CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

761066Orig1s000

STATISTICAL REVIEW(S)

US Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI

BLA NO.:	761066
DATE RECEIVED BY OB:	May 30, 2017
DRUG NAME:	SB4 (proposed biosimilar to Enbrel, Amgen)
DOSAGE FORM:	50 mg single-use prefilled syringe for subcutaneous injection
INDICATION:	Same as Enbrel
APPLICANT:	Samsung Bioepis Co., Ltd.
REVIEW FINISHED:	March 22, 2018
STATISTICAL REVIEWER:	Yu-Yi Hsu, PhD
SECONDARY REVIEWER:	Meiyu Shen, PhD
PROJECT MANAGER:	Brandi Wheeler

Concur: Yi Tsong, Ph.D., Division Director, CDER/OTS/OB/DB VI

Distribution:

CDER/OTS/OB/DBVI,	Yi Tsong
CDER/OTS/OB/DBVI,	Meiyu Shen
CDER/OTS/OB/IO,	Lillian Patrician
CDER/OPQ/OBP/DBRRI,	Brian Janelsins
CDER/OPQ/OBP/DBRRI,	Rachel Novak
CDER/OND/TBBS,	Anne Rowzee
CDER/OND/TBBS,	Sue Lim

Table of Contents

1 Executive summary and recommendation 3
2 Overview 4
3 Applicant's statistical equivalence testing 4
4 Analytical similarity 4
4.1 Data analyzed 4
4.2 Statistical method 7
4.3 Equivalence testing for TNF-α binding assay 8
4.4 Equivalence testing for TNF-α neutralization assay 9
5 Conclusion 10
6 Reference 11

List of Tables

Table 1 Results of Equivalence Testing for the TNF-α binding assay 3
Table 2 Results of Equivalence Testing for the TNF-α neutralization assay 3
Table 3 Number of lots tested for Tier-1 quality attributes 4
Table 4 Descriptive statistics for the TNF-α binding assay data grouped by type of lot 5
Table 5 Descriptive statistics for the TNF-α neutralization assay data grouped by type of lot 5
Table 6 Descriptive statistics for the TNF-α binding assay data grouped by drug substance manufacturing process 5
Table 7 Descriptive statistics for the TNF-α neutralization assay data grouped by drug substance manufacturing process 5
Table 8 The descriptive statistics TNF-α binding assay 9
Table 9 The descriptive statistics of TNF-α neutralization assay 10

List of Figures

Figure 1 Plot of TNF-α binding assay from SB4 lots against manufacture date 6
Figure 2 Plot of TNF-α binding assay from SB4 lots against manufacture date 6
Figure 3 TNF-α binding assay data 8
Figure 4 Equivalence margins and 90%CI of TNF-α binding assay 9
Figure 5 TNF-α neutralization assay 10
Figure 6 Equivalence margins and 90%CI of TNF-α neutralization assay 10

1 Executive summary and recommendation

SB4 is a proposed biosimilar to US-licensed Enbrel (etanercept). The CMC statistical reviewer in the Office of Biostatistics analyzed the comparative results of two critical Quality Attributes (QAs): TNF-α neutralization and TNF-α binding, which were recommended for equivalence testing analysis by the Office of Biotechnology Products (OBP). Tier 1 statistical equivalence testing was conducted using equivalence margins of ±1.5 times of the estimated comparator standard deviation ($\hat{\sigma}_R$), where the subscript R represents the comparator product. A three-way comparison, including equivalence tests between SB4 and US-Enbrel, SB4 and EU-Enbrel, and EU-Enbrel and US-Enbrel, is used to support the use of clinical data generated from EU-Enbrel. Samples from 24 lots of SB4 (Samsung), 51 lots of US-licensed Enbrel (US-Enbrel), and 54 lots of EU-authorized Enbrel (EU-Enbrel) were used for evaluating the similarity of TNF-α binding assay using the equivalence testing. The results are summarized in Table 1.

Table 1 Results of Equivalence Testing for the TNF-α binding assay

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI[1]		
SB4 (24)	US-Enbrel (51)	3.23	(1.84, 4.61)	(-4.94, 4.94)	Pass
SB4 (24)	EU-Enbrel (54)	0.56	(-1.12, 2.23)	(-7.16, 7.16)	Pass
EU-Enbrel (54)	US-Enbrel (51)	2.67	(1.35, 3.99)	(-4.94, 4.94)	Pass

1. The 90% confidence interval is adjusted by the sample size imbalance

Samples from 24 lots of SB4, 47 lots of US-Enbrel, and 44 lots of EU-Enbrel were used for evaluating the similarity of TNF-α neutralization assay using the equivalence testing. The results are summarized in Table 2.

Table 2 Results of Equivalence Testing for the TNF-α neutralization assay

Test Product (Number of lots)	Comparator (Number of lots)	Mean Difference, %		Equivalence Margin, %	Conclusion
		Estimate	90% CI[1]		
SB4 (24)	US-Enbrel (47)	1.72	(-1.55, 4.98)	(-11.09, 11.09)	Pass
SB4 (24)	EU-Enbrel (44)	-2.38	(-6.02, 1.27)	(-14.17, 14.17)	Pass
EU-Enbrel (44)	US-Enbrel (47)	4.09	(1.12, 7.06)	(-11.09, 11.09)	Pass

1. The 90% confidence interval is adjusted by the sample size imbalance

The 24 SB4 lots tested for Tier-1 QAs were manufactured from two processes that are considered comparable. Please see OBP's review for details of processes comparability assessment.

As shown in Tables 1 and 2, the results from statistical equivalence testing of the TNF-α binding assay and the TNF-α neutralization assay support a demonstration that the proposed biosimilar SB4 is similar to US-Enbrel and support the analytical portion of the scientific bridge to justify the relevance of EU-Enbrel data for the comparative clinical study.

2 Overview

SB4 is a proposed biosimilar to US-licensed Enbrel (etanercept) submitted under Section 351(k) of the Public Health Service Act. Following the guidance[1], the applicant submitted analytical similarity assessments using a three-way bridge approach to demonstrate analytical similarity between SB4, US-Enbrel and to establish a quality portion of scientific bridge between US-Enbrel and EU-Enbrel.

Two Tier-1 critical quality attributes, TNF-α neutralization and TNF-α binding, were tested using samples from multiple biosimilar, US-Enbrel and EU-Enbrel lots. Our comments regarding the applicant's statistical equivalence testing is provided in Section 3. Our independent statistical analyses and descriptions of received data are summarized in Section 4. Our conclusion is provided in Section 5.

3 Applicant's statistical equivalence testing

In this submission, the applicant followed the FDA's recommendation to conduct Tier 1 statistical equivalence testing with the margin defined as $1.5\sigma_R$, where σ_R is estimated by the comparator product lots, for TNF-α binding assay and TNF-α neutralization assay. Hence, the applicant's analyses are valid. The CMC statistical reviewer performs an independent statistical analysis to confirm the applicant's analyses in the next section.

4 Analytical similarity

To evaluate analytical similarity, the FDA recommended the applicant to apply a tiered approach in the FDA responses to IND meeting. That is, product QAs amenable to statistical evaluation are assigned to three tiers based on their criticality and other factors. The quality attributes with potential highest risk in product quality, efficiency, safety, and PK/PD are generally assigned to Tier 1, in which analytical similarity is assessed by statistical equivalence test. More details are described in the FDA Draft Guidance on Analytical Similarity (2017)[2]. This review focuses on the Tier 1 statistical equivalence testing

4.1 Data analyzed

The applicant tested multiple lots of SB4, US-Enbrel and EU-Enbrel for Tier-1 QAs. The numbers of lots tested for Tier-1 QAs are listed in Table 3.

Table 3 Number of lots tested for Tier-1 quality attributes

Quality attribute	SB4	US-Enbrel	EU-Enbrel
TNF-α binding assay	24	51	54
TNF-α neutralization assay	24	47	44

The applicant tested 24 SB4 lots for both Tier-1 QAs, including 8 drug product (DP) lots and 16 drug substance (DS) lots. The 8 DP lots are manufactured from 8 different DS lots that are not included in the similarity assessment. In other words, the 24 lots used for Tier-1 similarity assessment are independent since they are corresponding to 24 different DS lots.

Figure 1 and Figure 2 present the scatter plots of Tier-1 QAs. The sample means calculated from DP and DS lots are similar for both Tier-1 QAs. Table 4 and Table 5 shows the descriptive statistics of the Tier-1 QAs grouped by lot type (DS or DP).

Table 4 Descriptive statistics for the TNF-α binding assay data grouped by type of lot

Subgroup	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
DP	8	104.38	4.57	99	112
DS	16	104.06	2.08	100	108
All (DP+DS)	24	104.17	3.03	99	112

Table 5 Descriptive statistics for the TNF-α neutralization assay data grouped by type of lot

Subgroup	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
DP	8	99.75	5.01	90	105
DS	16	100.69	8.50	87	121
All (DP+DS)	24	100.38	7.41	87	121

The manufacture process of SB4 drug substance lots changed after August 2015. The group of 24 SB4 lots tested for Tier-1 QAs consists 14 lots (9 DS and 5 DP) manufactured using process-A, and 10 lots (7 DS and 3 DP) manufactured using process-B. Table 6 and Table 7 show descriptive statistics of Tier-1 QAs grouped by process. The OBP reviewer reviewed comparability of lots manufactured from two process, and conclude that lots from process A and process B are comparable. Therefore, data yield from lots manufactured using both process can be combined in equivalence tests.

Table 6 Descriptive statistics for the TNF-α binding assay data grouped by drug substance manufacturing process

Group	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
Process A	14	104.93	3.52	99	112
Process B	10	103.10	1.85	101	108
All (A+B)	24	104.17	3.03	99	112

Table 7 Descriptive statistics for the TNF-α neutralization assay data grouped by drug substance manufacturing process

Group	Sample size	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
Process A	14	99.93	5.84	90	111
Process B	10	101.00	9.51	87	121
All (A+B)	24	100.38	7.41	87	121

