. Have or have had a substance abuse (alcohol or drug) problem within the previous 3 years prior to Screening.
Appendix 2. Randomisation Scheme and Blinding

Randomisation Scheme

Randomisation was implemented using Interactive Web Response System (IWRS) with a block size of 4 at the site level. Within each block the patients were allocated to the treatment group at 1:1 ratio. There was no stratification factor for the randomisation.

Blinding

Patients, Investigators, joint assessors and other study staff remained blinded throughout the study period. Patients were assigned to either SB4 or ETN through IWRS, and none of the study staff had access to the treatment code. At each study visit, the Investigator or designee connected to IWRS and obtained number of codes which indicated the prefilled syringes to be dispensed. To ensure blinding of the treatments, SB4 and ETN solutions were identical in appearance, packaging and labelling.

After the database lock for the 24-week interim report, a limited number of individuals of the Sponsor were unblinded for reporting purposes. The process of unblinding and measures to keep the blinding of other study staff were documented.

There were no cases of unblinding due to medical emergency during the study.
Appendix 3. Serum Measurement and ADA Detection Assay

**Serum Measurement**

To quantify SB4 and ETN in human serum, a validated enzyme-linked immunosorbent assay (ELISA) with anti-TNF Receptor II antibodies and anti-TNF Receptor II biotinylated antibody (R&D systems, Minneapolis, MN, USA) was used. The quantification range was 160.00-4000.00 ng/mL. The inter precision and accuracy was 8.8% and 6.3%, respectively.

**ADA Detection Assay**

MSD electrochemiluminescence (ECL) bridging assay (Meso Scale Discovery, Rockville, MD, USA) with acid dissociation was used to establish the cut points and to determine ADA in human RA serum.

A single assay format with labelled versions of the biosimilar candidate was used to minimise bioanalytical bias associated with inter-assay variability and the possibilities of inconstant false-positive/false-negative results due to labelling of multiple antigens (to minimise preparing biotinylated and sulfo versions of both SB4 and ETN).

The tiered approach for ADA determination was used. After the screening assay, the confirmatory assay was performed for ADA determination. The cut point for a screened positive signal was set with 5% false-positive rate and for a confirmed positive it was set with 0.01% false-positive rate.
Appendix 4. Time-response Model

The exponential growth model is a parsimonious representation of the data with parameters that are interpretable from a clinical perspective, so that it is decided to use the time-response modeling to show the similarity of the time course of the treatment effects between reference drug and experimental drug. For modeling with the historical trials, the following exponential distribution is assumed for the ACR20 response rate at time $t$ for treatment arm $j$ in the $i$-th study.

$$f(t) = (\theta_j + \eta_i)(1 - e^{-\beta_j t^e}) + \epsilon_{ij}$$

where $\theta_j$ is a fixed parameter describing the change from baseline of the response, $\beta_j$ denotes the slope of the change from baseline, and $\eta_i$ is assumed to be a study level random variable. In order to fit the model for each treatment group, the initial parameter estimates are chosen from the prior fitted model, and the final parameter estimates are optimised using a simple Newton’s method until a sufficiently accurate value is reached.

The 2-norm can be viewed as the response difference between the two treatments over time course and calculated as follows.

$$\|f(t) - g(t)\|_2 = \left[\int (f(t) - g(t))^2 dt\right]^{1/2}$$

where $f(t)$ and $g(t)$ represent the ACR20 response time course for each treatment group.

With the fitted models of treatment groups using the historical data, the 2-norm of the difference between treatment groups at Week 24 and its 95% CI are estimated as 223.23 [166.56, 279.90]. The equivalence margin of the time-response modeling was determined as 83.28 which is the half of the lower bound of the 95% CI. Therefore, the equivalence will be concluded if the upper limit of the 95% CI for the 2-norm of the difference between SB4 and Enbrel® treatment groups is less than 83.28.
### Appendix 5. Protocol Deviations

