Protocol Number: EE-301

Official Title: A Phase 3, Randomized, Double-Blind, Two-Phase, Multicenter Study to Evaluate the Efficacy and Safety of Vonoprazan 20 mg Compared to Lansoprazole 30 mg for Healing in Patients with Erosive Esophagitis and to Evaluate the Efficacy and Safety of Vonoprazan (10 mg and 20 mg) Compared to Lansoprazole 15 mg for the Maintenance of Healing in Patients with Healed Erosive Esophagitis

Document Date: 13 August 2021

NCT Number: NCT04124926

Phathom Pharmaceuticals, Inc.

EE-301

A Phase 3, Randomized, Double-Blind, Two-Phase, Multicenter Study to Evaluate the Efficacy and Safety of Vonoprazan 20 mg Compared to Lansoprazole 30 mg for Healing in Patients with Erosive Esophagitis and to Evaluate the Efficacy and Safety of Vonoprazan (10 mg and 20 mg) Compared to Lansoprazole 15 mg for the Maintenance of Healing in Patients with Healed Erosive Esophagitis

13Aug2021

Statistical Analysis Plan

Version 3.0

Prepared by:

Signature Page

Study Title	A Phase 3, Randomized, Double-Blind, Two-Phase, Multicenter Study to Evaluate the Efficacy and Safety of Vonoprazan 20 mg Compared to Lansoprazole 30 mg for Healing in Patients with Erosive Esophagitis and to Evaluate the Efficacy and Safety of Vonoprazan (10 mg and 20 mg) Compared to Lansoprazole 15 mg for the Maintenance of Healing in Patients with Healed Erosive Esophagitis
Protocol Number	EE-301

Approved by:

[redacted]
Senior Director, Biostats & Programming

[redacted]

TABLE OF CONTENTS

DOCUMENT HISTORY – CHANGES COMPARED TO FINAL VERSION 1.0 6
LIST OF ABBREVIATIONS 9
1.0 INTRODUCTION 10
2.0 OBJECTIVES AND ESTIMANDS 10
2.1 Healing Phase 10
2.2 Maintenance Phase 11
3.0 INVESTIGATIONAL PLAN 11
3.1 Overall Study Design and Plan 11
3.2 Study Endpoints 12
3.2.1 Healing Phase 12
3.2.2 Maintenance Phase 13
3.3 Treatments 14
3.3.1 Healing Phase 14
3.3.2 Maintenance Phase 14
3.4 Dose Adjustment/Modifications 15
4.0 GENERAL STATISTICAL CONSIDERATIONS 15
4.1 Sample Size 16
4.1.1 Healing Phase 16
4.1.2 Maintenance Phase 16
4.2 Randomization and Blinding 17
4.3 Assessment Windows 17
4.3.1 Study Day 17
4.3.2 Unscheduled Endoscopy Assessments 18
4.3.3 Visit Window for Analysis 18
4.4 Analysis Set 19
4.4.1 Screened Set 19
4.4.2 Randomized Set 20
4.4.3 Modified Intent-to-Treat (MITT) Set 20
4.4.4 Per Protocol (PP) Set 20
4.4.5 Safety Set 21
4.5 Missing Data Handling for Endoscopy Assessments 21
5.0 SUBJECT DISPOSITION 22
5.1 Disposition 22
5.1.1 Screened and Screen Failure Subjects 22
5.1.2 Healing Phase Randomized Subjects 22
5.1.3 Maintenance Phase Randomized Subjects 23
5.1.4 Healing Phase Treatment Period 23

5.1.5 Maintenance Phase Treatment Period....24
5.2 Protocol Deviations....24
6.0 DEMOGRAPHICS AND BASELINE CHARACTERISTICS....24
6.1 Demographics and Baseline Characteristics....24
6.2 Medical History....25
6.2.1 General Medical History....25
7.0 TREATMENTS AND MEDICATIONS....26
7.1 Prior and Concomitant Medications....26
7.2 Study Treatments....27
7.2.1 Extent of Exposure....27
7.2.2 Treatment Compliance....28
8.0 EFFICACY ANALYSIS....29
8.1 Primary Efficacy Endpoint....33
8.1.1 Primary Analysis....33
8.1.2 Assumption Testing....34
8.1.3 Sensitivity and Supportive Analyses....34
8.1.4 Subgroup Analysis....36
8.2 Secondary Efficacy Endpoint....37
8.2.1 Subgroup Analysis....39
8.2.2 Sensitivity Analysis....39
8.3 Exploratory Endpoints....40
9.0 SAFETY ANALYSIS....42
9.1 Adverse Events....42
9.1.1 Adverse Events of Special Interests (AESI)....44
9.2 Clinical Laboratory Evaluations....45
9.3 Vital Signs....47
9.4 Physical Examination....47
9.5 Electrocardiogram....47
9.6 Gastric Biopsy....48
10.0 INTERIM ANALYSIS/OTHER ANALYSES....48
10.1 Interim Analysis....48
10.2 Coronavirus Pandemic....48
11.0 CHANGES IN THE PLANNED ANALYSIS....49
12.0 REFERENCES....50
13.0 APPENDICES....50
13.1 Schedule of Events....51
13.2 Imputation Rules for Missing Date Information....54
13.2.1 Rules for Concomitant Medication Start Date Imputation....54

13.2.2 Rules for Concomitant Medication End Date Imputation 54
13.2.3 Rules for AE Start Date Imputation....55
13.3 LA Classification of Esophagitis Grading 56
13.4 Symptom Diary....56
13.4.1 Data Handling for Subject Daily Diary 56
13.4.2 Baseline Summary of Subject Diary....57
13.4.3 Sustained Resolution of Heartburn 58
13.5 PAGI-QOL and PAGI-SYM Scoring 58
13.5.1 PAGI-QOL....58
13.5.2 PAGI-SYM 59
13.5.3 EQ-5D-5L 61
13.6 Estimands and Sensitivity Analysis....62
13.6.1 Healing Phase....62
13.6.2 Maintenance Phase....70

LIST OF IN-TEXT TABLES

Table 1 Analysis Visit Windows 19
Table 2 Adverse Events of Special Interest – Search Criteria 45
Table 3 Schedule of Events....51
Table 4 LA Grade Classification 56
Table 5 Definitions of Heartburn Severity (Daytime/Nighttime)....56
Table 6 Intercurrent Event Types....62
Table 7 Objectives, Endpoints, and Estimands....63
Table 8 Summary of Estimands and Strategies for Managing Intercurrent Events....64
Table 9 Summary of Statistical Methods and Sensitivity Analyses 67
Table 10 Intercurrent Event Types....70
Table 11 Objectives, Endpoints, and Estimands....71
Table 12 Summary of Estimands and Strategies for Managing Intercurrent Events....72
Table 13 Summary of Statistical Methods and Sensitivity Analyses 75

LIST OF IN-TEXT FIGURES

Figure 1 Study Scheme....12
Figure 2 Prior and Concomitant Medications 26
Figure 3 Testing Hierarchy....31

DOCUMENT HISTORY – CHANGES COMPARED TO FINAL VERSION 1.0

Version	Date	Changes
Final version 1.0	30-Oct-2019	Final version
Final version 2.0	25-Nov-2020	1. Information was incorporated into Section 2, Section 8.1, and Section 8.1.3. to reflect the sensitivity and supportive analyses described in Appendix 13.6. There was no change to content. 2. Changed rules on how actual treatment received is defined for safety analyses using the Safety Set in Section 4. 3. Changed significance level for non-inferiority tests from 1-sided at the 5% significance level to 1-sided at the 2.5% significance level in Section 4. 4. In Section 4.4, changed the definition of the MITT Set for both the Healing and Maintenance Phase to exclude subjects with no post-baseline endoscopy due to COVID-19 related reasons. Since these subjects will be excluded from the MITT Set, it was clarified in Section 4.5 that missing data for these subjects will not be imputed. 5. Updated definition of Per Protocol Sets in Section 4.4.4 to exclude subjects with major protocol deviations regardless of study outcome for each phase (healed or not healed during the Healing Phase and remained healed or recurred during the Maintenance Phase). 6. In Section 5 “Subject Disposition”, subsections were reorganized and the summary of protocol deviations was clarified. 7. Section 7 Treatments and Medications was updated to clarify the summarization of medications, including that ATC level 4 instead of level 2 coding will be used. 8. Updated wording describing hypothesis testing during the Maintenance Phase in Section 8 to reflect that the Hochberg procedure is an alpha-exhaustive testing procedure. The secondary endpoints for the Maintenance Phase will only be tested if both vonoprazan doses are found to be non-inferior to lansoprazole for the primary endpoint. 9. Changed the sequential testing hierarchy for the Maintenance Phase in Figure 3 to include the superiority tests for the primary endpoint and to reorder the secondary endpoints. 10. Added subgroup analyses for race and ethnicity for the primary endpoint in each phase. 11. In Section 8.2, clarified that the difference in means and the 95% CI for the percentage of 24-hour heartburn free days will be computed using Welch’s t-test. Added sensitivity analysis of 24-hour heartburn free days in which the days after a subject discontinues treatment will be imputed as non-responder days. For each phase, included the justification of the NI margin for the percentage of 24-hour heartburn free days from the protocol into the SAP for completeness and added

Version	Date	Changes
		subgroup analyses by baseline adjudicated EE grade category (A/B, C/D) for the percentage of 24-hour heartburn free days. 12. Added the details of SMQs and CQs to identify Adverse Events of Special Interest in Section 9.1.2. 13. Updated that clinical laboratory results will be summarized for both SI and US conventional units, clarified the criteria for total bilirubin as 2xULN, and added additional criteria for abnormal liver function tests in Section 9.2. 14. Added criteria for abnormal vital signs and abnormal ECG values in Section 9.3 and 9.5, respectively. 15. It was decided not to conduct an analysis at the end of the Healing Phase. Therefore, paragraphs in Section 4.2 and Section 10 "Final Analysis of the Healing Phase" that described this potential analysis were deleted and replaced with a new Section 10.1 indicating that no interim analysis is planned for the study. 16. Added Section 10.2 to describe the handling of impacts from COVID-19 pandemic. 17. Added Section 11 Changes in the Planned Analysis to describe changes from the planned analysis in the protocol including updating the definition of the MITT set to address the impact of COVID-19 pandemic and removing wording related to a potential analysis at the end of the Healing Phase. 18. In Section 13.4.1, added clarifications on the analysis of the diary data including that maximum severity will be considered as 0 (0=None) if the presence of heartburn is answered as "no" in a diary. 19. Throughout the document, additional edits were made for clarity and formatting with no change to content.
Final version 3.0	13-Aug-2021	1. Visit windows for the Week 2 visit during the Healing Phase and for the Week 24 visit during the Maintenance Phase were updated in Table 1. 2. In Section 4.4, changed the definition of the MITT Set for both the Healing and Maintenance Phase to remove the phrase that subjects with no post-baseline endoscopy due to COVID-19 related reasons will be excluded from the MITT Set (which had been added in Amendment 1). These subjects will now be included in the MITT Set and missing data due to COVID-19 will be imputed using a missing at random assumption as described in Section 4.5. This change was also reflected in revisions to Section 10.2 Coronavirus Pandemic and Section 11.0 Changes in the Planned Analysis. 3. Significant protocol deviations are to be summarized by protocol deviation subtype category.

Version	Date	Changes
		4. Additional exploratory endpoints for daytime and nighttime heartburn were added to Section 8.3. 5. Additional terms were added to the custom query for the Adverse Events of Special Interest of Bone Fracture in Section 9.1.2. 6. The MedDRA version was updated to version 23. 7. A summary of the number and percentage of subjects with elevated gastrin values was added. 8. The summarization of subjects with abnormal vital signs values was corrected to include subjects with at least 1 post-baseline abnormal vital sign value that met prespecified criteria and with a vital sign value worse than the baseline value. 9. Updated Section 11 Changes in the Planned Analysis to describe the imputation of missing data due to the COVID-19 pandemic. 10. The summary of concomitant medications was clarified. 11. Details were provided for the calculation of study drug compliance. 12. Appendix 13.6 was updated to reflect the missing data imputation for COVID-19 pandemic related reasons, including adding a new intercurrent event of unable to perform endoscopy assessment due to COVID-19 related reasons. 13. Additional edits were made for clarity and formatting with no change to content.

LIST OF ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
COVID-19	coronavirus disease 2019
CQ	customized queries
CRF	case report form
CYP2C19	cytochrome P450 2C19
ECG	electrocardiogram
EE	erosive esophagitis
EQ-5D-5L	EuroQoL-5 Dimensions-5 Levels
HCG	human chorionic gonadotropin
HIV	human immunodeficiency virus
ICF	informed consent form
LA Classification	Los Angeles Classification
MedDRA	Medical Dictionary for Regulatory Activities
MITT	modified intent-to-treat
PAGI-QoL	Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life
PAGI-SYM	Patient Assessment of Gastrointestinal Disorders-Symptom Severity Index
PPI	proton pump inhibitor
PPS	per protocol set
PT	preferred term
QD	once daily
SAE	serious adverse event
SAP	statistical analysis plan
SMQ	standardized MedDRA queries
SOC	system organ class
TEAE	treatment-emergent adverse event
WHO	World Health Organization

1.0 INTRODUCTION

Vonoprazan belongs to a new class of acid-inhibitory agents called "potassium-competitive acid blockers" and is being developed for healing of all grades of erosive esophagitis (EE) and relief of heartburn and maintenance of healing of all grades of EE and relief of heartburn.

