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

BLA/Serial Number:	761042 / 0
Drug Name:	GP2015
Indication(s):	Same indications as Enbrel
Applicant:	Sandoz
Dates:	Submitted: 7/30/2015 BsUFA: 5/27/2016
Review Priority:	Standard review
Biometrics Division:	Division of Biometrics III
Statistics Reviewer:	Kathleen Fritsch, Ph.D.
Concurring Reviewer:	Mohamed Alosh, Ph.D.
Medical Division:	Division of Dermatology and Dental Products Division of Pulmonary, Allergy, and Rheumatology Products
Clinical Team:	Gary Chiang, M.D. / David Kettl, M.D. (DDDP) Rachel Glaser, M.D. / Nikolay Nikolov, M.D. (DPARP)
Project Manager:	Leila Hann / Jessica Lee

Keywords: biosimilar, stratification, post-hoc analysis

Table of Contents

1 Executive Summary

GP2015 is a proposed biosimilar to US-licensed Enbrel (etanercept). As part of the development program, the applicant conducted a comparative clinical study of GP2015 versus EU-approved etanercept in subjects with moderate to severe psoriasis (Study GP15-302). The primary endpoint was the proportion of subjects at Week 12 achieving at least a 75% reduction from baseline in PASI (PASI 75). The proportion of subjects achieving PASI 75 at Week 12 was similar on both the GP2015 and EU-etanercept arms (70.5% vs. 71.5% in the full analysis population) and the exact 90% confidence intervals for both the full analysis population and the per protocol population were within the pre-specified margin of ± 18%. See Table 1. Thus the comparative clinical study met its similarity criterion. The results of the supportive endpoints based on the mean percent change in PASI and the Investigator's Global Assessment were consistent with the primary endpoint.

Table 1 - Exact Confidence Intervals for the Risk Difference of PASI 75 Response Rates

Population	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
FAS	186/264 70.5%	191/267 71.5%	-1.1%	(-8.3%, 6.0%)
PPS	176/239 73.6%	182/241 75.5%	-1.9%	(-9.4%, 5.6%)

FAS = full analysis set, PPS = per protocol set
Source: reviewer analysis

Adverse events were similar on both arms. Five subjects developed anti-drug antibodies during the first 12 weeks of treatment. All 5 subjects were on the EU-etanercept arm (N=267). At week 12, 98 subjects were switched from EU-etanercept to GP2015. No subjects developed anti-drug antibodies 6 weeks after transitioning from EU-etanercept to GP2015.

The randomization in Study 302 was stratified on prior systemic therapy (no prior systemic therapy, any prior systemic therapy including biologic immunomodulation agents but no prior treatment with a TNF antagonist, or prior treatment with a TNF antagonist) and weight (< 90 kg vs. ≥ 90 kg). However, during the blinded review of the data after the Week 12 database lock, the applicant determined that the stratification had been incorrectly performed for many subjects and that the stratification classification used for the randomization did not agree with the data collected for the study. The applicant conducted two medical reviews of the data to re-classify subjects based on the data recorded in the case report form. One review was conducted before unblinding the data and finalizing the statistical analysis plan (SAP) for the Week 12 database lock. The second review was conducted at the time of the Week 30 database lock. For the Week 12 report, subjects who received UVA or UVB phototherapy, but no systemic treatments for psoriasis were considered to have had prior systemic therapy (that is, included in the 'Any prior therapy' category). At the Week 30 database lock, the applicant removed subjects who had received UVA or UVB phototherapy, but no systemic treatments for

psoriasis from the 'Any prior therapy' category and placed them in the 'No prior therapy', category. Several other subjects were re-classified for other reasons (for example, vitamins, analgesics, and antihistamines were no longer considered systemic therapies for psoriasis). The only rationale provided by the applicant for the reclassification for the Week 30 report was that "it was identified that some patients were incorrectly classified."

The prior therapy classification is relevant to the analyses because, although the protocol stated that the PASI 75 endpoint would be evaluated with exact confidence intervals, the statistical analysis plan stated that the endpoint would be analyzed with a covariate-adjusted confidence interval using estimates from a logistic regression model with terms for treatment group, body weight classification, and prior therapy classification. Thus the results depend upon which version of the prior therapy classification is used in the model (randomization classification, Week 12 classification, or Week 30 classification). The results using all three versions of the prior therapy classification are similar. However, changing the prior therapy groupings twice, including making changes after the initial study report had been finalized, raises concerns with post-hoc changes to the database. Therefore this reviewer recommends using the analysis for the primary endpoint that is most consistent with the original protocol: exact confidence intervals that do not use the stratification factors.

2 Introduction

2.1 Overview

GP2015 is being developed as a proposed biosimilar to US-licensed Enbrel (etanercept) under Section 351(k) of the Public Health Service Act. Section 351(i) of the PHS Act defines biosimilarity to mean "that the biological product is highly similar to the reference product notwithstanding minor differences in clinically inactive components" and that "there are no clinically meaningful differences between the biological product and the reference product in terms of the safety, purity, and potency of the product." As part of their development program, the applicant has conducted a comparative clinical study of GP2015 versus EU-approved etanercept (Study GP15-302) in subjects with psoriasis. The design details for Study 302 are summarized in Table 2.

Table 2 – Comparative Clinical Study Overview

Study Number	GP15-302
Study Design	GP2015 vs. EU-etanercept (12 weeks), followed by randomization to continued treatment or switching between treatments (weeks 12-30), followed by maintenance of the last assigned treatment through Week 52
Inclusion criteria	Adult subjects with active, clinically stable plaque psoriasis who were candidates for systemic therapy with at least 10% BSA, PASI ≥ 10, and IGA ≥ 3.
Treatment regimen	50 mg twice weekly for 12 weeks (50 mg weekly after week 12)
Primary endpoint	PASI 75 at Week 12
Secondary endpoints	Percent change in PASI from baseline to Week 12, the mean averaged treatment effect of percent change in PASI between Week 2 and 12
Treatment arms and Sample Size	GP2015 – 264 EU-etanercept - 267
Study location	Bulgaria, Czech Republic, Estonia, Germany, Hungary, Poland, Romania, Russia, Slovakia, South Africa, United Kingdom, and Ukraine

With the original BLA submission (July 30, 2015), the applicant submitted the 12-week clinical study report for Study 302 based on the database lock after all subjects completed 12 weeks of treatment. The study report also includes immunogenicity data following the first switch in Treatment Period 2 (Week 18). Three months into the review cycle (November 9, 2015), the applicant submitted an updated clinical study report that included data up through the Week 30 database lock. Some of the Week 12 results presented in the 30-week clinical study amendment report differ from those reported in the 12-week clinical study report. The reasons for these differences are discussed in Section 3.2.3 below, and primarily related to how subjects were classified with regard to prior systemic therapies for psoriasis. Unless otherwise noted, this review will present results from the 12-week database lock that was submitted with the original application.

In addition to Study 302, the applicant conducted three pharmacokinetic studies in healthy volunteers comparing GP2015 to either US-etanercept or EU-etanercept. Study GP15-102 is a crossover study of 1 dose of GP2015 and 1 dose of US-etanercept. Study GP15-101 and Study GP15-104 are crossover studies of 1 dose of GP2015 and 1 dose of EU-etanercept. The applicant also submitted Study GP15-105 which was a pre-specified cross-study comparison of the data from Studies GP15-101 and GP15-102 (no new data). This review will focus only on the comparative clinical study GP15-302.

The design and statistical analysis of the comparative clinical study (Study 302) was discussed with the applicant at two Pre-IND meetings. These meetings were held on July 9, 2012 and December 19, 2012. Both meetings were classified as Type 2 Biosimilar Biologic Product Development (BPD) meetings. At the December meeting, the Agency agreed that a primary endpoint of PASI 75 at Week 12 with a similarity margin of 18% may be appropriate. Protocol 302 was discussed at the Pre-IND meetings, but was not

submitted to the Agency for review under an IND and was conducted entirely outside the U.S. Thus, while the definitions of the primary and secondary endpoints and the similarity margins were discussed at the meetings, other specific details of the statistical analysis plan were not reviewed by the Agency.

Shortly before the submission of this BLA (July 13, 2015), an IND for GP2015 was opened in the US with a protocol for a comparative clinical study in subjects with rheumatoid arthritis.

2.2 Data Sources

This reviewer evaluated the applicant's clinical study report, clinical summaries, and proposed labeling. The submission was submitted in eCTD format and was entirely electronic. Both SDTM and analysis datasets were submitted. Two sets of data for Study 302 were submitted. One set was associated with the 12-week clinical study report (based on the database lock after all subjects completed Week 12) that was submitted with the original application. The originally submitted analysis datasets are archived at \\cdsesub1\evsprod\bla761042\0000\m5\datasets\gp15-302\ . On November 9, 2015, the sponsor submitted the 30-week clinical study report (based on the database lock after all subjects completed Week 30) and associated SDTM and analysis datasets. These datasets are archived at \\cdsesub1\evsprod\bla761042\0003\m5\datasets\gp15-302\ . On March 7, 2016, the applicant submitted an additional dataset that listed the subject's prior therapies for psoriasis. This dataset is archived at \\cdsesub1\evsprod\bla761042\0015\m5\datasets\gp15-302\tabulations\legacy\ppm.xpt .

3 Statistical Evaluation

3.1 Data and Analysis Quality

In general, the databases for the studies required minimal data management prior to performing analyses. However, two requests for additional information were made during the review cycle. In the first request, the Agency requested statistical programs for creating the estimates and confidence intervals for the primary and key secondary analyses in Study 302, as the statistical analysis plan did not contain sufficient detail regarding the applicant's models to replicate the analyses without the statistical programs. In the second request, the Agency requested additional datasets in a sufficiently usable form that included information on the recorded prior therapies for psoriasis (which was used to define a key factor in the analyses). The applicant submitted the requested materials.