Figure 1 Plot of TNF-α binding assay from SB4 lots against manufacture date

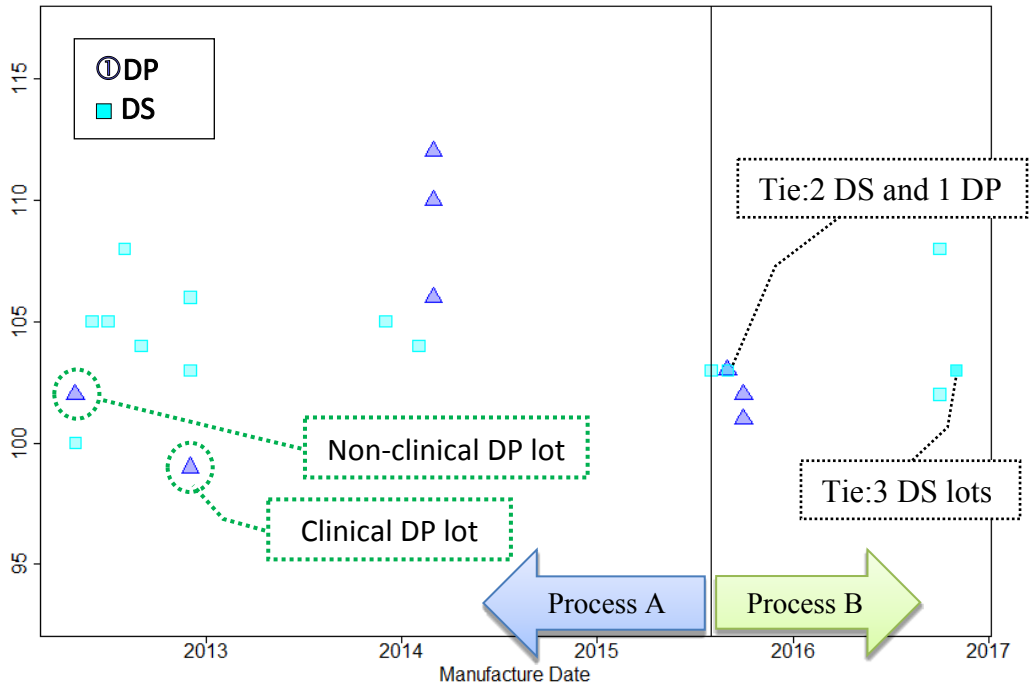
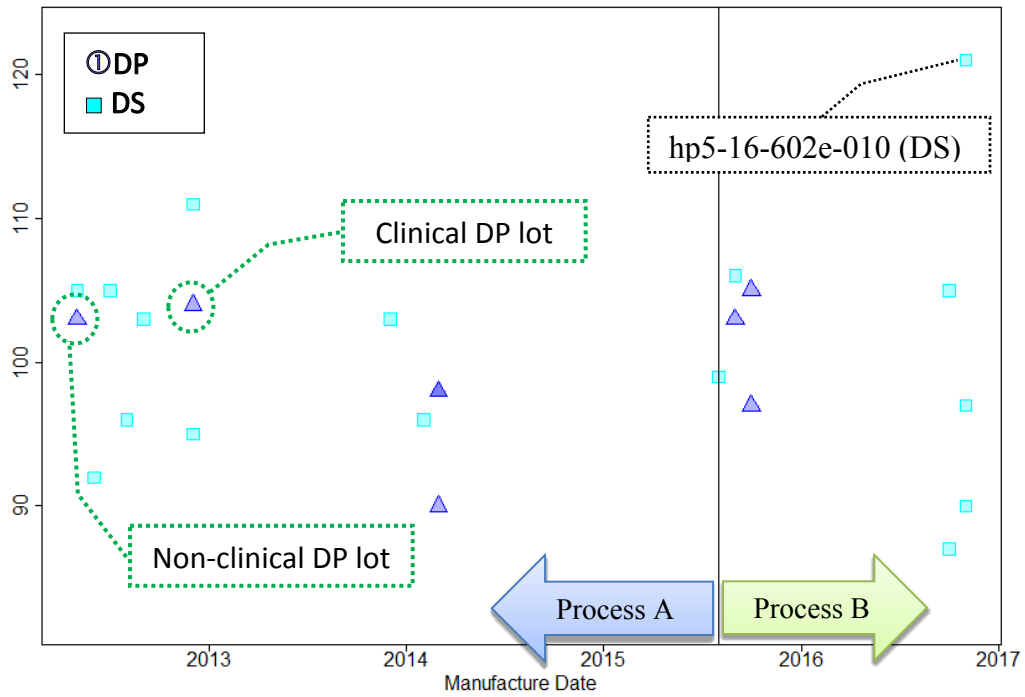


Figure 2 Plot of TNF-α binding assay from SB4 lots against manufacture date



The scatter plots for SB4 are shown in

Figure 1 and Figure 2. Ties are observed in the TNF-α binding assay dataset as shown in

Figure 1. Three DP lots (50mg/mL) were used in clinical studies and analytical similarity assessment. The three clinical lots were manufactured in Dec 2012, Feb 2013 and Apr 2013 from the same DS lot. One clinical lot (manufacture date: Dec 2012) is used for Tier-1 equivalence test. The Tier-1 QA data from the clinical lot and the DP lot used in non-clinical studies are labeled

Figure 1 and Figure 2. As shown in Figure 1, the point corresponding to clinical lot is at the lower end of the range of data. As shown in Figure 2, the point corresponding to the PD lots used in clinical and non-clinical studies are around the center of range of data.

For both Tier-1 QAs, the data are presented as relative activity to the research reference standard (RRS). There are three US sourced Enbrel selected as RRS during the development stage with expiration dates Feb 2014, Sep2014 (extended to Sep 2016) and Dec 2016 (extended to Dec 2017). The characteristics of the RRS lots are listed in Module 3.2.S.5.4. The three RRS used for testing Tier-1 QAs are comparable according to discussions with OBP in an internal meeting. Therefore, the measurements related to the three RRS can be combined for tier-1 equivalence test.

The OBP reviewer also conducted additional investigation regarding the working reference standard (WRS) and the related issues. The WRS was not used in testing Tier-1 QAs. Therefore, the concern about WRS does not impact the equivalence testing results.

4.2 Statistical method

Let μ_T and μ_R be the population mean of the QA for the test product and the population mean of the QA for the reference product, respectively. Let $\sigma_R$be the standard deviation of the QA of interest for the reference product. To conclude the equivalence in the QA of interest between the test product and the reference product, we test the following null
and alternative hypotheses:

$$H_0{:}\mu_T - \mu_R \le \theta_1 \text{ or } \mu_T - \mu_R \ge \theta_2$$
$$H_1{:}\theta_1 < \mu_T - \mu_R < \theta_2$$

where $\theta_1 = -1.5\sigma_R$, $\theta_2 = 1.5\sigma_R$, and θ_1 and θ_2 are equivalence margins.

The null hypothesis is rejected with significance level not exceeding α=0.05 if the 90% Confidence Interval (CI) of the mean difference, $\mu_T - \mu_R$, of the QA of interest falls within $\left(-1.5\sigma_R, 1.5\sigma_R\right)$. In other words, we conclude equivalence of the QA of interest between the test product and the reference product if null hypothesis is rejected. Since the margin is conducted

with unknown parameter, we replace σ_R by the sample standard deviation of the reference product.

Let X_{Tj} be the observed value of the QA of interest for Lot j of the test product (the proposed biosimilar product), and X_{Rj} be the observed value of the QA of interest for Lot j of the reference product. Since the two products are manufactured by two manufacturers, two products are independent.

Let $\overline{X}_i = \Sigma_{j=1}^{n_i} X_{Tj}/n_i$ and $S_i^2 = \Sigma_{j=1}^{n_i} \left(X_{ij} - \overline{X}_i\right)^2/\left(n_i - 1\right)$, where n_i is the number of lots of the product i, and $i \in \{T,R\}$. Assuming unequal variances of the test product and the reference product, the $(1-2\alpha) * 100\%$ CI of the mean difference in the QA of interest can be calculated as $\left(\overline{X}_T - \overline{X}_R - t_\alpha(v)\sqrt{S_T^2/n_T^* + S_R^2/n_R^*}, \overline{X}_T - \overline{X}_R + t_\alpha(v)\sqrt{S_T^2/n_T^* + S_R^2/n_R^*}\right)$, where $t_\alpha(v)$ is the $1-\alpha$ t-distribution quantile and $v = \frac{\left(S_T^2/n_T^* + S_R^2/n_R^*\right)^2}{\frac{\left(S_T^2/n_T^*\right)}{\left(n_T - 1\right)} + \frac{\left(S_R^2/n_R^*\right)}{\left(n_R - 1\right)}}$ is the degrees of freedom calculated by Welch-Satterthwaite's approximation with adjustment for imbalance sample size. The adjusted value $n_R^* = 1.5n_T$ if $n_R > 1.5n_T$ and $n_R^* = n_R$ otherwise, and $n_T^* = 1.5n_R$ if $n_T > 1.5n_R$ and $n_T^* = n_T$ otherwise.

4.3 Equivalence testing for TNF-α binding assay

The descriptive statistics TNF-α binding assay are summarized at Table 8. Figure 3 shows TNF-α binding assay from SB4, US-Enbrel and EU-Enbrel. The average of data from SB4 is slightly higher than the average of data from US-Enbrel. The results of 3-way equivalence testing are listed on Table 1. Figure 4 shows relationship between equivalence margins and confidence intervals of three pairwise comparisons. The first panel of Figure 4 shows that the confidence interval of difference between means TNF-α binding assay from SB4 and US-Enbrel is within the equivalence margin, $1.5\hat{\sigma}_R$.

Figure 3 TNF-α binding assay data

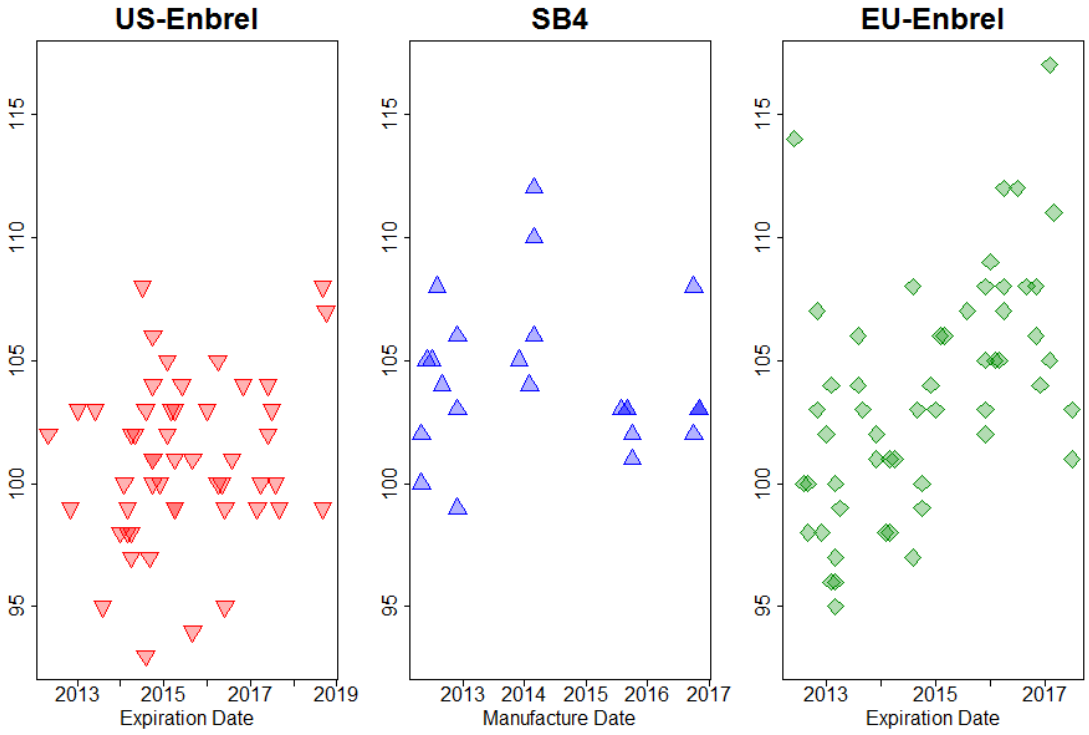
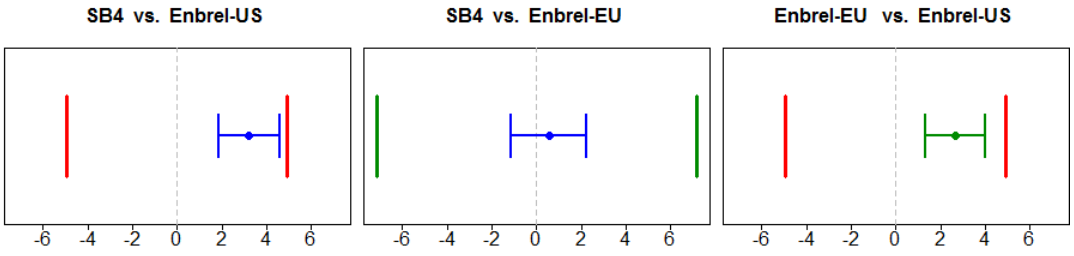


Table 8 The descriptive statistics TNF-α binding assay

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
SB4	24	104.17	3.03	99	112
US-Enbrel	47	100.94	3.29	93	108
EU-Enbrel	44	103.61	4.78	95	117

Figure 4 Equivalence margins and 90%CI of TNF-α binding assay



4.4 Equivalence testing for TNF-α neutralization assay

The descriptive statistics of TNF-α neutralization assay are summarized at Table 9. Figure 5 shows data of TNF-α neutralization assay from SB4, US-Enbrel and EU-Enbrel. The data from

SB4 have range that is slightly higher than the range of data from US-Enbrel. The 3-way equivalence testing results are listed on Table 2. Figure 6 shows relationship between equivalence margins and confidence intervals. All confidence intervals are within prespecified equivalence margins.

