

STATISTICAL ANALYSIS PLAN

NCT Number: NCT02736409

Study Title: A Phase 3, Multicenter, Double-blind Extension Study to Evaluate Maintenance of Efficacy of Oral Budesonide Suspension (OBS) and Long-term Treatment Effect of OBS in Adolescent and Adult Subjects (11 to 55 Years of Age, Inclusive) with Eosinophilic Esophagitis (EoE)

Study Number: SHP621-302

SAP Version and Date:

Version 1.0: 04 December 2019

STATISTICAL ANALYSIS PLAN

Protocol No.:	SHP621-302
Protocol Title:	A Phase 3, Multicenter, Double-blind Extension Study to Evaluate Maintenance of Efficacy of Oral Budesonide Suspension (OBS) and Long-term Treatment Effect of OBS in Adolescent and Adult Subjects (11 to 55 Years of Age, Inclusive) with Eosinophilic Esophagitis (EoE)
Drug:	SHP621, oral budesonide suspension (OBS), budesonide oral suspension (BOS)
Sponsor:	Shire ViroPharma, Incorporated (Shire) 300 Shire Way, Lexington, MA 02421 USA
Version No. and Date	Final Version 1.0, Date 04-December-2019

Version No:	Document History Description of Update	Author(s)	Effective Date
0.1	First draft	████████, PPD statistician	08-Mar-2019
0.2	Second draft	████████, PPD statistician	11-Apr-2019
0.3	Version 0.3	████████, PPD statistician	07-May-2019
0.4	Version 0.4	████████, PPD statistician	05-Jun-2019
0.5	Version 0.5	████████, PPD statistician	21-Jun-2019
0.6	Version 0.6	████████, PPD statistician	25-Jun-2019
1.0	Version 1.0	████████, PPD statistician	04-Dec-2019

PREPARED BY

PPD, Inc.
3900 Paramount Pkwy
Morrisville, NC, 27560

Issued by: _____
[REDACTED] [REDACTED] (BIOS)
[REDACTED]
I approve this document
04 Dec 2019 16:19:08 -05:00
DocuSign[®]
[REDACTED], Biostatistics
PPD

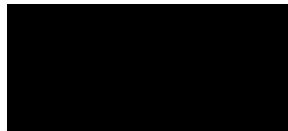
Date: ____/____/____

Reviewed by: _____
[REDACTED] [REDACTED] (BIOS)
[REDACTED]
I approve this document
04 Dec 2019 15:47:23 -05:00
DocuSign[®]
[REDACTED], Biostatistics
PPD

Date: ____/____/____

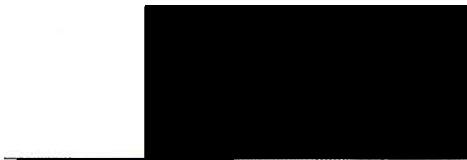
Upon review of this document, including table, listing, and figure shells, the undersigned approves the statistical analysis plan. The analysis methods and data presentation are acceptable, and the table, listing, and figure production can begin.

APPROVED BY



[Redacted], Statistical & Quantitative
Sciences

Takeda



[Redacted], Statistical & Quantitative
Sciences

Takeda

For non-commercial use only

TABLE OF CONTENTS

PREPARED BY	2
APPROVED BY	3
TABLE OF CONTENTS	4
LIST OF TABLES	7
LIST OF FIGURES	7
ABBREVIATIONS	8
1. INTRODUCTION	10
2. STUDY DESIGN	11
2.1 General Study Design	11
2.2 Randomization	13
2.3 Blinding	14
2.3.1 Blinding of Investigational Product	14
2.3.2 Blinding of Study Data	14
2.4 Schedule of Assessments	14
2.5 Determination of Sample Size	19
2.6 Multiplicity Adjustments for Type I Error Control	19
3. OBJECTIVES	20
3.1 Primary Objectives	20
3.2 Secondary Objectives	20
4. SUBJECT POPULATION SETS	21
4.1 Screened Set	21
4.2 Randomized Set	21
4.3 FAS	21
4.3.1 SHP621-301 BOS Full Responder FAS	21
4.3.2 SHP621-301 BOS Non-responder FAS	21
4.3.3 SHP621-301 Placebo FAS	21
4.4 Safety Analysis Set	22
4.5 Per-protocol Set	22
4.6 Treatment Group	23
5. SUBJECT DISPOSITION	24
6. PROTOCOL DEVIATIONS	25
7. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS	26
7.1 General Medical History	26
7.2 Interim EoE Medical History	26
8. EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE	27

8.1	Exposure to Investigational product.....	27
8.2	Measurement of Treatment Compliance	27
9.	PRIOR AND CONCOMITANT MEDICATION	29
10.	EFFICACY ANALYSES	30
10.1	Primary Efficacy Endpoint(s) and Analysis.....	30
10.1.1	Definition of Primary Efficacy Endpoint.....	30
10.1.2	Primary Analysis of Primary Efficacy Endpoint	31
10.1.3	Sensitivity and Other Analyses of Primary Efficacy Endpoint	31
10.2	Key Secondary Efficacy Endpoint(s) and Analysis	32
10.2.1	Definition of Key Secondary Efficacy Endpoints	32
10.2.2	Primary Analysis of Key Secondary Efficacy Endpoint.....	34
10.2.3	Sensitivity Analyses of Key Secondary Efficacy Endpoint	34
10.3	Other Secondary Efficacy Endpoint(s) and Analysis.....	35
11.	SAFETY ANALYSES	40
11.1	Adverse Events.....	40
11.2	Clinical Laboratory Variables	41
11.3	Vital Signs	42
11.4	Electrocardiogram (ECG).....	43
11.5	Other Safety Variables	43
12.	CLINICAL PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES.....	44
13.	OTHER ANALYSES	45
14.	INTERIM ANALYSIS	47
15.	CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL	48
16.	DATA MONITORING/REVIEW COMMITTEE	49
17.	COMPUTER METHODS	50
18.	CHANGES TO ANALYSES SPECIFIED IN PROTOCOL	51
19.	DATA HANDLING CONVENTIONS.....	52
19.1	General Data Reporting Conventions	52
19.2	Derived Efficacy Endpoints	52
19.3	Derived Safety Endpoints.....	53
19.4	Missing Date of Investigational Product	53
19.5	Missing Date Information for Prior or Concomitant Medications/Procedures	53
19.5.1	Incomplete Start Date	53
19.5.2	Incomplete Stop Date.....	54
19.6	Missing Date Information for Adverse Events	55
19.6.1	Incomplete Start Date	55

19.6.2	Incomplete Stop Date.....	55
19.7	Missing Severity Assessment for Adverse Events.....	55
19.8	Missing Relationship to Investigation Product for Adverse Events	55
19.9	Character Values of Clinical Laboratory Variables	55
19.10	AESI Categories.....	56
20.	REFERENCES	58
21.	APPENDICES	59

LIST OF TABLES

Table 1:	Schedule of Assessments.....	15
Table 2:	Examples for Coding of Special Character Values for Clinical Laboratory Variables	56

LIST OF FIGURES

Figure 1:	Study Design Flow Chart	13
-----------	-------------------------------	----

ABBREVIATIONS

λ_z	first order rate constant associated with the terminal (log-linear) portion of the curve
ACTH	adrenocorticotropic hormone
AE	adverse event
AESI	AEs of special interest
ALT	alanine aminotransferase
AST	aspartate aminotransferase
BMD	bone mineral density
BMI	body mass index
BOS	budesonide oral suspension
CI	confidence interval
DSQ	Dysphagia Symptom Questionnaire
DXA	(DEXA) dual-energy X-ray absorptiometry
eCRF	electronic case report form
ePRO	electronic patient-reported outcome
EGD	esophagogastroduodenoscopy
EoE	eosinophilic esophagitis
ET	early termination
ERES	EoE Endoscopic Reference Score
FAS	Full Analysis Set
GCP	Good Clinical Practice
HPF	high-powered field
HRQoL	health-related quality of life
hs	at bedtime
ICH	International Conference on Harmonisation
IWRS	Interactive web-based response system
MedDRA	Medical Dictionary for Regulatory Activities
OBS	oral budesonide suspension

PP	per-protocol
pc	after meals
qAM	every morning
SAE	serious adverse event
SAP	statistical analysis plan
SAS [®]	statistical analysis system
SOC	system organ class
TEAE	treatment-emergent adverse event
VAS	Visual Analogue Scale
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the SHP621-302 statistical analyses of efficacy, safety/tolerability, and health related quality of life (HRQoL) data as described in study Protocol Amendment 2, dated 19 Dec 2016 and SHP621-302 Protocol Amendment 2 – Administrative Change, dated 20 February 2019. Specifications for tables, figures, and listings are contained in a separate document.

For non-commercial use only

2. STUDY DESIGN

2.1 General Study Design

This is a Phase 3, multicenter, double-blind study to evaluate the efficacy, safety and tolerability of Oral Budesonide Suspension (OBS) (herein referred to as BOS: Budesonide Oral Suspension) treatment administered twice daily (qAM, pc, and hs) for 36 weeks. The study is being conducted in adolescents and adults, aged 11-55 years, inclusive, with EoE and dysphagia who completed the SHP621-301 induction study.

This study will comprise 3 periods: 4-week screening period, 36-week double-blind treatment period, and a 4-week safety follow-up (see [Figure 1](#)). All subjects who have completed the SHP621-301 induction study will be eligible to enroll in this extension study. Approximately 200 subjects (88%) who were randomized in the SHP621-301 study are estimated to complete the SHP621-301 study and will be enrolled in this treatment extension study based on enrollment observed in the Phase 2 study (MPI 101-06). Randomization into the randomized withdrawal period will be stratified by treatment assignment and treatment response in the SHP621-301 study. Subjects who are full responders (defined as ≤ 6 eos/HPF across all available esophageal levels and at least a 30% reduction in DSQ at the final treatment evaluation visit [Visit 4]) in the SHP621-301 induction study will be eligible to enter the randomized withdrawal period to receive either OBS 2 mg twice daily or placebo at a 1:1 ratio. The randomization of subjects who are full responders will also be stratified by age (adults and adolescents) to ensure approximately equal number of subjects within each age group are assigned to the treatment groups.

Subjects who sign informed consent (or assent as applicable for subjects < 18 years) will be screened (Visit 0); assessments from the SHP621-301 final treatment evaluation visit (Visit 4) will be used as the Visit 0 screening assessments of SHP621-302. Subjects who meet eligibility criteria at the screening visit (Visit 0) will enter the 36-week double-blind period. Eligible subjects will receive investigational product based on treatment assignment in SHP621-301 for up to 4 weeks prior to enrollment in study SHP621-302. This 4-week screening period is required to allow for blinded treatment response information (reduction in DSQ from baseline and eosinophilic count as determined by a central reader) to be collected during the SHP621-301 final treatment evaluation visit and to be entered into the interactive web-based response system (IWRS). Only the unblinded data team who are independent from the blinded study team and not involved with the day-to-day conduct of the study will have restricted access to blinded treatment response information. Once information is available in IWRS, subjects will return for the randomization visit (Visit 1) to receive investigational product.

During the 36-week treatment extension study, subjects who were assigned to, and fully responded to BOS treatment in the SHP621-301 study (≤ 6 eos/HPF across all available esophageal levels and at least a 30% reduction in DSQ at the final treatment evaluation visit) will be randomized to receive either BOS 2 mg twice daily or placebo at a 1:1 ratio. Subjects who were assigned to BOS treatment in the SHP621-301 study and did not respond or partially responded, will receive BOS 2 mg twice daily. Subjects who were assigned to placebo in the SHP621-301 study will receive BOS 2 mg twice daily. At Week 12, all subjects will undergo an upper esophagogastroduodenoscopy (EGD) with biopsy to evaluate eosinophil counts.