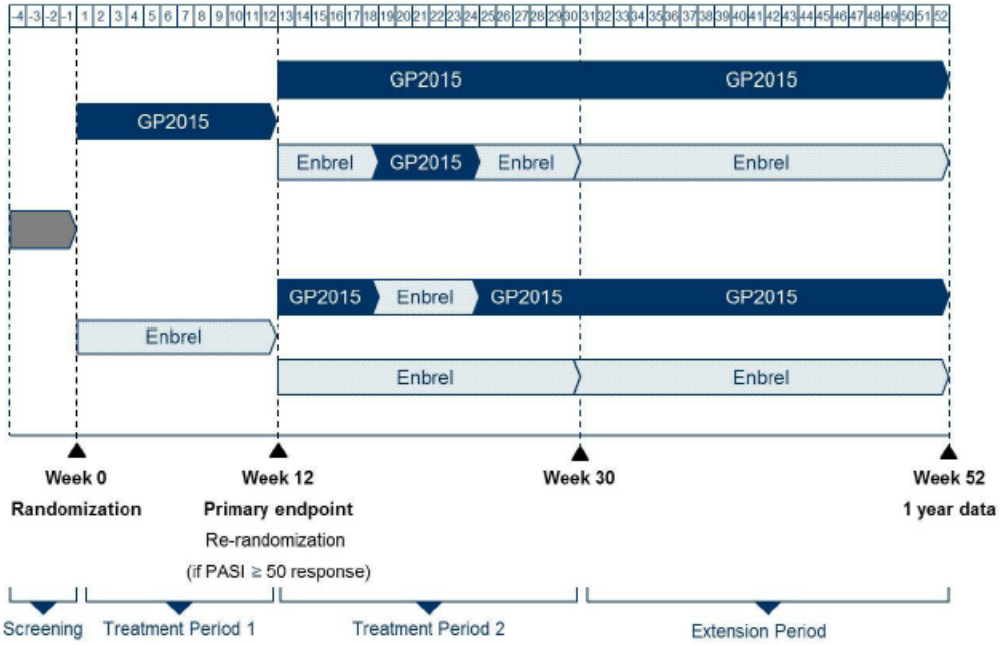
3.2 Evaluation of Efficacy

3.2.1 Study Design and Statistical Analysis

Study 302 was a randomized, double-blind comparative clinical study of GP2015 and EU-etanercept in subjects with moderate to severe chronic plaque psoriasis. The study enrolled subjects age 18 and older with clinically stable chronic plaque psoriasis involving at least 10% body surface area (BSA), PASI $\geq$ 10, and Investigator's Global Assessment (IGA) $\geq$3. Subjects must have previously received phototherapy or systemic

therapy or were candidates for such therapy in the opinion of the investigator. The study enrolled 531 subjects, 264 randomized to the GP2015 arm and 267 randomized to the EU-etanercept arm. Subjects were enrolled at 71 centers in 12 countries (mostly in Eastern Europe). Subjects received subcutaneous injection of 50 mg twice weekly for the first 12 weeks followed by 50 mg once weekly thereafter. The primary timepoint for efficacy assessment was Week 12. After Week 12, subjects who achieved at least PASI 50 response were re-randomized to either maintain the originally randomized treatment through Week 52 or to switch between treatments. Subjects on the switching arms received the other treatment for 6 weeks, then the original treatment for 6 weeks, and finally the other treatment through the end of the study at Week 52. See Figure 1.

Figure 1 – Study 302 Design



Source: pg. 35 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf.

Randomization for Treatment Period 1 was stratified by body weight (< 90 kg vs. ≥ 90 kg) and prior psoriasis therapy (no prior systemic therapy, any prior systemic therapy including biologic immunomodulating agents but no prior treatment with a TNF antagonist, or prior treatment with a TNF antagonist).

Subjects were evaluated at screening, baseline, and Weeks 2, 4, 8, and 12 in Treatment Period 1, Weeks 18, 24, and 30 in Treatment Period 2, and Weeks 36, 42, 48, and 52 in the Extension Period. Efficacy was assessed using the PASI scale, BSA, and IGA. The IGA scale was defined as follows:

Table 3 – Investigator's Global Assessment (IGA)

0	Clear	No signs of psoriasis Post-inflammatory hyperpigmentation could be present
1	Almost Clear	Normal to pink coloration of lesions No thickening No to minimal (focal) scaling
2	Mild	Pink to light red coloration Just detectable to mild thickening Predominantly fine scaling
3	Moderate	Dull bright red, clearly distinguishable erythema Clearly distinguishable to moderate thickening Moderate scaling
4	Severe	Bright to deep dark red coloration Severe thickening with hard edges Severe / coarse scaling covering almost all or all lesions

Source: pg 52 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf

The primary efficacy endpoint was PASI 75 at Week 12. The version of the protocol discussed at the December 19, 2012 meeting and the originally implemented version of the protocol dated February 4, 2013, specified that the difference in PASI 75 response would be analyzed with a 90% confidence interval. Protocol Amendment 1 (dated September 18, 2013) changed the confidence level to 95% citing advice from national European Health authorities. All versions of the protocol stated that an exact confidence interval for the difference in response rates would be calculated and compared with a similarity margin of 18%. The statistical analysis plan (SAP) modified several details of the planned analysis for the PASI 75 endpoint. The proposal to calculate an exact, not covariate-adjusted confidence interval was changed to a proposal to calculate a covariate-adjusted confidence interval based on estimates from a logistic regression model that included the stratification factors. This updated analysis used estimates from a logistic regression with terms for treatment group, body weight stratum, and prior therapy stratum, along with standard errors calculated with the delta method. At the blinded data review meeting, the study team noted that many subjects had discrepancies between the prior therapies recorded in the clinical database and the classification used in the randomization. Therefore the applicant modified the SAP to state that the body weight stratum and prior therapy stratum classifications used in the logistic regression model were to be derived from the information in the clinical database, rather than the classification entered into the IRT (interactive response technology) system at randomization.

The protocol stated that the primary analysis population was the per protocol population set (PPS), which excluded subjects with major protocol deviations. The per protocol population excluded subjects who had deviations in the inclusion/exclusion criteria that affect efficacy outcomes, missing PASI scores at baseline or Week 12, more than 4 doses

(or 2 consecutive doses) missed, the Week 12 visit more than 6 days from the planned visit day, or had taken prohibited medications that may impact efficacy. Missing data was not imputed for the per protocol population (except that dropouts due to unsatisfactory therapeutic effect were to be imputed as non-responders). Supportive analyses were also conducted with the full analysis set (FAS; all randomized subjects). Missing response data in the full analysis set was to be imputed as non-response.

The key secondary endpoint was the percent change in PASI. The protocol proposed two analyses. One analysis used a mixed-effect model repeated measures (MMRM) analysis during Treatment Period 1. The model fit factors for treatment group, weight classification, and prior systemic therapy classification, and a covariate for baseline PASI score. The model used an unstructured covariance matrix. A 95% confidence interval for the difference in adjusted means was calculated. A second analysis calculated the average treatment effect for each subject during Treatment Period 1 and then analyzed the subject mean values with ANCOVA with terms for treatment group, body weight classification, prior systemic therapy classification, and baseline PASI as a covariate. Limited details for these analyses were included in the protocol.

Additional secondary endpoints included percent change in PASI and observed PASI scores at each visit (Weeks 2, 4, 8, and 12), IGA response (0 or 1), Dermatology Life Quality Index (DLQI), EuroQol 5-Dimension Health Status Questionnaire (EQ-5D), and Health Assessment Questionnaire-Disability Index (HAQ-DI).

3.2.2 Randomization in Treatment Period 2

Subjects who achieved at least PASI 50 at Week 12 in Treatment Period 1 were to be re-randomized for Treatment Period 2 to either maintain the same treatment or switch between treatments. No stratification factors were used for the re-randomization. The re-randomization ratio for Treatment Period 2 was modified in amendments to the protocol during the course of the study. In the originally implemented version of protocol for Study 302 (dated February 4, 2013), the re-randomization scheme was 1:1 randomization (same treatment : switching) for subjects from each arm in Treatment Period 1. The first subject was randomized on July 8, 2013. The protocol was amended approximately 2 months later (September 18, 2013). In this amendment, the protocol stated that the re-randomization in the second treatment period would be changed to 3:1 randomization with 75% of subjects randomized to remain on the same treatment and 25% of subjects randomized to switch between treatments. However, it does not appear that the 3:1 randomization scheme was ever implemented. In a subsequent amendment (dated May 8, 2014; approximately 10 months after the first subject was randomized), the protocol changed the re-randomization ratio for the second treatment period to 6:1 (same treatment: switching) in order to get 'as close as possible' to 3:1 randomization overall after complete enrollment.

The randomization numbers for the second treatment period come in two styles: for subjects enrolled early in the study the re-randomization numbers are 7-digit integers, while for subjects enrolled late in the study the re-randomization numbers are 4-digit decimal numbers between 0 and 1. Based on these two groups of numbers, of the 504

subjects re-randomized in Treatment Period 2, 362 subjects (72%) were randomized under the 1:1 randomization scheme (integer re-randomization numbers) and 142 subjects (28%) were randomized under the 6:1 randomization scheme (decimal re-randomization numbers). In the end, approximately 60% of subjects were randomized to maintain the original treatment and 40% of subjects were randomized to switch treatments. See Table 4.