Vonoprazan has been studied in a number of acid-related diseases including EE healing and maintenance, gastric ulcer/duodenal ulcer healing, and for the prevention of recurrence of a gastric or duodenal ulcer during nonsteroidal anti-inflammatory drugs or aspirin administration and has received regulatory approval in Japan and other countries in Asia and Latin America for these indications. Study EE-301 will further examine the effectiveness and safety of vonoprazan in subjects in the US and Europe in the healing and maintenance of healed EE.

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the statistical analyses of efficacy and safety data as described in the Protocol Amendment 2, dated 01Oct2019.

2.0 OBJECTIVES AND ESTIMANDS

2.1 Healing Phase

The primary objectives of this phase are as follows:

- To assess the efficacy of vonoprazan 20 mg compared to lansoprazole 30 mg in healing of EE over 8 weeks in subjects with endoscopically proven EE.
- To assess the safety of vonoprazan 20 mg compared to lansoprazole 30 mg in subjects with endoscopically proven EE.

The secondary objectives of this phase are as follows:

- To assess the efficacy of vonoprazan 20 mg compared to lansoprazole 30 mg in healing of EE over 2 weeks in subjects with endoscopically proven EE.
- To assess the efficacy of vonoprazan 20 mg compared to lansoprazole 30 mg in healing of EE in subjects with endoscopically proven EE LA Classification Grades C or D over 2 weeks and 8 weeks.
- To assess relief of heartburn of vonoprazan 20 mg compared to lansoprazole 30 mg in subjects with endoscopically proven EE over 8 weeks and within the first 3 days of treatment.

The exploratory objectives of this phase are as follows:

- To assess the efficacy of vonoprazan 20 mg compared to lansoprazole 30 mg in subject symptoms by subject daily diary.

The estimands of the Healing Phase are described in Table 7 of Section 13.6.1.2.

2.2 Maintenance Phase

The primary objectives of this phase are as follows:

- To assess the efficacy in maintenance of healing of vonoprazan 10 mg and vonoprazan 20 mg compared to lansoprazole 15 mg in subjects with healed EE.
- To assess the safety of vonoprazan 10 mg and vonoprazan 20 mg compared to lansoprazole 15 mg in subjects with healed EE.

The secondary objectives of this phase are as follows:

- To assess the efficacy in maintenance of healing of vonoprazan 10 mg and vonoprazan 20 mg compared to lansoprazole 15 mg in subjects with healed EE with baseline LA Classification Grades C or D.
- To assess the efficacy of vonoprazan 10 mg and vonoprazan 20 mg compared to lansoprazole 15 mg in subject symptoms by subject daily diary.

The exploratory objectives of this phase are as follows:

- To assess the efficacy of vonoprazan 10 mg and vonoprazan 20 mg compared to lansoprazole 15 mg in subject symptoms by subject daily diary.

The estimands of the Maintenance Phase are described in Table 11 of Section 13.6.2.2.

3.0 INVESTIGATIONAL PLAN

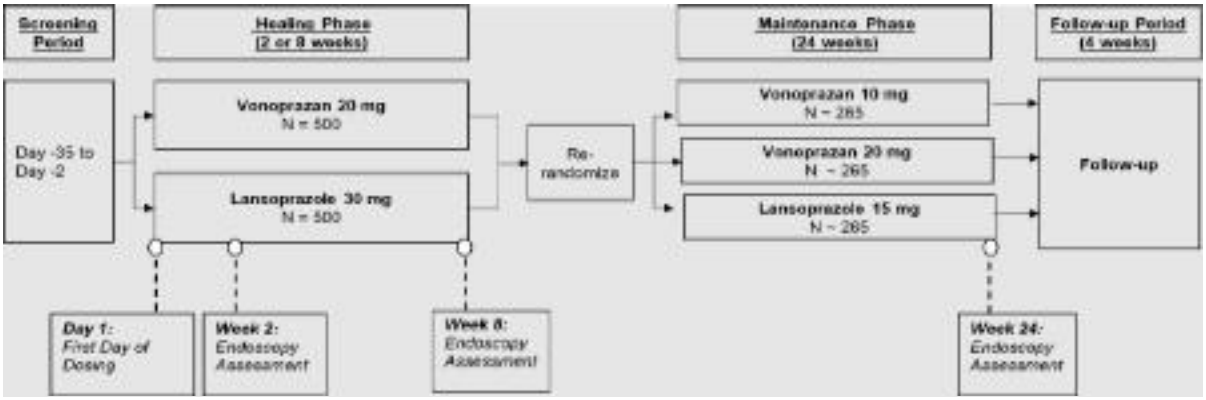
3.1 Overall Study Design and Plan

This is a Phase 3, 2-phase multi-center, double-blind, noninferiority study of vonoprazan versus lansoprazole for the healing of all grades of EE and relief of heartburn and the maintenance of healing of all grades of EE and relief of heartburn. Subjects with EE with the LA classification Grade A/B or C/D serving as a stratification at randomization (see Table 4 for LA Classification definitions) will be randomized to receive vonoprazan 20 mg or lansoprazole 30 mg given QD for up to 8 weeks for the healing of EE. All subjects with endoscopic healing of EE at 2 or 8 weeks after the start of the study will enter a continuous 24-week Maintenance Phase. Subjects will be rerandomized to receive either vonoprazan 10 mg, vonoprazan 20 mg, or lansoprazole

15 mg for 24 weeks. The subjects without confirmed endoscopic healing of EE at Week 8 will not be allowed to enter the Maintenance Phase. Subjects will complete the Patient Assessment of Gastrointestinal Disorders-Symptom Severity Index (PAGI-SYM), Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life (PAGI-QoL), and the EuroQoL-5 Dimensions-5 Levels (EQ-5D-5L) during each site visit and an electronic diary twice daily to record the presence and maximum severity of daytime and nighttime heartburn symptoms throughout the study.

A schematic diagram of the overall study design is presented in Figure 1.

Figure 1 Study Scheme



In Healing Phase, if endoscopic healing of EE is confirmed at Week 2, the subject will enter the Maintenance Phase. If endoscopic healing is not confirmed at Week 2 then the subject will continue to receive treatment until Week 8.

3.2 Study Endpoints

3.2.1 Healing Phase

The primary efficacy endpoint of the Healing Phase is as follows:

- The percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy

The secondary efficacy endpoints of the Healing Phase are as follows:

- The percentage of 24-hour heartburn–free days over the Healing Phase as assessed by the daily diary
- The percentage of subjects who have complete healing of EE at Week 2 as assessed by endoscopy for subjects with baseline LA Classification Grades C or D

- The percentage of subjects with onset of sustained resolution of heartburn by Day 3 (sustained resolution is defined as at least 7 consecutive days with no daytime or nighttime heartburn as assessed by the daily diary)
- The percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy for subjects with baseline LA Classification Grades C or D
- The percentage of subjects who have complete healing of EE at Week 2 as assessed by endoscopy

The exploratory efficacy endpoints of the Healing Phase are as follows:

- The percentage of days without daytime heartburn over the Healing Phase as assessed by the daily diary
- The percentage of days without nighttime heartburn over the Healing Phase as assessed by the daily diary
- The mean severity of daytime and nighttime heartburn over the Healing Phase as assessed by the daily diary
- The mean severity of daytime heartburn over the Healing Phase as assessed by the daily diary
- The mean severity of nighttime heartburn over the Healing Phase as assessed by the daily diary
- The time to sustained resolution of heartburn (sustained resolution is defined as at least 7 consecutive days with no daytime or nighttime heartburn as assessed by the daily diary)
- The change from baseline to the end of the Healing Phase for each subscale and the total score of the PAGI-SYM questionnaire
- The change from baseline to the end of the Healing Phase for each subscale and the total score of the PAGI-QoL questionnaire
- The change from baseline to the end of the Healing Phase for the EQ-5D-5L

3.2.2 Maintenance Phase

The primary efficacy endpoint of the Maintenance Phase is as follows:

- The percentage of subjects who maintain complete healing of EE after 24 weeks as assessed by endoscopy

The secondary efficacy endpoints of the Maintenance Phase are as follows:

- The percentage of 24-hour heartburn–free days over the Maintenance Phase as assessed by the daily diary
- The percentage of subjects who maintain complete healing of EE after 24 weeks as assessed by endoscopy for subjects with baseline LA Classification Grades C or D

The exploratory efficacy endpoints of the Maintenance Phase are as follows:

- The percentage of days without daytime heartburn over the Maintenance Phase as assessed by the daily diary
- The percentage of days without nighttime heartburn over the Maintenance Phase as assessed by the daily diary
- The mean severity of daytime and nighttime heartburn over the Maintenance Phase as assessed by the daily diary
- The mean severity of daytime heartburn over the Maintenance Phase as assessed by the daily diary
- The mean severity of nighttime heartburn over the Maintenance Phase as assessed by the daily diary
- The percentage of 24-hour heartburn-free days over the Follow-up Phase as assessed by the daily diary
- The change from maintenance baseline to the end of the Maintenance Phase for each subscale and the total score of the PAGI-SYM questionnaire
- The change from maintenance baseline to the end of the Maintenance Phase for each subscale and the total score of the PAGI-QoL questionnaire
- The change from maintenance baseline to the end of the Maintenance Phase for the EQ-5D-5L

3.3 Treatments

3.3.1 Healing Phase

Subjects who meet all of the inclusion criteria and none of the exclusion criteria will be randomized on Day -1 of the Healing Phase to receive blinded treatments for the Healing Phase using a 1:1 allocation ratio.

The blinded treatments for the Healing Phase are:

- Vonoprazan 20 mg, taken orally for up to 8 weeks, once daily
- Lansoprazole 30 mg, taken orally for up to 8 weeks, once daily

3.3.2 Maintenance Phase

Upon verification of healing, subjects will be randomized to receive blinded treatments for the Maintenance Phase using a 1:1:1 allocation ratio.

The blinded treatments for the Maintenance Phase are:

- Vonoprazan 10 mg, taken orally for 24 weeks, once daily
- Vonoprazan 20 mg, taken orally for 24 weeks, once daily
- Lansoprazole 15 mg, taken orally for 24 weeks, once daily

3.4 Dose Adjustment/Modifications

No dose adjustments or modifications are allowed for this study.

4.0 GENERAL STATISTICAL CONSIDERATIONS

In general, descriptive statistics will be presented by treatment group and by visit, as applicable. For continuous variables, summary statistics for the raw value and change from baseline at each time-point will include the number of subjects (n), arithmetic mean, standard deviation (SD), median, minimum and maximum.

Categorical variables will be summarized using subject counts and percentages. Percentages will be calculated using the total subjects per treatment unless otherwise specified.

The efficacy analyses will be conducted on the Modified Intent-to-Treat Set and Per Protocol Set using the planned treatment. See Section 4.4 for the Modified Intent-to-Treat Set and Per Protocol Set definitions.

The safety analyses will be conducted on the Safety Set using the actual treatment the subject received. If a subject receives more than one type of study drug during a phase, the planned treatment will be assigned as the actual treatment received for this subject in the safety summary tables for this phase. If a subject receives only one type of study drug during a phase, but this study drug type is inconsistent with the planned treatment, the actual treatment received will be assigned as the actual treatment for this subject in the safety summary tables for this phase. See Section 4.4 for the Safety Set definition.

Statistical tests for noninferiority will be 1-sided and will be conducted at the 2.5% significance level. All the other statistical tests will be 2-sided and will be conducted at the 5% significance level, unless otherwise specified. P-values will be reported to 4 decimal places, with p-values less than 0.0001 reported as “<0.0001”.

SAS® version 9.4 or higher will be used to perform all statistical analyses or procedures.

4.1 Sample Size

4.1.1 Healing Phase

A sample size of 500 subjects per treatment group provides at least 90% power to achieve noninferiority using a Farrington Manning test with a noninferiority margin of 10%, assuming EE healing rates by Week 8 of 80% and 80% for vonoprazan and lansoprazole, respectively.

A fixed noninferiority margin of -10% is justified based on historical results from 2 randomized, double-blind, placebo-controlled studies that evaluated EE healing rates after 8 weeks with lansoprazole, the active comparator in the current study. In these studies, the EE healing rates with placebo were compared to lansoprazole 15 mg, lansoprazole 30 mg, or lansoprazole 60 mg (Earnest et al 1998) or to lansoprazole 15 mg, lansoprazole 30 mg, or omeprazole 20 mg (Castell et al 1996). The difference in the EE healing rate between lansoprazole 30 mg and placebo reported by Earnest was 40% (87% vs 47%) and by Castell was 55% (87% vs 32%) with lower limits of the 2-sided 95% confidence interval of 26% and 48%, respectively. Using the most conservative assumption of the lower bound of 26% as the estimate of the treatment effect of lansoprazole 30 mg compared to placebo, a -10% noninferiority margin would assure that vonoprazan retains at least 60% of the treatment effect of lansoprazole 30 mg.

The current study will enroll approximately 30% of subjects with more severe EE (Grade C or Grade D). The historical EE healing rates are from studies in which the percentage of subjects with more severe EE (Grade 3 or Grade 4) ranged from 35% (Castell et al 1996) to 63% (Earnest et al 1998). The grading scale used in the historical data was different but does provide information that the percentage of subjects with more severe EE that were enrolled in these studies is similar or higher than the 30% of Grade C or Grade D subjects planned for the current study. Therefore, these rates are considered relevant to informing the noninferiority margin for the primary endpoint for the Healing Phase and a fixed noninferiority margin of -10% is justified.