Figure 5 TNF-α neutralization assay

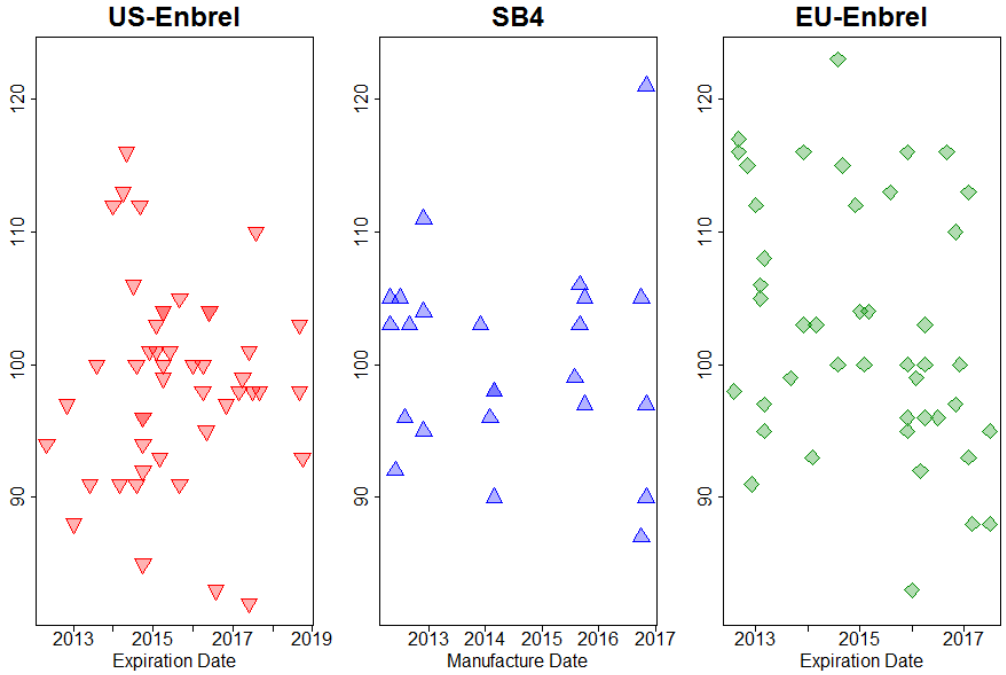
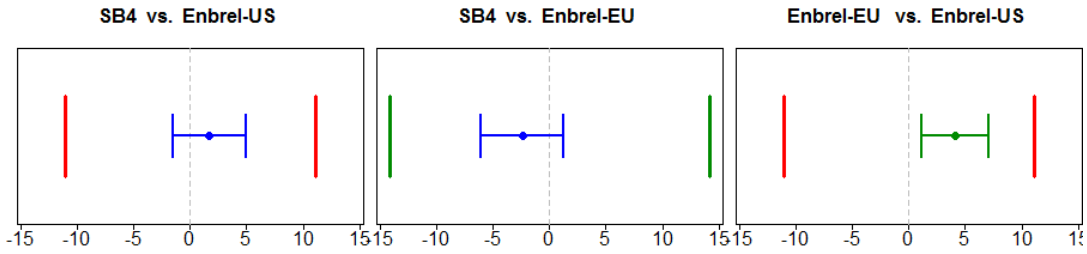


Table 9 The descriptive statistics of TNF-α neutralization assay

Product	Number of lots	Sample mean, %	Sample standard deviation, %	Minimum, %	Maximum, %
SB4	24	100.38	7.41	87	121
US-Enbrel	51	98.66	7.39	82	116
EU-Enbrel	54	102.75	9.45	83	123

Figure 6 Equivalence margins and 90%CI of TNF-α neutralization assay



5 Conclusion

The results from the statistical equivalence analyses for the TNF-α binding assay and the TNF-α neutralization assay support a demonstration that the proposed biosimilar SB4 is similar to US-Enbrel. In addition, the results support the analytical portion of the scientific bridge to justify the relevance of EU-Enbrel data from the comparative clinical study.

6 Reference

[1] Scientific considerations in demonstrating biosimilarity to a reference product (2015)
https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf
[2] Statistical Approaches to Evaluate Analytical Similarity Guidance for Industry (2017)
https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM576786.pdf

---

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

YU YI HSU
03/23/2018

MEIYU SHEN
03/23/2018

YI TSONG
03/24/2018

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	BLA761066/001
Drug Name:	SB4 (proposed Enbrel biosimilar)
Indication(s):	All Enbrel indications
Applicant:	Samsung Bioepis
Date(s):	Receipt Date: May 25, 2017 PDUFA Goal Date: March 25, 2018
Review Priority:	Standard
Biometrics Division:	DBII
Statistical Reviewer:	Ginto Pottackal, PhD.
Concurring Reviewers:	Robert Abugov, PhD.
Medical Division:	Division of Pulmonary, Allergy, and Rheumatology Products
Clinical Team:	Rachel Glaser, MD., Nikolay Nikolov, MD.
Project Manager:	Brandi Wheeler

Keywords: BLA review, clinical studies, missing data, biosimilar, etanercept

Table of Contents

1 EXECUTIVE SUMMARY 5

2 INTRODUCTION 6

2.1 BACKGROUND 6
2.2 HISTORY OF PRODUCT DEVELOPMENT 6
2.3 SPECIFIC STUDIES REVIEWED 7
2.4 DATA SOURCES 7

3 STATISTICAL EVALUATION 7

3.1 DATA AND ANALYSIS QUALITY 7
3.2 EVALUATION OF EFFICACY 8
3.2.1 Study Design and Endpoints 8
3.2.2 Randomized, Double-blind Period 9
3.2.3 Transition-Extension Period 10
3.2.4 Data Sets 11
3.3 STATISTICAL METHODOLOGIES 12
3.3.1 Planned Analysis 12
3.3.2 Similarity Margin 13
3.4 DEMOGRAPHIC AND BASELINE CHARACTERISTICS 14
3.4.1 Double-blind Period 14
3.4.2 Transition Extension Period 16
3.5 PATIENT DISPOSITION 17
3.6 RESULTS AND CONCLUSIONS 18
3.6.1 Primary Efficacy Analysis 18
3.6.2 Secondary Efficacy Analysis 21
3.6.3 Transition Extension Period 22
3.6.4 Constancy Assumption 23
3.6.5 Assay Sensitivity 25
3.6.6 Potential Impacts of Missing Data 26
3.7 EVALUATION OF SAFETY 27

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS 28

5 SUMMARY AND CONCLUSIONS 29

5.1 STATISTICAL ISSUES 29
5.2 COLLECTIVE EVIDENCE 29

6 APPENDIX 30

BIOSIMILARITY MARGIN: 30
ADDITIONAL TABLES AND FIGURES: 30

7 BIBLIOGRAPHY 31

LIST OF TABLES

Table 1: Design for Clinical Study SB4-G31-RA 8
Table 2: Similarity Margin Meta-Analysis: ACR20 in Active RA Despite Treatment with Methotrexate 14
Table 3: Data Sets (Double-blind Period) 15
Table 4: Baseline Demographic Characteristics (Double-blind Period) 15
Table 5: Baseline Disease Characteristics (Double-blind Period) 16
Table 6: Demographic Characteristics (Transition Extension Period) 17
Table 7: Reasons for Withdrawal through Week 52 18
Table 8: Primary Analysis of ACR20 Response Rate at Week 24 (Per-protocol Set) 18
Table 9: Analysis of ACR20 Response Rate at Week 24 (Full Analysis Set) 19
Table 10: Reasons for Non- Response, ACR20 Primary Endpoint at Week 24 19
Table 11: Mean Changes from Baseline in ACR Components at Week 24 (Double-blind period) 20
Table 12 : Analysis of ACR20 Response Rate at Week 52 (Double-blind Period) 21
Table 13 : ACR-N at Week 24 (Double-blind period) 21
Table 14 : Change in DAS28 Score at Week 24 (Double-blind period) 22
Table 15 : Change from Baseline Value of Modified Total Sharp Score (mTSS) at Week 52 22
Table 16: ACR Responses Over Time in Transition Extension Period 23
Table 17: Key Characteristics of Study SB4-G31-RA and Two Historical Randomized, Placebo-Controlled Clinical Trials[1] of Etanercept in RA 24
Table 18: Analysis of ACR20 Response Rate at Week 12 (Per-protocol Set) 25
Table 19 : Tipping Point Analysis for Week 24 ACR20[1] Response Rate 27
Table 20 : Analysis of ACR50 Response in Double-blind Period 30
Table 21 : Analysis of ACR70 Response in Double-blind Period 30
Table 22 : ACR-N at Week 24 in Response in Double-blind Period 30

LIST OF FIGURES

Figure 1: Study Design of the Randomized, Double-blind Period 9
Figure 2: Entire Study Design, Including Open-Label Extension Period..... 11
Figure 3: ACR20 Response Probabilities over Time (Double-blind period) 20
Figure 4: Estimated differences between SB4 and EU-Enbrel in the probability of remaining in the study and achieving an ACR20 response at Week 24, stratified by selected subgroups, in study SB4-G31-RA. 28

1 EXECUTIVE SUMMARY

Enbrel (etanercept) is a US licensed biologic response modifier (BRM) for the treatment of five long-term inflammatory conditions: rheumatoid arthritis (RA), polyarticular juvenile idiopathic arthritis (JIA), psoriatic arthritis (PsA), ankylosing spondylitis (AS), and moderate to severe plaque psoriasis in adult patients. Etanercept inhibits binding of both TNFα and TNFβ (lymphotoxin alpha [LTα]) to cell surface tumor necrosis factor receptors (TNFRs), rendering TNF biologically inactive. Enbrel has been approved for use in RA at doses of 50 mg/week or 25 mg/twice weekly and for use with or without methotrexate (MTX).

Samsung Bioepis has proposed SB4 as a biosimilar to US-Enbrel for the treatment of rheumatoid arthritis (RA) and other related indications. SB4 is a dimer of a chimeric protein produced in Chinese hamster ovary suspensions. The applicant conducted a comparative clinical study (SB4-G31-RA) to evaluate the efficacy, safety, pharmacokinetics and immunogenicity of SB4 compared to EU-Enbrel in patients with rheumatoid arthritis. The applicant claims that the results from this trial show similar efficacy between the products.

The primary endpoint of the study was the proportion of patients who remained in the study and achieved an American College of Rheumatology 20% improvement (ACR20) at week 24. The adjusted treatment difference in ACR20 response rate between the SB4 and EU-Enbrel treatment groups was -1.7% and the 90% confidence interval of the adjusted treatment difference was (-4.4%, 7.7%), which was contained within the similarity margins [−12%, +12%] recommended by FDA. The ACR20 response probabilities over time comparing the SB4 and EU-Enbrel up to week 24 also supported similarity.

The finding of similar efficacy is highly credible notwithstanding the number of dropouts. Of 596 total randomized patients, 551 (92.4%) patients completed week 24 and 505 (84.7%) completed week 52. By week 52, 91 (15.3%) patients had withdrawn from the study: 51 (17.2%) patients from SB4 and 29 (9.8%) patients from EU-Enbrel. We conducted tipping point analyses to explore the sensitivity of results to violations in assumptions about the missing data. Confidence intervals for the differences between SB4 and EU-Enbrel failed to rule out concerning losses in efficacy only under the assumption that patients who dropped out on SB4 had much worse outcomes than dropouts on EU-Enbrel. Given the similar proportions of patients and distributions of reasons for early withdrawal on the two treatment arms, in addition to the similar baseline characteristics between dropouts on the two arms, an assumption of such large differences between the outcomes in dropouts on the two treatments seems implausible.

2 INTRODUCTION

2.1 Background

Rheumatoid Arthritis (RA) is an autoimmune disease characterized by inflammation in the synovium of joints, malaise, morning stiffness, and fatigue. If not treated, RA may lead to significant disabilities including bone erosion and joint deformity, over 10-20 years. Increased levels of tumor necrosis factor alpha (TNF-α) have been detected in RA patients, indicating that it may have a role in inducing inflammatory response.