Table S1. Summary of major protocol deviations

<table>
<thead>
<tr>
<th>Protocol deviations</th>
<th>SB4 50 mg</th>
<th>ETN 50 mg</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=299</td>
<td>N=297</td>
<td>N=596</td>
<td></td>
</tr>
<tr>
<td>With at least one major protocol deviation</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>73 (24.4)</td>
<td>72 (24.2)</td>
<td>145 (24.3)</td>
<td></td>
</tr>
<tr>
<td>Excluded from Per-protocol set</td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>40 (13.4)</td>
<td>35 (11.8)</td>
<td>75 (12.6)</td>
<td></td>
</tr>
<tr>
<td>Concomitant medication criteria</td>
<td>9 (3.0)</td>
<td>14 (4.7)</td>
<td>23 (3.9)</td>
</tr>
<tr>
<td>Eligibility and entry criteria</td>
<td>7 (2.3)</td>
<td>5 (1.7)</td>
<td>12 (2.0)</td>
</tr>
<tr>
<td>Investigational product compliance</td>
<td>9 (3.0)</td>
<td>2 (0.7)</td>
<td>11 (1.8)</td>
</tr>
<tr>
<td>Study procedures criteria</td>
<td>16 (5.4)</td>
<td>16 (5.4)</td>
<td>32 (5.4)</td>
</tr>
</tbody>
</table>
Appendix 6. ACR20/50/70 Response Rates

Figure S1. ACR20/50/70 response rates by visit in PPS
Appendix 7. Sub-group Analysis

Table S2. Analysis of ACR20 response rate at Week 24 by overall 24-week ADA status (PPS)

<table>
<thead>
<tr>
<th>24-week ADA status</th>
<th>Treatment</th>
<th>n/n’ (%)</th>
<th>Adjusted Difference (SE)</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>SB4 50 mg (N=2)</td>
<td>2/2 (100.0)</td>
<td>22.14% (37.095%)</td>
<td>(−54.79%, 99.07%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ETN 50 mg (N=29)</td>
<td>21/29 (72.4)</td>
<td>(37.095%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>SB4 50 mg (N=245)</td>
<td>191/245 (78.0)</td>
<td>−3.57% (3.846%)</td>
<td>(−11.12%, 3.99%)</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>ETN 50 mg (N=205)</td>
<td>167/205 (81.5)</td>
<td>(3.846%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ADA, anti-drug antibody; CI, confidence interval; SE, standard error.
ADA status was defined as positive if patient had a positive test result at least once up to Week 24.
The adjusted difference and its 95% CI were estimated by analysis of covariance model with treatment and region as factors and baseline C-reactive protein value as covariate.

Table S3. Injection site reaction by overall 24-week ADA status

<table>
<thead>
<tr>
<th>Overall 24-week ADA status</th>
<th>SB4 (N=299)</th>
<th>ETN (N=297)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/n’ (%)</td>
<td>E</td>
</tr>
<tr>
<td>Positive</td>
<td>0/2 (0.0)</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>11/297 (3.7)</td>
<td>22</td>
</tr>
</tbody>
</table>

ADA, anti-drug antibody; E, events
n’: number of patients with available overall 24-week ADA assessment results. Percentages were based on n’.
n: number of patients who have injection site reactions counted by the high-level group term (HLGT) of administration site reaction
Overall 24-week ADA result was defined as positive for patients with at least one ADA positive result up to Week 24 after Week 0.
Appendix 8. Pharmacokinetic Results

Figure S2. Mean (standard deviation) serum trough concentration (C\text{trough}) profile

Figure S3. Mean (standard deviation) serum concentration profiles at Week 8
Appendix 9. Immunogenicity Results

Table S4. Incidence of ADA by Visit and Treatment Group

<table>
<thead>
<tr>
<th>Timepoint</th>
<th>SB4 (N=299)</th>
<th>ETN (N=297)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/n' (%)</td>
<td>n/n' (%)</td>
</tr>
<tr>
<td>Week 0</td>
<td>0/299 (0.0)</td>
<td>0/297 (0.0)</td>
</tr>
<tr>
<td>Week 2</td>
<td>0/298 (0.0)</td>
<td>1/295 (0.3)</td>
</tr>
<tr>
<td>Week 4</td>
<td>1/299 (0.3)</td>
<td>32/291 (11.0)</td>
</tr>
<tr>
<td>Week 8</td>
<td>1/298 (0.3)</td>
<td>6/288 (2.1)</td>
</tr>
<tr>
<td>Week 12</td>
<td>0/294 (0.0)</td>
<td>1/280 (0.4)</td>
</tr>
<tr>
<td>Week 16</td>
<td>0/290 (0.0)</td>
<td>0/277 (0.0)</td>
</tr>
<tr>
<td>Week 24</td>
<td>0/288 (0.0)</td>
<td>0/272 (0.0)</td>
</tr>
<tr>
<td>Week 24 overall</td>
<td>2/299 (0.7)</td>
<td>39/297 (13.1)</td>
</tr>
</tbody>
</table>

ADA, anti-drug antibody

n': number of patients with available overall 24-week ADA assessment results. Percentages were based on n'. Overall 24-week ADA result was defined as positive for patients with at least one ADA positive result up to Week 24 after Week 0.