4.1.2 Maintenance Phase

All subjects who have complete healing of EE during the Healing Phase will enter the Maintenance Phase. It is expected that approximately 800 subjects will enter the Maintenance Phase, but the actual number of subjects who enter will depend on the observed healing rates during the Healing Phase. During the study if less than 800 subjects are projected to enroll into the Maintenance Phase, additional subjects may be enrolled to ensure that a sufficient number of subjects enter the Maintenance Phase.

For the maintenance of EE healing after 24 weeks, a sample size of 265 subjects per treatment group provides at least 90% power to achieve noninferiority with a noninferiority margin of 10%

and at least 90% power to achieve superiority using the Farrington Manning test, assuming maintenance of EE healing rates of 82% and 70% for vonoprazan and lansoprazole, respectively.

Two randomized, double-blind studies were conducted that evaluated maintenance of EE healing after a 12-month maintenance treatment period with lansoprazole 15 mg QD, lansoprazole 30 mg QD, or placebo (PREVACID PI 2018). In these 2 studies, the differences in the Month 6 maintenance of EE healing rate between lansoprazole 15 mg and placebo were 54% (81% vs 27%) and 59% (72% vs 13%) with lower limits of the 2-sided 95% confidence interval of 39% and 43%, respectively. Using the most conservative assumption of the lower bound of 39% as the estimate of the treatment effect of lansoprazole 15 mg compared to placebo, a -10% noninferiority margin would assure that vonoprazan retains at least 75% of the treatment effect of lansoprazole 15 mg. In these historical maintenance studies, the percentage of subjects with Grade 3 or Grade 4 EE was approximately 45% which is similar or higher than the 30% of subjects with Grade C or Grade D EE planned for the current study. Therefore, these rates are considered relevant to informing the noninferiority margin for the primary endpoint for the Maintenance Phase and a fixed noninferiority margin of -10% is justified.

4.2 Randomization and Blinding

Subjects will be randomized to receive vonoprazan 20 mg QD or lansoprazole 30 mg QD using a 1:1 allocation ratio in the Healing Phase of the study. Upon verification of healing, subjects will be randomized to receive either vonoprazan 10 mg QD, vonoprazan 20 mg QD, or lansoprazole 15 mg QD for 24 weeks in the Maintenance Phase of the study using a 1:1:1 allocation ratio. The randomization for both the Healing and Maintenance Phases will be stratified by baseline LA classification (Grades A/B and Grades C/D).

4.3 Assessment Windows

4.3.1 Study Day

The date of first dosing day is defined as Day 1 for both the Healing and Maintenance Phases. When study day is used for display or in comparisons the following algorithm will be used:

- study day = date of assessment - Day 1 +1,
  if date of assessment ≥ Day 1.
- study day = date of assessment - Day 1,
  if date of assessment < Day 1.

Note that the date of randomization is defined as Day -1. There is no Day 0 for study day.

4.3.2 Unscheduled Endoscopy Assessments

During the Healing Phase, if the investigator deems that an additional endoscopy is warranted based on the subject's symptoms prior to the window of the subject's next scheduled per protocol endoscopy, an unscheduled endoscopy will be permitted whenever required as part of routine clinical care. If the subject's EE is found to have healed at this unscheduled endoscopy visit, the subject will be considered to have completed the Healing Phase. If the subject is found to still have EE (of any grade) at the unscheduled endoscopy visit, the subject will be allowed to continue in the study within the Healing Phase and will undergo their subsequent per-protocol endoscopy.

During the Maintenance Phase, if the investigator deems that an endoscopy is warranted, based on subject symptoms, an unscheduled endoscopy will be permitted whenever required. If the subject's EE is found to have relapsed at this unscheduled endoscopy visit, the subject will be discontinued from the Maintenance Phase of the study, considered a treatment failure, and will enter the Follow-up Phase. If the subject is found to have maintained healing of EE at the unscheduled endoscopy visit, the subject will be allowed to continue in the study until Week 24.

All subjects who discontinue study drug or withdraw from the study prematurely will undergo all end-of-study assessments, including the endoscopy, whenever possible.

For purposes of reporting unscheduled endoscopy assessments during the Healing Phase, a "By Week 8" will be identified for endoscopy assessments as the last valid endoscopy assessment obtained after baseline and while on study drug of this phase. The number and percentage of subjects who have unscheduled endoscopy assessments will be reported for each phase by treatment group and overall. All scheduled and unscheduled endoscopy assessments will be listed.

4.3.3 Visit Window for Analysis

Visit windows will be defined for by-visit summary and analysis purposes. Summary data (such as AEs and concomitant medications) that are not reported by visit will not use visit windows. Both scheduled and unscheduled assessments will be considered as valid assessments for analysis. Visit labels will be assigned to each post-baseline record based on the windows for study day relative to the date of first dose. For the Healing Phase, if an assessment on Day -1 is missing, the closest visit with non-missing assessment on or before the date of first dose will be used as baseline. For the Maintenance Phase, the closest visit in the Healing Phase with non-missing assessment on or before the date of re-randomization will be used as baseline. All by-visit summary and analysis will be based on the analysis visit windows in Table 1.

Table 1 Analysis Visit Windows

Phase	Nominal Visit (recorded on eCRF)	Analysis Visit	Target Study Day of Visit	Analysis Visit Window
Healing	Visit 1	Screening	-35 to -2	NA
Healing	Visit 2	Baseline	-1	NA
Healing	Visit 3	Week 2	15	8 to 22 days
Healing	Visit 4	Week 8	57	23 to 71 days
Maintenance	Visit M1	Week 4	29	26 to 57 days
Maintenance	Visit M2	Week 12	85	58 to 99 days
Maintenance	Visit M3	Week 16	113	100 to 127 days
Maintenance	Visit M4	Week 20	141	128 to 155 days
Maintenance	Visit M5	Week 24	169	156 or greater

Unscheduled assessments will also be assigned to analysis visits based on the analysis visit window. When data is summarized by assigned analysis visit based on study day, visits will be referenced in summary tables by analysis visits only. Listings will present both nominal visits as recorded on the eCRF, and the analysis visits. After all the records have been assigned to an analysis visit window based on study day, if there are multiple valid records for an assessment within an assigned analysis visit, only one of these records will be used for summary statistics and analyses. The record to be used is determined using the following hierarchy (in descending order):

Safety assessments, including laboratory tests, ECG and vital signs:

- the record closest to the target visit day
- the latest visit in the analysis visit window

Endoscopy assessments:

- the latest visit in the analysis visit window

4.4 Analysis Set

4.4.1 Screened Set

All subjects who signed the informed consent form (ICF) before entering the Healing Phase. Screen failures are defined as subjects who were not randomized into the Healing Phase.

4.4.2 Randomized Set

4.4.2.1 Healing Phase

All subjects randomly assigned to receive study drug regardless of whether or not they received a dose of study drug during the Healing Phase.

4.4.2.2 Maintenance Phase

All subjects randomly assigned to receive study drug regardless of whether or not they received a dose of study drug during the Maintenance Phase.

4.4.3 Modified Intent-to-Treat (MITT) Set

4.4.3.1 Healing Phase

The MITT set will be defined as all subjects randomized into the Healing Phase who have documented EE at baseline and receive at least 1 dose of study drug during the Healing Phase. All analyses using the MITT set will group subjects according to the randomized treatment.

4.4.3.2 Maintenance Phase

The MITT set will be defined as all subjects randomized into the Maintenance Phase who have healed EE at the end of the Healing Phase and receive at least 1 dose of study drug during the Maintenance Phase. All analyses using the MITT set will group subjects according to the randomized treatment.

4.4.4 Per Protocol (PP) Set

4.4.4.1 Healing Phase

The PP Set will consist of all MITT subjects who have all of the following:

- at least 80% and no greater than 120% compliance with study treatment
- have not taken a proton pump inhibitor or H2 receptor antagonist during the Healing Phase
- endoscopy performed by Week 8
- have no other major protocol deviations that could affect the primary analysis of the Healing Phase

Per definition, the PP Set will only include subjects who received the study drug to which they were randomly assigned. All analyses using the PP Set will group subjects according to the randomized treatment.

4.4.4.2 Maintenance Phase

The PP Set will consist of all MITT subjects who have all of the following:

- at least 80% and no greater than 120% compliance with study treatment
- have not taken a proton pump inhibitor or H2 receptor antagonist during the Maintenance Phase
- endoscopy performed at Week 24
- have no other major protocol deviations that could affect the primary analysis of the Maintenance Phase.

Per definition, the PP Set will only include subjects who received the study drug to which they were randomly assigned. All analyses using the PP Set will group subjects according to the randomized treatment.

4.4.5 Safety Set

4.4.5.1 Healing Phase

The safety set will be defined as all randomized subjects who receive at least 1 dose of study drug during the Healing Phase. All analyses using the safety set will group subjects according to the treatment actually received.

4.4.5.2 Maintenance Phase

The safety set will be defined as all randomized subjects who receive at least 1 dose of study drug during the Maintenance Phase. All analyses using the safety set will group subjects according to the treatment actually received.

4.5 Missing Data Handling for Endoscopy Assessments

If the post-baseline endoscopy was missing due to the COVID-19 pandemic, the missing data will be imputed using a missing at random (MAR) assumption. Subjects with missing endoscopy data not due to the COVID-19 pandemic will be imputed as non-responders.

For the primary endpoint during the Healing Phase, subjects with missing data not due to the COVID-19 pandemic, ie, those who do not have a post-baseline endoscopy by Week 8, will be considered “not healed” for the primary analysis.

For the primary endpoint during the Maintenance Phase, subjects with missing data not due to the COVID-19 pandemic, ie, those who do not have endoscopy results at Week 24, will be considered “recurred” for the primary analysis.

The analyses of the secondary endpoints assessed by endoscopy will follow the same missing data handling as the primary endpoints for both Healing Phase and Maintenance Phase.

For subgroup analyses for the primary endpoint during both the Healing Phase and the Maintenance Phase, all missing endoscopy results will be imputed as non-responders, given the complexity of implementing the multiple imputation for missing data due to COVID-19 related reasons within each subgroup for the relatively low number of subjects with missing results within each subgroup.

5.0 SUBJECT DISPOSITION

5.1 Disposition

5.1.1 Screened and Screen Failure Subjects

The number of screened and screen failure subjects will be presented for overall subjects included in the Screened Set.

The following will be summarized for the Screened Set:

- The total number of screened subjects.
- The number and percentage of screen failures.
- The number and percentage of each primary reason for screen failures.

Subjects who fail to fulfill all inclusion/exclusion criteria will be listed for the Screened Set.

5.1.2 Healing Phase Randomized Subjects

The number of subjects included in each analysis set will be presented by treatment group and overall for all randomized subjects.

The following will be summarized for the Healing Phase Randomized Set:

- The number of randomized subjects.
- The number and percentage of subjects who completed the treatment period of the Healing Phase.
- The number and percentage of subjects who discontinued from the treatment period of the Healing Phase.
- The number and percentage of subjects who entered the Maintenance Phase.
- The number and percentage of subjects who did not enter the Maintenance Phase.
- The number and percentage of subjects who completed the participation of the entire study after Discontinuing from the Healing Phase.

- The number and percentage of subjects who completed or discontinued from the follow-up period after Discontinuing from the Healing Phase.
- Reasons for not entering the Maintenance Phase.
- Reasons for discontinuation from the study participation.

Percentages will be based on the number of subjects in the Randomized Set.

Subject disposition data for subjects who completed the study will be listed for the Randomized Set. Disposition data will be listed separately for subjects who discontinued during the Healing Phase for the Randomized Set.

5.1.3 Maintenance Phase Randomized Subjects

The number of subjects included in each analysis set will be presented by treatment group and overall for all randomized subjects.

The following will be summarized for the Randomized Set:

- The number of randomized subjects.
- The number and percentage of subjects who completed the treatment period of the Maintenance Phase.
- The number and percentage of subjects who discontinued from the treatment period of the Maintenance Phase.
- The number and percentage of subjects who completed the participation of the entire study.
- The number and percentage of subjects who completed or discontinued from the follow-up period of the Maintenance Phase.
- Reasons for discontinuation from the study participation.

All percentages will be based on the number of subjects in the Randomized Set.

Subject disposition data for the treatment period of the Maintenance Phase will be listed for the Randomized Set. Disposition data for the treatment period will be listed separately for subjects who discontinued from the study for the Randomized Set.

5.1.4 Healing Phase Treatment Period

The following will be summarized for the Healing Phase Safety Set:

- The number and percentage of subjects in the Safety Set, MITT Set, and PP Set.
- The number and percentage with each reason for exclusion from the MITT Set and the PP Set.
- Reasons for discontinuation from the study drug.

Subject disposition data for the treatment period of the Healing Phase will be listed for the Safety Set.

5.1.5 Maintenance Phase Treatment Period

The following will be summarized for the Maintenance Phase Safety Set:

- The number and percentage of subjects in the Safety Set, MITT Set, and PP Set.
- The number and percentage with each reason for exclusion from the MITT Set and the PP Set.
- Reasons for discontinuation from the study drug.

Subject disposition data for the treatment period of the Maintenance Phase will be listed for the Safety Set.

5.2 Protocol Deviations

Protocol deviations will be recorded within the [redacted]Clinical Trial Management System (CTMS) and will undergo a blinded review prior to database lock and unblinding. Significant protocol deviations are defined as the subset of deviations which are considered to affect primary efficacy and safety assessments, the safety or mental integrity of a subject, or the scientific value of the trial.