Upon receipt of blinded EGD and DSQ data, subjects on placebo in the randomized withdrawal period who have relapsed (≥ 15 eos/HPF from at least 2 levels of the esophagus and 4 days of dysphagia over the 2-week period prior to scheduled visit) will have their treatment assignment changed to BOS 2 mg twice daily at the next scheduled visit. The treatment assignment change will be performed in a blinded manner in IWRS at the subsequent study visit. If either criteria for relapse are not met, the subject will remain assigned to placebo.

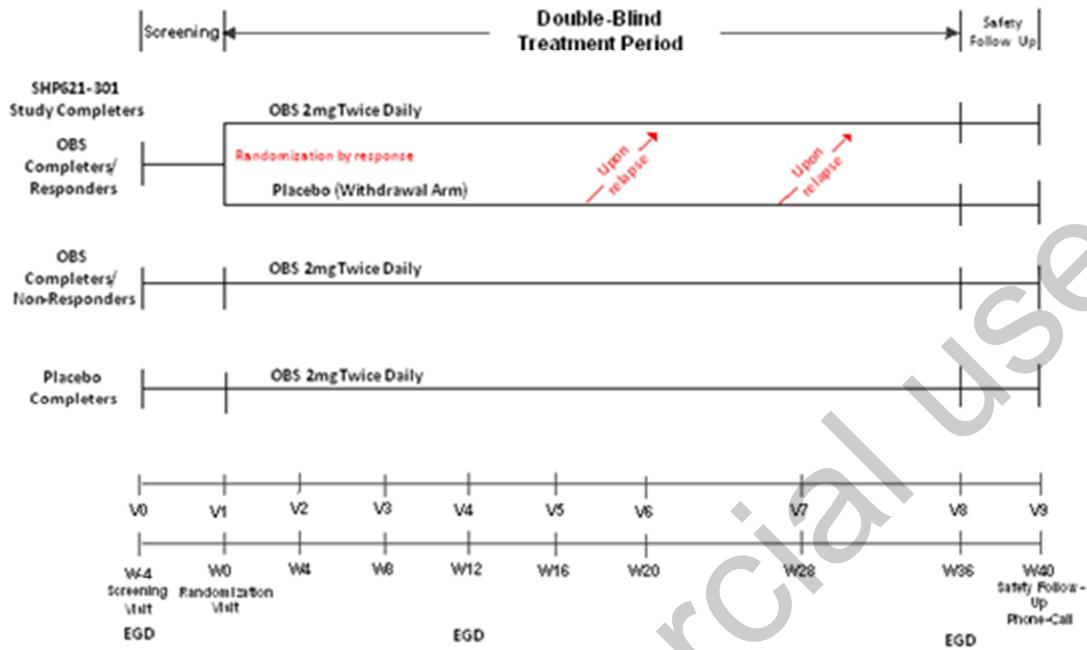
In the protocol, all subjects who are assigned to treatment at the beginning of the 36-week double-blind period are described as "randomized", consistent with the processes for double-blind treatment assignment described, while not true randomization for patients who received placebo in the SHP621-301 study or received BOS 2 mg BID and did not meet the criteria for full response in the SHP621-301 study and did not meet the criteria for full response.

Subjects who report increased or worsening dysphagia symptoms to the investigator prior to the Week 12 EGD may have an unscheduled EGD to confirm whether relapse has occurred. If an unscheduled EGD is performed prior to Week 12, an EGD does not have to be repeated at Week 12. If a subject on placebo meets the criteria for relapse, the subject's treatment assignment will be changed in a blinded manner from placebo to BOS 2 mg twice daily at the subsequent study visit.

Subjects will have efficacy and safety assessments at each visit during the 36-week period. Subjects who fail to meet all eligibility criteria at Visits 0 or 1 will be considered screen failures. Subjects cannot be rescreened once they have been designated as a screen failure. Subjects who discontinue will not be replaced.

Subjects will be required to visit the site up to 8 times during the 36-week period. At the end of the 36-week double-blind treatment period (Visit 8, Week 36), subjects who complete the study may have the opportunity to enroll in an open-label, continuation study (SHP621-303). Subjects who fail screening, who discontinue at any time during the extension study (SHP621-302), or who do not enroll in the continuation study will have a follow-up phone call 4-weeks post last dose of investigational product to query for SAEs, AEs and concomitant treatments.

Figure 1: Study Design Flow Chart



Abbreviations: EGD=esophagogastroduodenoscopy; OBS=oral budesonide suspension

2.2 Randomization

Subjects will be randomized via a computer-generated randomization schedule at the randomization visit (Visit 1) following a screening period and confirmation of study eligibility. Subjects who fully responded to BOS treatment in the SHP621-301 study (≤ 6 eos/HPF across all available esophageal levels and at least a 30% reduction in DSQ at the final treatment evaluation visit) will enter a randomized withdrawal period to receive either BOS 2 mg twice daily (qAM, pc, and hs) or placebo twice daily (qAM, pc, and hs) at a 1:1 ratio. Subjects who did not respond or partially responded to BOS treatment in the SHP621-301 study will receive BOS 2 mg twice daily. Subjects who were assigned to placebo in the SHP621-301 study will receive BOS 2 mg twice daily.

Randomization of subjects in the randomized withdrawal period will also be stratified by age (adults and adolescents) to ensure approximately equal number of subjects within each age group are assigned to the treatment groups. Individual subject treatment is automatically assigned by the IWRS.

2.3 Blinding

2.3.1 Blinding of Investigational Product

This is a double-blind study. Blinding was achieved by means of identical appearance for BOS and placebo. Placebo consists of all components of the investigational product solution with the exception of budesonide.

2.3.2 Blinding of Study Data

To protect the integrity of the study blind in the SHP621-302 (treatment extension study of SHP621-301), the post-randomization central histology and Dysphagia Symptom Questionnaire (DSQ) subject level data are segregated in separate case report forms that are not available to the blinded study team, study sites, and subjects until the final database lock of SHP621-302.

Although treatment assignment will remain blinded to the blinded study team, study sites, and subjects as documented in the unblinded cross functional team charter, these subject level data would have potentially identified which treatment a subject was randomized to (eg, in individual subjects with substantial decreases in esophageal eosinophils from screening). An Unblinded Data Team (UBDT) was established to handle the processing, review, and validation of all histology and DSQ data to ensure consistency in the conduct of these activities and collection of data in electronic case report forms. The UBDT has restricted access to post-randomization histology and DSQ data during the conduct of the SHP621-301 and SHP621-302 studies, and the UBDT operates independently from the blinded study team, who are involved in study oversight or day-to-day conduct, and investigators at study sites.

2.4 Schedule of Assessments

[Table 1](#) presents the schematic of the study design.

Table 1: Schedule of Assessments

Procedures	Screening ^a	Double-blind Treatment Period								Safety Follow-up Contact ^b
	Visit 0	Randomization/ Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8 or ET ^c	Visit 9
Week	-4	0	4	8	12	16	20	28	36	40
Window	≤4 weeks	--	±3 days	±6 days	±6 days	±3 days				
Informed consent/assent	X									
Medical history review	X	X								
Inclusion/exclusion criteria review	X	X								
Vital signs ^b ; height ^c , and weight assessment ^d	X	X	X	X	X	X	X	X		
EGD with endoscopy score (ERFS) and biopsy ^e	X				X				X	
Issue/Retrieve DSQ handset	X									X
DSQ completion	Once-daily completion									
DSQ compliance assessment	X	X	X	X	X	X	X	X		
	X				X				X	
	X				X				X	
	X				X				X	
	X				X				X	

Table 1: Schedule of Assessments

Table 1: Schedule of Assessments

Procedures	Screening ^a	Double-blind Treatment Period								Safety Follow-up Contact ^b
	Visit 0	Randomization/ Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8 or ET ^c	Visit 9
Week	-4	0	4	8	12	16	20	28	36	40
Window	≤4 weeks	--	±3 days	±6 days	±6 days	±3 days				
Review of adverse events ^d	X	X	X	X	X	X	X	X	X	X

Abbreviations: ACTH=adrenocorticotrophic hormone; DSQ=Dysphagia Symptom Questionnaire; DXA=dual-energy X-ray absorptiometry; EGD=esophagogastroduodenoscopy;

EREFS=EoE Endoscopic Reference Score; hs=at bedtime; IWRS=interactive web-based response system;
 ; pc=after meals;

- ^a The assessments from the SHP621-301 final treatment evaluation visit (Visit 4) will be used as the screening assessments (Visit 0) of this treatment extension study.
- ^b Vital signs will be assessed after the subject has been in a supine position for at least 5 minutes immediately prior to the assessment and will include blood pressure (systolic and diastolic), heart rate, respirations, and temperature.
- ^c Height to be collected at screening (Visit 0) and Visit 8 for all subjects. Height to be collected at Visit 4 for adolescents only (11-17 years, inclusive). Height measurements for adolescents should be measured in triplicate using stadiometers.
- ^d Weight measurements for adolescents (11-17 years, inclusive) should be measured in duplicate.
- ^e Endoscopy must include esophageal biopsies; gastric and duodenal biopsies may be done at the discretion of the investigator. Endoscopies at Visit 4 and Visit 8 should occur at or within 7 days of the scheduled visit. Unscheduled endoscopies may be performed at the discretion of the investigator.
- ^f
- ^g Tanner staging assessments will be performed for all subjects ≥11 years of age until investigator confirms subject is post puberty.
- ^h Clinical laboratory tests will include the following: alkaline phosphatase, aspartate aminotransferase, alanine aminotransferase, total bilirubin, total protein, albumin, glucose, blood urea nitrogen, creatinine, sodium, potassium, chloride, calcium, carbon dioxide, hemoglobin, hematocrit, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, mean corpuscular volume, erythrocyte count, leukocyte count, neutrophils, lymphocytes, monocytes, eosinophils, basophils, and platelet count). All subjects must fast overnight prior to collection.
- ⁱ Urinalysis parameters will include glucose, protein, specific gravity, pH, nitrite, bilirubin, ketones, hemoglobin, urobilinogen, and leukocyte esterase.
- ^j The serum pregnancy test will be performed for all female subjects at screening (Visit 0) and Visit 8. Urine pregnancy tests will be performed at all other visits.

Table 1: Schedule of Assessments

Procedures	Screening ^a	Double-blind Treatment Period								Safety Follow-up Contact ^p
	Visit 0	Randomization/ Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8 or ET ^o	Visit 9
Week	-4	0	4	8	12	16	20	28	36	40
Window	≤4 weeks	--	±3 days	±6 days	±6 days	±3 days				

^k Dual-energy X-ray absorptiometry scans should be performed using the same machine and software as used in the SHP621-301 study. Dual-energy X-ray absorptiometry scan at Visit 8 should occur at or within 7 days of the scheduled visit.

^l Randomization will occur via IWRS at Visit 1 once the subject's eligibility is confirmed.

^m Study medication is supplied at the SHP621-301 final treatment visit based on treatment assignment in SHP621-301.

ⁿ Adverse event assessments at each visit and physical examination must include specific assessments for signs of glucocorticoid excess (eg, moon facies, acne, hirsutism, mood swings, insomnia, and depression).

^o If subject discontinues study prematurely during the treatment period, the evaluations listed for Visit 8 are to be performed as completely as possible.

^p A safety follow-up contact by phone will be performed 4 weeks following the last dose of study medication for subjects who fail screening, who discontinue early, or who do not enroll in the continuation study. A safety follow-up contact by phone is not required for subjects who roll over into SHP621-303 immediately upon completion of SHP621-302.