Table 4 – Re-randomization in Treatment Period 2

	Period 1				
	GP2015		EU-Etanercept		
Period 2	Same	Switch	Same	Switch	Total
1:1 randomization	94	92	88	88	362 (72%)
6:1 randomization	58	8	66	10	142 (28%)
Total	152 (30%)	100 (20%)	154 (31%)	98 (19%)	504

Source: reviewer analysis.

3.2.3 Prior Therapy Stratification

For the initial randomization for Treatment Period 1, the randomization was stratified by prior therapy (no prior systemic therapy, any prior systemic therapy including biologic immunomodulation agents but no prior treatment with a TNF antagonist, or prior treatment with a TNF antagonist) and weight (< 90 kg vs. ≥ 90 kg). At the blinded data review meeting, only 7 subjects were stratified into the 'prior treatment with a TNF antagonist' category at randomization. Thus the applicant decided to combine the 'TNF antagonist' category with the 'any prior systemic therapy including biologic immunomodulation agents but no prior treatment with a TNF antagonist' category.

The protocol provided limited details regarding the proposed analyses, but the protocol stated that the primary endpoint would be analyzed with an exact confidence interval. The SAP proposed analyses based on logistic regression that incorporated the prior therapy and weight strata. However, during the blinded review of the data, the applicant determined that the stratification had been incorrectly performed for many subjects and that the stratification classification entered into the IRT did not agree with the data collected for the study. In the database, of the 531 randomized subjects, 110 subjects were listed as having a protocol deviation related to the stratification (99 subjects were listed as having a protocol deviation related to the assignment of the prior therapy stratum only, 7 subjects were listed as having a protocol deviation related to assignment to the weight stratum only, and 4 subjects who had deviations related to both stratum variables). Due to the large number of stratification deviations identified at the blinded data review meeting, the applicant finalized the SAP to state that the 'actual' value for the stratification variables (as determined by medical review of the data recorded on the CRF) would be used in the analyses rather than the value entered in to the IRT.

A significant number of subjects (191 or 36%) were reclassified after medical review of the prior systemic therapies data for the Week 12 study report, including additional subjects who had not been flagged as protocol deviations related to stratification. Of note, the applicant classified subjects who received UVA or UVB phototherapy, but no

systemic treatments for psoriasis in the 'Any' prior therapy category after reclassification. The applicant reclassified 109 subjects (21%) from 'Any' or 'TNF' to 'No', and 82 subjects (15%) from 'No' to 'Any'. See Table 5.

Table 5 – Changes from the Prior Therapy Randomization Stratification to the Week 12 Study Report Analysis Classification

	Week 12 Report Classification		
Randomization (IRT) Classification	Any prior therapy (including TNF)	No prior therapy	Total
Any prior therapy (excluding TNF)	128	108	236
No prior therapy	82	206	288
TNF prior therapy	6	1	7
Total	216	315	531

Source: Reviewer analysis

The initial BLA submission contained the data for Study 302 after the Week 12 database lock. More than 3 months into the review cycle (on November 9, 2015), the applicant submitted an updated study report containing data through Week 30. The November 9, 2015 submission contains two study report documents: an 'original' Week 30 report and an 'amended' Week 30 report. The original Week 30 report contains the same Week 12 analyses as the Week 12 report, and just adds additional analyses for the data collected between Week 12 and Week 30. However, the amended Week 30 report includes Week 12 results that are different from those presented in the Week 12 report. The primary reason the results in the Week 30 amendment differ is that the applicant unlocked the database and defined a newly revised version of the 'actual' prior therapy stratum which was then used in the efficacy analyses. The Week 30 definition of prior therapy classification differs from the Week 12 report definition primarily in handling of subjects who received UVA or UVB phototherapy, but no systemic treatments for psoriasis. The Week 12 analysis considered 49 subjects who had had phototherapy but no systemic treatment in the 'Any' systemic therapy stratum, while the Week 30 analysis reclassified these subjects to the 'No' systemic therapy stratum. Of these 49 subjects who had only phototherapy, 40 had been classified by the investigators at randomization as being in the 'No' prior systemic therapy stratum, while 9 had been classified at randomization into the 'Any' prior therapy stratum. Eight other subjects were reclassified for the Week 30 analysis (5 from 'Any' to 'No' and 3 from 'No' to 'Any') for various reasons (vitamins, analgesics, and antihistamines were no longer considered systemic therapies for psoriasis, while photochemotherapy and oral prednisolone were).

The applicant did not provide a clear rationale for why the prior therapy classification issue was re-considered for the Week 30 amendment and the analyses for Week 12 results modified (the 'original' Week 30 report does not include any of these changes, and they were only applied to the amended report). The only rationale provided by the applicant regarding why the prior therapy classification variable was redefined post-hoc for the Week 30 study report amendment is that "it was identified that some patients were incorrectly classified regarding the stratification variable 'prior systemic therapy', which might affect the efficacy analyses already presented in the CSR." (pg 13 of

\\cdsesub1\evsprod\bla761042\0003\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body-2.pdf).

In addition to the changes to the 'actual' prior therapy stratification variable, the amended Week 30 study report included a change to the reported Week 12 PASI score for one subject, after the treating site updated the subject's information. Subject 4202/012 (randomized to etanercept) had their Week 12 PASI score changed from 0.4 to 1.0. The corresponding percent change in PASI was changed from -98.01% to -95.02%. The PASI 75 classification was unchanged (success), so this data change has no impact on the primary analysis, but it has a very minor impact on the percent change in PASI endpoint.

Reviewer Comment
The applicant has not provided a rationale for re-opening the database for Study 302 and changing the value of a key analysis variable after the data had already been analyzed, finalized in a study report, and submitted to the FDA. This review will consider the analyses submitted with the application in the Week 12 study report to be the applicant's primary analysis. The analyses presented in the Week 30 study report amendment will only be discussed in terms of sensitivity analyses. Post-hoc analyses cannot be considered reliable, and any analysis based on the Week 30 amendment prior therapy classification cannot be considered pre-specified and could be impacted by bias.

After the applicant's re-review of the data for the Week 30 study report amendment, relative to the Week 12 study report, 54 subjects (10%) were reclassified from 'Any' to 'No', while 3 subjects (0.6%) were reclassified from 'No' to 'Any'. The majority of subjects who were classified as 'Any' in the Week 12 report and 'No' for the Week 30 amendment had phototherapy without other systemic treatments (49 out the 54 subjects). The remaining 5 subjects had reported use of vitamins, NSAIDS, or antihistamines that the applicant no longer considered as systemic treatments for psoriasis. See Table 6. The 3 subjects who were classified as 'No' in the Week 12 report and 'Any in the Week 30 amendment had been previously treated with photochemotherapy or oral prednisolone.

Table 6 – Changes from the Week 12 Prior Therapy Analysis Classification to the Week 30 Study Report Amendment Analysis Classification

	Week 30 Report Classification		
Week 12 Report Classification	Any prior therapy (including TNF)	No prior therapy	Total
Any prior therapy (including TNF)	162	54	216
No prior therapy	3	312	315
Total	165	366	531

Source: Reviewer analysis

3.2.4 Subject Disposition

Study 302 randomized 531 subjects: 264 to GP2015 and 267 to EU-etanercept. Three percent of GP2015 and 4.5% of EU-etanercept subjects discontinued during Treatment

Period 1. The most common reason for study discontinuation was 'subject decision.' A greater number of EU-etanercept subjects than GP2015 subjects (1.9% vs. 0.8%) discontinued due to patient decision.

Table 7 – Disposition of Subjects in Treatment Period 1 (Study 302)

	GP2015	EU-etanercept
Subjects Randomized	264	267
Discontinued Treatment Period 1	8 (3.0%)	12 (4.5%)
Adverse event	4 (1.5%)	3 (1.1%)
Death	--	1 (0.4%)
Lost to follow-up	1 (0.4%)	--
Non-compliance with study treatment	--	1 (0.4%)
Physician decision	--	1 (0.4%)
Protocol deviation	1 (0.4%)	--
Subject decision	2 (0.8%)	5 (1.9%)
Injection site reaction	--	1 (0.4%)

Source: pg 81 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf.

Most subjects who completed Treatment Period 1 continued on to Treatment Period 2 (Weeks 12 through 30). Five subjects (2 on GP2015 and 3 on EU-etanercept) were not randomized into Treatment Period 2 because they did not achieve at least PASI 50, and 9 subjects (4 GP2015 and 5 EU-etanercept) did not enter Treatment Period 2 for other reasons. Note, however, that 3 subjects entered Treatment Period 2 even though they did not achieve at least PASI 50. The most common reasons for discontinuation in Treatment Period 2 were subject decision and adverse events. See Table 8.

Table 8 – Disposition of Subjects in Treatment Period 2 (Study 302)

	Treatment in Period 1			
	GP2015 N=264		EU-etanercept N=267	
Completed Treatment Period 1	256		255	
Treated in Treatment Period 2	250		247	
Not treated in Treatment Period 2	6		8	
<PASI 50	2		3	
Study termination[a]	--		2	
Other reasons	4		3	
	Treatment Sequence in Period 2			
	G/G/G	E/G/E	E/E/E	G/E/G
Subjects treated	150	100	151	96
Completed Treatment Period 2	143	96	142	91
Discontinued Treatment Period 2	7 (4.7%)	4 (4.0%)	9 (6.0%)	5 (5.2%)
Adverse Event	1 (0.7%)	--	2 (1.3%)	4 (4.2%)
Physician decision	1 (0.7%)	--	--	--
Protocol deviation	--	--	1 (0.6%)	--
Lack of efficacy	1 (0.7%)	1 (1.0%)	--	--
Subject decision	3 (2.0%)	1 (1.0%)	4 (2.6%)	1 (1.0%)
Study termination[a]	1 (0.7%)	2 (2.0%)	2 (1.3%)	--

G = GP2015; E =EU-etanercept
[a] Center terminated due to war situation in Ukraine
Source: pg 216-217 of \\cdsesub1\evsprod\bla761042\0003\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body-1.pdf and reviewer analysis

Approximately 10% of subjects on each treatment arm were excluded from the per protocol population. The reasons for exclusion were reasonably balanced across the treatment arms. The most common reasons for exclusion were not completing Treatment Period 1 and having the visit more than 6 days from the planned Week 12 visit day (visit window exclusion). See Table 9.