For each phase, the number and percentage of subjects with subject-specific significant protocol deviations will be summarized by CTMS activity subtype, treatment group and overall for the Randomized Set. Individual subject protocol deviations, both significant and non-significant, will be presented in a by-treatment, by-subject data listing using the Randomized Set.

6.0 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

6.1 Demographics and Baseline Characteristics

Demographic variables collected at Screening, such as age, sex, race, ethnicity, height (cm), weight (kg), and body mass index (BMI) will be summarized for both phases. Continuous variables, including age (years), body mass index, weight, height, mean severity of daytime and nighttime heartburn will be summarized using descriptive statistics for each treatment group and overall. The following categorical variables will be summarized by reporting the number and percentage of subjects in each category for each treatment group and overall using the MITT Set, PP Set, and Safety Set.

- Age group at Screening (<45, ≥45 - <65, ≥65 - <75, ≥75)
- Age group 2 at Screening (≥18 - ≤64, ≥65 - ≤84, ≥85)

- Sex (Male, Female)
- BMI category (<25, ≥25 - <30, ≥30)
- LA Classification Adjudicator Grade (A, B, C, D)
- LA Classification Adjudicator Grade (A/B, C/D)
- LA Classification Investigator Grade (A, B, C, D)
- LA Classification Investigator Grade (A/B, C/D)
- Serum gastrin (<200 pg/mL, ≥200 pg/mL)
- Pepsinogen I/II level (≤2, >2 - ≤3, >3)
- CYP2C19 status (extensive metabolizer, poor metabolizer)
- Smoking status (never smoked, current smoker, ex-smoker)
- Alcohol use (drink every day, drink a couple of days per week, drink a couple of days per month, never drink)
- Country

In addition, the following variables will be summarized based on symptoms reported in the daily diary from the last 7 days prior to Day-1 of the Healing Phase:

- Mean severity of:
  - Daytime/Nighttime Heartburn (0, >0 - ≤1, >1 - ≤2, >2 - ≤3, >3 - ≤4)
  - Heartburn during the day (0, >0 - ≤1, >1 - ≤2, >2 - ≤3, >3 - ≤4)
  - Heartburn at night (0, >0 - ≤1, >1 - ≤2, >2 - ≤3, >3 - ≤4)

- Number of days with:
  - Daytime or nighttime heartburn (≥0 - ≤3, >3 - ≤5, >5 - ≤7)
  - Daytime heartburn (≥0 - ≤3, >3 - ≤5, >5 - ≤7)
  - Nighttime heartburn (≥0 - ≤3, >3 - ≤5, >5 - ≤7)

Refer to Section 13.4.2 for calculations of baseline summary of symptom diary variables. Demographic and baseline characteristics data will be listed using the Randomized Set.

6.2 Medical History

6.2.1 General Medical History

Medical history will be coded using Version 23 of the Medical Dictionary for Regulatory Activities (MedDRA). The primary system organ class (SOC), preferred term (PT), verbatim term, start date, and stop date or indication of ongoing, will be collected on the eCRF. The number and percentage of subjects with medical history coded to each MedDRA SOC and PT will be summarized by treatment group and overall using the Safety Set.

Each subject's medical history will be listed by MedDRA SOC and verbatim term using the Safety Set.

7.0 TREATMENTS AND MEDICATIONS

7.1 Prior and Concomitant Medications

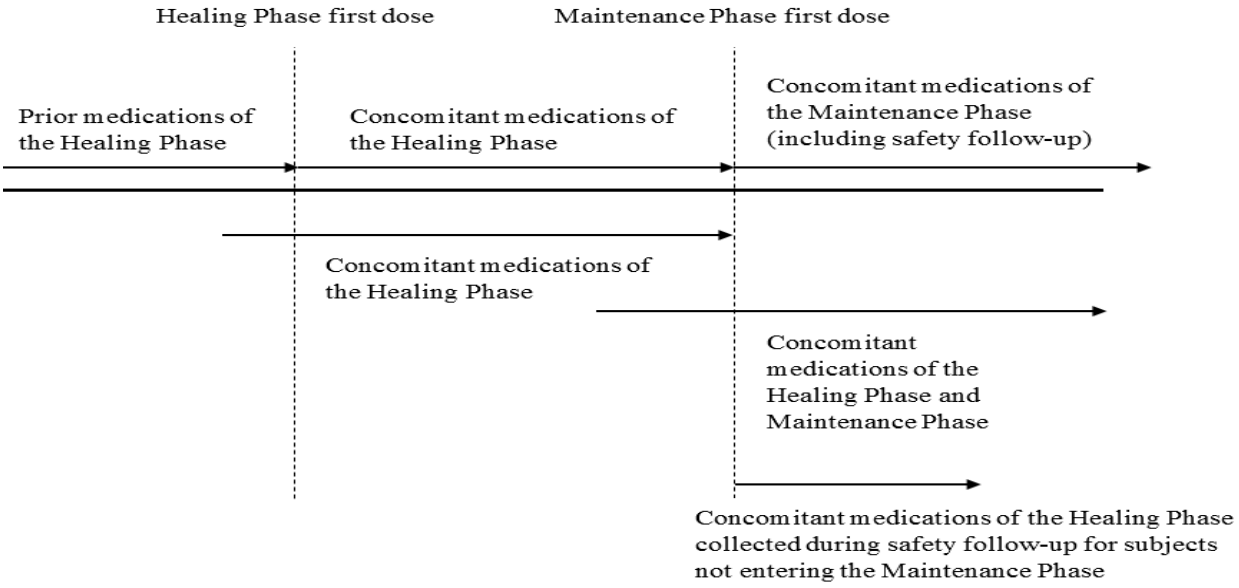
Any prior and concomitant medication used during the study will be recorded and coded using WHODRUG Version B3-March 2019. Summaries of all medications by drug class (ATC Level 4 coding) and preferred term will be provided separately for prior medications and concomitant medications for each treatment group and overall. All summaries will be performed using the Safety Set.

For the Healing Phase, prior medications are those with the start and stop dates prior to the first dose of Healing Phase study drug. Concomitant medications are those with start dates prior to the first dose of Healing Phase and continuing after the first dose of Healing Phase or with start dates between the first dose of Healing Phase and the first dose of Maintenance Phase.

For the Maintenance Phase, concomitant medications are those with start dates prior to the first dose of Maintenance Phase and continuing after the first dose of Maintenance Phase or with start dates on or after the first dose of Maintenance Phase.

See Figure 2 for details on the categorization of prior and concomitant medications.

Figure 2 Prior and Concomitant Medications



For the Healing Phase, the prior medications and concomitant medications will be summarized separately by treatment group and overall. For subjects who did not enter the Maintenance Phase, concomitant medications with start dates after the last dose of treatment in the Healing Phase will be summarized by treatment group and overall.

For the Maintenance Phase, the concomitant medications with start dates on or after the first dose of the Maintenance Phase will be summarized by treatment group and overall. The concomitant medications that are ongoing at the end of the Healing Phase, i.e. with start dates before the first dose of the Maintenance Phase, will be summarized by treatment group and overall. All concomitant medications, including those with start dates before, on or after the first dose of the Maintenance Phase, will also be summarized by treatment group and overall.

In instances where a medication start date is incomplete, it will be conservatively imputed to determine whether or not the medication was prior or concomitant. If the start date is missing, then it will be assumed to be concomitant. Imputation details for missing concomitant medication start and end date are presented in Section 13.2.

For each phase, all prior and concomitant medications will be listed by treatment and subject for the Safety Set.

7.2 Study Treatments

7.2.1 Extent of Exposure

Descriptive summary statistics including the number of subjects, mean, standard deviation, median, minimum, and maximum for the duration of study drug exposure (days) will be presented by treatment group and overall.

For the Healing Phase, treatment duration of exposure will be categorized and summarized as follows: ≥1 to ≤14 days, >15 to ≤28 days, >28 to ≤56 days, >56 days. For the Maintenance Phase, treatment duration of exposure will be categorized and summarized as follows: ≥1 to ≤28 days, >29 to ≤56 days, >57 to ≤84 days, >85 to ≤112 days, >113 to ≤140 days, >141 to ≤168 days and >169 days.

Treatment exposure will be calculated as the number of days from first to last dose date of each phase:

Exposure = Date of last dose – Date of first dose + 1

The date of last dose for each phase is recorded on the End of Treatment (EOT) page of that phase on the eCRF. If the date of last dose of the Healing Phase is missing, it will be imputed as follows:

If a subject entered the Maintenance Phase,

- Imputed date of last dose = the date of Maintenance Phase randomization – 1

if a subject did not enter the Maintenance Phase,

- Imputed date of last dose = the date of first dose of the Healing Phase + 27 for subjects if the latest available kit was dispensed on Day -1.
- Imputed date of last dose = the latest kit dispense date + 27 for subjects if the latest available kit was dispensed after date of first dose.

If the date of last dose of the Maintenance Phase is missing, it will be imputed as the latest kit dispense date + 83. The date of first dose for each phase is recorded on the Drug Accountability page on the eCRF. If the date of first dose for the Healing Phase is missing, it will be imputed using the date of randomization + 1. If the date of first dose for the Maintenance Phase is missing, it will be imputed using the date of last available visit during the Healing Phase + 1. All summaries for the treatment period will be performed using the Safety Set.

The exposure summary, including person-time (overall total exposure in years), will also be presented for the treatment types across the two phases:

- Vonoprazan to vonoprazan
- Vonoprazan to lansoprazole
- Lansoprazole to vonoprazan
- Lansoprazole to lansoprazole

All data for treatment exposure during the treatment period will be listed using the Safety Set.

7.2.2 Treatment Compliance

For each phase, treatment compliance will be calculated as:

> Compliance (%) = (total actual capsules taken / total expected capsules) × 100, where
> total expected capsules = (date of last dose of this phase – date of first dose of this phase + 1) and
> total actual capsules taken = total number of capsules dispensed for this phase – total number of capsules returned for this phase

For subjects with at least one kit not returned, the compliance will be imputed using the following rules for each phase:

- If no kits are returned among the kits dispensed during a phase, the compliance is imputed as 100% for this phase.

- If only partial kits are returned among the kits dispensed during a phase, the returned amount for kits not returned will be imputed as 7 capsules, so that the taken amount is 35-7=28 capsules for each kit that is not returned.

- If a subject has returned amount imputed for at least one kit, which results in a treatment compliance of >120%, the treatment compliance will be set to 100% for this phase.

Compliance rate will be summarized for each treatment group and overall.

Overall compliance information will be used to categorize subjects as being either compliant or not. A subject is considered compliant if the overall study drug compliance is greater than or equal to 80% and less than or equal to 120%.

Treatment compliance for the treatment period will be summarized for the Safety Set. Summary statistics for treatment compliance percentages and compliance categories (<80, 80 - ≤100, 100 - ≤120, >120) will be summarized for each treatment group and overall. Individual subject compliance information will be listed using the Safety Set.

8.0 EFFICACY ANALYSIS

Healing Phase:

The hypothesis-testing of the primary and secondary endpoints in the Healing Phase will be adjusted for multiple comparisons using a fixed-sequence testing procedure to control the Type 1 error rate at the 0.05 level. The statistical tests will be performed at α =0.05 in the sequential order shown in Figure 3 until a test is not significant. If the primary hypothesis for the comparison between the treatment groups is statistically significant (α=0.05) in the MITT Population, then the primary objective of the trial will have been achieved and the next hypothesis can be tested; otherwise, testing will stop. The testing of each subsequent hypothesis is conditioned upon all the previous hypotheses being rejected at the 0.05 level of significance. If a hypothesis is not rejected at the 0.05 level of significance, then all remaining hypotheses are deemed not statistically significant. Only p-values that are significant according to this sequential order in Figure 3 are inferential and statistically significant. All the other p-values are descriptive.