The current submission proposes SB4 as a biosimilar to Enbrel (etanercept), a US licensed biologic response modifier (BRM) for the treatment of five long-term inflammatory conditions: rheumatoid arthritis (RA), polyarticular juvenile idiopathic arthritis (JIA), psoriatic arthritis (PsA), ankylosing spondylitis (AS), and moderate to severe plaque psoriasis in adult patients. The applicant provided reports on direct physico-chemical and biological comparisons between SB4 and EU-Enbrel. Similarity between EU- and US-Enbrel was established using analytical and PK bridging studies.

The submission includes results from several nonclinical, analytical, and clinical studies to evaluate the claim of no clinically meaningful differences between SB4 and US-Enbrel. This review focuses on the safety and efficacy evaluation of SB4 in a comparative clinical study in RA.

2.2 History of Product Development

In November 1998, FDA approved Enbrel for the treatment of RA patients who had an inadequate response to one or more disease-modifying antirheumatic drugs (DMARDS). Since the initial approval, additional indications include; polyarticular juvenile idiopathic arthritis, psoriatic arthritis (PsA), ankylosing spondylitis (AS), plaque psoriasis and pediatric plaque psoriasis.

The applicant conducted *in vitro* and *in vivo* non-clinical studies, followed by a phase I study in healthy male patients to compare the PK, safety, tolerability and immunogenicity. From the phase I study (SB4-G11-NHV), the applicant reported that the PK profiles of SB4 and reference product (Enbrel) were similar in healthy male patients. To evaluate similarity in efficacy, safety, PK and immunogenicity, the applicant conducted a randomized, double-blind, parallel group, multicenter phase III study in patients with moderate to severe RA despite MTX therapy.

On September 23, 2011, the applicant requested a pre-IND meeting with FDA to discuss the acceptability of available CMC and nonclinical data to support the proposed development program for SAIT104 (former product name of SB4) as a biosimilar to US-approved Enbrel. In the BPD Type 4 meeting held on May 26, 2016, the applicant provided details of the studies involved in the developmental program and requested FDA's recommendations on PK and efficacy assessments. In particular, the applicant agreed to FDA's recommendation to include all randomized patients in the primary analysis and report 90% confidence interval for the difference in ACR20 responses and a similarity margin with a lower bound of −12%.

2.3 Specific Studies Reviewed

The statistical review will focus on the efficacy results from study SB4-G31-RA, a randomized, double-blind, parallel group, multicenter clinical study to evaluate the efficacy, safety, PK and immunogenicity of SB4 compared to EU-Enbrel in patients with moderate to severe RA despite MTX therapy. The study consisted of 6 weeks of screening period, 52 weeks of active treatment and 4 weeks of safety follow-up period. The study was followed by an open-label, extension period to evaluate the long-term safety, tolerability, immunogenicity and efficacy of SB4 in patients with RA treated previously with SB4 or EU Enbrel. There was a 48 week active treatment period and a 4 week safety follow up period in this phase of the study.

2.4 Data Sources

Data and reports to the CDER electronic data room can be found at \\cdsesub1\evsprod\BLA761066\0001\m5

3 STATISTICAL EVALUATION

The applicant developed a statistical analysis plan (SAP) based on the final version of protocol SB4-G31-RA. The SAP provided details of the statistical methods to be used in the analysis of efficacy, safety, pharmacokinetics, and immunogenicity data.

3.1 Data and Analysis Quality

The applicant submitted data files of acceptable quality and it was possible to reproduce the primary analysis dataset, and in particular, the primary endpoint results, from the original data. The final statistical analysis plan (SAP) was submitted and relevant analysis decisions were made prior to unblinding. It contains details regarding the datasets, original variables, and derived variables used for analysis. The applicant submitted SAS programs and SAS macros for primary and secondary analyses.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

Study SB4-G31-RA was a randomized, double-blind, parallel group, multicenter clinical study to evaluate the efficacy, safety, pharmacokinetics and immunogenicity of SB4 compared to EU-Enbrel in patients with moderate to severe active rheumatoid arthritis despite methotrexate therapy (Table 1). The study consisted of male or female patients aged 18–75 years, who had been diagnosed as having RA according to the revised 1987 ACR criterial for at least 6 months prior to screening. Active disease was defined by the presence of six or more swollen joints, six or more tender joints, and at least two of the following: morning stiffness lasting at least 45 minutes, an erythrocyte sedimentation rate (ESR) greater than 28 mm/h, and a serum C-reactive protein (CRP) concentration greater than 1.0 mg/dL. Patients were on methotrexate for at least 6 months prior to randomization, with a stable dose of MTX 10-25 mg/week for at least 4 weeks. Patients treated previously with any biological agents including any tumor necrosis factor-alpha (TNF-α) inhibitor were excluded. The study was conducted in 73 investigator sites in Czech Republic, Lithuania, Poland, Ukraine, Bulgaria, Hungary, UK, Colombia, Mexico and South Korea.

Table 1: Design for Clinical Study SB4-G31-RA

Study Phase	Patients	Study Design	Primary Endpoints	Study Population
SB4-G11-NHV	Healthy male patients	Randomized, single-blind, three-part, two-period, two-sequence (1:1 ratio), single-dose, cross-over study; Active-controlled (vs. EU Enbrel, US Enbrel)	AUCinf, Cmax	N=138 Part A: N=46 (n=23 per arm) Part B: N=46 (n=23 per arm) Part C: N=46 (n=23 per arm)
SB4-G31-RA	Patients with moderate to severe rheumatoid arthritis (RA) despite MTX therapy	Randomized, double-blind, parallel group, multicenter clinical Study	ACR20 response rate at week 24	Randomized, Double-blind Period: N=596 (SB4: 299; Enbrel: 297) Transition extension Period: 245 RA patients

Source: Reviewer

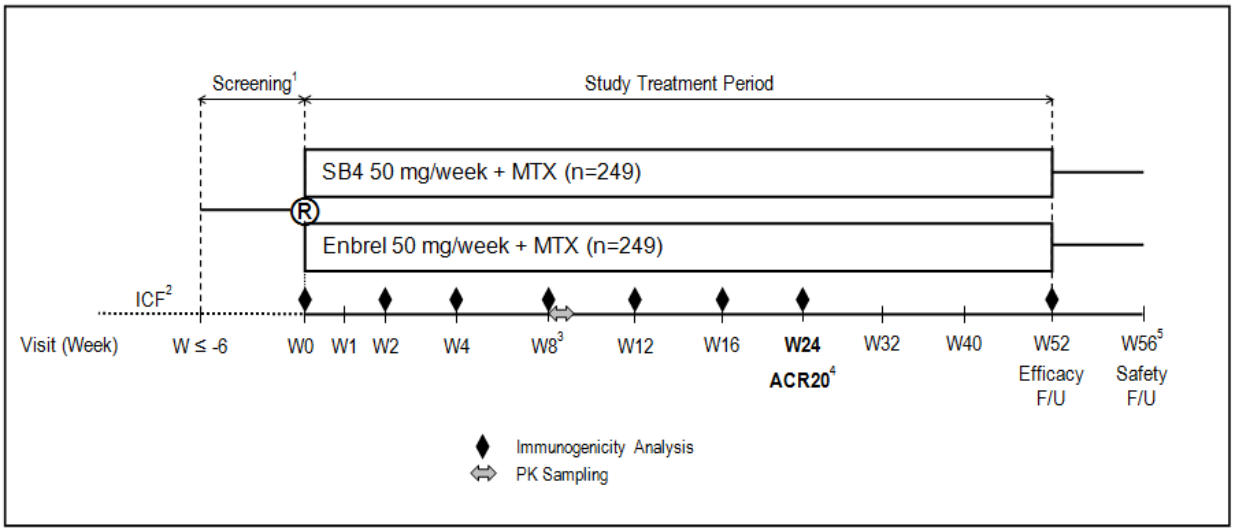
Efficacy, safety, PK, and immunogenicity were evaluated in the randomized double-blind phase of the study. The study comprises a 6-week screening period, a 52-week active treatment and a 4-week safety follow-up period. The study was followed by an open-label, extension period to evaluate the long-term safety, tolerability, immunogenicity and efficacy of SB4 in patients with RA treated previously with SB4 or EU-Enbrel. The extension period comprised a 48-week active treatment and a 4-week safety follow-up period.

3.2.2 Randomized, Double-blind Period

The initial period consisted of 6 weeks of screening and 52 weeks of active treatment stages. A total of 596 patients from 70 centers with moderate to severe RA were randomized in a 1:1 ratio to receive 50 mg of either SB4 or EU-Enbrel via self-administered subcutaneous injection once-weekly and a stable dose of 10-25 mg/week MTX were given during the treatment. Patients were required to take folic acid 5-10 mg/week while taking MTX during the study period.

A graphical representation of the randomized, double-blind period is given in Figure 1 below.

Figure 1: Study Design of the Randomized, Double-blind Period



(Source: Applicant)

The primary efficacy endpoint was the American College of Rheumatology 20% improvement (ACR20) response rate at week 24. An ACR20 responder was defined as a patient with:

- at least a 20% improvement from baseline in swollen joint count (66 joint count)
- at least a 20% improvement from baseline in tender joint count (68 joint count)
- at least a 20% improvement from baseline in at least three of the following five criteria:
    - subject pain assessment using a 100 mm visual analogue scale (VAS)
    - subject global assessment using a 100 mm VAS
    - physician global assessment using a 100 mm VAS
    - patients assessment of disability using the Health Assessment Questionnaire - Disability Index (HAQ-DI)
    - acute phase reactant level (CRP)

Secondary efficacy endpoints included:

- ACR20 response at week 52.
- ACR 50% response criteria (ACR50) and ACR 70% response criteria (ACR70) response at week 24 and week 52.
- numeric index of the ACR response (ACR-N) at week 24 and week 52.
- area under the curve (AUC) of ACR-N up to week 24.
- disease activity score based on a 28 joint count (DAS28) score at week 24 and week 52.
- European League Against Rheumatism (EULAR) response at week 24 and week 52.
- AUC of the change in DAS28 from baseline up to week 24.
- major clinical response (ACR70 response for 6 consecutive months) at week 52.
- change from baseline in modified total sharp score (mTSS) at week 52.

Patients could withdraw from the study at any time for any reason. Withdrawal of informed consent, loss to follow-up and death were considered valid reasons for study termination of randomized patients. Those who discontinued administration of the investigational product (IP) prior to week 52 were asked to return for an early termination (ET) visit procedures and to have a follow-up telephone interview, but were not followed up to obtain key efficacy and safety assessments. Safety and efficacy scores were recorded at the ET visit. Patients who withdrew from the study with missing ACR20 response at week 24/week 52 were considered non-responders at week 24/week 52 in analyses in the full analysis set.