Table 9 – Per Protocol Population Exclusions

	GP2015 N=264	EU-etanercept N=267
Subjects excluded from Per Protocol Population	25 (9.5%)	26 (9.7%)
Compliance to study drug administration	3 (1.1%)	2 (0.7%)
Exclusion criteria	5 (1.9%)	2 (0.7%)
Inclusion criteria	2 (0.8%)	3 (1.1%)
Prohibited medication	3 (1.1%)	5 (1.9%)
Visit window	6 (2.3%)	7 (2.6%)
Did not complete Treatment Period 1	10 (3.8%)	12 (4.5%)

Note: Subjects may have had more than one reason for exclusion
Source: reviewer analysis.

3.2.5 Baseline Characteristics

The baseline demographics were generally balanced across the treatment groups in Study 302. The mean age was about 42 years, with about 5% of subjects age 65 and older. The majority of subjects were male (62%) and white (99%). See Table 10.

Table 10 – Baseline Demographics

	GP2015 N=264	EU-etanercept N=267
Age (years)		
Mean	42.1	42.7
Range	18 - 78	19-75
18 to 64 years	249 (93.3%)	255 (96.6%)
65 + years	18 (6.7%)	9 (3.4%)
Gender		
Female	107 (40.5%)	95 (35.6%)
Male	157 (59.5%)	172 (64.4%)
Race		
White	263 (99.6%)	264 (98.9%)
Black	1 (0.4%)	--
Asian	--	1 (0.4%)
Unknown	--	1 (0.4%)
Weight (kg)		
Mean (SD)	86.3 (21.12)	85.9 (18.72)
<90 kg	160 (60.6%)	161 (60.3%)
≥90 kg	104 (39.4%)	106 (39.7%)

Source: pg 85-86 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf and reviewer analysis.

To be enrolled in the study, subjects were to have clinically stable chronic plaque psoriasis involving at least 10% body surface area (BSA), PASI ≥ 10, and Investigator's Global Assessment (IGA) ≥3. At baseline, subjects had a mean PASI score of 22.5 and 31% BSA. Approximately 70% had an IGA score of moderate. More than half of subjects reported having no prior systemic therapy. The baseline disease characteristics were balanced across treatment arms. See Table 11.

Table 11 – Baseline Disease Characteristics

	GP2015 N=264	EU-etanercept N=267
PASI		
Mean (SD)	22.5 (8.93)	22.5 (9.52)
Range	9.4 – 55.2	10.2 – 55.2
BSA		
Mean (SD)	30.5 (13.77)	30.9 (14.8)
Range	9.5 - 77	8.7 - 76
IGA		
Mild	--	1 (0.4%)
Moderate	191 (72.3%)	186 (69.7%)
Severe	73 (27.7%)	80 (30.0%)
Prior systemic therapy (randomization strata)		
No	143 (54.2%)	145 (54.3%)
Any (except TNF)	117 (44.3%)	119 (44.6%)
TNF	4 (1.5%)	3 (1.1%)
Prior systemic therapy (actual use)[a]		
No	153 (58.0%)	162 (60.7%)
Any (including TNF)	111 (42.0%)	105 (39.3%)

[a] From the Week 12 Study report
Source: pg 88-89 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf and reviewer analysis.

3.2.6 Primary Efficacy Endpoint

The primary efficacy endpoint was PASI 75 at Week 12. The protocol specified that the primary analysis would be based on an exact 95% confidence interval for the difference in response rates between GP2015 and EU-etanercept using the per protocol population. The analysis based on the full analysis population was to be supportive. The SAP modified the proposed analysis and specified that the data would be analyzed with a logistic regression model with terms for treatment group, body weight stratum, and prior systemic therapy stratum. The confidence interval would be calculated using the delta method. The details of the delta method calculation are presented in the Appendix. Because the blinded data review identified that many subjects' stratification values for prior therapies and weight did not match the data recorded on the CRF, the SAP stated that 'actual' values for both stratification variables would be used in the logistic regression model, rather than the values used to stratify the randomization. In addition, because so few subjects had previously used other TNF-alpha inhibitors, the SAP stated that the subjects who had previously received TNF-alpha inhibitors would be grouped with the subjects who had previously received other prior therapies.

Based on FDA advice that a 90% confidence interval would be acceptable for a U.S. regulatory submission, the applicant also presented the 90% confidence level results. The similarity margin in each case was ±18%.

Missing data was not imputed for the per protocol population (except that dropouts due to unsatisfactory therapeutic effect were to be imputed as non-responders). However, no subjects dropped out during Treatment Period 1 due to unsatisfactory therapeutic effect. The analysis based on the full analysis population (all randomized subjects) was conducted as a supportive analysis. Missing response data in the full analysis set was imputed as non-response.

The results of the applicant's per protocol analysis and the full analysis set analysis are similar, and the 90% and 95% confidence intervals based on both the per protocol set and the full analysis set were within the pre-specified margin of ±18%. The applicant's results from the Week 12 study report are presented in Table 12. Although the per protocol and full analysis set analyses are generally both considered for similarity and non-inferiority analyses, the full analysis set preserves the randomization and is generally preferred as the primary analysis.

Table 12 – Applicant's PASI 75 Response Rates (Primary Endpoint) [Week 12 Study Report]

	GP2015	EU-etanercept
Per Protocol Population	N=239	N=241
Adjusted response rate	73.3%	75.8%
Difference (GP2015-etanercept)	-2.5%	
90% Confidence interval	(-8.8%, 3.9%)	
95% Confidence interval	(-10.0%, 5.1%)	
Full Analysis Set	N=264	N=267
Adjusted response rate	70.3%	71.7%
Difference (GP2015-etanercept)	-1.4%	
90% Confidence interval	(-7.7%, 5.0%)	
95% Confidence interval	(-9.0%, 6.3%)	

Note: Confidence intervals computed using a logistic regression model with terms for treatment group, 'actual' body weight stratum, and 'actual' prior systemic therapy classification
Source: pg 329 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf and pg 529 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5353-rep-analys-data-more-one-stud\all-studies\statistical-overview-report.pdf

At the blinded data review meeting for the 12-week database, the applicant noted that the stratification classification at randomization did not match the data recorded in the CRF for many subjects, particularly with regard to the prior therapies. Thus the applicant defined 'actual' values for the stratification variables for use in the analyses. At the time of the 30-week data analysis, the applicant re-opened the issue of how to classify the prior therapies. The decision to re-define the 'actual' prior therapy classification for the 30-week study report, appears to be driven by a rethinking of the handling of subjects who had received prior phototherapy (or antihistamines, analgesics, or vitamins), but not any other systemic drugs for psoriasis. For the Week 30 report amendment, these subjects were reclassified as *not* having had prior systemic therapy. Table 13 presents the point estimates and 90% confidence intervals using three sets of 'stratification' factors in

the logistic regression analysis: the 'actual' prior therapy and weight classifications used in the Week 12 report, the 'actual' prior therapy and weight classifications used in the Week 30 amendment, and the 'randomization' prior therapy and weight classifications used in the randomization. Although the different definitions lead to point estimates (adjusted for model factors) that differ slightly, all analyses lead to 90% confidence intervals within the range of ±9%. The largest-in-magnitude bound (8.8%) corresponds the analysis that the applicant considers the primary analysis in the Week 12 report (actual prior therapy and weight using the per protocol population). Thus, all of the variations among the various stratification variable definitions lead to the conclusion of similarity for this endpoint.

Table 13 –Analyses using Various Stratification Variable Definitions

	Pop.	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
Week 12 Report 'actual' PT and weight	FAS	70.3%	71.7%	-1.4%	(-7.7%, 5.0%)
	PPS	73.3%	75.8%	-2.5%	(-8.8%, 3.9%)
Week 30 Amend. 'actual' PT and weight	FAS	70.4%	71.6%	-1.2%	(-7.5%, 5.2%)
	PPS	73.4%	75.7%	-2.3%	(-8.6%, 4.1%)
Randomization strata for PT and actual weight	FAS	70.4%	71.6%	-1.1%	(-7.5%, 5.3%)
	PPS	73.5%	75.7%	-2.2%	(-8.6%, 4.2%)
Randomization strata for both PT and weight	FAS	70.5%	71.5%	-1.1%	(-7.5%, 5.3%)
	PPS	73.6%	75.6%	-2.0%	(-8.4%, 4.4%)

PT = prior therapy, FAS = full analysis set, PPS = per protocol set.
Note: Confidence intervals computed using a logistic regression model with terms for treatment group, body weight classification, and prior systemic therapy classification
Source: reviewer analysis.