Maintenance Phase:

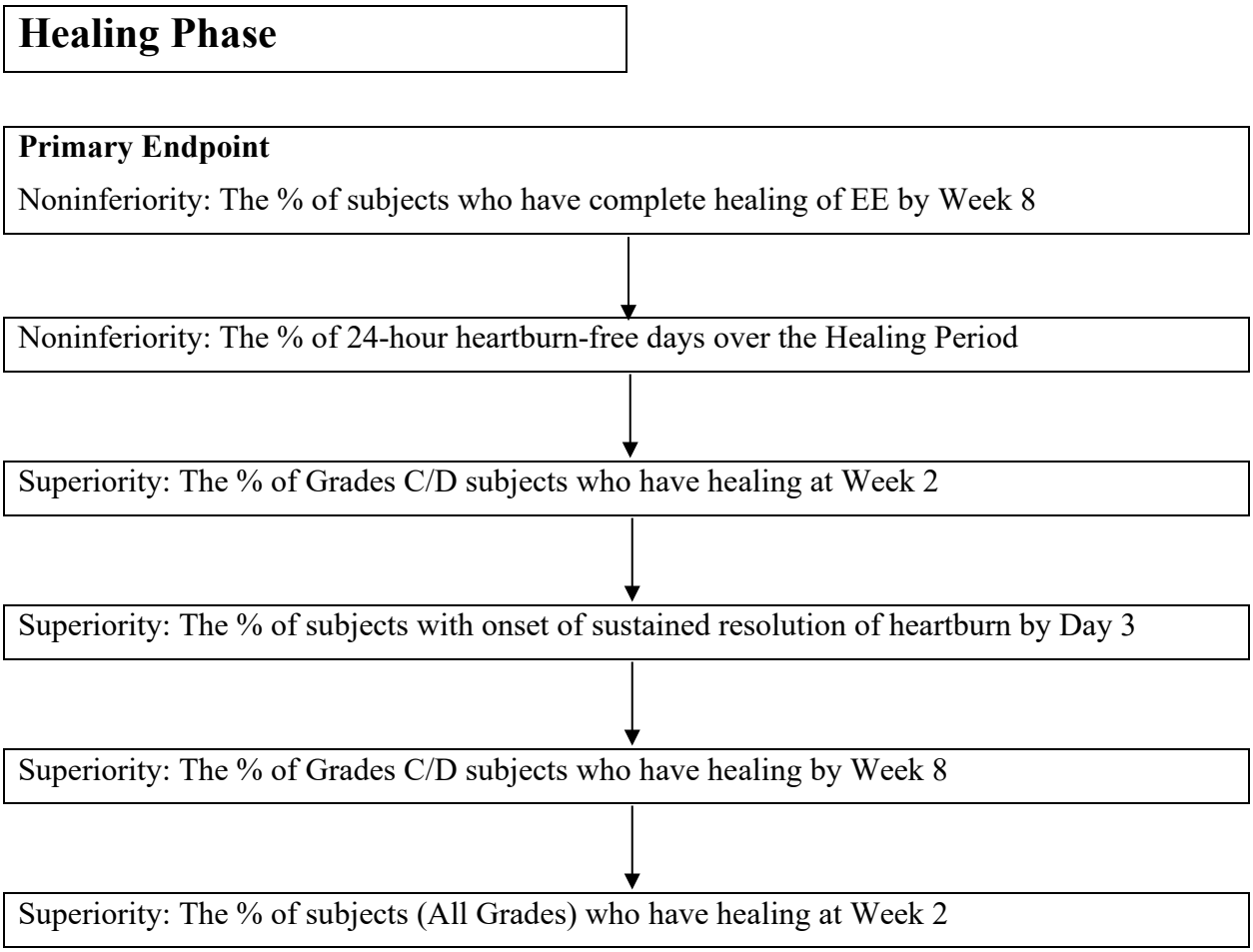
The hypothesis-testing of the primary endpoint of the Maintenance Phase will be adjusted using Hochberg's multiple comparisons method to control the overall 0.05 level of significance to test the non-inferiority of each dose group of vonoprazan to lansoprazole. If both vonoprazan dose groups have p-value ≤0.05, both dose groups are found to be non-inferior to lansoprazole 15 mg. Otherwise, if one dose group has p-value > 0.05 and the other dose group has p-value ≤0.025, the dose group with p-value ≤0.025 is found to be non-inferior to lansoprazole 15 mg.

If non-inferiority is declared for both vonoprazan dose groups from the primary efficacy analysis, the comparisons to lansoprazole 15 mg for the secondary efficacy endpoints will be performed using a fixed-sequence testing procedure shown in Figure 3 to control the Type 1 error rate at the 0.05 level until a test is not significant.

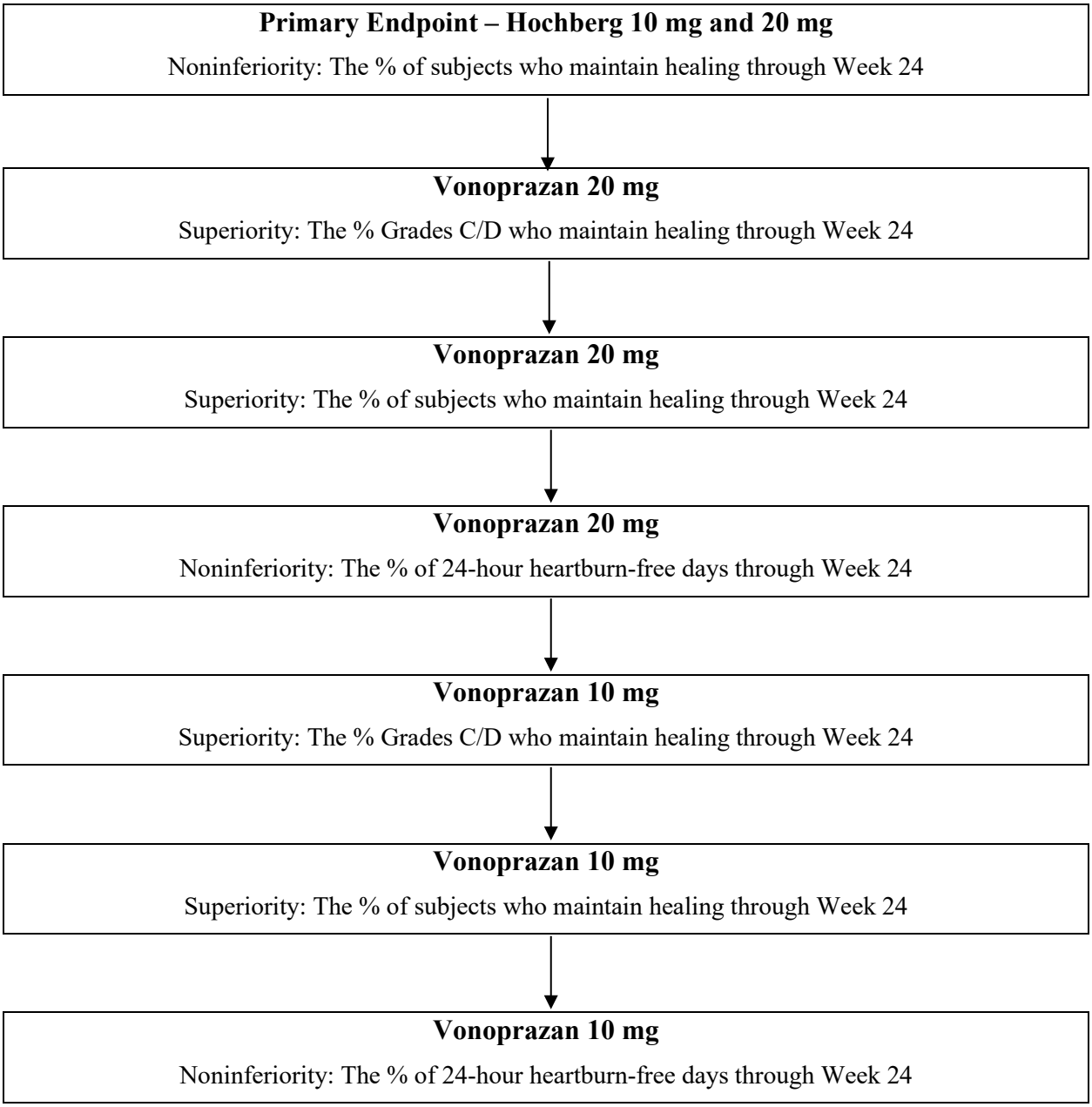
If non-inferiority is not declared for either vonoprazan dose group from the primary efficacy analysis, the testing procedure will end.

During the Maintenance Phase, additional comparisons will be made between the two vonoprazan dose groups for each primary and secondary efficacy endpoint with no adjustment for multiple comparisons.

Figure 3 Testing Hierarchy



Maintenance Phase



8.1 Primary Efficacy Endpoint

The primary efficacy endpoint of the Estimand 1 (primary) of Healing Phase is the percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy.

The primary efficacy endpoint of the Estimand 1 (primary) Maintenance Phase is the percentage of subjects who maintain complete healing of EE after 24 weeks as assessed by endoscopy.

See Section 13.6 for the primary endpoints of the additional estimands defined for the Healing Phase and Maintenance Phase.

8.1.1 Primary Analysis

Healing Phase:

For the Healing Phase, the frequency of EE healing rate by Week 8 will be summarized for each treatment group. A subject will be considered to have "complete healing of EE by Week 8" if the subject demonstrates healing at the Week 2, an unscheduled visit between Week 2 and Week 8 or Week 8 endoscopy.

Non-inferiority of vonoprazan 20 mg to lansoprazole 30 mg will be evaluated with a Farrington and Manning test with a non-inferiority margin of 10% for the difference in EE rates by Week 8 between treatments (vonoprazan 20 mg - lansoprazole 30 mg). The point estimate and 2-sided 95% CI of the difference in endoscopic healing rate between vonoprazan 20 mg and lansoprazole 30 mg will be calculated via the Miettinen and Nurminen method. The primary analysis will be performed using the MITT Set.

If non-inferiority is shown, superiority of vonoprazan 20 mg to lansoprazole 30 mg by Week 8 will also be assessed via the Farrington and Manning test of the null hypothesis difference ≤ 0 versus the alternative hypothesis difference >0 as an exploratory analysis.

All efficacy data will be listed using the Randomized Set.

The number of subjects who had endoscopies at each visit, including both scheduled and unscheduled visits, will be summarized by analysis visit for each treatment group.

Efficacy endpoints assessed by results from endoscopy will be based on the assessment from the central adjudication. The differences between central adjudication and investigators assessments of healing/not healed and LA Grade will be summarized in a table.

Maintenance Phase:

For the Maintenance Phase, the frequency of subjects who maintain complete healing will be summarized for Week 24 for each treatment group. The maintenance of healing during the 24-week Maintenance Phase will be calculated along with 2-sided 95% CIs for each treatment group. Non-inferiority of each dose group of vonoprazan to lansoprazole (i.e. vonoprazan 10 mg to lansoprazole 15 mg and vonoprazan 20 mg to lansoprazole 15 mg) will be evaluated with a Farrington and Manning test with a non-inferiority margin of 10% for the difference in maintenance of healing rates between treatments (each dose group of vonoprazan – lansoprazole 15 mg) at Week 24. The point estimate and 2-sided 95% CI of the difference in the maintenance of healing rate between each dose group of vonoprazan and lansoprazole 15 mg will be calculated via the Miettinen and Nurminen method. The primary analysis will be performed using the MITT Set.

If non-inferiority is shown for a vonoprazan dose group (vonoprazan 10 mg or vonoprazan 20 mg), superiority compared to lansoprazole 15 mg will also be assessed following the testing hierarchy in Section 8.0 via the Farrington and Manning test of the null hypothesis difference ≤0 versus the alternative hypothesis difference >0.

All efficacy data will be listed using the Randomized Set.

The number of subjects who had endoscopies at each visit, including both scheduled and unscheduled visits, will be summarized by analysis visit for each treatment group.

Efficacy endpoints assessed by results from endoscopy will be based on the assessment from the central adjudication. The differences between central adjudication and investigators assessments of healing/not healed and LA Grade will be summarized in a table.

8.1.2 Assumption Testing

Given the large sample sizes for both phases, the test statistic for the Farrington and Manning test is expected to be normally distributed. Testing for normality with Farrington and Manning test might be added as an ad hoc analysis at the end of the study.

8.1.3 Sensitivity and Supportive Analyses

To explore how intercurrent events affect the robustness of each estimand, the following sensitivity and supportive analyses will be performed. See Section 13.6 for more details on estimands, intercurrent event types, and associated analyses for the Healing Phase and Maintenance Phase.

Data summaries will parallel those described for the primary analysis of the primary efficacy endpoints for all sensitivity and supportive analyses.

8.1.3.1 Tipping Point Analysis

To explore how missing data affects the robustness of the primary analysis, a missing data sensitivity analysis using a tipping point method will be performed for non-inferiority for both Healing and Maintenance Phases using different missing data strategies for Estimand 1. The tipping point analysis begins with most conservative imputation for missing data (i.e., most heavily slanted against the active treatment non-inferior to active control).

Implementation of the tipping point approach for the Healing Phase will involve the following steps:

1. The missing data for lansoprazole 30 mg subjects is assumed to be responders (ie, healed). Missing data for vonoprazan 20 mg subjects is assumed to be non-responders (ie, not healed).
2. The data set is analyzed using the same Farrington and Manning method as the primary analysis to see if the p-value is ≤0.05; and if so, the tipping point analysis will be stopped at this point.
3. Repeat step #1 switching an imputed responder for one lansoprazole 30 mg subject to an imputed non-responder.
4. Repeat step #2 to obtain the p-value to see if the p-value is ≤0.05.

Repeat steps #3 and #4 increasing the number of imputed lansoprazole 30 mg non-responders one at a time until the p-value is ≤0.05 or all missing data has been switched from an imputed responder to a non-responder for all lansoprazole 30 mg subjects with an imputed result. The number of non-responders that achieves this p-value will be considered the "tipping point."

Implementation of the tipping point approach for the Maintenance Phase will follow the same steps as above, starting with considering all missing data for lansoprazole 15 mg subjects as responders (ie, maintained healing) and missing data for each dose group of vonoprazan as non-responders (ie, did not maintain healing).

8.1.3.2 Sensitivity Analysis on Observed Data Only

For each phase, the primary analysis will be repeated for Estimand 1 on observed data only. Subjects without post-baseline endoscopy results will be excluded from these analyses.

8.1.3.3 Multiple Imputation Analyses

For each phase, 3 types of multiple imputation analyses will be performed for each estimand. The details of multiple imputation methods are summarized in Table 9 for the Healing Phase, and Table 13 for the Maintenance Phase.

8.1.3.4 Sensitivity Analysis on Prohibited Medications

To explore how use of the prohibited medications of proton pump inhibitors or H2 receptor antagonists affects the robustness of the primary analysis, a sensitivity analysis will be performed for non-inferiority for both Healing and Maintenance Phases for Estimand 1 considering subjects as non-responders at timepoints after any PPI/H2RA use.

8.1.3.5 Supportive Analysis Using PP Set

Supportive analyses will be repeated using PP Set for Estimand 1, on the primary efficacy endpoints for non-inferiority (for both Healing and Maintenance Phases) and superiority (for Maintenance Phase if shown).