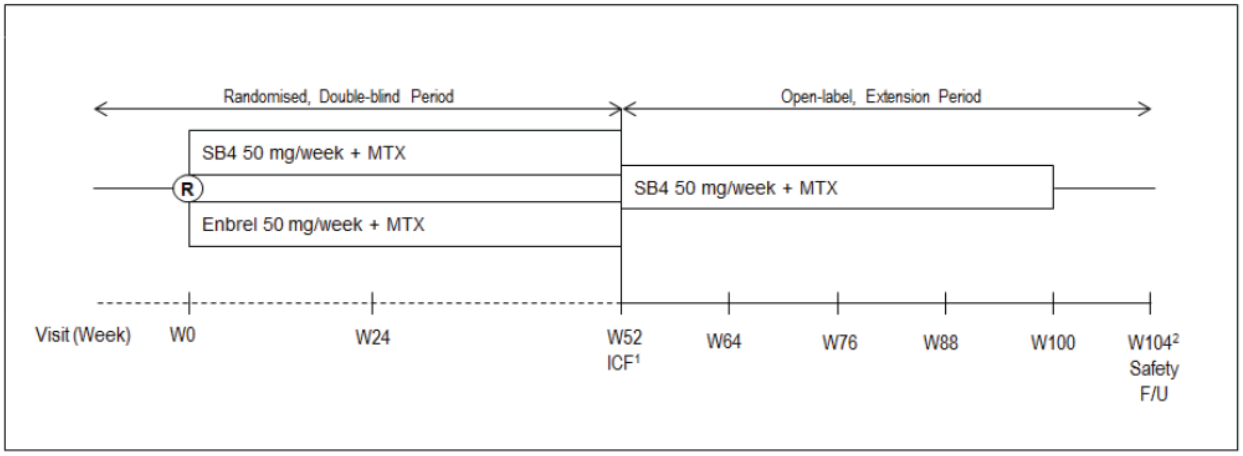
3.2.3 Transition-Extension Period

The long-term safety, tolerability, immunogenicity and efficacy of SB4 in patients with RA treated previously with SB4 or EU-Enbrel were evaluated in an open label study consisted of 48 weeks of active treatment and 4 weeks of safety follow-up. The transition-extension period of ranged from week 52 to week 100. A total of 245 patients who had completed the scheduled

week 52 visit of the randomized, double-blind period, and who were willing to participate were enrolled from Czech Republic and Poland.

A graphical representation of the transition-extension period is given in the figure (Figure 2) below.

Figure 2: Entire Study Design, Including Open-Label Extension Period



Source: Applicant

3.2.4 Data Sets

The following datasets were used in different analyses:

The randomized set (RAN) consisted of all enrolled patients who received a randomization number at the randomization visit.

The full analysis set (FAS) consisted of all patients in the RAN. Patients were analyzed according to the treatment they were assigned at randomization. However, patients who did not qualify for randomization and were inadvertently randomized into the study were excluded from the FAS, provided these patients did not receive any IP during that study phase.

The per-protocol set 1 (PPS1) consisted of all FAS patients who completed the week 24 visit and had an adherence (from baseline to week 24) within the range of 80–120% for both the expected number of IP administrations and the expected sum of MTX doses without any major protocol deviations that affected the efficacy assessment. The applicant defined the PPS1 as the primary analysis set. Major protocol deviations that led to exclusion from this set were pre-specified prior to unblinding the treatment codes for analyses.

The per-protocol set 2 (PPS2) consisted of all FAS patients who completed the week 52 visit and had an adherence (from baseline to week 52) within the range of 80-120% for both the expected

number of IP administrations and the expected sum of MTX doses without any major protocol deviations that affected the efficacy assessment.

The safety set (SAF) consisted of all patients who received at least 1 dose of double-blind IP during the study period. Patients were analyzed according to the treatment received.

The pharmacokinetic population (PK) consisted of all patients in the SAF who had at least 1 post-dose PK sample collected.

3.3 Statistical Methodologies

3.3.1 Planned Analysis

ACR20 response rate at week 24 was the primary endpoint of the study. In order to evaluate similarity, the applicant compared ACR20 response rates between the SB4 and EU-Enbrel treatment arms. The null hypothesis was defined as either 1) SB4 is inferior to EU-Enbrel or 2) SB4 is superior to EU-Enbrel based on a pre-specified similarity margin. According to the statistical analysis plan, biosimilarity would be concluded if the two-sided 95% confidence interval of the difference in ACR20 response rate was contained within the similarity margins of [-15%, 15%]. The applicant also carried out an analysis using a 90% confidence interval with a similarity margin of [-12%, 12%] based on FDA recommendations at a type 4 meeting on May 26, 2016.

A randomization based non-parametric ANCOVA (Koch, 1998)[3] was used to analyze the primary endpoint by adjusting for effects of region (pooled centers) and baseline CRP value. The primary efficacy analysis for ACR20 response was performed in the per-protocol population (PPS1). No missing data was imputed. In addition to the primary analysis in the PPS1, the applicant performed the same analysis in the full analysis set to explore the robustness of the results.

Similar statistical analyses were performed on secondary endpoints such as ACR50 and ACR70 response at week 24 in the PPS1 population and at week 52 in the PPS2 population. Continuous ACR-N at week 24 and week 52 and the AUC of ACR-N up to week 24 were analyzed using an ANOVA model with treatment group and region as factors.

Change from baseline DAS28 at week 24 and week 52, and the AUC of DAS28 up to week 24 were analyzed using an ANCOVA model, with treatment group and region as factors, and using DAS28 baseline value as a covariate. Change from baseline value of Modified Total Sharp score (mTSS) at week 52, an endpoint assessing radiographic progression, was analyzed using an

ANCOVA model. In addition, the applicant reported 95% confidence intervals for the adjusted difference in rates for binary endpoints and difference in means for continuous endpoints using the full analysis set.

The applicant conducted sensitivity analyses using three different approaches: an analysis of available data without imputation (excluding patients with missing data at week 24/week 52), a non-responder imputation analysis (considering patients with missing ACR20 response to be non-responders), and a pattern mixture analysis with a multiple imputation approach that assumes data is missing at random (MAR) except for patients who withdraw from the study with a primary reason of lack of efficacy.

In addition to the proposed sensitivity analysis methods, FDA recommended that the applicant conduct additional analyses that more systematically and comprehensively explore the space of plausible missing data assumptions. Specifically, FDA recommended the inclusion of tipping point analyses that vary assumptions about the missing outcomes on the two treatment arms. As a response, the applicant conducted tipping point analyses and included the results in the summary of clinical efficacy report. The applicant's analyses were based on single imputation, which failed to take into account the uncertainty in the imputation process; the statistical team therefore evaluated sensitivity using multiple imputation based tipping point analyses.

3.3.2 Similarity Margin

The determination of similarity margins was a critical design aspect of this comparative clinical study because it determined the null hypothesis being tested in the primary analysis, i.e., the differences in efficacy that the study will need to rule out at an acceptable significance level.

The applicant initially proposed to conduct the primary efficacy analysis by comparing the 95% confidence interval (CI) of the difference of two proportions with the pre-specified similarity margins of [−15%, 15%]. However, FDA recommended similarity margins of [-12%, 12%] at a type 4 meeting on Dec 14, 2015. FDA also recommended use of a two-sided 90% CI because it generally expects the type I error probability for each margin to be controlled at the overall 5% level in comparative clinical studies. The applicant agreed with this recommendation and performed additional analyses to calculate 90% CIs for the difference in ACR20 in the FAS and PPS. As the double-blind period of the study had already been completed, results from the revised analysis were not included in the clinical study report but were instead reported in the Integrated Summary of Effectiveness and Summary of Clinical Efficacy reports.

The lack of a priori agreement between the applicant and FDA on a similarity margin is not of concern in this case because the primary analysis successfully ruled out the ±12% margin

recommended by FDA. Our selection of ±12% similarity margins was based on discussions with clinicians aimed at weighing the clinical importance of different losses in effect against the feasibility of different study sizes. In a comparative clinical study designed with 90% power to reject absolute differences greater than 12% in magnitude, observed differences larger than approximately 6% will result in failure to establish similarity, as the 90% confidence interval for the estimated difference will not rule out the 12% margin. Therefore, the comparative clinical study will be able to rule out losses in ACR20 response greater than 12% with high (at least 95%) statistical confidence, and will be able to rule out losses greater than around 6% with moderate (at least 50%) statistical confidence. The lower similarity margin (-12%) also corresponds to the retention of approximately 50% of conservative estimates of treatment effect sizes relative to placebo for etanercept (e.g., see Table 2).

Table 2: Similarity Margin Meta-Analysis: ACR20 in Active RA Despite Treatment with Methotrexate

Study	Week	MTX + Placebo		MTX + Etanercept		Difference in % Response
		N	% Response	N	% Response	
Weinblatt et al.	12	30	33%	59	66%	33%
Lan et al.	12	29	34%	29	90%	55%
Meta-Analysis (fixed effects[1]): Difference (95% CI)						42.2% (27.4%, 57.0%)
Meta-Analysis (random effects[2]): Difference (95% CI)						44.0% (21.8%, 66.3%)
Heterogeneity p-value						0.13

[1] Based on Mantel-Haenszel weights
[2] Based on DerSimonian-Laird approach
MTX methotrexate

3.4 Demographic and Baseline Characteristics

3.4.1 Double-blind Period

In the study, 777 patients were screened, of which 584 were randomized. A total of 596 patients were included in the FAS, including patients who were not randomized but inadvertently received treatment, 483 (81.0%) patients satisfied the criteria for the PPS1, 440 (73.8%) patients satisfied the criteria for the PPS2 and 79 (13.7%) patients were analyzed for PK. Missing ACR responses were treated as non-responders in the FAS and no missing data were imputed in the PPS1 and PPS2. Table 3 summarizes the number of patients included in each analysis set

Table 3: Data Sets (Double-blind Period)

	Treatment				*All*	
	SB4		**EU-Enbrel**			
	N	**(%)**	**N**	**(%)**	**N**	**(%)**
Double Blind Phase						
Full Analysis Set Population	299	100	297	100.00	596	100
PPS1 Population	247	82.6	236	79.5	483	81.0
PPS2 Population	224	74.9	216	72.7	440	73.8
Safety Population	299	100	297	100.00	596	100
PKS Population	41	13.7	38	12.8	79	13.7

Source: Reviewer

The demographic characteristics by treatment arm are given in Table 4. Patients in the two treatment arms had comparable demographic characteristics in the baseline. There were 299 patients in the SB4 arm and 297 patients in EU-Enbrel arm. Majority of the study population was older than 65 years (86.4%). Around 92% of patients were whites, 84% were female and the mean age was 52 years. There were no obvious imbalances in demographic and disease characteristics between the SB4 and EU-Enbrel groups.

Table 4: Baseline Demographic Characteristics (Double-blind Period)

	SB4 N=299		**EU-Enbrel** N=297		**Total** N=596	
Age	52.1	11.7	51.6	11.6	51.8	11.7
Age group						
< 65 years	253	84.6	262	88.2	515	86.4
>= 65 years	46	15.4	35	11.8	81	13.6
Sex						
Male	50	16.7	44	14.8	94	15.8
Female	249	83.3	253	85.2	502	84.2
Race						
White	279	93.3	273	91.9	552	92.6
Other	4	1.3	4	1.4	8	1.3
Asian	11	3.7	13	4.4	24	4.0
American Indian	5	1. 7	7	2.4	12	2.0
Weight (kg)	72.5	15.9	71.0	14.6	71.8	15.3
Height (cm)	164.4	8.8	164.4	8.6	164.4	8.7
BMI (kg/m2)	26.8	5.5	26.3	5.3	26.6	5.4

Source: Reviewer program: SB4_Subject_Level.sas

Cell contents are mean (standard deviation) or frequency (percent)

The baseline disease characteristics were comparable between the SB4 and EU-Enbrel treatment arm (Table 5).

Table 5: Baseline Disease Characteristics (Double-blind Period)

	SB4 N=299		**EU-Enbrel** N=297		**Total** N=596	
	Mean	SD	Mean	SD	Mean	SD
C-reactive protein (mg/L)	14.5	19.9	12.8	16.0	13.6	18.1
DAS28	6.5	0.9	6.5	0.9	6.5	0.9
DAS28-ESR	46.4	22.1	46.3	22.6	46.3	22.3
HAQ-DI-Total	1.5	0.6	1.5	0.6	1.5	0.6
Physician global assessment	62.2	15.1	63.2	14.8	62.7	14.9
Subject global assessment	61.7	19.0	63.1	17.7	62.4	18.4
Swollen joint count (28)	11.2	5.4	10.9	5.0	11.0	5.2
Swollen joint count (66)	15.4	7.5	15.0	7.3	15.2	7.4
Subject pain assessment	61.8	20.2	62.3	19.2	62.1	19.7
Swelling joint count	15.3	7.4	14.9	7.3	15.1	7.4
Tender joint count	23.5	11.8	23.5	12.6	23.5	12.2
Total joints for tender count	67.9	0.5	67.9	0.7	67.9	0.6
Total joints for swollen count	65.9	0.4	65.9	0.7	65.9	0.6

Source: Reviewer program: SB4_Subject_Level.sas

3.4.2 Transition Extension Period

Results given in Table 6 shows that the demographic characteristics in the transition extension period were similar across between the SB4 and EU-Enbrel group.