2.5 Determination of Sample Size

Approximately 200 subjects (88%) who were to be randomized in the SHP621-301 study were estimated to complete the SHP621-301 study and will enroll in this treatment extension study (SHP621-302) based on enrollment observed in the Phase 2 study (MPI 101-06).

The primary efficacy measure of the study will be the proportion of subjects who relapse during the double-blind randomized withdrawal period, defined as having an eosinophil count of ≥ 15 eos/HPF from at least 2 of 3 levels of the esophagus (as determined by a central reader) and at least 4 days of dysphagia in the 2-week period prior to the scheduled visit (as determined by the DSQ). To be considered as a subject with relapse, both criteria must be met. To determine whether a subject meets the criterion for dysphagia, refer to Section 7.2.1.2 of the protocol.

Based on enrollment in the Phase 2 study (MPI 101-06), approximately 26% of subjects, or approximately 40 subjects, who were assigned to BOS treatment in the SHP621-301 study were anticipated to respond fully after 12 weeks in the SHP621-301 study. For this extension study, approximately 38 subjects were to be required to achieve more than 80% power at the significance level of 0.05 (2-sided) to detect a 50 % point difference between relapse proportions of 20% and 70% in the BOS and placebo groups respectively using the Fisher's exact test with equal allocation 1:1 to treatment groups (19 subjects in the BOS group and 19 subjects in the Placebo group).

2.6 Multiplicity Adjustments for Type I Error Control

For the SHP621-301 BOS Full Responder FAS, hierarchical testing with fixed sequence procedure will be used to control the Type I error for the multiple endpoints tested at one-sided $\alpha = 0.025$. For a test to be considered statistically significant within the testing hierarchy, it must be statistically significant at the one-sided 0.025 level, and all previous tests within the testing hierarchy must be statistically significant at the one-sided 0.025 level, the fixed sequence of testing hierarchy is as follows:

1. Primary Endpoint: Relapse during the double-blind randomized withdrawal period.
2. Secondary Endpoint 1: Histologic response [peak eosinophil count of ≤ 6 /high-powered field (HPF) across all available esophageal levels] at Visit 8 (Week 36)
3. Secondary Endpoint 2: Dysphagia symptom response [$\geq 30\%$ reduction in the Dysphagia Symptom Questionnaire (DSQ) combined score (questions 2+3) from baseline of the SHP621-301 study] at Visit 8 (Week 36)
4. Secondary Endpoint 3: Change in total endoscopy score from baseline of the SHP621-301 study, as measured by the EREFS classification, at Visit 8 (Week 36)

3. OBJECTIVES

3.1 Primary Objectives

The primary objective of the study is:

- To evaluate the maintenance of efficacy over 36 weeks, as measured by the peak eosinophilic count and the DSQ score, through a randomized withdrawal design for subjects who responded to 12 weeks of BOS treatment (2 mg twice daily) with a peak count of ≤ 6 eos/HPF across all available esophageal levels at the final treatment visit and a $\geq 30\%$ reduction in DSQ score from baseline during the SHP621-301 induction study.

3.2 Secondary Objectives

The key secondary objective of this study is:

- To evaluate the long-term treatment response over an extended period of 36 weeks for subjects who were randomized to BOS treatment but did not respond after 12 weeks in the SHP621-301 induction study (subject did not have a peak count of ≤ 6 eos/HPF across all available esophageal levels at the final treatment visit and/or a $\geq 30\%$ reduction in DSQ score from baseline).

Additional secondary objectives of the study are:

- To evaluate the response to BOS treatment over 36 weeks for subjects who received placebo in the SHP621-301 induction study).
- To evaluate the effect of reinitiating BOS treatment for subjects who relapse after being randomized to placebo in the randomized withdrawal period (treatment-withdrawal-treatment reinitiation).
- To assess endoscopically identified esophageal features as measured by the EoE Endoscopic Reference Score (ERFS).
- To evaluate other responding criteria based on histology and DSQ.
- To evaluate the long-term safety and tolerability of BOS treatment.

3.3 Exploratory Objectives

The exploratory objective of this study is:

- [REDACTED]
- [REDACTED]

4. SUBJECT POPULATION SETS

4.1 Screened Set

The screened set will consist of all subjects who have signed informed consent and have conducted screening assessments.

4.2 Randomized Set

The randomized set will include all randomized subjects. Note: Randomized subjects include all subjects who were full responders to BOS in study SHP621-301 and are randomized to BOS or placebo during the randomized withdrawal period in study SHP621-302 and subjects who were either non-responders to BOS or assigned to placebo in SHP621-301 and assigned to BOS in study SHP621-302.

4.3 FAS

The full analysis set will include all randomized subjects who received at least 1 dose of SHP621-302 investigational product. The FAS is summarized in [Appendix 6](#).

4.3.1 SHP621-301 BOS Full Responder FAS

The SHP621-301 BOS full responder FAS will include all randomized subjects who received at least 1 dose of investigational product in the SHP621-302 study and who were determined to be full responders to the BOS treatment during the induction study SHP621-301 (subjects who randomized to BOS treatment and had a peak count of ≤ 6 eos/HPF across all available esophageal levels at the final treatment visit in SHP621-301 and $\geq 30\%$ reduction in DSQ score from baseline in SHP621-301 at the final treatment visit in SHP621-301). The SHP621-301 BOS full responder FAS will be used in efficacy data analyses in which subjects will be analyzed according to their assigned randomized treatment, regardless of the treatment actually received.

4.3.2 SHP621-301 BOS Non-responder FAS

The SHP621-301 BOS non-responder FAS will include all randomized subjects who received at least 1 dose of investigational product in the SHP621-302 study and who did not respond or partially responded to assigned BOS treatment in the induction study SHP621-301 (subject who randomized to BOS treatment and did not have peak count of ≤ 6 eos/HPF across all available esophageal levels at the final treatment visit in SHP621-301 and/or $\geq 30\%$ reduction in DSQ score from baseline in SHP621-301 at the final treatment visit in SHP621-301).

4.3.3 SHP621-301 Placebo FAS

The SHP621-301 placebo FAS will include all randomized subjects who received at least 1 dose of investigational product in the SHP621-302 study and who were randomized to receive placebo in the induction study SHP621-301.

4.4 Safety Analysis Set

The safety analysis set will consist of all subjects who received at least 1 dose of investigational product. The Safety Set will be used for all safety analyses in which subjects will be analyzed according to the treatment they received.

4.5 Per-protocol Set

The per-protocol (PP) set will include all subjects in the FAS excluding subjects with major protocol deviations. The PP set will be identified prior to unblinding the treatment assignments by a team consisting of, at a minimum, a physician and a statistician from Shire.

Major protocol deviations that may lead to exclusion from the PP set, are as follows, but not limited to:

1. Violations of inclusion and/or exclusion criteria
2. Compliance with study medication

A subject, who has less than 70% or greater than 130% overall compliance during the double-blind treatment period with his/her assigned treatment

3. Study Treatment Administration/Dispensing

A subject who had incorrect IP treatment kit ID assigned/dispensed

4. Prohibited concomitant medications

- Significant use of CYP450 3A4 inhibitors (e.g., ketoconazole, grapefruit juice) during the treatment period
- Significant use of excluded medication while participating in the study
- Initiation of swallowed topical corticosteroid for EoE or systemic corticosteroid for any condition within 4 weeks of EGD
- Significant changes in uses of PPIs, H2 antagonists, antacids, or leukotriene inhibitors for any condition during the treatment period

5. Accidental Unblinding

Subject unblinding occurs when a subject's randomization drug code is broken (accidentally or not per protocol) or their post-randomization pathology or DSQ results are unblinded prior to completion of double-blind treatment period

4.6 Treatment Group

The majority of the data displays will be presented by 4 different treatment groups in the extension study SHP621-302 (Refer to [Appendix 6](#)):

- SHP621-301 BOS Full responders: BOS-BOS and BOS-PBO
- SHP621-301 BOS Non-responders: BOS-BOS
- SHP621-301 Placebo: PBO-BOS

In addition, BOS-PBO-BOS will be used separately in selected efficacy analyses.

SHP621-301 BOS Full responders:

- The responder BOS-BOS treatment group will include subjects who were randomized to the BOS treatment and fully responded to treatment in the induction study SHP621-301 and were randomized to BOS treatment in the extension study SHP621-302.
- The responder BOS-PBO treatment group will include subjects who were randomized to the BOS treatment and fully responded to treatment in the induction study SHP621-301 and were randomized to placebo treatment in the extension study SHP621-302.

SHP621-301 BOS Non-responders: BOS-BOS

- The non-responder BOS-BOS treatment group will include subjects who were randomized to the BOS treatment and either partially responded or did not respond to treatment in the induction study SHP621-301 and were randomized to BOS treatment in the extension study SHP621-302.

The PBO-BOS treatment group will include subjects who were randomized to the placebo treatment in the induction study SHP621-301 and assigned to BOS treatment in the extension study SHP621-302.

The BOS-PBO-BOS treatment group will include subjects who were randomized to the BOS treatment and fully responded to treatment in the induction study SHP621-301 and randomized to placebo treatment in the extension study SHP621-302, but relapsed during the randomized withdrawal and reinitiated treatment with BOS (intermittent therapy) in the extension study SHP621-302. This treatment group will be used only for selected efficacy analyses of these relapsed subjects.

5. SUBJECT DISPOSITION

Listing of Screen Failures (subjects who failed to meet all eligibility criteria at Visits 0 or 1) will be presented along with reasons for screen failure.

The total number of subjects screened and the number of subjects who screen failed will be summarized. The number of subjects included in each analysis set (ie, Randomized, Safety, FAS, SHP621-301 BOS Full Responder FAS, SHP621-301 BOS Non-responder FAS, SHP621-301 Placebo FAS, and PP) will be summarized by treatment group. The number and percentage of subjects who completed and who prematurely discontinued the study will also be presented for each treatment group and overall for the FAS. Reasons for premature discontinuation from the study as recorded on the study completion page of the electronic case report form (eCRF) will also be summarized (number and percentage) by treatment group. All subjects who prematurely discontinued during the study will be listed by discontinuation reason for FAS Subjects.

Supporting listings of the discontinuation data from end of screening phase and end of study will be provided along with the information on adverse events (AEs) that caused discontinuation (if subjects discontinued due to AEs).

The number of subjects screened, randomized, and those who completed the study will be tabulated by site. In addition, the duration of enrollment, in days, will be summarized for each site, and overall. Duration of enrollment will be calculated as (last date of contact for any subject at that site - the first date of informed consent for any subject at that site +1).

6. PROTOCOL DEVIATIONS

The number and percentage of subjects reporting any protocol deviations and the incidence of each deviation type will be summarized by treatment group and overall for the FAS and individual responses will be presented in a data listing for all screened subjects. Protocol deviations will be summarized by deviation level (major or minor) and by type (International Conference on Harmonisation [ICH]/Good Clinical Practice [GCP] or Protocol Deviation) within each deviation level. Full list of deviations will be documented in a separate file.

For non-commercial use only

7. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Baseline characteristics for endpoints assessed by EGD including biopsy will be determined using assessments at the Screening Visit (Visit 0). Other demographic and baseline characteristics will be determined using assessments at the Randomization Visit 1 (Week 0) or the last non-missing observation prior to the first dose of investigational product, whichever is later. Descriptive summaries of demographic and baseline characteristics will be presented by treatment group for the FAS and Safety Set.