8.1.4 Subgroup Analysis

For both phases, the primary efficacy endpoints will be analyzed separately for the following subgroups using the same method as the primary:

- Age group at Screening (<45, ≥45 - <65, ≥65 - <75, ≥75)
- Sex (Male, Female)
- Baseline BMI category (<25, ≥25 - <30, ≥30)
- Region (United States, Europe)
- Baseline Adjudicated EE grade (A, B, C, D)
- Baseline Adjudicated EE grade category (A/B, C/D)
- CYP2C19 status (extensive metabolizer, poor metabolizer)
- Smoking status (never smoked, current smoker, ex-smoker)
- Alcohol use (drink every day, drink a couple of days per week, drink a couple of days per month, never drink)
- Race (white, black or African American, Asian, American Indian or Alaska native, native Hawaiian or other Pacific islander, other)
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)

In addition, for the Maintenance Phase, the primary efficacy endpoint will be analyzed separately for the following subgroups:

- Healing Phase treatment (vonoprazan 20 mg, lansoprazole 30 mg)
- Healing Phase treatment duration (Week 2, Week 8)

The subgroup analyses will be performed using the MITT Set.

8.2 Secondary Efficacy Endpoint

Healing Phase

For the Healing Phase, the following secondary endpoints will be analyzed for superiority of vonoprazan 20 mg compared to lansoprazole 30 mg

1. The percentage of subjects who have complete healing of EE at Week 2 as assessed by endoscopy
2. The percentage of subjects who have complete healing of EE at Week 2 as assessed by endoscopy for subjects with baseline LA Classification Grades C or D
3. The percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy for subjects with baseline LA Classification Grades C or D
4. The percentage of subjects with onset of sustained resolution of heartburn by Day 3 (sustained resolution is defined as at least 7 consecutive days with no daytime or nighttime heartburn as assessed by the daily diary)

Superiority of vonoprazan 20 mg to lansoprazole 30 mg will be assessed via the Farrington and Manning test of the null hypothesis difference ≤ 0 versus the alternative hypothesis difference >0.

For the Healing Phase, the following secondary endpoint will be analyzed for non-inferiority of vonoprazan 20 mg compared to lansoprazole 30 mg.

5. The percentage of 24-hour heartburn–free days over the Healing Phase as assessed by the daily diary

A 2-sided 95% CI will be calculated for the difference between the treatment groups in the mean percentage of 24-hour heartburn–free days (each dose group of vonoprazan minus lansoprazole). The 95% confidence interval will be computed from Welch's t-test. The difference in the means will be compared from Welch's t-test as an exploratory analysis. If the lower bound of this CI is greater than -15%, noninferiority will be concluded. If noninferiority is shown, the percentage of 24-hour heartburn–free days will also be compared between treatment groups using a Wilcoxon rank-sum test as an exploratory analysis.

A fixed noninferiority margin of -15% is justified based on results from a randomized, double-blind study in 1284 subjects that evaluated EE healing with placebo, lansoprazole 15 mg, lansoprazole 30 mg, or omeprazole 20 mg (Castell et al 1996). In this study, a twice daily patient diary was used to collect daytime and nighttime heartburn symptoms. The diary was very similar to the diary being used in the current study, except that it used a 4-point severity scale (none, mild, moderate, severe) rather than a 5-point severity scale (none, mild, moderate, severe, very severe). Given that the secondary endpoint is based on presence or absence of any degree of heartburn, these minor scale differences are not relevant for the purpose of establishing the noninferiority margin. The difference in the mean percentage of 24-hour heartburn–free days between lansoprazole 30 mg (72.6%) and placebo (31.6%) was 41.0%; the lower limit of the 2-sided 95% CI was 35.8%. Hence, even with a conservative assumption of this lower bound as the true treatment effect of lansoprazole 30 mg, a noninferiority margin of -15% assures that vonoprazan retains at least 58% of the treatment effect of lansoprazole 30 mg compared with placebo.

Endpoints #4 and #5 will use data collected from the subject symptom diary. For each subject, the percentage of 24-hour heartburn-free days over the Healing Phase will be calculated using all days with at least 1 morning or evening diary entry during the Healing Phase.

See Section 13.4.1 for data handling details for the subject daily diary and Section 13.4.3 for data handling details on sustained resolution of heartburn.

Maintenance Phase

For the Maintenance Phase, the following secondary endpoint will be analyzed for superiority of each dose group of vonoprazan compared to lansoprazole 15 mg.

1. The percentage of subjects who maintain complete healing of EE after 24 weeks as assessed by endoscopy for subjects with baseline LA Classification Grades C or D

Superiority compared to lansoprazole 15 mg will be assessed via the Farrington and Manning test of the null hypothesis difference ≤ 0 versus the alternative hypothesis difference >0.

For the Maintenance Phase, the following secondary endpoint will be analyzed for non-inferiority of each dose group of vonoprazan compared to lansoprazole 15 mg.

2. The percentage of 24-hour heartburn–free days over the Maintenance Phase as assessed by the daily diary

The secondary endpoint will use data collected from the subject symptom diary. A 2-sided 95% CI will be calculated for the difference between the treatment groups in the mean percentage of 24-hour heartburn–free days (each dose group of vonoprazan minus lansoprazole). The 95% confidence interval will be computed from Welch's t-test. The difference in the means will be compared from Welch's t-test as an exploratory analysis. If the lower bound of this CI is greater than -15%, noninferiority will be concluded. If noninferiority is shown, the percentage of 24-hour heartburn–free days will also be compared between treatment groups using a Wilcoxon rank-sum test as an exploratory analysis.

A fixed noninferiority margin of -15% is justified based on results from a randomized, double-blind study that evaluated maintenance of EE healing over 6 months with placebo, dexlansoprazole 30 mg QD, and dexlansoprazole 60 mg QD (Metz et al 2009). This study used the same twice daily patient diary to collect daytime and nighttime heartburn symptoms that will be used in the current study.

Data for the relief of heartburn based on a daily patient diary from a maintenance of EE healing study that compared the efficacy of lansoprazole 15 mg QD to placebo are not available, as the maintenance of EE healing studies for lansoprazole demonstrated superior relief of symptoms including heartburn based on the investigator's assessment. However, data are available from a study comparing dexlansoprazole, the R-enantiomer of lansoprazole, to placebo, and results for the approved dexlansoprazole maintenance dose of 30 mg are considered relevant to inform the noninferiority margin for the approved lansoprazole maintenance dose of 15 mg.

The difference in the mean percentage of 24-hour heartburn–free days between dexlansoprazole 30 mg (83.3%) and placebo (36.0%) was 47.3%; the lower limit of the 2-sided 95% CI was 40.3%. Hence, even with a conservative assumption of this lower bound as the true treatment effect of lansoprazole 15 mg, a noninferiority margin of -15% assures that vonoprazan retains at least 62% of the treatment effect of lansoprazole 15 mg compared with placebo.

All secondary endpoints will be analyzed using the MITT Set.

8.2.1 Subgroup Analysis

For both phases, the percentage of 24-hour heartburn–free days will be analyzed separately for subjects with Baseline Adjudicated EE grade category A/B and C/D using the same method as the secondary endpoint.

8.2.2 Sensitivity Analysis

To explore how missing data affects the robustness of the secondary analysis on the percentage of 24-hour heartburn–free days over each phase, a sensitivity analysis will be performed for each

phase, in which the days after a subject discontinues treatment will be imputed as non-responders, i.e. having heartburn on those days. The percentage will be calculated from the number of planned days in the treatment period of each phase:

- Healing Phase: the treatment duration or 56 days, whichever is greater
- Maintenance Phase, the treatment duration in this phase, or 168 days, whichever is greater.

A summary of the percentage of missing diaries during each phase will be summarized by treatment group. A study day with no diary will be counted as 2 missing diaries, a study day with only either daytime or nighttime diary will be counted as 1 missing diary. The percentage will be calculated from the number of planned days in the treatment period of each phase.

8.3 Exploratory Endpoints

For the exploratory efficacy endpoints, all comparisons between treatment groups will be performed with no adjustment to the alpha level.

Healing Phase

For the Healing Phase, the following exploratory endpoints will be compared between treatment groups using a Wilcoxon rank-sum test.

- The percentage of days without daytime heartburn over the Healing Phase as assessed by the daily diary
- The percentage of days without nighttime heartburn over the Healing Phase as assessed by the daily diary
- The mean severity of daytime and nighttime heartburn over the Healing Phase as assessed by the daily diary
- The mean severity of daytime heartburn over the Healing Phase as assessed by the daily diary
- The mean severity of nighttime heartburn over the Healing Phase as assessed by the daily diary
- The percentage of 24-hour heartburn–free days over the first 14 days of the Healing Phase as assessed by the daily diary
- The percentage of days without daytime heartburn over the first 14 days of the Healing Phase as assessed by the daily diary
- The percentage of days without nighttime heartburn over the first 14 days of the Healing Phase as assessed by the daily diary
- The mean severity of daytime and nighttime heartburn over the first 14 days of the Healing Phase as assessed by the daily diary

- The mean severity of daytime heartburn over the first 14 days of the Healing Phase as assessed by the daily diary
- The mean severity of nighttime heartburn over the first 14 days of the Healing Phase as assessed by the daily diary
- The change from baseline to the end of the Healing Phase for each subscale and the total score of the PAGI-SYM questionnaire
- The change from baseline to the end of the Healing Phase for each subscale and the total score of the PAGI-QoL questionnaire
- The change from baseline to the end of the Healing Phase for the EQ-5D-5L index and visual analogue scale (VAS) scores

For the following endpoint, cumulative incidences of events will be estimated using the Kaplan-Meier method. Standard errors (SE) and 2-sided 95% CIs will be calculated using the Greenwood's formula. Log-rank tests will be performed. The median and its 2-sided 95% CIs, the mean, and the SE will be calculated for the number of days to events. See Section 13.4.3 for details on the definitions of events and censoring.

- The time to sustained resolution of heartburn

Maintenance Phase

For the Maintenance Phase, the following exploratory endpoints will be compared between each two treatment groups (i.e. vonoprazan 10 mg vs. lansoprazole 15 mg, vonoprazan 20 mg vs. lansoprazole 15 mg, vonoprazan 10 mg vs vonoprazan 20 mg) using a Wilcoxon rank-sum test. The maintenance baseline is defined as the last visit during the Healing Phase prior to the first dose of the study drug in the Maintenance Phase.

- The percentage of days without daytime heartburn over the Maintenance Phase as assessed by the daily diary
- The percentage of days without nighttime heartburn over the Maintenance Phase as assessed by the daily diary
- The mean severity of daytime and nighttime heartburn over the Maintenance Phase as assessed by the daily diary
- The mean severity of daytime heartburn over the Maintenance Phase as assessed by the daily diary
- The mean severity of nighttime heartburn over the Maintenance Phase as assessed by the daily diary
- The percentage of 24-hour heartburn–free days, the percentage of days without daytime heartburn, and the percentage of days without nighttime heartburn during each month of the Maintenance Phase as assessed by the daily diary

- The mean severity of daytime and nighttime heartburn, the mean severity of daytime heartburn, and the mean severity of nighttime heartburn during each month of the Maintenance Phase as assessed by the daily diary
- The change from maintenance baseline to the end of the Maintenance Phase for each subscale and the total score of the PAGI-SYM questionnaire
- The change from maintenance baseline to the end of the Maintenance Phase for each subscale and the total score of the PAGI-QoL questionnaire
- The change from maintenance baseline to the end of the Maintenance Phase for the EQ-5D-5L index and visual analogue scale (VAS) scores

The following endpoints will be summarized by treatment group for all Maintenance Phase subjects who enter the Follow-up Phase.

- Percentage of 24-hour heartburn-free days over the Follow-up Phase as assessed by the daily diary
- Percentage of days with use of antacids over the Follow-up Phase as assessed by the daily diary
- Percentage of days with use of H2-receptor antagonists over the Follow-up Phase as assessed by the daily diary
- Percentage of days with use of either antacids or H2-receptor antagonists over the Follow-up Phase as assessed by the daily diary

All exploratory endpoints will be analyzed using the MITT Set. Refer to Section 13.4 for data handling details on symptom diary, PAGI-SYM, PAGI-QoL and EQ-5D-5L.

9.0 SAFETY ANALYSIS

For each phase, safety will be assessed by summarizing the incidence of AEs and changes in clinical laboratory tests, gastrin and pepsinogen I/II levels, and vital signs. In addition, ECG and gastric biopsy will be summarized in the Maintenance Phase. For all safety analyses, baseline for the Healing Phase will be the last visit prior to the first dose of study drug in the Healing Phase. Baseline for the Maintenance Phase will be the last visit during the Healing Phase prior to the first dose of study drug in the Maintenance Phase.

All safety analyses will be conducted for each treatment group and overall using the Safety Set.

9.1 Adverse Events

Adverse events will be coded using Version 23 of the Medical Dictionary for Regulatory Activities (MedDRA). A treatment-emergent AE (TEAE) is defined as any event that occurs after the first dose of study drug in that phase or any event at baseline that worsens in either intensity or frequency after the first dose of study drug in that phase. For subjects who did not enter the Maintenance Phase, an AE that starts after more than 30 days of the last dose of study drug in the

Healing Phase will not be counted as a TEAE of the Healing Phase. For subjects who entered the Maintenance Phase, an AE that starts after more than 30 days of the last dose of study drug in the Maintenance Phase will not be counted as a TEAE of the Maintenance Phase. A subject with multiple adverse events within a primary SOC or preferred term is only counted once towards the total for that SOC and/or preferred term. For the AE severity and relationship summaries, if a subject reported more than one adverse event with the same preferred term, the adverse event with the greatest severity or relationship will be presented. If a subject reported more than one adverse event within the same primary system organ class, then the subject will be counted only once with the greatest severity or relationship at the system organ class level. For table summaries, if severity is missing then 'severe' is assumed. If relationship is missing, relationship to study drug is assumed to be 'related'.