It is not clear why so many of the stratification values entered by the investigators into the IRT system do not match the data recorded about prior therapies, and whether there was lack of clarity in the options or if it was due to poor system design. The applicant has also proposed two different versions of the 'actual' prior therapy classification: one in the Week 12 report and one in the Week 30 report amendment. It is also not clear why the applicant decided to re-open the issue for the Week 30 amendment, when unlocking the database and re-defining a key variable after the data have been thoroughly analyzed could introduce bias into the results, or give the appearance of seeking a more 'favorable' set of results. In fact, the original analysis specified in the protocol (as opposed to the SAP) did not rely on the stratification factors: the originally proposed analysis was an exact confidence interval. The exact binomial 90% confidence intervals are presented in Table 14. The unadjusted estimates for the PASI 75 response rates are very similar to the adjusted estimates from the logistic regression model, and the confidence interval is slightly wider. However both the FAS and PPS confidence intervals are within a range of ±10% and thus within the pre-specified 18% similarity margin. Because of concerns that the analysis based on the logistic regression model was only defined in the SAP and not in the protocol, and that the stratification variables were modified multiple times for the applicant's study reports (including the definition used for the Week 30 report amendment that was defined after the data had been fully analyzed), this reviewer

recommends considering the protocol-specified exact confidence intervals as the primary analysis.

Table 14 – Exact Confidence Intervals for the Risk Difference of PASI 75 Response Rates

Population	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
FAS	186/264 70.5%	191/267 71.5%	-1.1%	(-8.3%, 6.0%)
PPS	176/239 73.6%	182/241 75.5%	-1.9%	(-9.4%, 5.6%)

FAS = full analysis set, PPS = per protocol set
Source: reviewer analysis.

3.2.7 Missing Data Handling for the Primary Endpoint

Missing data was not imputed for the per protocol population. For the primary endpoint of PASI 75 in the full analysis set, PASI 75 response missing data was imputed as non-response. The applicant did not propose any alternate methods for handling missing data as sensitivity analyses. Thus, to assess whether the handling of missing data had any impact on the results, this reviewer conducted sensitivity analyses where all the subjects with missing data on one arm were treated as failures and all of the subjects with missing data on the other arm were treated as successes. Twenty subjects had missing Week 12 PASI results, 8 on the GP2015 arm and 12 on the EU-etanercept arm. One analysis was conducted where the GP2015 subjects with missing data were treated as failures, and the EU-etanercept subjects with missing data were treated as successes. The second analyses treated GP2015 subjects with missing data as successes, and the EU-etanercept subjects with missing data as failures. The exact confidence interval results are presented in Table 15. Because the PASI 75 response rate (based on observed cases) on the GP2015 arm was lower than on the EU-etanercept arm, and fewer GP2015 subjects had missing data, the sensitivity analysis where the GP2015 subjects were treated as failures and the EU-etanercept subjects were treated as success leads to the larger in magnitude point estimate for the treatment difference (-5.6%). The corresponding confidence intervals using this extreme method of imputation for missing data still fall within the similarity margin of ± 18%. Thus, the conclusion of similarity for this study remains the same even when subjects with missing data on the two arms are handled in opposite ways.

Table 15 – Sensitivity Analyses for PASI 75 Response Rates (Exact Confidence Intervals; FAS)

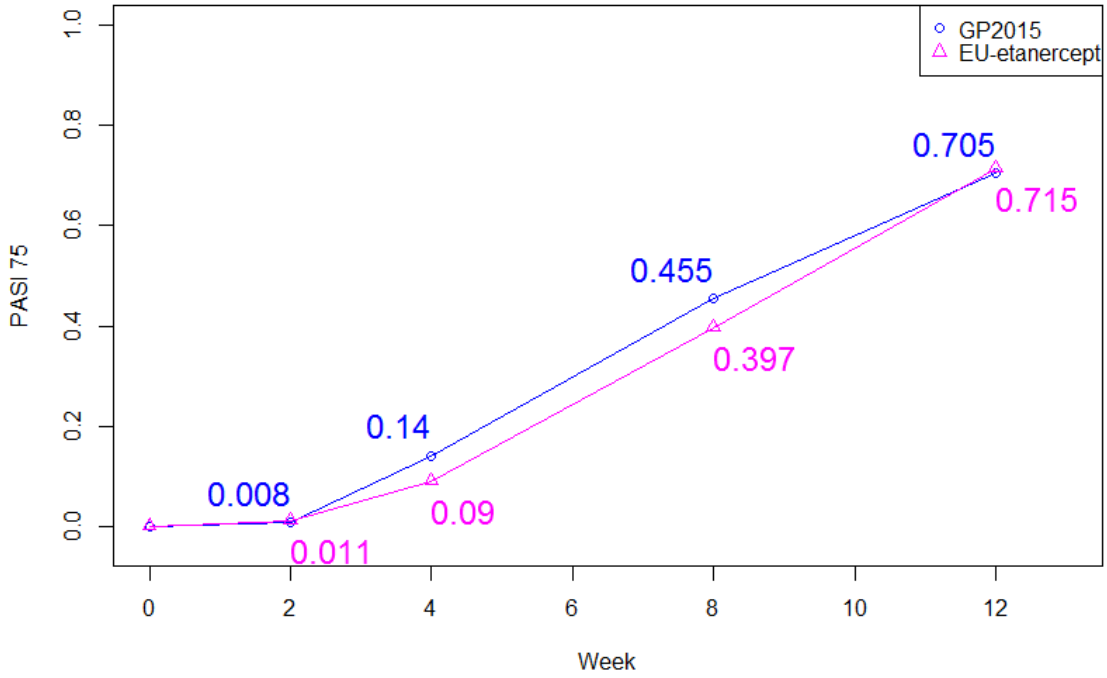
Population	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
GP2015 Missing as Failure/ EU-etanercept Missing as Success	186/264 70.5%	203/267 76.0%	-5.6%	(-12.8%, 1.5%)
GP2015 Missing as Success/ EU-etanercept Missing as Failure	194/264 73.5%	191/267 71.5%	2.0%	(-5.3%, 9.0%)

Source: reviewer analysis.

3.2.8 PASI 75 Response over Time

In Treatment Period 1, PASI 75 response was evaluated at Week 2, 4, 8, and 12. The response rates over time were similar for subjects treated with GP2015 and EU-etanercept. See Figure 2.

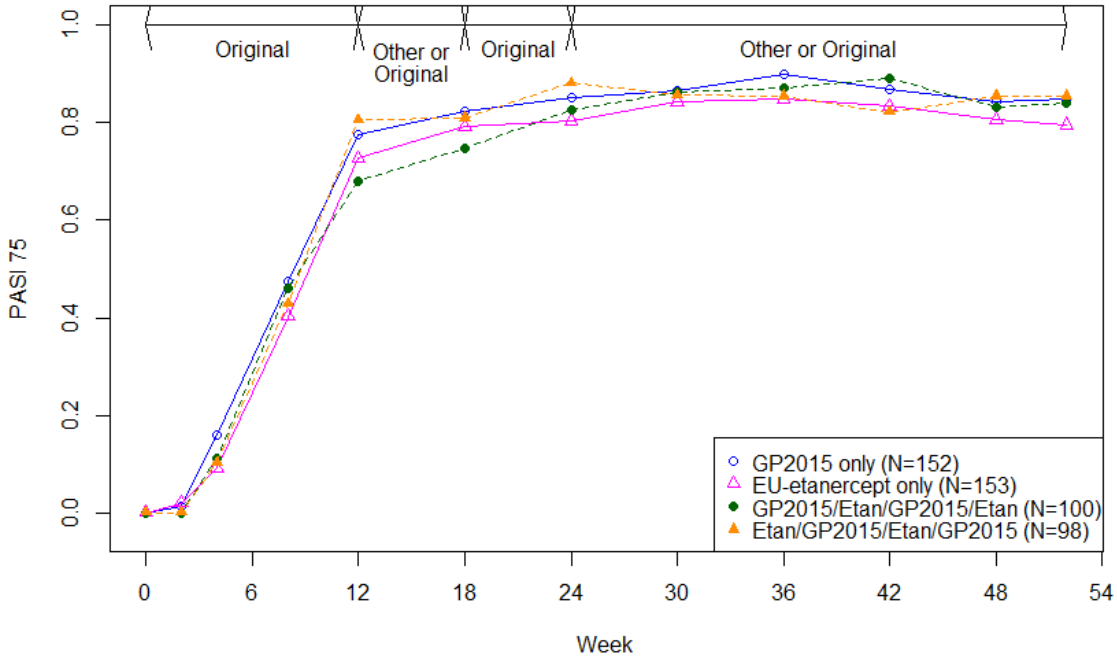
Figure 2 – PASI 75 Response Rates in Treatment Period 1 (FAS, Missing as Failure)



Source: reviewer analysis.

At Week 12, subjects with at least a PASI 50 response were randomized to either remain on the original treatment through the end of the study or switch to between treatments. Subjects randomized to the switching arms used the other treatment between Weeks 12 and 18, the original treatment between Week 18 and 24, and the other treatment between Weeks 24 and 52. Note that this data was from the Week 30 database lock, and therefore may contain only partial data from Weeks 36 to 54. The PASI 75 response rates were similar from Weeks 18 to 52 across all four arms (GP2015 only, etanercept only, GP2015/etanercept/GP2015/etanercept, and etanercept/GP2015/etanercept/GP2015). See Figure 3.

Figure 3 - PASI 75 Response Rates in Treatment Period 2 (Subjects Re-randomized in Treatment Period 2, Observed Cases)



Source: reviewer analysis

3.2.9 Stratification Subgroups

The randomization was stratified by weight and prior systemic therapy. As discussed in Section 3.2.3, medical review of the data indicated that the stratum selected by the investigator during randomization did not necessarily match the information recorded on the CRF. In addition, the applicant considered certain types of psoriasis therapies (e.g. phototherapy or systemic antihistamines) as prior systemic therapies for psoriasis at the Week 12 database lock, but not as prior systemic therapies at the Week 30 database lock. In addition, the protocol originally stated that unstratified analyses would be conducted, but the SAP stated that stratified analyses would be conducted. Thus, this section presents the PASI 75 response rates by the various prior therapy and weight classifications.