Subject demographic and baseline characteristics including age, age group (<18 years, ≥ 18 years), sex, race, ethnicity, weight, height, body mass index (BMI), Tanner stage, peak eosinophil count at three esophageal levels and overall, DSQ combined scores, histopathologic epithelial features combined total score ratio for grade and stage, and total endoscopic reference score will be summarized by treatment group for the FAS and Safety Set. If there are multiple measurements of weight and height collected on the same date (e.g. adolescent subjects), the average of multiple measurement will be used as baseline. Continuous variables will be summarized by descriptive statistics including number of subjects, mean, standard deviation (SD), median, Q1, Q3, minimum, and maximum. Categorical variables will be summarized by the number of subjects in each category and the percentage of subjects out of the total in the respective analysis set.

Height and weight will be used to calculate BMI using the formula below:

$$\text{BMI} = \text{weight [kg]} / (\text{height [m]})^2$$

Listings will be created to show all the demographics and baseline characteristics for all subjects in the Randomized Set.

7.1 General Medical History

The investigator will record all new clinically or medically relevant information which arose after the recording of the medical history in the antecedent study. Adverse events recorded during the SHP621-301 study may be added as medical history at the investigator's discretion. General medical history data will be summarized by treatment group for the FAS. General medical history findings will also be listed for all subjects for the FAS. Medical history will be coded using MedDRA Version 18.0.

7.2 Interim EoE Medical History

The following information associated with EoE history will be recorded in the eCRF at the screening visit:

- Any changes in diet since the Screening visit in the SHP621-301 study (yes/no answer).
- Any changes in medical treatment for EoE since the Screening visit in the SHP621-301 study (yes/no answer).

EoE history will be summarized by treatment group for the FAS and will also be listed for all the subjects in the FAS.

8. EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE

8.1 Exposure to Investigational product

For all treated subjects, including those who terminated early, treatment exposure is the duration between the first dose date and last dose date:

Duration of exposure (days)=(Last dose date – First dose date +1).

In addition, total dose administered (mL) and average daily dose will be calculated and summarized by treatment group:

Actual Average Daily Dose (mL/day)=Sum of the doses (mL)/Duration of exposure (days).

Continuous variables will be summarized by descriptive statistics (n, mean, SD, Q1, Q3, minimum, median, and maximum) by treatment group for total dose, duration of exposure and average daily dose. Categorical variables will be summarized by the number of subjects in each category and the percentage of subjects out of the total in the respective analysis set for categories duration of exposure in weeks (<=1, >1 - <=4, >4 - <=8, >8 - <=12, >12 - <=16, >16 - <=20, >20 - <=28, >28 - <=36, >36) for all safety subjects (separated by population sets: SHP621-301 BOS full responder FAS, SHP621-301 BOS non-responder FAS, SHP621-301 placebo FAS, and overall). For SHP621-301 BOS non-responder FAS and SHP621-301 placebo FAS subjects, the exposure is BOS treatment. For SHP621-301 BOS full responder FAS, after all subjects undergo an EGD with biopsy to evaluate eosinophil counts at Week 12, subjects on placebo in the randomized withdrawal period who have relapsed will have their treatment assignment changed to BOS 2 mg twice daily at the next scheduled visit. The exposure summary will be summarized by SHP621-301 BOS full responder FAS, SHP621-301 BOS non-responder FAS, and SHP621-301 placebo FAS. An overall exposure for all subjects in Safety Set will also be presented. The exposure summary for subjects who relapse on placebo will be summarized in a separate table.

A listing will be created by subject and treatment group and will provide exposure for the Safety Set.

8.2 Measurement of Treatment Compliance

Compliance with study medication will be assessed across all study visits. Subjects will be instructed to bring any remaining study medication and empty bottles to each study visit. Designated site staff will evaluate compliance by questioning the subject and evaluating the amount of study medication remaining. The subject will be questioned regarding any discrepancies.

When a bottle is returned, the site will measure the amount of investigational product remaining in the bottle in centimeters and the remaining amount will be recorded as 9.375 cm if the measured amount in the bottle is >9.375 cm. The volume taken in mL for each bottle can be determined in centimeters based on the equation below:

$$\text{Volume in mL} = 210 \text{ cm}^3 - [22.4 \text{ cm}^2 \times \text{height cm}].$$

The percent compliance at each study visit will be determined as follows:

(volume of investigational product taken in all returned bottles)/(expected volume of investigational product to be taken).

Subjects will be considered compliant with study medication if they received no less than 70% and no more than 130% of the intended dosing as assessed during the treatment period.

Study drug exposure and overall compliance will be summarized, separately and combined, by treatment group for the Safety Set after subject randomization. The number and percentage of subjects whose compliance is <70% or >130% as well as between 70% and 130% will be summarized. The overall approximate compliance rates will be summarized using descriptive statistics by treatment group for Safety Set.

9. PRIOR AND CONCOMITANT MEDICATION

World Health Organization-Drug Dictionary (WHODRUG) as of March 2015 will be used to classify prior and concomitant medications by therapeutic class.

Concomitant medication is defined as any medication with a start date prior to the date of the first dose of investigational product (Visit 1) and continuing after the first dose of investigational product (Visit 1) or with a start date between the dates of the first dose of investigational product (Visit 1) and the Safety Follow-up Contact, or 31 days after last dose of SHP621-302 investigational product for subjects who do not have a Safety Follow-up Contact. The definition of concomitant medication is similar to the treatment-emergent adverse event (TEAE) definition in Section 11.1.

Medical/surgical procedures performed prior and during the treatment period will be recorded on the eCRF, along with the date and reason for the procedure. The prior medical/surgical procedure is defined as any procedure with the start date prior to the first dose of investigational product (Visit 1). The concomitant medical/surgical procedure is defined as any procedure with the start date between the dates of the first dose of investigational product (Visit 1) and the last dose of investigational product.

Both concomitant medication usage and medical/surgical procedures will be summarized by the number and proportion of subjects in each treatment group receiving each medication/procedure for the FAS. Procedures can be counted both as prior and concomitant. Multiple medication/procedure received by a subject in the same category will be counted only once.

All concomitant medications and medical/surgical procedures will be listed.

10. EFFICACY ANALYSES

The primary efficacy analyses will be performed on the SHP621-301 BOS full responder FAS and presented by treatment group (Placebo or BOS). The key secondary analysis will be performed on the SHP621-301 BOS non-responder FAS. Other secondary efficacy analyses will be performed on the FAS and presented by treatment group. The most recent non-missing measurement (scheduled or unscheduled) collected prior to first dose of SHP621-302 study drug administered in the double-blind phase in this study will be used as the baseline for all efficacy analyses. The most recent non-missing measurement (scheduled or unscheduled) collected prior to first dose of study drug administered in double-blind phase in SHP621-301 will also be used as the baseline for some of the key secondary and other secondary efficacy analyses that are specified to include SHP621-301 baseline.

10.1 Primary Efficacy Endpoint(s) and Analysis

10.1.1 Definition of Primary Efficacy Endpoint

The primary efficacy endpoint for subjects in SHP621-301 BOS full responder FAS is relapse during the double-blind randomized withdrawal period. Relapse, a binary response (either with a relapse or not), is defined as meeting both the eosinophil histology relapse criterion and the dysphagia symptom relapse criterion as follows:

- Having an eosinophil count of ≥ 15 eos/HPF from at least 2 of 3 levels of the esophagus, as determined by a central reader's evaluation of EGD biopsy tissues,

And

- Having at least 4 days of dysphagia (with answer 'Yes' for question 2 in DSQ) in the 2-week period prior to the scheduled visit, as determined by the DSQ

In determining whether a subject meets the dysphagia symptom relapse criterion and considering the potential for missing diary entries, determination of relapse will occur as follows:

- If at least 4 days of dysphagia are reported on the DSQ in the 2-week period prior to the visit, the subject would meet the criterion for dysphagia symptom relapse.
- If less than 4 days of dysphagia are reported and at least 8 diary entries are recorded within the 2-week period, the subject would not meet the criterion for dysphagia symptom relapse.
- If less than 4 days of dysphagia are reported and less than 8 diary entries are recorded within the 2-week period, the subject would be designated as meeting the criterion for dysphagia symptom relapse due to missing diary data.

To derive the primary endpoint of relapse during the double-blind randomization withdrawal period, the eosinophil histology relapse criterion will be evaluated first using the EGD assessments (scheduled or unscheduled). If the unscheduled EGD assessment meets the histology relapse criteria, the DSQ relapse criteria will be first evaluated at the scheduled visit prior to the unscheduled EGD assessment.

If the scheduled EGD assessment meets the histology relapse criteria, the DSQ relapse criteria will be first evaluated at the same scheduled visit of the EGD assessment.

If the DSQ relapse criteria is not met at the first evaluation visit, then the subsequent scheduled visits will be evaluated until the DSQ relapse criteria is met at one scheduled visit or not met at any visit.

For subjects who were prematurely withdrawn from the study (early termination), if there is no EGD assessment at the early termination (ET) visit, the latest EGD assessment after randomization will be used at the ET visit to evaluate the relapse.

For subjects who have a missing EGD result at Visit 8 (Week 36) and do not meet eosinophil histology relapse prior to Visit 8 (Week 36) but meet the dysphagia symptom relapse criterion at Visit 8 (Week 36), the primary efficacy endpoint (relapse) will be categorized as missing ([Appendix 7](#)).

10.1.2 Primary Analysis of Primary Efficacy Endpoint

The primary efficacy endpoint will be analyzed as the proportion of subjects who relapse during the double-blind randomized withdrawal period in the SHP621-301 BOS full responder FAS using a one-sided Fisher's Exact test comparing BOS 2 mg twice daily against placebo. The primary test of treatment effect will be one-sided and conducted at the significance level of 0.025. The proportion of subjects with relapse for each treatment group and the corresponding 95% confidence interval (CI) will be estimated. In addition, the difference in the proportion of subjects with relapse between the 2 treatment groups and its 95% exact confidence interval ([Chan and Zhang, 1999](#)) will be estimated. Subjects with missing primary efficacy endpoint (refer to [Appendix 7](#)) will be classified as relapsers in the primary efficacy analysis.

The null hypothesis states that there is no difference in relapse proportions between BOS 2 mg twice daily and placebo, with the one-sided alternative of a smaller relapse proportion in BOS group compared to placebo group. Relapse proportions at each adjacent double-blind visit interval (Visit 1 to 2, Visit 2 to 3, ..., Visit 7 to 8) will also be assessed by applying the one-sided Fisher's Exact test to the observed data at each double-blind visit interval.

10.1.3 Sensitivity and Other Analyses of Primary Efficacy Endpoint

The following sensitivity and supportive analyses will be performed for the primary endpoint to evaluate the robustness of the results from the primary analysis methods.

- I. The primary analysis will be repeated for subjects in both the Per-Protocol set and the SHP621-301 BOS full responder FAS.
- II. Sensitivity analysis will be performed using the SHP621-301 BOS full responder FAS by classifying subjects with missing primary efficacy endpoint (refer to [Appendix 7](#)) as non-relapsers.
- III. Sensitivity analysis will be performed using the SHP621-301 BOS full responder FAS by classifying subjects who meet the Histology relapse criteria and meet the dysphagia symptom relapse criterion only because of missing diary data as non-relapsers.