The number and percentage of subjects with TEAEs will be summarized in the following ways:

- by primary system organ class, preferred term and treatment group
- by primary system organ class, preferred term, maximum severity and treatment group
- by primary system organ class, preferred term, relationship to study drug and treatment group

The number and percentage of subjects with TEAEs related to study drug will be summarized in the following ways:

- by primary system organ class, preferred term and treatment group
- by primary system organ class, preferred term, maximum severity and treatment group

The most common TEAEs (≥5% of subjects in any treatment group) and TEAEs related to study drug (≥2% of subjects in any treatment group) will be presented by preferred term in descending frequency starting from the most common event.

The most common non-serious TEAEs (>5% of subjects in any treatment group) will be presented by primary system organ class, preferred term and treatment group along with the number and proportion of subjects reporting at least one of these most frequent non-serious TEAEs.

The number and proportion of subjects as well as the number of events (except deaths) with the following types of events will be summarized by primary system organ class, preferred term and treatment:

- Adverse events leading to treatment discontinuation

- Serious Adverse Events (SAEs)
- Deaths
- Adverse events of special interest (AESI)

In the Maintenance Phase, TEAEs, TEAEs related to study drug, SAEs and AEs leading to treatment discontinuation will also be presented by primary system organ class and preferred term for the treatment types across the two phases:

- Vonoprazan to vonoprazan
- Vonoprazan to lansoprazole
- Lansoprazole to vonoprazan
- Lansoprazole to lansoprazole

All adverse events in the Healing Phase will be included in a listing using the Screened Set. All adverse events in the Maintenance Phase will be included in a listing using the Randomized Set. All adverse events over the Follow-Up Phase will be included in a listing using the Safety Set. In addition, the following select adverse events will be displayed in separate listings for each phase:

- Deaths
- Serious adverse events
- Adverse events leading to treatment discontinuation
- Adverse events related to study drug
- Adverse events of special interest

9.1.1 Adverse Events of Special Interests (AESI)

The number and percentage of subjects with TEAEs and SAEs that are in one of the AESI categories presented in Table 2 will be summarized by AESI category, primary system organ class and preferred term for each treatment group and overall. The search criteria that will be used to identify AESIs are specified in the table.

Table 2 Adverse Events of Special Interest – Search Criteria

Adverse Event of Special Interest	Search Criteria
Clostridium difficile enteric infection	Pseudomembranous colitis SMQ (Narrow)
Bone Fracture	Bone Fracture Custom Query (PTs defined below) Acetabulum fracture Ankle fracture Atypical femur fracture Atypical fracture Avulsion fracture Bone fissure Bone fragmentation Chance fracture Clavicle fracture Comminuted fracture Complicated fracture Compression fracture Craniofacial fracture Epiphyseal fracture Facial bones fracture Femoral neck fracture Femur fracture Fibula fracture Foot fracture Forearm fracture Fracture Fracture blisters Fracture displacement Fracture malunion Fracture nonunion Fracture of clavicle due to birth trauma Fractured coccyx Fractured ischium Fractured sacrum Greenstick fracture Hand fracture Hip fracture Humerus fracture Ilium fracture Impacted fracture Jaw fracture Limb fracture Lower limb fracture Fractured skull depressed Lumbar vertebral fracture Metaphyseal corner fracture Multiple fractures Open fracture Osteoporotic fracture Patella fracture Pathological fracture Pelvic fracture Pubis fracture Radius fracture Rib fracture Sacroiliac fracture Scapula fracture Skull fracture Skull fractured base Spinal compression fracture Spinal fracture Spinal fusion fracture Sternal fracture Stress fracture Subchondral insufficiency fracture Thoracic vertebral fracture Tibia fracture Torus fracture Traumatic fracture Ulna fracture Upper limb fracture Wrist fracture
Severe cutaneous adverse reactions	Severe cutaneous adverse reactions SMQ (Narrow)
Hepatotoxicity	• Drug related hepatic disorders - comprehensive search (SMQ) (Narrow) • Cholestasis and jaundice of hepatic origin (SMQ) (Broad) • Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions SMQ (Broad) • Hepatitis, non-infectious (SMQ) (Broad) • Liver related investigations, signs and symptoms (SMQ) (Narrow)
Gastric cancer	Gastric Neoplasms Malignant MedDRA High Level Term

MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term; SMQ: Standardized MedDRA Queries.

9.2 Clinical Laboratory Evaluations

Descriptive statistics for clinical laboratory values (hematology, chemistry and urinalysis laboratory tests; in SI units and in US conventional units) and serum gastrin and pepsinogen I/II

levels will be presented by treatment group and overall for each phase. Changes from baseline will also be presented for quantitative variables by treatment group and overall for each phase. For categorical variables (ie, normal or abnormal findings, or qualitative clinical laboratory tests), shift tables for the change from baseline to each post-baseline time point will be presented by treatment group and overall.

For each period, the number and percentage of subjects with at least one post-baseline serum gastrin value >500 pg/mL and >1000 pg/mL and with the test value higher than baseline value, if available, will be presented. A supportive listing of subjects with such post-baseline elevations will be provided including the subject ID, baseline, and post-baseline values.

Abnormal liver function tests are defined as liver test values that meet at least one of the criteria listed below. For each phase, the number and percentage of subjects with at least one post-baseline abnormal liver function test and with the test value higher than baseline value, if available, will be presented by treatment group and overall. A supportive listing of subjects with such post-baseline elevations will be provided including the subject ID, baseline, and post-baseline values.

- ALT > 3xULN
- ALT > 5xULN
- ALT > 10xULN
- ALT > 3xULN and Total Bilirubin > 2xULN

- AST > 3xULN
- AST > 5xULN
- AST > 10xULN
- AST > 3xULN and Total Bilirubin > 2xULN

- Total Bilirubin > 2xULN

- AST > 3xULN or ALT > 3xULN
- AST > 5xULN or ALT > 5xULN
- AST > 10xULN or ALT > 10xULN
- (AST > 3xULN or ALT > 3xULN) and Total Bilirubin > 2xULN

- AST > 3xULN and ALT > 3xULN
- AST > 5xULN and ALT > 5xULN
- AST > 10xULN and ALT > 10xULN
- AST > 3xULN and ALT > 3xULN and Total Bilirubin > 2xULN

- Alkaline phosphatase > 1.5xULN

- ALT > 3xULN and Alkaline phosphatase > 1.5xULN
- AST > 3xULN and Alkaline phosphatase > 1.5xULN

- Alkaline phosphatase > 3xULN
- ALT > 3xULN and Alkaline phosphatase > 3xULN
- AST > 3xULN and Alkaline phosphatase > 3xULN

9.3 Vital Signs

Descriptive statistics for vital signs, including body temperature, systolic blood pressure, diastolic blood pressure and pulse rate, will be presented by treatment group and overall for each phase. Changes from baseline will also be presented by treatment group and overall for each phase.

Abnormal vital sign values are defined as vital sign values that meet one of the criteria listed below. For each phase, the number and percentage of subjects with at least one post-baseline abnormal vital sign value and with the value worse than the baseline value, if available, will be presented by treatment group and overall. A supportive listing of subjects with such post-baseline elevations will be provided including the subject ID, baseline, and post-baseline values.

- Systolic blood pressure (mmHg):
  - <50
  - >180
- Diastolic blood pressure (mmHg):
  - <50
  - >100
- Heart rate (bpm):
  - <50
  - >120

9.4 Physical Examination

All data collected from the physical examinations assessments must be available in the source documents but will not be added to the analysis database.

9.5 Electrocardiogram

ECG parameters, including heart rate, RR interval, PR interval, QRS interval, QT interval, QTc Fridericia (QTcF) will be collected on eCRF. Descriptive statistics for these ECG parameters will be presented by treatment group and overall for the Maintenance Phase only. Changes from baseline will also be presented for quantitative variables by treatment group and overall for the Maintenance Phase. For ECG interpretations (within normal limits, abnormal but not clinically

significant, or abnormal and clinically significant), shift tables for the change from baseline to each post-baseline time point will be presented by treatment group and overall.

Abnormal QTcF values are defined as ECG values that meet at least one of the criteria listed below, the number and percentage of subjects with at least one of the post-baseline abnormal values and with post-baseline value higher than baseline value, if available, will be presented by treatment group and overall in the Maintenance Phase. A supportive listing of subjects with such post-baseline elevations will be provided including the subject ID, baseline, and post-baseline values.

- Absolute QTcF interval prolongation:
  - QTc interval > 450
  - QTc interval > 480
  - QTc interval > 500
- Change from baseline in QTcF interval:
  - QTc interval increases from baseline >30
  - QTc interval increases from baseline >60
  - QTc interval > 450 with increase from baseline >30

9.6 Gastric Biopsy

A shift table showing the change in gastric biopsy results from baseline to the final visit will be produced for each biopsy by treatment group.

10.0 INTERIM ANALYSIS/OTHER ANALYSES

10.1 Interim Analysis

No interim analysis is planned.

10.2 Coronavirus Pandemic

Shortly after this study began enrolling subjects, the SARS-COV-2 virus, which causes COVID-19 was declared a global pandemic by the World Health Organization. In accordance with guidance issued by regulatory agencies, study data collection has been amended for subjects to capture visits missed/delayed due to COVID-19 related reasons, and assessment completed via alternative method due to COVID-19 related reasons.

The following adjustments will be made to the efficacy analysis: Missing data for subjects without a post-baseline endoscopy at a visit due to COVID-19 related reasons will be imputed using a missing at random assumption (Section 4.5).

COVID-19 impacts on individual subjects collected on COVID-19 CRF pages will be listed for the Randomized Set. Protocol deviations related to COVID-19 will be marked in the protocol deviation listing for the Randomized Set.

The anticipated impact of COVID-19 is widely regarded as unknown. If the impact of COVID-19 on the conduct of this study is observed to be significant, further summaries and listings of the impact will be explored.

11.0 CHANGES IN THE PLANNED ANALYSIS

The following changes to the analysis specified in Protocol Amendment 2, dated 01Oct2019 have been made.

The definition of the Per Protocol Set was updated to exclude subjects with major protocol deviations regardless of study outcome (healed or not healed during the Healing Phase and remained healed or recurred during the Maintenance Phase).

The protocol includes description of a potential analysis of the Healing Phase data to be performed after all subjects have completed the Healing Phase. It has been decided not to conduct this analysis so reference to this potential analysis was removed in Amendment 1 to the SAP.

12.0 REFERENCES

Castell DO, Richter JE, Robinson M, et al. Efficacy and safety of lansoprazole in the treatment of erosive reflux esophagitis. The Lansoprazole Group. Am J Gastroenterol. 1996;91(9):1749-57.

Earnest DL, Dorsch E, Jones J, et al. A placebo-controlled dose-ranging study of lansoprazole in the management of reflux esophagitis. Am J Gastroenterol. 1998;93(2):238-43.

Metz DC, Howden CW, Perez MC, et al. Clinical trial: dexlansoprazole MR, a proton pump inhibitor with dual delayed-release technology, effectively controls symptoms and prevents relapse in patients with healed erosive oesophagitis. Aliment Pharmacol Ther. 2009;29(7):742-54.

PREVACID (lansoprazole) delayed-release Capsules, PREVACID SoluTab (lansoprazole) delayed-release orally disintegrating tablets. Full Prescribing Information including Medication Guide. Deerfield, IL: Takeda Pharmaceuticals America, Inc., Revised Jun 2018.