For the three prior therapy classifications used by the applicant (used for the randomization, used in the Week 12 report, and used in the Week 30 report amendment), the within-subgroup point estimates for the PASI 75 treatment differences ranged from -3.5% to +1.0%, and the corresponding confidence intervals were within the bounds of -15.1% to +10.8%. For the two classification groupings used in the Week 12 report and the Week 30 report amendment, subjects who had had prior therapies tended to have slightly higher response rates than those who did not, while for the randomization classification the response rates in the two groups were similar. However, in all three cases the treatment differences were similar. See Table 16.

Table 16 – Week 12 PASI 75 Response Rates by Prior Therapy Classification

	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
Randomization Stratum				
Any	83/121 68.6%	88/122 72.1%	-3.5%	(13.8%, 7.3%)
No	103/143 72.0%	103/145 71.0%	1.0%	(-8.7%, 10.7%)
Actual (Week 12 Report)				
Any	84/111 75.7%	80/105 76.2%	-0.5%	(-11.7%, 10.8%)
No	102/153 66.7%	111/162 68.5%	-1.9%	(-11.2%, 7.5%)
Actual (Week 30 Report)				
Any	61/82 74.4%	64/83 77.1%	-2.7%	(-15.1%, 10.4%)
No	125/182 68.7%	127/184 69.0%	-0.3%	(-8.9%, 8.3%)

Source: reviewer analysis

Similarly, the applicant stratified the randomization by weight (<90 kg, ≥ 90 kg) and defined an 'actual' weight classification for the subjects where the weight stratum classification did not agree with the recorded weight at baseline (11 subjects). The subgroup results from the two classifications are similar, though the subjects in the lighter stratum had higher response rates than those in the heavier stratum. See Table 17.

Table 17 – Week 12 PASI 75 Response Rates by Weight Classification

	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
Randomization Stratum				
<90 kg	122/162 75.3%	127/164 77.4%	-2.1%	(-11.2%, 7.0%)
≥ 90 kg	64/102 62.8%	64/103 62.1%	0.6%	(-10.6%, 12.4%)
Actual				
<90 kg	120/160 75.0%	126/161 78.3%	-3.3%	(-12.3%, 6.1%)
≥ 90 kg	66/104 63.5%	65/106 61.3%	2.1%	(-9.0%, 13.7%)

Source: reviewer analysis

3.2.10 Secondary Endpoint—Percent Change in PASI

The key secondary endpoint was the percent change in PASI. The protocol proposed two analyses. Both analyses evaluated the average percent change in PASI throughout Treatment Period 1, using the observations from Weeks 2, 4, 8, and 12. One analysis used

a mixed-effect model repeated measures (MMRM) analysis during Treatment Period 1. A second analysis calculated the average treatment effect (ATE) for each subject during Treatment Period 1 and then analyzed the computed subject mean values. Both analyses used similarity margins of ± 15% and 95% confidence intervals. Missing data were not imputed for either analysis. This review will present the 95% confidence intervals as specified in the protocol.

The MMRM model fit factors for treatment group, weight stratum, and prior systemic therapy stratum, and a covariate for baseline PASI score, including visit-by-treatment interaction for the repeated measures analysis. The model used an unstructured covariance matrix. The ATE model analyzed the subject mean values with ANCOVA with terms for treatment group, body weight classification, prior systemic therapy classification, and baseline PASI as a covariate. No interaction terms were included in the model.

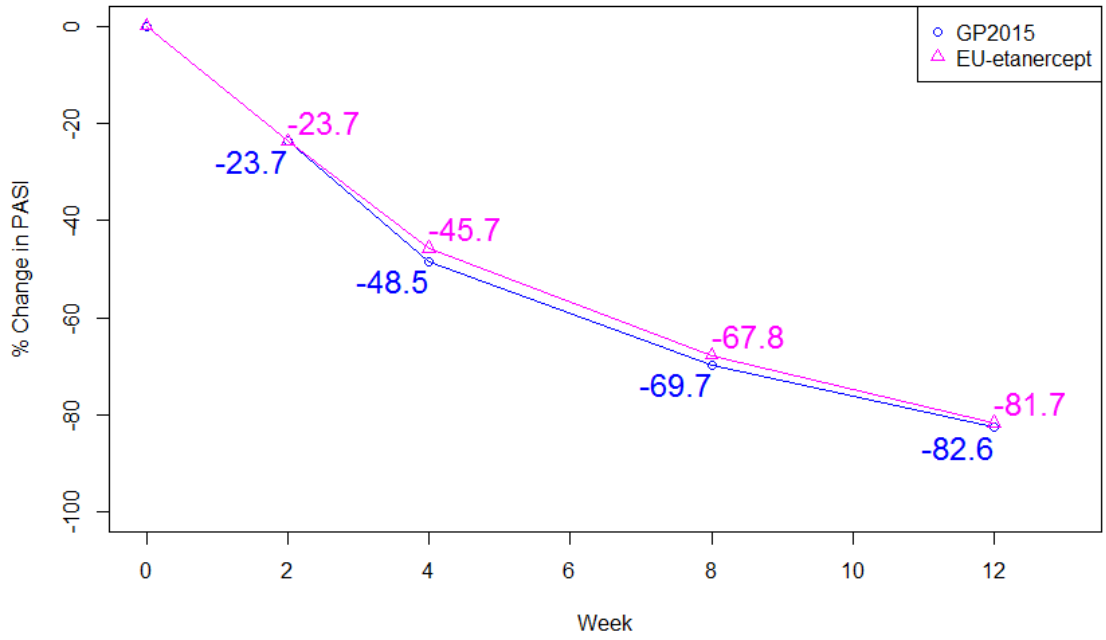
Both the repeated measures analysis and the analysis of the average treatment effect yielded similar results for the average percent change in PASI across Weeks 2, 4, 8, and 12. Point estimates for the two analyses in both the FAS and PPS populations for both treatments ranged from 50 to 56% with treatment differences ranging from -0.57% to 2.05%. All confidence intervals were within the pre-specified margin of 15%. See Table 18. The mean percent change in PASI values by visit (FAS, observed cases) are presented in Figure 4.

Table 18 – Average Percent Change in PASI during Treatment Period 1

	GP2015	EU-etanercept	Difference	95% Conf. Int.
PPS	N=239	N=241		
MMRM	-55.89	-55.32	-0.57	(-3.41, 2,26)
ATE	-52.84	-52.05	-0.78	(-3.51, 1.94)
FAS	N=264	N=266		
MMRM	-55.84	-54.29	-1.55	(-4.32, 1.22)
ATE	-52.18	-50.12	-2.05	(-4.88, 0.77)

PPS = Per protocol set; FAS = Full analysis set; MMRM = mixed-effect model repeated measurement; ATE = average treatment effect
Source: pg 612, 617, 621 623 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf

Figure 4 – Percent Change in PASI by Visit during Treatment Period 1 (FAS, Observed Cases)



Source: reviewer analysis.

Although the key secondary endpoints assessed the average treatment effect across Treatment Period 1, the supportive endpoint of percent change in PASI at Week 12, which is related to the primary endpoint of PASI 75 is also of interest. The applicant did not specify an analysis method for percent change in PASI at individual timepoints, except to present point estimates. The applicant also did not specify any method for handling missing data for this endpoint. Therefore this reviewer computed 90% confidence intervals to be consistent with the primary analysis. As no method of handling missing data was specified, both observed cases and results for relatively extreme differential imputation are presented (imputing missing values as 0% improvement on one arm and 100% improvement on the other). Confidence intervals are computed using an ANOVA model with baseline PASI score as a covariate. The estimated treatment difference for the observed cases analysis (FAS) is -0.93%. The results in the per protocol population are similar. See Table 15. While these differential imputations shift the point estimates for the treatment differences by 3 to 4%, the 90% confidence intervals remain relatively narrow—within ±8%. Thus the analyses of percent change in PASI outcomes for GP2015 and EU-etanercept are similar for the MMRM, ATE and Week 12 analyses, and support the findings of the primary analysis.

Table 19 – Sensitivity Analyses for Percent change in PASI at Week 12

Population	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.[a]
PPS	-82.97	-82.21	-0.76	(-2.86, 1.34)
FAS/Observed cases	-82.59	-81.66	-0.93	(-3.03, 1.17)
FAS/GP2015 missing as 0/ EU-etanercept missing as -100	-78.31	-80.72	2.40	(-0.11, 4.91)
FAS/GP2015 missing as -100/ EU-etanercept missing as 0	-83.11	-77.99	-5.12	(-7.79, -2.45)

[a] ANOVA model with baseline as covariate.
Source: reviewer analysis.

3.2.11 Secondary Endpoint – Investigator's Global Assessment

Another supportive endpoint was the Investigator's Global Assessment (IGA) which evaluated subjects on a 5-point scale from 0 = clear to 4 = severe. Success on the IGA was defined as a score of clear or almost clear with at least two grades reduction from baseline. Although the applicant did not impute missing data for the IGA success endpoint, for consistency with the PASI 75 analyses, this reviewer imputed missing data as non-response. For the various analyses, the treatment difference on the IGA success endpoint is in the range of 3.1% to 3.8%, and the confidence intervals are within the boundary of ±12%. See Table 20. Whereas the point estimates for GP2015 were slightly lower than EU-etanercept for the PASI 75 endpoint at Week 12, for the IGA success at Week 12 endpoint, the point estimates for GP2015 are slightly higher than those for EU-etanercept. The results of the IGA success analysis are consistent with the results for the primary endpoint.