IV. Sensitivity analysis will be performed using the SHP621-301 BOS full responder FAS based on the prorated number of dysphagia days within the 2-week period if less than 4 days of dysphagia are reported and at least 8 diary entries are recorded within the 2-week period. The prorated number of dysphagia days will be based on the rate of dysphagia days over non-missing diaries as below. If the prorated number of dysphagia days is 4 or more, the subject would meet the criteria for dysphagia symptom relapse, otherwise, if the prorated number of dysphagia days is less than 4, the subject would not meet the criteria for dysphagia symptom response.

$$\text{Number of Dysphagia Days} = \frac{\text{Number of Dysphagia Days Observed within 2 weeks}}{\text{Number of Diaries Reported within 2 weeks}} \times 14$$

10.2 Key Secondary Efficacy Endpoint(s) and Analysis

10.2.1 Definition of Key Secondary Efficacy Endpoints

The key secondary efficacy endpoints are the long-term treatment responses, binary responses over an extended period of 36 weeks in SHP621-301 BOS non-responder FAS set.

Long-term treatment response from the baseline of the SHP621-301 study is defined as subjects who met the following criteria:

- Histologic response, defined as a peak eosinophil count of ≤ 6 eos/HPF across all available esophageal levels at the final treatment period evaluation (Visit 8)
AND
- Dysphagia symptom response, defined as $\geq 30\%$ reduction in the DSQ combined score (questions 2+3) from baseline of the SHP621-301 study to the final treatment period evaluation (Visit 8)

Long-term treatment response from the baseline of this extension study is defined as subjects who met the following criteria:

- Histologic response, defined as a peak eosinophil count of ≤ 6 eos/HPF across all available esophageal levels at the final treatment period evaluation (Visit 8)
AND
- Dysphagia symptom response, defined as $\geq 30\%$ reduction in the DSQ combined score (questions 2+3) from baseline of this extension study to the final treatment period evaluation (Visit 8)

Histology Response:

An EGD with endoscopy biopsy will be performed at the Week 12 visit (Visit 4), and Week 36 visit (Visit 8) or at the ET Visit; the peak eosinophil count per HPF across all available esophageal levels at the final treatment period evaluation (Visit 8) will be used as a key secondary measure of efficacy.

If at any time, the subject reports increased or worsening dysphagia symptoms to the investigator, the investigator may choose to perform an unscheduled EGD. If an unscheduled EGD is performed prior to Week 12, an EGD does not have to be repeated at Week 12. If an unscheduled EGD is performed between Week 12 and Week 36, the Week 36 EGD should still be performed. If the subject has a peak eosinophil count of ≤ 6 /HPF across all available levels at the final treatment period evaluation (Visit 8), histology response variable will be set to 1 and the subject will be considered a histology responder. Otherwise, histology response will be set to 0 and subject will be treated as a non-responder. Subjects who discontinue during the treatment period or who do not have a peak eosinophil count available at Visit 8 (week 36) will be classified as histology non-responders.

Dysphagia Symptom Response:

Subject's dysphagia symptoms will be evaluated using a DSQ electronic patient-reported outcome (ePRO) device. The questionnaire will be completed each evening by subjects during the screening period and during the 36-week double-blind treatment period. Each evening before bedtime, subjects will be asked to indicate if they experienced dysphagia symptoms (e.g., food passing slowly or food sticking) during that day. [Appendix 1](#) shows the daily DSQ consisting of four questions with corresponding potential point scoring in parentheses. Subjects are prompted to respond to Questions 1 first, if subjects respond positively to Question 1, then they are prompted to respond to Question 2, if subjects respond positively to Question 2, they are prompted to respond to Questions 3 and 4.

Calculations of DSQ combined score at each scheduled visit will be based on daily ePRO entries during a 2-week interval prior to each study visit. The DSQ combined score will be calculated by summing the scores of responses to questions 2 and 3 only. Responses to Questions 1 and 4 will not be included to calculate the DSQ combined score.

Percent change from baseline of the SHP621-301 study to this extension study's final treatment period evaluation Visit 8 (Week 36) in DSQ combined score will be calculated as (DSQ combined score at Week 36 – DSQ combined Score at SHP 621-301 Baseline) / DSQ combined score at SHP 621-301 Baseline *100. Percent change from baseline of this extension study to the final treatment period evaluation Visit 8 (Week 36) in DSQ combined score will be calculated similarly.

Dysphagia symptom response from the baseline of the SHP621-301 will be set to 1 (responder) if a subject achieved a minimum of 30% reduction in DSQ combined score from baseline of the SHP621-301 study to the final treatment period evaluation Visit 8 (Week 36), ie percent changes of the DSQ combined scores from baseline of the induction SHP621-301 study is ≤ -30 .

Dysphagia symptom response from the baseline of this extension study will be set to 1 (responder) if a subject achieved a minimum of 30% reduction in DSQ combined score from baseline of this extension study to the final treatment period evaluation Visit 8 (Week 36). Otherwise, if the percent change from either baseline is > -30 or a subject discontinues from the treatment phase early and the DSQ combined score at Visit 8 (Week 36) is incalculable, a subject will be classified as non-responder.

10.2.2 Primary Analysis of Key Secondary Efficacy Endpoint

The key secondary efficacy endpoint will be analyzed as the proportion of subjects who responded based on the histological data and the DSQ data at the final treatment period evaluation (Visit 8) for SHP621-301 BOS non-responder FAS. The proportion of subjects with long-term treatment response and the corresponding 95% confidence interval (CI) will be estimated and summarized.

10.2.3 Sensitivity Analyses of Key Secondary Efficacy Endpoint

The following sensitivity and supportive analyses will be performed for the key secondary endpoints to evaluate the robustness of the results from the key secondary analysis methods:

Sensitivity analyses will be performed using the SHP621-301 BOS non-responder FAS by classifying subjects into responders or non-responders by each of the following missing data handling methods.

Missing data handling method 1 for Dysphagia Symptom Response: Subjects who are missing 4 days or more of DSQ daily diary data in either week or 7 days in the last 2 weeks at a post-baseline visit, are treated as non-responders without applying the window shifting rule.

In the case of missing DSQ daily diary data, DSQ data for a given day will be considered missing if the diary is not reported or if the diary is reported but the answer to Question 1 is “No”. If fewer than 8 reported diaries are available within the 14-day period prior to the randomization visit (Visit 1), the window shift rule will be performed. The most recent 8 reported diaries in a consecutive 14-day period would be used to calculate the baseline DSQ combined score. Such 14-day periods cannot be shifted for more than 7 days (not including the study visit day) when calculating DSQ combined score at randomization visit (Visit 1). The DSQ combined score calculated during the 14-day period prior to randomization visit 1 (Week 0) will be the baseline DSQ combined score. For the 14-day period used to calculate the DSQ combined score at the post-randomization visits (Visit 2 to Visit 8/Termination Visit), 4 out of 7 days in each of the last two weeks are needed to calculate a DSQ combined score. The window shift rule for post-randomization data, which is described in [Appendix 1](#) will not be applied for this sensitivity analysis. Subjects who have missing data for at least 4 days in either of the last 2 weeks, or have missing data for at least 7 days in the last two weeks (ie, the 14-day period), will be treated as non-responders of Dysphagia Symptom Response, and the corresponding DSQ combined score at that visit will be incalculable.

Missing data handling method 2 for Dysphagia Symptom Response: Window shift rule is applied for post-baseline data (14-day periods cannot be shifted for more than 7 days at every visit).

In the case of missing DSQ daily diary data, DSQ data for a given day will be considered missing if the diary is not reported or if the diary is reported but the answer to Question 1 is “No”.

If fewer than 8 reported diaries are available within the 14-day period prior to the randomization visit (Visit 1) or any post-randomization visits (Visit 2 to Visit 8/Termination Visit), the window shift rule for randomization visit (Visit 1) or Visit 8 DSQ calculation will be performed (see [Appendix 1](#)) and the most recent 8 reported diaries in a consecutive 14-day period would be used to calculate the DSQ Combined Score. Such 14-day periods cannot be shifted for more than 7 days (not including the study visit day) when calculating DSQ Combined Score at randomization visit (Visit 1) or any post-randomization visits (Visit 2 to Visit 8). The DSQ combined score calculated during the 14-day period prior to randomization visit 1 (Week 0) will be the baseline DSQ combined score. If DSQ Combined Score at randomization visit (Visit 1) or any post-randomization visits (Visit 2 to Visit 8) is incalculable, a subject will be classified as non-responder.

10.3 Other Secondary Efficacy Endpoint(s) and Analysis

The following secondary efficacy endpoints will be analyzed for all subjects in FAS:

- Histologic response, defined as a peak eosinophil count of ≤ 6 eos/HPF across all available esophageal levels at Visit 4 (Week 12) and Visit 8 (Week 36)
- Dysphagia symptom response, defined as $\geq 30\%$ reduction in the DSQ combined score (questions 2+3) at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately
- Change in total endoscopy score, as measured by the EREFS classification, at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately
- Change in the peak eosinophilic count at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately
- Peak eosinophil count < 15 /HPF across all available esophagus levels at Visit 4 (Week 12) and Visit 8 (Week 36)
- Peak eosinophil count ≤ 1 /HPF across all available esophagus levels at Visit 4 (Week 12) and Visit 8 (Week 36)
- Change in the peak eosinophil count at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study at each assessment visit for each available esophageal level (proximal, mid-, and distal) separately
- Change in the histopathologic epithelial features combined total score (grade and stage) at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately
- Dysphagia symptom response (binary response), defined as a $\geq 50\%$ reduction in the DSQ combined score (questions 2+3), at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately
- Change in the DSQ combined score (questions 2+3) at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately

- Cumulative distribution function curves for the change and the percent change in the DSQ score at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately
- Overall binary response I, defined as a peak eosinophil count of $\leq 6/\text{HPF}$ across all esophageal levels and a reduction in the DSQ combined score of $\geq 30\%$ from baseline of the SHP621-301 study and from baseline of this extension study at Visit 4 (Week 12) and Visit 8 (Week 36) separately
- Overall binary response II, defined as a peak eosinophil count of $\leq 6/\text{HPF}$ across all esophageal levels and a reduction in the DSQ combined score of $\geq 50\%$ from baseline of the SHP621-301 study and from baseline of this extension study at Visit 4 (Week 12) and Visit 8 (Week 36) separately
- Change in the DSQ + pain score (questions 2+3+4) at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately
- Change in the DSQ pain score (question 4) at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study separately

For binary outcomes, the proportion of subjects who respond with corresponding 95% CI will be summarized by visit and treatment group. In addition, one-sided Fisher's exact test will be conducted for histologic response and dysphagia symptom response separately at Week 12 and Week 36 for subjects in SHP621-301 BOS full responder FAS.

Subjects who discontinue during the treatment period or who do not have a peak eosinophil count available or who has >-30 or missing DSQ percent change from baseline (SHP621-301 study or this extension study) will be classified as non-responders for overall binary response I and overall binary response II.

For subjects who relapse on placebo during the randomized withdrawal and reinitiate treatment with BOS 2 mg twice daily (intermittent therapy), separate descriptive analyses for peak eosinophil count and DSQ combined score will be conducted at each assessment visit. Changes of peak eosinophil count and DSQ combined score will be summarized over time from baseline of the SHP621-301 study and from the time of relapse. If relapse happens at the scheduled visit, the peak eosinophil count and DSQ combined score at this scheduled visit will be used as the baseline value at the time of relapse. If relapse happens at the unscheduled visit, the peak eosinophil count at this unscheduled visit will be used as the baseline at the time of relapse and the DSQ combined score at the scheduled visit prior to this unscheduled visit will be used as the baseline score at the time of relapse. The overall treatment response ($\leq 6 \text{ eos/HPF}$ across all available esophageal levels and at least a 30% reduction in DSQ from SHP621-301 baseline score) at the final treatment period evaluation (Visit 8) and the corresponding 95% CI will be summarized.