Table 6: Demographic Characteristics (Transition Extension Period)

	SB4 ⇒ SB4 N=126		EU-Enbrel ⇒ SB4 N=119		Total N=245	
Age	49.9	12.1	52.1	10.9	51.0	11.5
Age group						
< 65 years	112	88.9	108	90.8	220	89.8
>= 65 years	14	11.1	11	9.2	25	10.2
Sex						
Male	19	15.1	19	16.0	38	15.5
Female	107	84.9	100	84.0	207	84.5
Race						
White	126	100	118	99.2	244	99.6
Asian	0	0	1	.8	1	.4
Weight (kg)	72.1	16.2	71.5	14.7	72.8	15.5
Height (cm)	165.1	8.1	164.3	8.1	164.7	8.1
BMI (kg/m^2)	27.2	5.9	26.5	5.2	26.8	5.6

Source: Reviewer program: TE_SB4_Subject_Level.sast
Cell contents are mean (standard deviation) or frequency (percent)

3.5 Patient Disposition

The disposition of patients in the study did not contradict similarity between the two treatments (Table 7). There were 299 patients in the SB4 and 297 in EU-Enbrel group. Out of 596 total randomized patients, 551 (92.4%) patients completed week 24 and 505 (84.7%) patients completed week 52. Up to week 52, 91 (15.3%) patients had withdrawn from the study: 51 (13.4%) patients from SB4 and 29 (17.2%) patients from EU-Enbrel. Adverse events, investigator discretion, and withdrawal of patient consent were the major reasons for withdrawal from the study. Withdrawal from the study due to adverse events was comparable in both study arms. However, withdrawal due to investigator discretion was higher in SB4 group (37.5% vs. 19.6%) and withdrawal of consent was higher in the EU-Enbrel group (35.3% vs. 22.5%).

Table 7: Reasons for Withdrawal through Week 52

	SB4		EU- Enbrel		All	
	N	(%)	N	(%)	N	(%)
Randomized	299		297		596	
Completed week 52	259	86.6	246	82.2	505	84.7
Withdrew before week 52	40	13.4	51	17.2	91	15.3
Reason for Withdrawal						
Adverse event	13	32.5	17	33.3	30	33.0
Investigator discretion	15	37.5	10	19.6	25	27.5
Lack of efficacy	1	2.5	3	5.9	4	4.4
Protocol deviation	1	2.5	.	.	1	1.1
Subject lost to follow-up	1	2.5	3	5.9	4	4.4
Withdrew consent	9	22.5	18	35.3	27	29.7

Source: Reviewer program: SB4_Subject_Level.sas

3.6 Results and Conclusions

3.6.1 Primary Efficacy Analysis

The applicant's initial proposal was to conduct primary analysis on the per-protocol set. The results are given in Table 8. The proportion of patients who obtained an ACR20 response at week 24 was found to be similar between the groups, with a slightly higher rate in the EU-Enbrel group (80.5%) compared to the SB4 group (78.1%); the estimated absolute difference was -2.4% (95% CI: (-9.5%, 4.8%). Both the 90% and 95% confidence intervals were well contained within the FDA-recommended similarity margin of [-12%, 12%].

Table 8: Primary Analysis of ACR20 Response Rate at Week 24 (Per-protocol Set)

Treatment	n/N	%	Adjusted Difference	90% CI	95% CI
SB4 (N=247)	193/247	(78.1%)	**-2.4%**	(-8.4%, 3.7%)	(-9.5%, 4.8%)
EU Enbrel (N=236)	190/236	(80.5%)			

Source: Reviewer program SB4_Efficacy.sas

The analysis of ACR20 response based on FDA recommendation to evaluate similarity on the FAS population is given in Table 9 below. The estimated absolute difference was 1.7% with 90% CI: (-4.4%, 4.7%), which was well contained within the FDA-recommended similarity margin of [-12%, 12%].

Table 9: Analysis of ACR20 Response Rate at Week 24 (Full Analysis Set)

Treatment	n/N	%	Adjusted Difference	90% CI
SB4 (N=299)	220/299	(73.6%)	**1.7%**	(-4.4%, 7.7%)
EU-Enbrel (N=297)	213/297	(71.7%)		

Source: Reviewer program SB4_Efficacy.sas

More than half of the non-responders were patients who completed the study and did not satisfy the ACR20 response criteria. The remaining non-responders were patients who withdrew from the study prior to week 24. There were no large differences between SB4 and EU-Enbrel in the distributions of reasons for non-responses (Table 10).

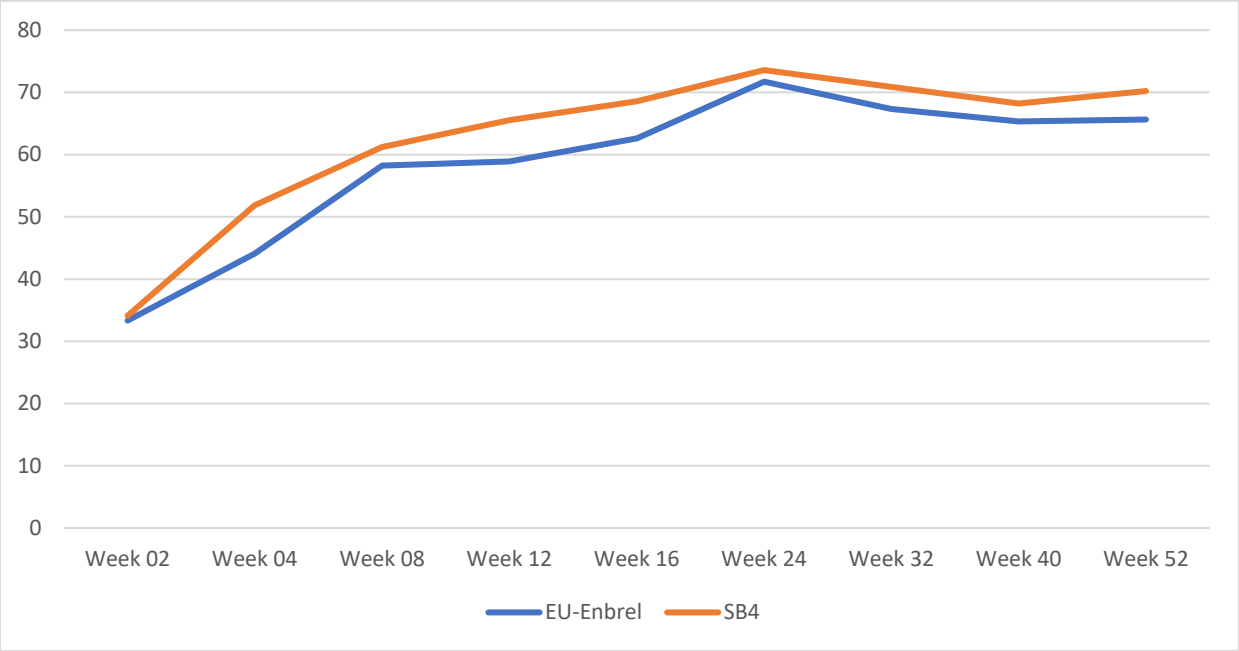
Table 10: Reasons for Non- Response, ACR20 Primary Endpoint at Week 24

	SB4		EU-Enbrel	
	N	**(%)**	**N**	**(%)**
Non-responder	79	26.4	84	28.3
ACR20 criteria not met	41	51.9	64	76.2
Withdraw from study	38	48.1	20	23.8
Adverse event	16	42.1	7	35.0
Investigator discretion	3	7.9	4	20.0
Lack of efficacy	3	7.9	1	5.0
Protocol deviation	0	0.0	1	5.0
Subject lost to follow-up	1	2.6	0	0.0
Withdrew consent	15	39.5	7	35.0

Source: Reviewer program SB4_Efficacy.sas

The proportions of patients remaining in the study and achieving ACR20 responses over time during the study period were similar between the treatment arms (Figure 3).

Figure 3: ACR20 Response Probabilities over Time (Double-blind period)



Source: Reviewer program SB4_Efficacy.sas

Mean changes from baseline in the components of the ACR composite endpoint were also similar between the arms in all randomized patients who completed the study (Table 11), as well as in the per-protocol population.

Table 11: Mean Changes from Baseline in ACR Components at Week 24 (Double-blind period)

	SB4		EU-Enbrel		Difference	95% CI
	N	Mean	N	Mean		
Swollen Joint count	288	-8.2	272	-8.0	0.2	(-0.7, 1.1)
Tender Joint Count	288	-9.5	272	-9.1	0.3	(-0.8, 1.5)
HAQ Score	287	-0.6	272	-0.6	0.0	(-0.1, 0.1)
Patient Pain	287	-29.4	272	-29.2	0.2	(-4.2, 4.6)
Patient Global	287	-30.0	272	-29.7	0.4	(-3.8, 4.6)
Physician Global	285	-40.5	266	-41.1	-0.6	(-4.0, 2.8)
ESR	288	-21.2	272	-22.6	-1.4	(-4.9, 2.2)
CRP	290	-9.6	275	-8.6	1.1	(-1.8, 3.9)

Source: Reviewer program SB4_Efficacy.sas

3.6.2 Secondary Efficacy Analysis

As shown below, comparative analyses of secondary endpoints also showed similar efficacy between SB4 and EU-Enbrel. Secondary endpoints in the study included ACR20 response at week 52, ACR50 and ACR 70 at week 24 and week 52, ACR-N at week 24 and week 52, area under the curve of ACR-N up to week 24, and disease activity score based on 28 joint counts (DAS 28 score) at week 24 and week 52, EULAR response at week 24 and week 52, AUC of the change in DAS28 from baseline up to week 24, major clinical responses, and modified total sharp score at week 52.

3.6.2.1 Analysis of ACR20 response rate at week 52

ACR20 response rate at week 52 was found to be similar between the treatment groups (Table 12). For patients in the in the full analysis set, it was 4.5% with a 90% CI of (-1.7%, 10.7%).

Table 12 : Analysis of ACR20 Response Rate at Week 52 (Double-blind Period)

Treatment	n/N	%	Adjusted Difference Rate	90% CI
SB4 (N=299)	210/299	(70.2%)	**4.5%**	(-1.7%, 10.7%)
EU Enbrel (N=297)	195/297	(65.7%)		

Source: Reviewer program SB4_Efficacy.sas

3.6.2.2 Additional Secondary Endpoints

In addition to the similar results obtained from the analysis of binary ACR20 response, the analysis of different continuous endpoints did not show obvious differences between SB4 and EU-Enbrel. As continuous endpoints may be more sensitive to detect differences in treatment effects, such results are reassuring. For example, the analysis of the continuous endpoint ACR-N at week 24 indicates that no obvious difference in the treatment effects between SB4 and EU Enbrel. From Table 13, the difference between the two treatment means was -1.8 with a 90% confidence interval (-6.4, 2.8).

Table 13 : ACR-N at Week 24 (Double-blind period)

Treatment	MEAN	Difference Between Means	90% Confidence Limit
SB4 (N=299)	46.0	-1.8	(-6.4, 2.8)
EU-Enbrel (N=297)	47.8		

Source: Reviewer program SB4_Efficacy.sas

Change from baseline values of DAS28 score at week 24 between SB4 and EU-Enbrel are compared using ANCOVA method and the results are given in the Table 14. The results show that 90% confidence limit is within the FDA recommended margin (-0.5, +0.5).

Table 14 : Change in DAS28 Score at Week 24 (Double-blind period)

Treatment	MEAN	Difference Between Means	90% Confidence Limit
SB4 (N=299)	-2.6	0.1	(-0.1, 0.2)
EU-Enbrel (N=297)	-2.5		

Source: Reviewer program SB4_Efficacy.sas

Table 15 presents results from the analysis of radiographic progression via the change from baseline value of modified total sharp score (mTSS) at week 52. In contrast to other endpoints measuring disease signs and symptoms, the mTSS is intended as a surrogate measure of irreversible disease progression. The result shows that the average score between SB4 and EU Enbrel does not vary significantly (Difference: -0.3; 90% CI: (-0.7,0.2)).