Table 20 – IGA Success at Week 12 (FAS)

Population	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
PPS	142/239 59.4%	134/241 55.6%	3.8%	(-3.7%, 11.4%)
FAS/Observed cases	149/256 58.2%	141/256 55.1%	3.1%	(-4.3%, 10.6%)
FAS/missing as failure	149/264 56.4%	141/267 52.8%	3.6%	(-3.6%, 10.9%)

Source: reviewer analysis.

3.2.12 Historical Etanercept Studies

Study 302 was a comparative clinical study of GP2015 and EU-etanercept; it did not include a placebo. Historical studies of etanercept (Leonardi (2003) and Papp (2005)) had Week 12 PASI 75 response rates for etanercept 50 mg twice weekly of approximately 49%. See Table 21. In contrast, in Study 302 the Week 12 PASI 75 response rate for EU-etanercept and GP2015 was approximately 70%. The reasons for the higher response rates in Study 302 relative to the historical studies are unclear. As

also noted in Table 21, the disease-related inclusion criteria were similar across all studies (PASI ≥ 10, BSA ≥ 10%, subjects have had or were candidates for prior phototherapy or systemic therapy; Study 302 also required subjects to have IGA ≥ 3). Study 302 did permit subjects who had prior use of a TNF-α inhibitor which the previous studies did not; but as classified at randomization, only 7 subjects in Study 302 reported using prior TNF- α inhibitors. The other major difference between Study 302 and the historical etanercept studies was location: the previous etanercept studies were conducted in the US, Canada, and Western Europe, while Study 302 was conducted in Europe and South Africa, with most centers in Eastern Europe.

Table 21 – Study Characteristics of Historical Etanercept Studies

	Leonardi (2003)	Papp (2005)	GP15-302
Selected inclusion criteria	PASI ≥ 10 BSA ≥ 10%	PASI ≥ 10 BSA ≥ 10%	PASI ≥ 10 BSA ≥ 10% IGA ≥ 3
Previous phototherapy or systemic therapy (or candidate for such therapy)?	Yes	Yes	Yes
Prior therapy restrictions	No etanercept, TNF-α, or biologic	No etanercept or TNF-α	No etanercept
Mild to moderate potency topical steroids for scalp, axilla, or groin permitted?	Yes	Yes	Screening period only
Region	US	US, Canada, Western Europe	Europe, South Africa
Baseline characteristics	BSA = 29.9%[a] PASI = 18.4[a]	BSA = 25.0%[b] PASI = 16.1[b]	BSA = 30.9%[a] / 28.8%[b] PASI = 22.5[a] / 20.0[b]
PASI 75 at Week 12 Etanercept (50 mg BIW) Etanercept (25 mg BIW) Placebo	 49% [N=164] 34% [N=162] 4% [N=166]	 49% [N=194] 34% [N=196] 3% [N=193]	 71.5% [N=267] NA NA

[a] Mean
[b] Median
Note: BIW = twice weekly

The applicant justified the choice of an 18% similarity margin noting that 18% maintains 60% of the observed treatment effects relative to placebo (45-46%) reported in Leonardi (2003) and Papp (2005). The applicant did not provide a rationale for the upper boundary. However, it is not clear that 60% retention of the observed treatment effect is a meaningful rationale for a particular margin. Another way of evaluating the relationship between a proposed margin and sample size is to consider the confidence interval widths for the expected response rate and sample size. Under the design characteristics used by the applicant (proposed sample size of approximately 546 subjects with an expected PASI 75 response rate of 49%), we can see from Table 22 that the 90% confidence interval would be the point estimate for the treatment difference plus or minus

approximately 7%. Thus the observed point estimate for the treatment difference could be approximately ±10% under these design assumptions and still be within the pre-specified margin of 18%. The Agency agreed with an 18% margin for the proposed design of this comparative clinical study at the December 19, 2012 Type 2 Biosimilar Biologic Product Development meeting.

Table 22 – Expected Confidence Interval Widths for Various Sample Sizes

Total Sample size	N=300	N=400	N=500	N=600
90% CI widths for response rates = 0.5	± 9.5%	± 8.2%	±7.4%	± 6.7%
Largest observed difference within 18% margin	8.5%	9.8%	10.6%	11.3%

Note: Confidence interval widths calculated using the normal approximation, equal response rates of 0.5 for each treatment group, and equal sample sizes for each treatment group (1/2 of the total sample size).
Source: reviewer analysis.

3.3 Evaluation of Safety

3.3.1 Extent of Exposure

The extent of exposure to study drug was similar for subjects randomized to GP2015 and EU-etanercept with approximately 80 days of study drug exposure on each arm. Approximately 87% of subjects on both arms took all scheduled doses. See Table 23.

Table 23 – Extent of Drug Exposure in Treatment Period 1

	GP2015 N=264	EU-etanercept N=267
Exposure Days		
Mean (SD)	80.6 (9.7)	79.2 (11.6)
Range	4 - 149	1 – 89
Number of Missed Doses		
0	229 (86.7%)	231 (86.5%)
1	17 (6.4%)	14 (5.2%)
2	4 (1.5%)	8 (3.0%)
3	4 (1.5%)	1 (0.4%)
4	3 (1.1%)	1 (0.4%)
5 or more	7 (2.7%)	12 (4.5%)

Source: pg 109-110 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf

3.3.2 Adverse Events

Similar rates of adverse events, serious adverse events, and study discontinuations due to adverse events, or treatment interruptions due to adverse events occurred on the GP2015 and EU-etanercept arms in Treatment Period 1. See Table 24.

Table 24 – Summary of Adverse Events in Treatment Period 1

	GP2015 N=264	EU-etanercept N=267
Any Adverse Events	99 (37.5%)	95 (35.6%)
Serious Adverse Events	4 (1.5%)	3 (1.1%)
Discontinued Study due to AE	5 (1.9%)	3 (1.1%)
Interrupted Treatment due to AE	3 (1.1%)	6 (2.2%)
Deaths	--	1 (0.4%)

Source: pg 112 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf

The preferred terms corresponding to special warnings and precautions in the Enbrel label were identified as adverse events of 'special interest'. These adverse events include infections, neoplasms, and immune system disorders. A slightly greater number of GP2015 subjects than EU-etanercept subjects had adverse events of special interest (3.4% vs. 1.9%). The difference was primarily due to a greater number of GP2015 subjects with neoplasms (5 vs. 1). See Table 25. For additional details on the safety evaluation, refer to the clinical review.

Table 25 – Adverse Events of Special Interest in Treatment Period 1

	GP2015 N=264	EU-etanercept N=267
AEs of Special Interest	9 (3.4%)	5 (1.9%)
Infections and infestations	3 (1.1%)	3 (1.1%)
Oral herpes	1 (0.4%)	2 (0.7%)
Herpes simplex	1 (0.4%)	1 (0.4%)
Tinea infection	1 (0.4%)	--
Neoplasms (benign, malignant, and unspecified)	5 (1.9%)	1 (0.4%)
Skin papilloma	1 (0.4%)	1 (0.4%)
Colon neoplasm	1 (0.4%)	--
Lipoma	1 (0.4%)	--
Malignant melanoma in situ	1 (0.4%)	--
Melanocytic nevus	1 (0.4%)	--
Immune system disorders	1 (0.4%)	--
Hypersensitivity	1 (0.4%)	--
Investigations	1 (0.4%)	--
White blood cell count decreased	1 (0.4%)	--
Skin and subcutaneous tissue disorders	--	1 (0.4%)
Swelling face	--	1 (0.4%)

Source: pg 123 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf

3.3.3 Immunogenicity

Five subjects, all in the EU-etanercept arm, showed confirmed positive binding anti-drug antibodies during Treatment Period 1. One subject was positive at Weeks 2 and 4, and the

other 4 subjects were positive at Week 4. No subjects had anti-drug antibodies at Week 18 after a portion of the subjects were randomized to switch to the other treatment at Week 12. See Table 26.

Table 26 – Anti-drug Antibody Response in Treatment Period 1 and the First Transition in Treatment Period 2.

Treatment Period 1	GP2015 N=264			EU-etanercept N=267		
	Positive	Negative	Missing	Positive	Negative	Missing
Baseline	--	260	4	--	259	8
Week 2	--	250	14	1	253	13
Week 4	--	258	6	5	250	12
Week 8	--	251	13	--	248	19
Week 12	--	251	13	--	250	17
Treatment Period 2	Continued Original Treatment N=301			Switched Treatments N=196		
	Positive	Negative	Missing	Positive	Negative	Missing
Week 18	--	261	40	--	187	9

Source: pg 2942, 2945 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-report-body.pdf

4 Findings in Special/Subgroup Populations

4.1 Gender, Race, Age, and Geographic Region

The PASI 75 response rates were generally consistent across gender and age groups. The study enrolled only 4 non-white subjects so race comparisons are not meaningful. See Table 27.