Refer to [Appendix 5](#) for the analysis for histopathologic epithelial features. The continuous endpoints will be analyzed as a change from baseline using an ANCOVA model that includes treatment group and age group as factors and baseline score as a covariate. The LS means and corresponding 95% CI will be summarized at each post baseline assessment visit.

For change in total endoscopy score for subjects in SHP621-301 BOS full responder FAS, both two-sided and one-sided p-values in ANCOVA analysis will be reported. In addition, descriptive summary statistics will be reported for DSQ combined score at each assessment visit, change and percent change at each assessment visit from baseline of the SHP621-301 study and from baseline of this extension study.

Refer to [Appendix 1](#) to derive DSQ + pain score, DSQ pain score, DSQ question 2 score and DSQ question 3 score at each assessment visit using the same rules applied for the DSQ combined score. Descriptive summary statistics for each score at baseline and every post-baseline visit will be reported.

The total endoscopy score is the sum of endoscopy scores of proximal and distal locations from the following 5 major feature categories: 1) exudates or plaques (grade 0–2); 2) fixed esophageal rings (grade 0–3); 3) edema (grade 0–2); 4) furrows (grade 0–2); and 5) strictures (grade 0–1).

The 10th, the 25th, the 50th (Median), the 75th, and the 90th percentile change scores and sample sizes and 95% confidence interval band will be included in the cumulative distribution function curves.

The following additional secondary efficacy figures will be plotted:

- The proportion of subjects who relapse during 36-week double-blind treatment period will be plotted by treatment group.
- The proportion of subjects who have a histology response as a peak eosinophil count of $\leq 6/\text{HPF}$, $< 15/\text{HPF}$ and $\leq 1/\text{HPF}$ across all available esophageal levels at Visit 4 (Week 12) and Visit 8 (Week 36) will be plotted by treatment group.
- The proportion of subjects who have a reduction of $\geq 30\%$ and 50% of DSQ from baseline will be plotted across visits by treatment group.
- The proportion of subjects who have an overall response, defined as a peak eosinophil count of $\leq 6/\text{HPF}$ across all esophageal levels and a reduction in the DSQ combined score of $\geq 30\%$ from baseline of the SHP621-301 study at Visit 4 (Week 12) and Visit 8 (Week 36), will be plotted by treatment group.
- The proportion of subjects who have an overall response, defined as a peak eosinophil count of $\leq 6/\text{HPF}$ across all esophageal levels and a reduction in the DSQ combined score of $\geq 30\%$ from baseline of this extension study at Visit 4 (Week 12) and Visit 8 (Week 36), will be plotted by treatment group.
- The proportion of subjects who have an overall response, defined as a peak eosinophil count of $\leq 6/\text{HPF}$ across all esophageal levels and a reduction in the DSQ combined score of $\geq 50\%$ from baseline of the SHP621-301 study at Visit 4 (Week 12) and Visit 8 (Week 36), will be plotted by treatment group.

- The proportion of subjects who have an overall response, defined as a peak eosinophil count of ≤ 6 /HPF across all esophageal levels and a reduction in the DSQ combined score of $\geq 50\%$ from baseline of this extension study at Visit 4 (Week 12) and Visit 8 (Week 36), will be plotted by treatment group.
- The mean DSQ scores and standard deviations by visit and treatment group will be plotted.
- A scatter plot with the maximum overall peak eosinophil count from the esophageal biopsies (x-axis) vs. DSQ scores at the final treatment period evaluation Week 36.
- A scatter plot with the change from baseline of the SHP621-301 study in maximum overall peak eosinophil count from the esophageal biopsies (x-axis) vs. change from baseline of the SHP621-301 study in DSQ scores at the final treatment period evaluation Week 36.
- A scatter plot with the change from baseline of this extension study in maximum overall peak eosinophil count from the esophageal biopsies (x-axis) vs. change from baseline of this extension study in DSQ scores at the final treatment period evaluation Week 36.
- A scatter plot with the percent change from baseline of the SHP621-301 study in maximum overall peak eosinophil count from the esophageal biopsies (x-axis) vs. the percent change from baseline DSQ scores of the SHP621-301 study at the final treatment period evaluation Week 36.
- A scatter plot with the percent change from baseline of this extension study in maximum overall peak eosinophil count from the esophageal biopsies (x-axis) vs. the percent change from baseline DSQ scores of this extension study at the final treatment period evaluation Week 36.

10.4 Exploratory Efficacy Endpoint(s) and Analyses

The exploratory endpoints are the following:

- Eosinophil histology relapse for subjects in the SHP621-301 BOS full responder FAS, defined as peak eosinophil count of ≥ 15 /high-powered field (HPF) from at least 2 of 3 levels of the esophagus) during the double-blind randomized withdrawal period
- [REDACTED]
- [REDACTED]

Similar to the primary analysis of the primary efficacy endpoint, eosinophil histology relapse [REDACTED] will be assessed using the one-sided Fisher's exact test assuming a smaller relapse proportion in BOS group compared to placebo group.

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

For non-Confidential Use Only

11. SAFETY ANALYSES

Safety analyses will be performed by treatment group using the Safety Set and the treatment group will be separated by SHP621-301 BOS full responder, SHP621-301 BOS non-responder and SHP621-301 placebo. Safety variables include AEs, physical examinations, stadiometry, vital signs (temperature, systolic and diastolic blood pressure, heart rate, and respiratory rate), weight and height assessments, dual-energy X-ray absorptiometry (DXA) scans for bone mineral density (BMD) and body composition measurements (for adolescents aged 11-17 years, inclusive), clinical laboratory tests (hematology, chemistry, urinalysis; serum pregnancy test, if appropriate), and adrenocorticotropic hormone (ACTH) stimulation tests. To account for the effects of puberty in adolescent subjects (11-17 years of age, inclusive), BMD z-scores will be adjusted for height z-score. Continuous safety parameters will be descriptively summarized by treatment group at baseline of this extension study and for each post-baseline visit. The last assessment prior to the first dose of investigational product will be used as baseline of this extension study. Both baseline of the induction SHP621-301 study and the baseline of this extension study will be used in the summary of clinical laboratory tests and the growth parameters (Height, Weight and BMI). Only the baseline of this extension study will be used for all other safety analyses.

11.1 Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), Version 18.0 or newer.

All AEs will be collected from the time of the informed consent through the completion of this extension study's post-treatment follow-up as described below. For eligible subjects from SHP621-301 who consent to participate in this extension study, AEs will be collected in the SHP621-301 clinical database up to the date of the first dose of SHP621-302 investigational product (ie, Visit 1 in the SHP621-302 study) or the Safety Follow-up Contact for subjects who screen fail prior to enrolling in extension study SHP621-302. AEs that have a start date on or after the first dose of SHP621-302 investigational product will be collected in the SHP621-302 database.

TEAEs are defined as AEs that start or deteriorate on or after the first dose of investigational product (Visit 1) through the Safety Follow-up Contact, or 31 days after the last dose of investigational product for subjects who do not have a Safety Follow-up Contact. This TEAE definition is a rule for programming TEAEs and will be applied only to the SHP621-302 database; and thus, events that occur after the Screening in the SHP621-303 study will not be considered as TEAEs in SHP621-302.

However, for any subjects who die during the study (ie, the date of death is between the date of first dose of investigational product and the date of study discontinuation entered by the site, inclusive), all AEs (including those resulting in death) that occur during the study will be considered as TEAEs irrespective of the last dose and will be included in the TEAE summaries.

An overall summary of the number of subjects with TEAEs will be presented, including the number and percentage of subjects with any TEAEs, serious TEAEs, severe TEAEs, life-threatening TEAEs, TEAEs related to investigational product, TEAEs related to EoE, deaths and hospitalizations due to TEAEs, TEAEs leading to discontinuation of investigational product, and TEAEs leading to study discontinuation.

The number, and percentage of subjects reporting TEAEs in each treatment group will be tabulated by system organ class (SOC) and preferred term; by SOC, preferred term, and maximum severity. Serious TEAEs, TEAEs considered related to investigational product, and TEAEs related to EoE will also be summarized by SOC and preferred term. If more than 1 AE occurs with the same preferred term for the same subject, the subject will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by severity and by relationship to investigational product.

The incidence of common TEAEs ($\geq 2\%$ of subjects in any treatment group) will be summarized by preferred term. Serious TEAEs, TEAEs leading to discontinuation of investigational product and TEAEs leading to discontinuation from the study will be summarized by SOC, preferred term and treatment group. Deaths will be summarized by preferred term and treatment group.

In addition, the analysis of AEs of Special Interest (AESI) categories will be summarized. The AESIs are defined in Section 19.10. The following analyses of AESIs will be performed for the Safety Set:

- Incidence of AESI
- Number and percentages of subjects by preferred terms
- Number and percentages of subjects by maximum severity
- Serious TEAEs
- Related serious TEAEs
- TEAEs leading to discontinuation of investigational product
- TEAEs leading to death

11.2 Clinical Laboratory Variables

Descriptive statistics for clinical laboratory values (in both US conventional and standard units) and changes from baseline of the SHP621-301 study and from baseline of this extension study at each assessment time point as well as shift tables from baseline of the SHP621-301 study and from baseline of this extension study to each visit for quantitative variables will be presented by treatment group for the following clinical laboratory variables. All laboratory data will also be listed.

Hematology hemoglobin, hematocrit, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, mean corpuscular volume, mean platelet volume, erythrocyte count, erythrocyte distribution width, leukocyte count, neutrophils, lymphocytes, monocytes, eosinophils, basophils, and platelet count.

Biochemistry	alkaline phosphatase, aspartate aminotransferase, alanine aminotransferase, total bilirubin, total protein, albumin, glucose, blood urea nitrogen, creatinine, sodium, potassium, chloride, calcium, carbon dioxide.
Urinalysis	glucose, protein, specific gravity, pH, nitrite, bilirubin, ketones, hemoglobin, urobilinogen, and leukocyte esterase.
Other tests	serum pregnancy, urine pregnancy, morning cortisol (6:00-9:00 AM collection), ACTH stimulation testing (serum cortisol collections at 30 and 60 minutes after synthetic ACTH injection in addition to baseline collection at 6:00-9:00 AM).

Pregnancy test results will only be listed. A serum β -hCG pregnancy test is performed on all female subjects at the screening visit (Visit 0) and the final treatment evaluation visit (Visit 8) or ET visit. A urine pregnancy test is performed on all female subjects at all other visits.

ACTH stimulation testing will be performed by measuring the levels of cortisol in the blood following the injection of a synthetic form of ACTH. The type of synthetic and route of administration will be per local lab discretion. Blood samples will be collected just prior to and approximately 30 and 60 minutes following the injection at baseline of SHP621-301 study, screening visit (Visit 0) and the final treatment evaluation visit (Visit 8)/ET.

Morning unstimulated cortisol testing results will be descriptively summarized by treatment group and visit. For ACTH stimulation testing results, the number and percentage of subjects with the highest of the two cortisol values at the 30 and 60 minute time points that meet a threshold criterion event of ≤ 18 mcg/dL will be summarized by treatment group at baseline of SHP621-301 study, screening visit (Visit 0) and the final treatment evaluation visit (Visit 8), and worsening of the shift from baseline of SHP621-301 study, screening visit (Visit 0) and the final treatment evaluation visit (Visit 8) will be presented. ACTH stimulation testing results will also be presented with a higher threshold criterion of ≤ 20 mcg/dL, i.e., the number and percentage of subjects with the highest of the two cortisol values at the 30 and 60 minutes time points of ≤ 20 mcg/dL.

The abnormal test results for cortisol, ALT, AST and glucose will be listed.