Table 15 : Change from Baseline Value of Modified Total Sharp Score (mTSS) at Week 52

Treatment	MEAN	Difference Between Means	90% Confidence Limit
SB4 (N=299)	0.70	-0.3	(-0.7, 0.2)
EU-Enbrel (N=297)	0.43		

Source: Reviewer program SB4_Efficacy.sas

Results from other secondary analyses are given in
Table 20, Table 21, and
Table 22 of the Appendix and do not indicate significant differences in efficacy between SB4 and EU-Enbrel groups.

3.6.3 Transition Extension Period

Patients who had completed the scheduled week 52 visit of the randomized, double-blind period, and who were willing to participate, were enrolled in the open-label, extension period. ACR20, ACR50, and ACR70 at week 76 and week 100 were the major secondary efficacy endpoints in the open-label phase of the study. The responses rates of these endpoints across various time points show comparable results between the different study arms (Table 16).

Table 16: ACR Responses Over Time in Transition Extension Period

	Week	SB4 ⇒ SB4 N=126	EU-Enbrel ⇒ SB4 N=119	Total N=245
ACR20	Week 52	99/125 (79.2%)	98/119 (82.4%)	197/244 (80.7%)
	Week 76	102/125 (81.6%)	90/117 (76.9%)	192/242 (79.3%)
	Week 100	95/122 (77.9%)	91/115 (79.1%)	186/237 (78.5%)
ACR50	Week 52	65/125 (52.0%)	64/119 (53.8%)	129/244 (52.9%)
	Week 76	74/125 (59.2%)	62/117 (53.0%)	136/242 (56.2%)
	Week 100	73/122 (59.8%)	70/115 (60.9%)	143/237 (60.3%)
ACR70	Week 52	48/125 (38.4%)	39/119 (32.8%)	87/244 (35.7%)
	Week 76	49/125 (39.2%)	44/117 (37.6%)	93/242 (38.4%)
	Week 100	52/122 (42.6%)	48/115 (41.7%)	100/237 (42.2%)

Source: Reviewer program: TE_SB4_Efficacy.sas

3.6.4 Constancy Assumption

To reliably evaluate whether the experimental treatment retains a certain proportion of the effect of the reference product versus placebo, the constancy assumption must be reasonable, i.e. that estimates of the effect of the reference product from historical, placebo-controlled trials used to calculate the similarity margins reflect those of the current comparative clinical study.

That the design and included patient population in SB4-G31-RA is similar to the two randomized, double-blind, parallel-group, placebo-controlled clinical trials used to define the similarity margins (Table 17) supports the constancy assumption. All three trials evaluated the impacts of treatment in TNF inhibitor naïve patients with active RA despite treatment with methotrexate, who were at least 18 years of age, manifested RA with at least six swollen and tender joint counts the time of enrollment, and received treatment as an add-on to a stable dose of methotrexate.

Table 17: Key Characteristics of Study SB4-G31-RA and Two Historical Randomized, Placebo-Controlled Clinical Trials[1] of Etanercept in RA

	Weinblatt et al[1]	Lan et.al[2]	SB4-G31-RA
Selected Inclusion/ Exclusion Criteria	≥ 18 Years ≥ 6 SJ, ≥6 TJ	≥ 18 Years ≥ 6 SJ, ≥6 TJ	≥ 18 Years ≥6 SJ, ≥6 TJ, ESR >28 mm/h, CRP >1.0 mg/dL
Anti-TNF experience allowed?	No	No	No
Concomitant DMARDs	Stable MTX	Stable MTX	Stable MTX
Region / Country	EU	TW	EU, AS, NA
Baseline Characteristics [2]	SJ: 18; TJ: 28; Disease Duration:13 yrs; HAQ: 1.5	SJ: 15; TJ: 14; Disease Duration: >1 year PGA:6.8, HAQ: 1.1	SJ: 15.2; TJ: 24; Disease Duration: 6.1 yrs; HAQ: 1.5
Time of ACR20 Evaluation	Week 12 Week 24	Week 12	Week12 Week 24
ACR20 Response on etanercept	Week 12: 66% Week 24: 71%	Week 12: 90%	Week 12: 63% Week 24: 81%
ACR20 Response on MTX only	Week 12: 33% Week 24: 27%	Week 12: 34%	
Withdrawal on etanercept	3%	6%	7%

Source: Reviewer

Abbreviations: SJ=swollen joint count; TJ=tender joint count; DMARD=disease-modifying anti-rheumatic drug; HAQ = Health Assessment Questionnaire; MTX = Methotrexate; NA=North America; EU=Europe; AS=Asia, TW= Taiwan

[1] Other studies exist which have not been evaluated for inclusion as historical studies for calculation of the similarity margin. See Appendix.

Although the similarity margin was derived from ACR20 response at week 12, conformance to that margin for assessment of similarity was evaluated at a different timepoint in the present study, at week 24 (Table 2), and this difference in time of measurement may violate the constancy assumption. However, this concern is allayed because similarity between results from the historic and present studies is maintained when the present study is evaluated at week 12 (Table 18). In particular, the estimated 90% confidence interval for the difference in ACR 20 response at week 12 between SB4 and EU-Enbrel is (-3.2%, 10.7%), which is contained within the FDA recommended margin of [-12%, 12%] to support similarity.

Table 18: Analysis of ACR20 Response Rate at Week 12 (Per-protocol Set)

Treatment	n/N	%	Adjusted Difference	90% CI
SB4 (N=247)	196/247	(68.4%)	**3.7%**	(-3.2%, 10.7%)
EU Enbrel (N=236)	152/236	(64.4%)		

Source: Reviewer program SB4_Efficacy.sas

The number of swollen and tender joints at baseline were similar in the historic and current trials, with baseline disease duration in the current trial between those of the two historic trials. At baseline, patients in Weinblatt et. al. had a median of 18 swollen joints and 28 tender joints, with a mean disease duration of 13 years. Patients in Lan et. al. had a mean of 15 swollen joints, 14 tender joints, and a mean disease duration of more than 1 year. Patients in study SB4-G31-RA had a mean of 15 swollen joints and 14 tender joints, and a mean disease duration of 6.1 years.

Completion rates were similar in the historic and current trial. In Weinblatt et. al, 57 patients (97%) in the treatment arm completed the study and 2 patients withdrew because of the adverse events unrelated to etanercept. In the Lan et. al, 27 patients (94%) completed the week 12 assessment. These high completion rates are comparable with that from SB4-G31-RA, in which all patients completed the study at week 12 and 93% completed at week 24.

ACR20 outcomes were similar in the etanercept arms in SB4-G31-RA and the historic studies. At week 24, the ACR20 response rate for the etanercept group was 71% in Weinblatt et. al and 81% in SB4-G31-RA. At week 12, the ACR20 response rate for etanercept was 66% in Lan et. al, and 63% in study SB4-G31-RA, again supportive of historic constancy.

In summary, the important aspects of the historical and current studies, including key inclusion criteria, prior medications, add-on medications, baseline disease severity, dropout rates, and ACR20 outcomes, were largely similar and support the assumption of constancy.

3.6.5 Assay Sensitivity

To reliably evaluate whether there are clinically meaningful differences between two products, a comparative clinical study must also have assay sensitivity, or the ability to detect meaningful differences if such differences exist. The absence of a placebo arm in the present active-controlled study makes it difficult to determine whether the assumptions of assay sensitivity and constancy have been met. As discussed in the ICH E10 guidelines and in the literature, historical evidence of sensitivity to drug effects and appropriate trial conduct may be used to support the presence of assay sensitivity and a conclusion that the treatments are similarly effective.

Unlike the historical studies, patient inclusion criteria for study SB4-G31-RA required high baseline ESR or CRP, markers of inflammation associated with high disease activity. Their inclusion supports high assay sensitivity.

That withdrawal rates were low (7%) in study SB4-G31-RA argues in favor of proper study conduct, consistent with continued collection of data regardless of treatment adherence, as specified in the study protocol.

The applicant planned to enroll an adequate number of patients in study SB4-G31-RA to detect a difference between two treatments with sufficient power (80%). Due to the high number of subjects recruited during the last phase of enrollment, the study as executed resulted provided increased power to detect the treatment difference.

Randomized treatment allocation of patients in to the two groups minimized bias and provided an assurance of comparability of the groups with respect to pertinent variables such as age, sex, severity of disease, duration of disease. Moreover, the ACR20 response rate of the etanercept arm of study SB4-G31-RA was comparable with that from the two historical studies used to compute the similarity margin.

In summary, the design, conduct, and within-group responses rates of study SB4-G31-RA strongly support its assay sensitivity.

3.6.6 Potential Impacts of Missing Data

The 2010 National Research Council (NRC) report *The Prevention and Treatment of Missing Data in Clinical Trials* recommends that "examining sensitivity to the assumptions about the missing data mechanism should be a mandatory component of reporting." As we noted before, up to week 24, 45 (7.6%) patients had withdrawn from the study: 16 (5.4%) patients from SB4 and 29 (9.8%) patients from EU-Enbrel. These patients were excluded in the primary analysis in the per-protocol set, and the results provided earlier therefore depend on the unverifiable assumption that the outcomes among protocol adherers were similar to those among non-adherers. Furthermore, in the key supportive analysis in the full analysis set, patients who dropped out were considered non-responders, such that the primary endpoint was a composite measure of treatment success defined by adherence to the treatment through week 24 and achieving an ACR20 response at week 24. Comparing treatments with respect to this composite measure of treatment success may confound differences between treatments in efficacy with differences in tolerability. For example, the composite measure could fail to identify clinically

meaningful differences in efficacy if the proposed biosimilar was better tolerated than the reference product but had lower efficacy in the subset of patients who adhered. Therefore, it is important to evaluate differences in the components of the composite primary endpoint. This includes an evaluation of ACR20 at week 24 in all randomized patients regardless of adherence (an evaluation of the de facto or intention-to-treat estimand), in addition to de facto evaluations of the components of ACR20. However, such evaluations are subject to some missing data (because patients who discontinued treatment were not followed up for assessment) and rely on strong and unverifiable assumptions, e.g., that outcomes in patients who withdrew early are missing at random. Therefore, we requested from the applicant, and conducted our own, tipping point analyses to explore the sensitivity of results to violations in assumptions about the missing data (i.e., to various missing-not-at-random assumptions).

Tipping point analysis indicates that missing data did not impact the conclusion that SB4 is similar to EU-Enbrel (Table 19). The axis values are the differences on each treatment arm between responses in patients who withdrew from the study early and responses in patients who completed the study, and correspond to response rates among patients missing data which range from less than 10% to greater than 90%. None of the scenarios for missing data resulted in loss of conformance to the 12% biosimilarity limits. Similar conclusions were drawn from the applicant's reported tipping point analyses.

Table 19 : Tipping Point Analysis for Week 24 ACR20[1] Response Rate

		Shift for EU-Enbrel[2]					
Shift for SB4[3]		-0.7	-0.5	-0.3	0	0.1	0.2
	-0.7	0.01 (-0.02, 0.05)	-0.01 (-0.05, 0.03)	-0.03 (-0.07, 0.02)	-0.06 (-0.1, -0.02)	-0.07 (-0.1, -0.03)	-0.08 (-0.1, -0.05)
	-0.5	0.02 (-0.02, 0.07)	0 (-0.05, 0.05)	-0.02 (-0.07, 0.03)	-0.05 (-0.09, 0)	-0.06 (-0.1, -0.01)	-0.07 (-0.1, -0.03)
	-0.3	0.03 (-0.02, 0.08)	0.01 (-0.04, 0.07)	-0.01 (-0.06, 0.05)	-0.04 (-0.09, 0.01)	-0.05 (-0.09, 0)	-0.06 (-0.1, -0.02)
	0	0.05 (0.01, 0.09)	0.03 (-0.02, 0.08)	0.01 (-0.04, 0.06)	-0.02 (-0.07, 0.02)	-0.03 (-0.07, 0.01)	-0.04 (-0.07, -0.01)
	0.1	0.05 (0.02, 0.09)	0.03 (-0.01, 0.08)	0.01 (-0.03, 0.06)	-0.02 (-0.06, 0.03)	-0.02 (-0.06, 0.01)	-0.03 (-0.06, -0.01)
	0.2	0.06 (0.03, 0.09)	0.04 (0, 0.08)	0.02 (-0.02, 0.06)	-0.01 (-0.04, 0.03)	-0.02 (-0.05, 0.01)	-0.03 (-0.05, -0.01)

Source: Reviewer program: tipping point.R

[1] Cell contents are estimated difference (90% confidence interval).