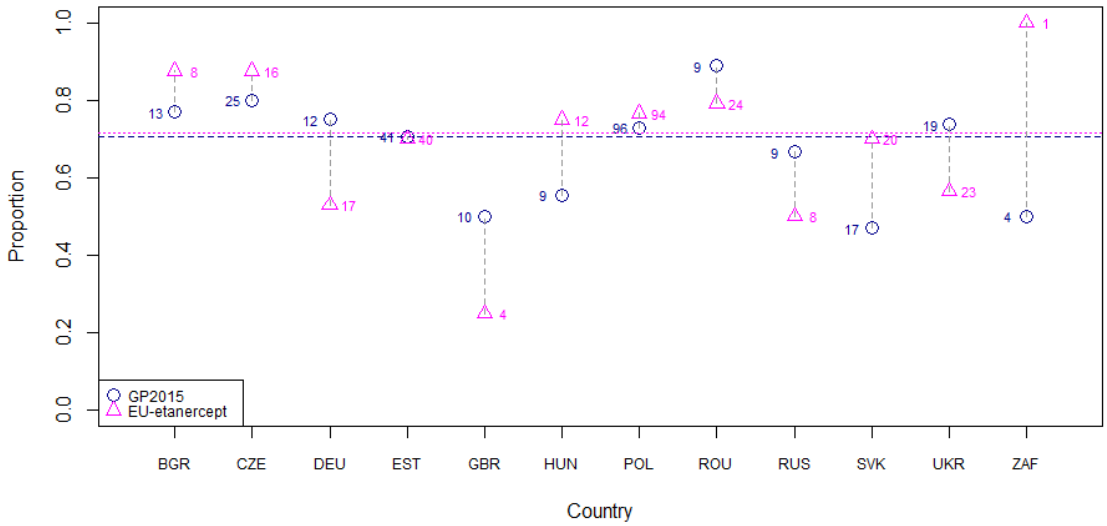
Table 27 – PASI 75 Response Rates by Gender, Race, and Age Group (FAS)

	GP2015 N=264	EU-etanercept N=267
Gender		
Female	79/107 (73.8%)	77/95 (81.1%)
Male	107/157 (68.2%)	114/172 (66.3%)
Race		
White	186/263 (70.7%)	188/264 (71.2%)
Non-White	0/1 (0%)	3/3 (100%)
Age		
< 65	181/254 (71.3%)	179/249 (71.9%)
≥ 65	5/10 (50.0%)	12/18 (66.7%)

Source: reviewer analysis

All subjects were enrolled in Europe, except for 5 subjects who were enrolled in South Africa. Many individual countries enrolled relatively few subjects, but in the two countries with the largest enrollment (Poland and Estonia), the PASI 75 response rates were consistent for GP2015 and EU-etanercept. See Figure 5.

Figure 5 – PASI 75 Response Rates by Country (FAS)



Note: BGR= Bulgaria, CZE = Czech Republic, DEU = Germany, EST = Estonia, GBR = United Kingdom, HUN = Hungary, POL = Poland, ROU = Romania, RUS = Russian Federation, SVK = Slovakia, UKR = Ukraine, ZAF = South Africa.
Source: reviewer analysis

4.2 Other Special/Subgroup Populations

See Section 3.2.9 for discussion of the results by the prior therapy and weight classifications.

5 Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

The randomization in Study 302 was stratified on prior systemic therapy (no prior systemic therapy, any prior systemic therapy including biologic immunomodulation agents but no prior treatment with a TNF antagonist, or prior treatment with a TNF antagonist) and weight (< 90 kg vs. ≥ 90 kg). However, during the blinded review of the data after the Week 12 database lock, the applicant determined that the stratification had been incorrectly performed for many subjects and that the stratification classification entered into the IVRS did not agree with the data collected for the study. The applicant conducted two medical reviews of the data to re-classify subjects based on the data recorded in the case report form. One review was conducted before unblinding the data and finalizing the statistical analysis plan (SAP) for the Week 12 database lock. The second review was conducted at the time the Week 30 report was created. For the Week 12 report, subjects who received UVA or UVB phototherapy, but no systemic treatments for psoriasis were considered to have had prior systemic therapy (that is, included in the 'Any prior therapy' category). At the Week 30 database lock, the applicant removed subjects who had received UVA or UVB phototherapy, but no systemic treatments for

psoriasis from the 'Any prior therapy' category and placed them in the 'No prior therapy', category. Several other subjects were re-classified for other reasons (for example, vitamins, analgesics, and antihistamines were no longer considered systemic therapies for psoriasis). The only rationale provided by the applicant for the reclassification for the Week 30 report was that "it was identified that some patients were incorrectly classified."

The prior therapy classification is relevant to the analyses because although the protocol stated that the PASI 75 endpoint would be evaluated with exact confidence intervals, the statistical analysis plan was finalized to state that the endpoint would be analyzed with a covariate-adjusted confidence interval based on estimates from a logistic regression model with terms for treatment group, body weight classification, and prior therapy classification. Thus the results depend upon which version of the prior therapy classification is used in the model (randomization classification, Week 12 classification, or Week 30 classification). The results using all three versions of the prior therapy classification are similar. However, changing the prior therapy groupings twice, including making changes after the initial study report had been finalized, raises concerns with post-hoc changes to the database. Therefore this reviewer recommends using the analysis for the primary endpoint that is most consistent with the original protocol: exact confidence intervals that do not use the stratification factors.

5.2 Conclusions and Recommendations

Study 302 is a comparative clinical study of GP2015 versus EU-approved etanercept in subjects with moderate to severe psoriasis. The primary endpoint was the proportion of subjects at Week 12 achieving at least a 75% reduction from baseline in PASI (PASI 75). The proportion of subjects achieving PASI 75 at Week 12 was similar on both the GP2015 and EU-etanercept arms (70.5% vs. 71.5% in the full analysis population) and the exact 90% confidence intervals for both the full analysis population and the per protocol population were within the pre-specified margin of ± 18%. See Table 28. Thus the comparative clinical study met its similarity criterion. The results of the supportive endpoints based on the mean percent change in PASI and the Investigator's Global Assessment were consistent with the primary endpoint.

Table 28- Exact Confidence Intervals for the Risk Difference of PASI 75 Response Rates

Population	GP2015 N=264	EU-etanercept N=267	Difference	90% Conf. Int.
FAS	186/264 70.5%	191/267 71.5%	-1.1%	(-8.3%, 6.0%)
PPS	176/239 73.6%	182/241 75.5%	-1.9%	(-9.4%, 5.6%)

FAS = full analysis set, PPS = per protocol set
Source: reviewer analysis

Adverse events were similar on both the GP2015 and EU-etanercept arms. Five subjects developed anti-drug antibodies during the first 12 weeks of treatment. All 5 subjects were on the EU-etanercept arm (N=267). At week 12, 98 subjects were switched from

EU-etanercept to GP2015. No subjects developed anti-drug antibodies 6 weeks after transitioning from EU-etanercept to GP2015.

Appendix

The applicant used the following procedure (excerpted from the SAP) to calculate confidence intervals for the PASI 75 endpoint using the delta method and estimates from the logistic regression model.

> With a data set of n patients, binary response vector $Y = (y_1, y_2, \ldots, y_n)'$, and, a logistic regression model assumes $\text{logit}[P(y_i = 1|x_i)] = \beta' x_i$, where $\text{logit}(p) = \ln[p/(1 - p)]$. If b denote the maximum likelihood estimate (MLE) of β, its estimated variance-covariance matrix is V and $X = (x_1, x_2, \ldots, x_n)'$ denote the covariate matrix.
>
> A new covariate matrix X_t from X by adjusting the column corresponding to treatment assignment will have to be created such that all patients are in the treated group. The vector of estimated probabilities of response to treatment, P_t, will be calculated from X_t and $b[P_t = \text{logit}-1(X_t\, b)]$. Similarly, assuming that each patient is assigned to control the above steps are repeated to get X_c and P_c. The estimated difference in proportions is $d = \Sigma_i(P_{ti} - P_{ci})/n$, where P_{ti} and P_{ci} are the ith elements of P_t and P_c respectively. At is defined as a vector with elements $A_{ti} = P_{ti} (1 - P_{ti})$. Similarly, A_c is defined with $A_{ci} = P_{ci} (1 - P_{ci})$. The delta method is then used to estimate the standard error of the estimation:
>
> $dt = (A_t' X_t)/n$
> $dc = (A_c' X_c)/n$
> $SE(d) = \sqrt{(dtVdt' + dcVdc' - 2dcVdt')}$
> The confidence interval of the estimation is obtained by $d \pm Z(1-\alpha/2)SE(d)$.

Source: pg 24 of \\cdsesub1\evsprod\bla761042\0000\m5\53-clin-stud-rep\535-rep-effic-safety-stud\ra-jia-psa-as-pso\5351-stud-rep-contr\gp15-302\gp15-302-statistical.pdf

References

CL Leonardi, JL Powers, RT Materson, BS Goffe, R Zitnik, A Wang, AB Gottlieb. "Etanercept as monotherapy in patients with psoriasis." *The New England Journal of Medicine*, 2003: 349:2014-2022.

KA Papp, S Tyring, M Lahfa, J Prinz, CEM Griffiths, AM Nakanishi, R Zitnik, PCM van de Kerkof. "A global phase III randomized controlled trial of etanercept in psoriasis: safety, efficacy, and effect of dose reduction." *British Journal of Dermatology*, 2005: 152:1304-1312.

Signatures/Distribution List

Primary Statistical Reviewer: Kathleen Fritsch, Ph.D.
Date: 4/28/2016

Statistical Team Leader: Mohamed Alosh, Ph.D.

cc:
DDDP/Marcus
DDDP/Kettl
DDDP/Chiang
DDDP/White
DPARP/Nikolov
DPARP/Glaser
OBIO/Patrician
DBII/Levin
DBIII/Wilson
DBIII/Alosh
DBIII/Fritsch

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

/s/

KATHLEEN S FRITSCH
04/28/2016

MOHAMED A ALOSH
04/28/2016