11.3 Vital Signs

Descriptive statistics for vital signs (systolic and diastolic blood pressure, heart rate, respiration rate, temperature, BMI and weight) and their changes from baseline at each post-baseline visit and at the end of study will be presented by treatment group. Weight is collected at every visit for all subjects. For adolescent subjects 11-17 years of age, inclusive, two weight measurements per visit will be recorded. The average weight per visit will be used in summary tables for these subjects.

A separate summary of height will be provided by visit and treatment group. Height will be collected at screening visit (Visit 0) and the final treatment evaluation visit (Visit 8) for all subjects. Height will be collected at Visit 4 for adolescent subjects only (11-17 years, inclusive).

For adolescent subjects 11-17 years of age, inclusive, stadiometers will be used to measure height, and three height measurements per visit will be recorded. The average height per visit will be used in summary tables for these subjects.

All vital signs data will also be listed. For the growth parameters (Height, Weight and BMI), the changes from baseline of the SHP621-301 study and from baseline of this extension study will be summarized by all subjects and by age group. The z-scores for growth parameters will be summarized descriptively at each visit for adolescents based on the subject's age at each visit. Z-scores will be derived using the CDC growth charts (Kuczmarski et al., 2002). A SAS program for the 2000 CDC Growth Charts will be used to derive the percentiles and z-scores (www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm).

11.4 Electrocardiogram (ECG)

Not applicable since ECG assessments are not planned for this study.

11.5 Other Safety Variables

Physical Examination

Physical examination assessments at each visit will include specific assessments for signs of glucocorticoid excess (e.g., moon faces, acne, hirsutism, mood swings, insomnia, and depression). Physical examination at the screening visit (Visit 0) and the final treatment evaluation visit (Visit 8)/ET will also include Tanner Staging Assessments for subjects <18 years of age. The number and percentage of subjects reporting symptoms described above will be presented by visit and treatment group.

Dual-energy X-ray Absorptiometry for Bone Mineral Density

DXA (also referred to as DEXA) scans for determination of BMD and body composition will be performed in subjects aged 11-17 years, inclusive, at screening visit (Visit 0) and the final treatment evaluation visit (Visit 8)/ET. The sites for DXA measurement will be the lumbar spine (L1-L4 preferred) and total body less head. To account for the effects of puberty in adolescent subjects (11-17 years of age, inclusive), BMD Z-scores will be adjusted for height in the following manner. The subject's height is compared to the 50th percentile on the CDC growth chart. The age on the CDC growth chart where the subject's height is equal to the 50th percentile height on the growth chart will be used in lieu of the subject's actual age to calculate a "height-adjusted" BMD Z-score. For example, if a 10-year-old boy is 125cm tall, according the CDC growth chart for boys aged 2-20 years, 125cm is the median height for a 7.5-year-old boy. When calculating the height-adjusted BMD Z-score, an age of 7.5 years will be used rather than the boy's actual age of 10 years. Mean Z-scores at scheduled assessments and mean within-subjects changes from baseline of the induction SHP621-301 study and from baseline of the extension SHP621-302 study at Visit 8 (Week 36)/ET will be calculated for the DXA Z-scores; corresponding two-sided 95% confidence intervals will be provided.

12. CLINICAL PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

Not applicable

For non-commercial use only

13. OTHER ANALYSES

[REDACTED]

FOIA
Request
Only

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

For internal use only

14. INTERIM ANALYSIS

No interim analysis is planned.

For non-commercial use only

15. CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL

The distribution-based imputation for subjects who were prematurely withdrawn from the study without the primary efficacy endpoint mentioned in the protocol was not included in SAP since the results may be bias due to the small size.

For non-commercial use only

16. DATA MONITORING/REVIEW COMMITTEE

Not applicable.

For non-commercial use only

17. COMPUTER METHODS

All statistical analyses will be performed using SAS[®] Version 9.4 or higher (SAS Institute, Cary, NC 27513) on a suitably qualified environment.

For non-commercial use only

18. CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

Not applicable.

For non-commercial use only

19. DATA HANDLING CONVENTIONS

19.1 General Data Reporting Conventions

Continuous variables will be summarized using the following descriptive statistics: n, mean, SD, median, Q1, Q3, minimum, and maximum. Categorical and count variables will be summarized by the number of subjects (n) and the percent of subjects in each category.

Unless specified otherwise, min/max will be presented to the same decimal places as the raw data. Percentage, mean and median will be presented to 1 more decimal place than the raw data. Standard deviation and standard error will be presented to 2 more decimal places than the raw data.

BMI should be rounded to 1 decimal place for reporting.

Derived questionnaire scores, and other similar efficacy parameters recorded as integers, should be rounded to 1 decimal place for reporting.

Averaged lab and vital sign results e.g. diastolic/systolic blood pressure and pulse (when taken in triplicate) should be rounded to 1 decimal place for reporting.

In addition, p-values will generally be presented to 3 decimal places; values less than 0.001 will be presented as <0.001.

19.2 Derived Efficacy Endpoints

Refer to Section 10.1 for derivation instructions of histology response and DSQ scores by visit. DSQ combined score and other DSQ scores will be calculated at study visit (Week 0, Week 4, Week 8, Week 12, Week 16, Week 20, Week 28, and Week 36) using the daily diary data in the selected 14-day period prior to actual scheduled visit date. For subjects who prematurely discontinue during the treatment period, the available daily DSQ dairy data collected after the last visit before the early termination will be included as candidate dairy data for the selected 14-day period for the calculation of DSQ scores at the next planned visit after early termination. The target date of the next planned visit will be determined by the number of days since randomization, the next planned visit date = the randomized date + 7*[the next planned visit week number-4] + 1 (e.g., for a subject who discontinued at Week 15, the next planned visit week would be Week 16), and the calculated planned date for Week 16 would be determined as: [1+12 weeks*7] days after the date of randomization [Week 4 visit]. In this example, the subject who discontinued at Week 15 would have diary data included in the calculation of Week 16 DSQ scores that was within 14 days of the calculated planned date of Week 16.

19.3 Derived Safety Endpoints

For safety parameters, the early termination visits will be mapped to the next scheduled visits. The last visits per subject will be summarized in “End of Study” visit, including both last visits of completers and early termination visits. Repeated or Unscheduled Assessments of Safety Parameters will be handled as follows:

If a subject has repeated assessments before the start of investigational product in the double-blind treatment period, the results from the final assessment made prior to the start of investigational product will be used as baseline. If end of study assessments are repeated or unscheduled, the last post-baseline assessment will be used as the end of study assessment for generating descriptive statistics excluding the unscheduled assessments. If a subject has repeated assessments between the start of investigational product in the treatment period and the end of study visit, the assessments of unscheduled visits will be excluded in the table summary; however, all the records will be included in the listings.

19.4 Missing Date of Investigational Product

When the date of the last dose of investigational product is missing for a subject in the Safety Set, all efforts should be made to obtain the date from the investigator. If the date is still missing, the last visit date when investigational product was returned will be used in the calculation of treatment duration.

19.5 Missing Date Information for Prior or Concomitant Medications/Procedures

For prior or concomitant medications, including rescue medications, incomplete (ie, partially missing) start date and/or stop date will be imputed. When the start date and the stop date are both incomplete for a subject, the start date will be imputed first.

19.5.1 Incomplete Start Date

The following rules will be applied to impute the missing numerical fields. If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

Missing day and month

- If the year of the incomplete start date is the same as the year of the date of the first dose of investigational product, then the day and month of the date of the first dose of investigational product will be assigned to the missing fields
- If the year of the incomplete start date is before the year of the date of the first dose of investigational product, then December 31 will be assigned to the missing fields
- If the year of the incomplete start date is after the year of the date of the first dose of investigational product, then 01 January will be assigned to the missing fields.

Missing month only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing day only

- If the month and year of the incomplete start date are the same as the month and year of the date of the first dose of investigational product, then the day of the date of the first dose of investigational product will be assigned to the missing day
- If either the year is before the year of the date of the first dose of investigational product or if both years are the same but the month is before the month of the date of the first dose of investigational product, then the last day of the month will be assigned to the missing day
- If either the year is after the year of the date of the first dose of investigational product or if both years are the same but the month is after the month of the date of the first dose of investigational product, then the first day of the month will be assigned to the missing day.

19.5.2 Incomplete Stop Date

The following rules will be applied to impute the missing numerical fields. If the date of the last dose of investigational product is missing, it will be replaced with the last visit date. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be equal to the start date.

Missing day and month

- If the year of the incomplete stop date is the same as the year as of the date of the last dose of investigational product, then the day and month of the date of the last dose of investigational product will be assigned to the missing fields
- If the year of the incomplete stop date is before the year of the date of the last dose of investigational product, then 31 December will be assigned to the missing fields
- If the year of the incomplete stop date is after the year of the date of the last dose of investigational product, then 01 January will be assigned to the missing fields.

Missing month only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing day only

- If the month and year of the incomplete stop date are the same as the month and year of the date of the last dose of investigational product, then the day of the date of the last dose of investigational product will be assigned to the missing day

- If either the year is before the year of the date of the last dose of investigational product or if both years are the same but the month is before the month of the date of the last dose of investigational product, then the last day of the month will be assigned to the missing day
- If either the year is after the year of the last dose of investigational product or if both years are the same but the month is after the month of the date of the last dose of investigational product, then the first day of the month will be assigned to the missing day.

19.6 Missing Date Information for Adverse Events

For AEs, the default is to only impute incomplete (ie, partially missing) start dates. Incomplete stop dates may also be imputed when calculation of the duration of an AE is required per the protocol. If imputation of an incomplete stop date is required, and both the start date and the stop date are incomplete for a subject, the start date will be imputed first.

19.6.1 Incomplete Start Date

Follow same rules as in Section 19.5.1.

19.6.2 Incomplete Stop Date

When required per the protocol, follow the same rules as in Section 19.5.2.

19.7 Missing Severity Assessment for Adverse Events

If severity is missing for an AE starting prior to the date of the first dose of investigational product, then a severity of “Mild” will be assigned. If the severity is missing for an AE starting on or after the date of the first dose of investigational product, then a severity of “Severe” will be assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be used in data listings.

19.8 Missing Relationship to Investigation Product for Adverse Events

If the relationship to investigational product is missing for an AE starting on or after the date of the first dose of investigational product, a causality of “Related” will be assigned. The imputed values for relationship to investigational product will be used for incidence summaries, while the actual values will be presented in data listings.

19.9 Character Values of Clinical Laboratory Variables

If the reported value of a clinical laboratory variable cannot be used in a statistical analysis due to, for example, an alpha character string being reported for a numerical variable, the appropriately determined coded value will be used in the statistical analysis. However, the actual values as reported in the database will be presented in data listings.