[2] Assumed difference in week 24 ACR20 response between completers and dropouts on EU-Enbrel. Response in EU-Enbrel completers was 0.76.

[3] Assumed difference in week 24 ACR20 response between completers and dropouts on SB4. Response in SB4 completers was 0.78.

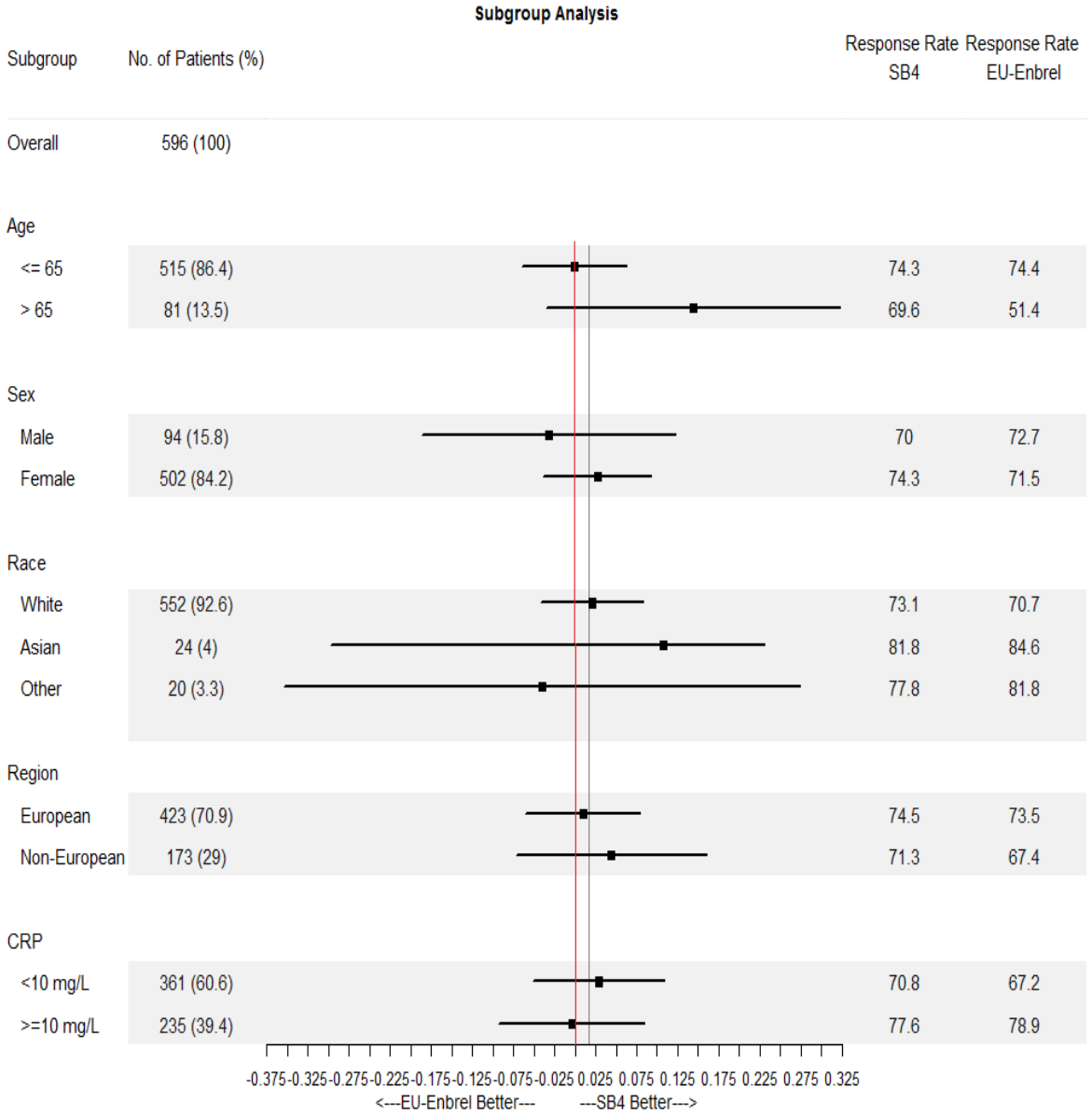
3.7 Evaluation of Safety

Dr. Rachel Glaser, the Medical Reviewer, conducted the complete safety evaluation. The details of the safety evaluation can be found in Dr. Glaser's report.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The subgroup analyses compared efficacy results across treatment arms within different subgroups defined by sex, region, age, race and CRP values. Figure 4 given below shows that estimated differences between subgroups were largely centered around similarity, and there were no striking trends between SB4 and EU-Enbrel. There were no U.S. sites in the study so subgroup analyses in the United States are not possible.

Figure 4: Estimated differences between SB4 and EU-Enbrel in the probability of remaining in the study and achieving an ACR20 response at Week 24, stratified by selected subgroups, in study SB4-G31-RA.



Source: Reviewer program: subgroup.R
Gray vertical line represents estimated difference in overall population, and red vertical line represents no difference

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

During this statistical review, we identified the following important issues:

- Margin selection and evidence of similarity

  Despite FDA recommendations for week 24 ACR20 similarity margins of ±12% evaluated at the two-sided 90% level of confidence, the applicant prespecified a primary analysis with similarity margins of ±15% evaluated at the two-sided 95% level of confidence. The lack of a priori agreement between the applicant and FDA on a similarity margin is not of concern in this case because the primary analysis successfully ruled out the ±12% margin recommended by FDA.

- Potential effect of missing data on the reliability of efficacy results

  As discussed in detail in 3.6.6, the amount of missing data was substantial: Up to week 24, 45 (7.6%)patients had withdrawn from the study: 16 patients (5.4%) from the SB4 treatment group and 29 patients (9.8%) from the EU-Enbrel treatment group. Tipping point sensitivity analyses, however largely supported the findings of similarity between SB4 and EU-Enbrel.

- Assay sensitivity and the constancy assumption

  As discussed in detail in 3.6.4, it is critical that a comparative clinical study have assay sensitivity, the ability to detect any meaningful differences between products and that constancy assumption holds. The totality of available information largely supports constancy as well as sufficient assay sensitivity in the current evaluation of similarity.

5.2 Collective Evidence

The collective evidence from this comparative clinical study in rheumatoid arthritis supports the conclusion of no clinically meaningful differences between SB4 and US-Enbrel. The adjusted treatment difference in ACR20 response rates between the SB4 and EU-Enbrel treatment groups in the per protocol population was -2.4% and the 90% confidence interval of the adjusted treatment difference was (-8.4%, 3.7%). For the FAS population the adjusted treatment difference was -1.7% with 90% CI (-4.4%, 7.7%). Both the CIs were contained within the similarity margins of [−12%, +12%] recommended by FDA. ACR20, ACR50, and ACR70 responses over time, mean changes from baseline in the components of the ACR composite endpoint and the disease activity score (DAS28), and other secondary efficacy endpoint results, showed no obvious differences between SB4 and EU-Enbrel. There was substantial missing data in important analyses, but tipping point analyses largely supported the finding of similarity. In addition, the totality of available information largely supports the assay sensitivity of study SB4-G31-RA as well as the constancy assumption.

6 APPENDIX

Biosimilarity margin:

In the course of this review, we identified other studies potentially relevant for computing biosimilarity margins for etanercept. *Moreland et. al*[4] and *Klareskog et. al*[5] reported results from similar etanercept studies in rheumatoid arthritis patients with baseline characteristics largely comparable to those from SB4-G31-RA. Because the ±12% similarity margins for ACR20 were based on feasibility as well as historical studies, we did not re-evaluate the prespecified margins for this study. However, we recommend consideration of these additional studies when prespecifying similarity margin in future etanercept clinical trials.

Additional Tables and Figures:

Table 20 : Analysis of ACR50 Response in Double-blind Period

Week	Treatment	n/N	%	Adjusted Difference Rate	90% CI
24	**SB4 (N=299)**	128/299	(42.8%)	**3.8%**	(-2.6%, 10.4%)
	EU-Enbrel (N=297)	116297	(39.1%)		
52	**SB4 (N=299)**	143/299	(47.8%)	**5.5%**	(-1.1%, 12.0%)
	EU Enbrel (N=297)	125/297	(42.1%)		

Source: Reviewer

Table 21 : Analysis of ACR70 Response in Double-blind Period

Week	Treatment	n/N	%	Adjusted Difference Rate	90% CI
24	**SB4 (N=299)**	69/299	(20.1%)	**3.3%**	(-2.1%, 8.7%)
	EU-Enbrel (N=297)	59297	(19.9%)		
52	**SB4 (N=299)**	91/299	(30.4%)	**3.5%**	(0.01%, 11.8%)
	EU-Enbrel (N=297)	73/297	(24.6%)		

Source: Reviewer

Table 22 : ACR-N at Week 24 in Response in Double-blind Period

Week	Treatment	n/N	%	Adjusted Difference Rate	90% CI
24	**SB4 (N=299)**	69/299	(20.1%)	**3.3%**	(-2.1%, 8.7%)
	EU-Enbrel (N=297)	59297	(19.9%)		
52	**SB4 (N=299)**	91/299	(30.4%)	**3.5%**	(0.01%, 11.8%)
	EU-Enbrel (N=297)	73/297	(24.6%)		

Source: Reviewer

7 BIBLIOGRAPHY

1. Weinblatt, M. E., Kremer, J. M., Bankhurst, A. D., Bulpitt, K. J., Fleischmann, R. M., Fox, R. I., ... & Burge, D. J. (1999). A trial of etanercept, a recombinant tumor necrosis factor receptor: Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. New England Journal of Medicine, 340(4), 253-259.

2. Lan, J. L., Chou, S. J., Chen, D. Y., Chen, Y. H., Hsieh, T. Y., & Young Jr, M. (2004). A comparative study of etanercept plus methotrexate and methotrexate alone in Taiwanese patients with active rheumatoid arthritis: a 12-week, double-blind, randomized, placebo-controlled study. Journal of the Formosan Medical Association= Taiwan yi zhi, 103(8), 618-623.

3. Koch, G. G. (1998). Issues for covariance analysis of dichotomous and ordered categorical data from randomized clinical trials and non-parametric strategies for addressing them. *Statistics in medicine*, 1863-1892.

4. Moreland, Larry W., Michael H. Schiff, Scott W. Baumgartner, Elizabeth A. Tindall, Roy M. Fleischmann, Ken J. Bulpitt, et al. Etanercept Therapy in Rheumatoid Arthritis: A Randomized, Controlled Trial. Ann Intern Med. 1999;130:478–486. doi: 10.7326/0003-4819-130-6-199903160-00004

5. Klareskog, Lars, et al. "Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial." The Lancet 363.9410 (2004): 675-681.

6. Goldman JA, Xia HA, White B, Paulus H. Evaluation of a modified ACR20 scoring system in patients with rheumatoid arthritis receiving treatment with etanercept. Annals of the Rheumatic Diseases. 2006;65(12):1649-1652. doi:10.1136/ard.2005.047266.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

GINTO J POTTACKAL
02/13/2018

ROBERT ABUGOV
02/13/2018