Table 2: Examples for Coding of Special Character Values for Clinical Laboratory Variables

Clinical Laboratory Test	Possible Results (in SI units)	Possible Results (in conventional units)	Coded Value for Analysis
Chemistry: ALT	<5 U/L	<5 U/L	0
Chemistry: AST	<5 U/L	<5 U/L	0
Chemistry: Total Bilirubin	<2 umol/L	<0.15 mg/dL	0
Urinalysis: Glucose	≥50 mg/dL	≥50 mg/dL	Positive
	≤0 mg/dL	≤0 mg/dL	Negative
Urinalysis: Ketones	>0 mg/dL	>0 mg/dL	Positive
Urinalysis: Protein	>0 mg/dL	>0 mg/dL	Positive
Urinalysis: pH	≥9 .0	≥9 .0	9.0

19.10 AESI Categories

The following AESIs are categorized with preferred terms (SOC where applicable) and will be summarized:

- Infections:
 - System Organ Classes of infections and infestations
 - Candidiasis related: oesophageal candidiasis, oral candidiasis, candidiasis, oropharyngeal candidiasis, tongue fungal infection, vulvovaginal mycotic infection, anal candidiasis.
- Potential systemic effects of corticosteroid use (including adrenal function):
 - Adrenal effects: adrenal suppression, adrenal insufficiency, ACTH stimulation test abnormal, and blood cortisol decrease or increase, blood cortisol abnormal, cushingoid
 - CNS/Mood effects: insomnia, mood swings, suicidal ideation, OCD, anxiety, depression, major depression, irritability, restlessness, sleep disorder, disturbance in attention, headache, dizziness, vertigo, lethargy, paraesthesia, hypoesthesia, psychomotor hyperactivity
 - Metabolic effects: blood glucose increased, blood glucose abnormal, glucose intolerance impaired, diabetes mellitus, intraocular pressure increased, acne, weight increased, weight fluctuation, hepatic steatosis, hirsutism, dermatitis acneiform, skin striae, skin texture abnormal, obesity, menopausal symptoms, metrorrhagia
 - Cardiac effects: oedema peripheral, peripheral swelling, hypertension, palpitations
 - Fractures: foot fracture, upper limb fracture, humerus fracture, rib fracture, wrist fracture

- GI effects:
 - Oesophagitis pain, oesophagitis, vomiting, nausea, diarrhea, abdominal pain, abdominal pain upper, constipation, abdominal distention, abdominal discomfort, hiatus hernia, dyspepsia, gastritis, gastritis erosive, chronic gastritis, dry mouth, erosive oesophagitis, and gastroesophageal reflux disease, mouth ulceration, gastric ulcer, duodenal ulcer, erosive duodenitis, mouth ulceration, esophageal ulcer, hepatic steatosis.

20. REFERENCES

Brooks, R. & de Charro, F. 1996. EuroQol: the current state of play. *Health policy*, 37, 53-72.
Chan, I. S. F. & Zhang, Z. 1999. Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions. *Biometrics*, 55, 1202-1209.



Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., Wei, R., Curtin, L. R., Roche, A. F. & Johnson, C. L. 2002. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat 11*, 1-190.



For non-commercial use only

21. APPENDICES

For non-commercial use only

APPENDIX 1 The Daily Dysphagia Symptom Questionnaire

1. Since you woke up this morning, did you eat solid food? Possible responses=yes, no.
2. Since you woke up this morning, has food gone down slowly or been stuck in your throat? Possible responses=yes (2), no (0).
3. For the most difficult time you had while swallowing food today, did you have to do anything to make the food go down or to get relief? Possible responses:
 - No, it got better or cleared up on its own (0)
 - Yes, I had to drink liquid to get relief (1)
 - Yes, I had to cough and/or gag to get relief (2)
 - Yes, I had to vomit to get relief (3)
 - Yes, I had to seek medical attention to get relief (4)
4. The following question concerns the amount of pain you have experienced when swallowing food. What was the worst pain you had while swallowing food today? Possible responses:
 - None, I had no pain (0)
 - Mild (1)
 - Moderate (2)
 - Severe (3)
 - Very Severe (4)

DSQ combined score=(Sum of points from questions 2+3 in the daily DSQ)×14 days/(Number of diaries reported with non-missing* data).

DSQ+pain score=(Sum of points from questions 2+3+4 in the daily DSQ)×14 days/(Number of diaries reported with non-missing* data).

DSQ pain score=(Sum of points from questions 4 in the daily DSQ)×14 days/(Number of diaries reported with non-missing* data).

DSQ question 2 score=(Sum of points from questions 2 in the daily DSQ)×14 days/(Number of diaries reported with non-missing* data).

DSQ question 3 score=(Sum of points from questions 3 in the daily DSQ)×14 days/(Number of diaries reported with non-missing* data).

If answer for Question 1 is “Yes” and answer for Question 2 is “No”, then scores for Question 3 and Question 4 are set to each 0 before calculating the above scores for any analysis method proposed in the efficacy analyses.

**Data for a given day will be considered missing if the diary is not reported or if the diary is reported but the answer to Question 1 is “No”. The number of diaries reported with non-missing data is also referred to as “reported diaries” in the following discussion.*

The DSQ combined score will be calculated summing the points from Questions 2 and 3 from each reported daily diary with non-missing data and dividing this by the number of reported daily diaries with non-missing data in the selected 14-day period. This quotient will then be multiplied by 14. If subjects respond negatively to Question 2 the DSQ combined score will be set to zero for that day.

The DSQ combined score calculated during the 14-day period prior to Baseline Visit 1 will be the Baseline DSQ Score. The DSQ combined score calculated during the 14-day period prior to the Final Treatment Period Evaluation will be the Final Treatment DSQ combined score.

For each 14-day period, at least 8 reported diaries are needed to calculate a DSQ combined score. If fewer than 8 reported diaries are available within the 14-day period, then the most recent 8 reported diaries in a consecutive 14-day period would be used to calculate the DSQ combined Score. In order to determine the most recent 14-day period with the minimum number of reported diaries, it may be necessary to make adjustments in some cases by shifting to earlier diary entries. Such 14-day periods cannot be shifted for more than 7 days (not including the study visit day) when calculating DSQ combined Score at Screening, Baseline Visit 1 (Week 0), Visit 2 (Week 4), Visit 3 (Week 8), Visit 4 (Week 12), Visit 5 (Week 16), Visit 6 (Week 20) and Visit 7 (Week 28), and cannot be shifted for more than 14 days prior to the Visit 8 (Week 36) (not including the visit day of Visit 8) for the DSQ combined Score at the Final Treatment Period Evaluation.

A document page with several horizontal black redaction bars. A large, faint, diagonal watermark reading "Confidential Use Only" is visible across the page.

A series of ten horizontal black bars of varying lengths, decreasing from top to bottom. The bars are positioned above a large, light gray diagonal watermark that reads "For non-commercial use only".

A document page featuring a large area of black redaction bars covering the upper half of the page. A diagonal watermark with the text 'For non-commercial use only' is visible across the page.

APPENDIX 5 Histopathologic Epithelial Features Combined Total Score (Grade and Stage)

At each of the esophageal levels (proximal, mid and distal), histopathologic features are scored for both Grade and Stage. The histopathologic epithelial features consist of eosinophil peak, basal layer hyperplasia, eosinophil abscesses, surface layering, dilated intercellular spaces, surface alteration, dyskeratotic epithelial cells, and lamina propria fibrosis. Each of the 8 Histopathologic features at each esophageal level has a possible score 0-3 points for both Grade (severity of abnormal histologic feature) and Stage (extend of abnormal histologic feature). Combined total scores include the following parameters for grade and stage respectively: 1, combined total score ratio (TSR)=(proximal TSR+mid TSR+distal TSR)/3; 2, combined score for each individual histopathologic feature (proximal+mid+distal)/N, where N is the number of non-missing sections for each individual feature.

APPENDIX 6 Full Analysis Set and Treatment Group

Analysis Set		SHP621-301		SHP621-302		
		Double-blind Treatment	Response	Treatment Assigned Randomly	Double-blind Treatment	Treatment Group
FAS	SHP621-301 BOS Full Responder FAS	BOS	Full	Yes	BOS	BOS-BOS
		BOS	Full	Yes	Placebo	BOS-PBO
		BOS	Full	Yes	Randomized to Placebo but relapsed during the randomized withdrawal period and reinitiated treatment with BOS	BOS-PBO-BOS
		SHP621-301 BOS Non-responder FAS	BOS	Partial or non-response	No	BOS
		SHP621-301 Placebo FAS	Placebo	Any	No	BOS

APPENDIX 7 Missing Primary Endpoint

The primary endpoint is defined as relapse during the double-blind randomized withdrawal period. Relapse, a binary response (either with a relapse or not), is defined as meeting both the eosinophil histology relapse criterion and the dysphagia symptom relapse criterion.

In determining whether a subject meets the dysphagia symptom relapse criterion and considering the potential for missing diary entries, determination of relapse will occur at scheduled visits as follows:

- If at least 4 days of dysphagia are reported on the DSQ in the 2-week period prior to the visit, the subject would meet the criterion for dysphagia symptom relapse.
- If less than 4 days of dysphagia are reported and at least 8 diary entries are recorded within the 2-week period, the subject would not meet the criterion for dysphagia symptom relapse.
- If less than 4 days of dysphagia are reported and less than 8 diary entries are recorded within the 2-week period, the subject would be designated as meeting the criterion for dysphagia symptom relapse due to missing diary data.

Based on the above dysphagia symptom relapse criterion, dysphagia symptom relapse will not be categorized as missing.

Eosinophil histology relapse criterion is shown below:

- Having an eosinophil count of ≥ 15 eos/HPF from at least 2 of 3 levels of the esophagus, as determined by a central reader's evaluation of EGD biopsy tissues

EGD assessments will be evaluated first to define the relapse of the primary endpoint. If the scheduled EGD assessment meets eosinophil histology relapse criterion, dysphagia symptom relapse criterion will be evaluated at the same EGD visit and the subsequent scheduled visits. If an unscheduled EGD assessment meets eosinophil histology relapse criterion, dysphagia symptom relapse criterion will be evaluated at the scheduled visit immediately prior to the EGD visit and the subsequent scheduled visits.

[Table A1](#) lists possible scenarios to derive the primary endpoint when the EGD assessment is missing at Visit 8 (Week 36) (i.e., EGD was not done or EGD results were not available):

Table A1 Scenarios to Determine the Primary Endpoint (Relapse) When the EGD Assessment is Missing at Visit 8 (Week 36)

Eosinophil Histology Relapse Determined Prior to Visit 8 (Week 36)	Visits to Evaluate Dysphagia Symptom Relapse	Dysphagia Symptom Relapse Occurs at Evaluated Visit	Primary Endpoint (Relapse)
No (missing EGD assessment or EGD results do not meet eosinophil histology relapse criteria)	Visit 8 (Week 36)	Yes	Missing

Table A1 Scenarios to Determine the Primary Endpoint (Relapse) When the EGD Assessment is Missing at Visit 8 (Week 36)

Eosinophil Histology Relapse Determined Prior to Visit 8 (Week 36)	Visits to Evaluate Dysphagia Symptom Relapse	Dysphagia Symptom Relapse Occurs at Evaluated Visit	Primary Endpoint (Relapse)
No (missing EGD assessment or EGD results do not meet eosinophil histology relapse criteria)	Visit 8 (Week 36)	No	No
Yes (determined at scheduled Week 12 or ET visit)	Scheduled visits on or after visit when eosinophil histology relapse criterion was met	Yes	Yes
Yes (determined at scheduled Week 12 or ET visit)	Scheduled visits on or after visit when eosinophil histology relapse criterion was met	No	No
Yes (determined at unscheduled visit)	Scheduled visit immediately prior to visit when eosinophil histology relapse criterion was met OR Subsequent scheduled visits after visit when eosinophil histology relapse criterion was met	Yes	Yes
Yes (determined at unscheduled visit)	Scheduled visit immediately prior to visit when eosinophil histology relapse criterion was met OR Subsequent scheduled visits after visit when eosinophil histology relapse criterion was met	No	No

The primary efficacy endpoint (relapse) will be categorized as missing when subjects have missing EGD result at Visit 8 (Week 36) and do not meet eosinophil histology relapse prior to Visit 8 (Week 36) (ie, missing EGD assessment or EGD results do not meet eosinophil histology relapse criteria) but meet the dysphagia symptom relapse criterion at Visit 8 (Week 36).