13.0 APPENDICES

13.1 Schedule of Events

Table 3 Schedule of Events

	Screening Period (a)	Healing Phase					Maintenance Phase						
Timing		**Healing Day -1 (b)**	**Healing Day 1(b)**	**Week 2 (c) Healing Day 15**	**Week 8 (c,m) Healing Day 57**		**Main Week 4 Main Day 29**	**Main Week 12 Main Day 85**	**Main Week 16 Main Day 113**	**Main Week 20 Main Day 141**	**Main Week 24 Main Day 169 Final Visit/ET (d)**	**Safety F/U (e)**	**Un-sched-uled Visit (f)**
Visit Windows (Days)	**Day-35 to -2**	**-**		**12 to 18**	**54 to 60**		**26 to 36**	**82 to 92**	**110 to 120**	**138 to 148**	**166 to 176**	**-**	**-**
Visit Number:	**1**	**2**		**3**	**4**		**M1**	**M2**	**M3**	**M4**	**M5**	**F/U**	**-**
Informed Consent	X												
Inclusion/Exclusion Criteria	X	X											
Demographic and medical history	X												
Smoking and Alcohol history	X												
Medication history	X												
Physical examination(g)	X	X		X	X		X	X			X	X	X
Vital signs	X	X		X	X		X	X			X	X	X
Weight and height	X												
Concomitant medications	X	X		X	X		X	X			X	X	X
Concurrent medical conditions	X												
FSH(h)	X												
Hepatitis B and C; HIV	X												
Urine drug screen	X												
Clinical laboratory test including hematology, serum chemistry, and urinalysis(i)	X			X	X		X	X			X		
^{13}C-UBT Breath Test	X												
CYP2C19 genotyping test				X									
Serum gastrin/pepsinogen I/II levels(j)	X			X	X		X	X			X	X	
Pregnancy test (serum HCG)(k)	X												
Pregnancy test (urine HCG)(k)		X		X	X		X	X			X		
Guidance on avoidance of pregnancy	X	X		X	X		X	X					
ECG	X										X		

	Screening Period (a)	Healing Phase					Maintenance Phase						
Timing		**Healing Day -1 (b)**	**Healing Day 1(b)**	**Week 2 (c) Healing Day 15**	**Week 8 (c,m) Healing Day 57**		**Main Week 4 Main Day 29**	**Main Week 12 Main Day 85**	**Main Week 16 Main Day 113**	**Main Week 20 Main Day 141**	**Main Week 24 Main Day 169 Final Visit/ET (d)**	**Safety F/U (e)**	**Un-sched-uled Visit (f)**
Visit Windows (Days)	**Day-35 to -2**	**-**		**12 to 18**	**54 to 60**		**26 to 36**	**82 to 92**	**110 to 120**	**138 to 148**	**166 to 176**	**-**	**-**
Visit Number:	**1**	**2**		**3**	**4**		**M1**	**M2**	**M3**	**M4**	**M5**	**F/U**	**-**
Subject's diary; distribute and/or review(l)	X(r)	X		X	X		X	X			X	X	
PAGI-SYM		X		X	X		X	X			X		
PAGI-QoL		X		X	X		X	X			X		
EQ-5D-5L		X		X	X		X	X			X		
Endoscopy	X			X	X						X		
Gastric biopsy(n)	X				X(s)						X		
Randomization		X		X(o)	X(o)								
Dispense study drug		X(p)		X(q)	X(q)			X					
First day of study drug administration			X										
Drug return/accountability/ review treatment compliance(q)				X	X		X	X			X		
Phone call to subject									X	X			
AE/pretreatment event assessment	X	X		X	X		X	X	X	X	X	X	X

Abbreviations: AE, adverse events; CYP2C19, cytochrome P450 2C19; ECG, electrocardiogram; EQ-5D-5L, EuroQol-5 Dimensions-5 Levels; FSH, follicle-stimulating hormone; F/U, follow-up; HCG, human chorionic gonadotropin; HIV, human immunodeficiency virus; Main, maintenance; PAGI-SYM, Patient Assessment of Gastrointestinal Disorders-Symptoms Severity Index; PAGI-QoL, Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life; UBT, urea breath test

(a) Visit window is Day -35 to Day -2 for assessment of *Helicobacter pylori*, serum gastrin and pepsinogen I/II values, ECG, and endoscopy in the Screening Period. However, given the invasive nature of an endoscopy, any endoscopic confirmation performed in a routine clinical setting before informed consent signing which would have included standard of care gastric mucosa sampling, endoscopy images available for adjudication and done within 7 days of randomization/dosing is acceptable to use for the purpose of fulfilling the Screening requirement, and repeat endoscopic examination during the Screening Period will not be required.

(b) The date of randomization is defined as Day -1. The date of first dosing day is defined as Day 1 for both the Healing and Maintenance Phases.

(c) If the subject is confirmed as healed and rerandomized to the Maintenance Phase, this visit is considered the baseline visit for the Maintenance Phase, and subjects who prematurely discontinue from the Healing Phase or who are not healed at Week 8 should undergo Maintenance Phase Week 24/Final Visit/ET procedures (including gastric biopsy).

(d) For early discontinuations, the assessments mentioned for Week 24 are required to be performed.

(e) A safety follow-up visit is to be scheduled 4 weeks after the last dose of study drug administration. Subjects whose healing is not confirmed after Week 8 and who discontinued the study after the Healing Phase will undergo a safety follow-up visit 4 weeks after the last dose of study drug administration.

(f) At an unscheduled visit, the following procedures are to be completed with additional procedures at the investigator's discretion: a brief physical examination, vital sign measurements, concomitant medication assessment, and adverse event assessment. If the visit results in premature termination, then all procedures outlined for Maintenance Phase Week 24 /Final visit should be performed.
(g) Full physical examination is performed at baseline; a brief physical examination is performed at all other visits.
(h) If menopause is suspected.
(i) See Protocol Section 6.6 for all protocol-required laboratory assessments.
(j) Gastrin and pepsinogen I and II results will not be reported to investigative sites or other blinded personnel until the study blind is broken.
(k) Only female subjects with childbearing potential.
(l) Subjects should be instructed to complete the electronic diary every morning upon waking (for nighttime symptoms) and every evening before bedtime (for daytime symptoms) on each day of the study.
(m) Visit to be performed if not healed at Week 2.
(n) Four gastric mucosal biopsies (1 each from the greater and lesser curvature of the body and 1 each from the greater and lesser curvature of the antrum) should be obtained at each time point a gastric biopsy is done.
(o) Subjects with confirmed endoscopic healing at Week 2 or Week 8 will be rerandomized to the Maintenance Phase.
(p) Subjects are to start dose administration from Day 1.
(q) Subjects are to visit the study site without taking the study drug at Week 2 and Week 8 visits. For subjects without confirmed endoscopic healing of erosive esophagitis at Week 2, the daily dose of the study drug is to be administered after completion of assessments scheduled on that day. Subjects may need to return to the site a few days after Week 2 and Week 8 visits to receive a new allotment of study drug. No study procedures will be performed.
(r) Complete at least 7 days of diary during screening period.
(s) Only if EE not healed (follow Week 24 Final Visit/ET procedures).

13.2 Imputation Rules for Missing Date Information

13.2.1 Rules for Concomitant Medication Start Date Imputation

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 double-blind study drug, then the day and month of the date of the first dose of double-blind study drug 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 double-blind study drug, then 31 December 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 double-blind study drug, 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 double-blind study drug, then the day of the date of the first dose of double-blind study drug will be assigned to the missing day.
- If either the year is before the year of the date of the first dose of double-blind study drug or if both years are the same but the month is before the month of the date of the first dose of double-blind study drug, 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 double-blind study drug or if both years are the same but the month is after the month of the date of the first dose of double-blind study drug, then the first day of the month will be assigned to the missing day.

13.2.2 Rules for Concomitant Medication End Date Imputation

The following rules will be applied to impute the missing numerical fields. If the date of the last dose of double-blind study drug is missing, then replace it 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 double-blind study drug, then the day and month of the date of the last dose of double-blind study drug 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 double-blind study drug, 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 double-blind study drug, 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 double-blind study drug, then the day of the date of the last dose of double-blind study drug will be assigned to the missing day.
- If either the year is before the year of the date of the last dose of double-blind study drug or if both years are the same but the month is before the month of the date of the last dose of double-blind study drug, 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 double-blind study drug or if both years are the same but the month is after the month of the date of the last dose of double-blind study drug, then the first day of the month will be assigned to the missing day.

13.2.3 Rules for AE Start Date Imputation

For AEs, incomplete (i.e., partially missing) start dates will be imputed and will follow the same rules as in Section 13.2.1. Incomplete stop dates will not be imputed.

13.3 LA Classification of Esophagitis Grading

Table 4 LA Grade Classification

Grade A	One or more mucosal breaks with a length of no longer than 5 mm that does not extend between the tops of 2 mucosal folds
Grade B	One or more mucosal breaks with a length of longer than 5 mm that does not extend between the tops of 2 mucosal folds
Grade C	One or more mucosal breaks that are continuous between the tops of 2 or more mucosal folds, which involves less than 75% of the circumference
Grade D	One or more mucosal breaks, which involves at least 75% of the circumference

A mucosal break is defined as "an area of slough or erythema with a sharp line of demarcation from adjacent normal mucosa"

13.4 Symptom Diary

13.4.1 Data Handling for Subject Daily Diary

Subjects will document the presence (yes/no) and maximum severity (1=Mild, 2=Moderate, 3=Severe, or 4=Very Severe) of daytime and nighttime heartburn symptoms in their diary two times per day, once in the morning for nighttime heartburn and once in the evening for daytime heartburn. If the presence of heartburn is answered as "no" in a diary, the question of maximum severity will not be answered. In this situation, the maximum severity will be considered as 0 (0=None) in the analysis. See Table 5 for details on definitions of heartburn severity.

Table 5 Definitions of Heartburn Severity (Daytime/Nighttime)

Definitions of Daytime Heartburn Severity (Daytime=Awake Time)
None - No heartburn
Mild - Occasional heartburn, can be ignored, does not influence daily routine
Moderate - Heartburn cannot be ignored and/or occasionally influences daily routine
Severe - Heartburn present most of day and/or regularly influences daily routine
Very Severe - Constant heartburn and/or markedly influences daily routine
Definitions of Nighttime Heartburn Severity (Nighttime=Sleep Time)
None - No heartburn
Mild - Occasional heartburn, can be ignored, does not influence sleep
Moderate - Heartburn cannot be ignored and/or occasionally influences sleep
Severe - Heartburn present most of night and/or regularly influences sleep
Very Severe - Constant heartburn and/or markedly influences sleep

For analysis purposes, diary entries will be assigned to a study day based on the day the collection interval started for that entry. For example, the Day 1 diary entries will include the

evening diary completed on Day 1 (the collection interval started when the subject awoke on Day 1) and the morning entry completed on Day 2 (the collection interval started when the subject went to bed on Day 1).

For each subject, the percentage of days with neither daytime nor nighttime heartburn during treatment will be calculated using all days with at least 1 morning or evening diary entry during the treatment period (Day 1 to the day of the last dose of study drug during the phase). For example, if a subject completed at least 1 diary entry on 44 of 50 days, but missed both entries on 6 days, 44 days will be used as the denominator in the analysis. All entries on that day will need to be heartburn-free in order for the day to be counted as a day with neither daytime nor nighttime heartburn. This will also apply when more than 2 diary entries are assigned to the same study day.

If a subject has only one diary entry on a day and that entry does not indicate heartburn, the day will be considered heartburn-free for the analysis.

Only evening diary entries will be used for assessing daytime heartburn and only morning diary entries will be used for assessing nighttime heartburn. For each phase, the mean severity of daytime and nighttime heartburn during treatment for each subject will be calculated by taking the mean of the average severity on all days with at least 1 morning or evening diary entry. For each day, the highest recorded severity will be determined separately for all morning diary entries and all evening diary entries. The average severity for the day will be determined by taking the mean of the highest recorded morning and evening diary entries on that day. These daily averages will then be averaged for all days with at least 1 morning or evening diary entry to obtain the endpoint value for each subject.

The mean severity of daytime heartburn during treatment for each subject will be calculated by determining the highest recorded severity for all evening diary entries per day, then taking the mean of all days with at least 1 evening diary entry during that phase.

The mean severity of nighttime heartburn during treatment for each subject will be calculated by determining the highest recorded severity for all morning diary entries per day, then taking the mean of all days with at least 1 morning diary entry during that phase.

13.4.2 Baseline Summary of Subject Diary

For the Healing Phase, the baseline diary variables will be summarized using diaries from the last 7 days prior to Day-1, ie, Days -8 to -2, inclusive. Only days with at least 1 morning or evening diary entry during the specified days will be included in these summaries.

Summary statistics (n, mean, SD, minimum, 25th percentile, median, 75th percentile, and maximum) will be generated for the mean severity of daytime and nighttime heartburn, mean severity of daytime heartburn, and mean severity of nighttime heartburn. The number of days with daytime or nighttime heartburn, number of days with daytime heartburn, and number of days with nighttime heartburn will be summarized by treatment group in categories of (≥0 - ≤3, >3 - ≤5, >5 - ≤7).

13.4.3 Sustained Resolution of Heartburn

Sustained resolution of heartburn will be defined as the first occurrence of 7 consecutive days with neither daytime nor nighttime heartburn during treatment. The time to sustained resolution of heartburn will be defined as the time from the first dose of study drug to the first day of the 7 consecutive days.

Subjects included in the analysis who do not achieve sustained resolution of heartburn will be censored on the first day of the last grouping of 7 consecutive days. For example, a subject whose last diary entries for 7 consecutive days during treatment correspond to Study Days 22 to 28 and who does not achieve sustained resolution by that time will be censored as of Study Day 22.

For the secondary endpoint of the percentage of subjects with onset of sustained resolution of heartburn by Day 3, subjects who do not have sustained resolution of heartburn for 7 consecutive days due to missing diary data will be considered as not having sustained resolution of heartburn.

13.5 PAGI-QOL and PAGI-SYM Scoring

13.5.1 PAGI-QOL

The PAGI-QOL questionnaire includes questions that ask about how some of the gastrointestinal problems the subject may be experiencing may have affected his quality of life. The final version of the questionnaire consists of 30 items, each with response options based on a 6-point Likert scale and with a recall period of the previous 2 weeks. The items are grouped into 5 subscales and a total score, as described in the following table.

PAGI-QOL Subscale	Description	Location on PAGI-QOL eCRF
Daily Activities	10 items related to avoiding or having difficulties with daily activities	Questions 1-10
Clothing	2 items, 1 related to feeling constricted and 1 related to frustration felt about not being able to dress as wanted	Questions 11-12
Diet and Food Habits	7 items related to restrictions made and induced frustrations	Questions 13-19
Relationship	3 items describing the impact of the disease on relationships with their partner, relatives, and friends	Questions 20-22
Psychological Well-being and Distress	8 items describing disease impact on feelings or emotional state	Questions 23-30
PAGI-QOL Total Score	Mean of PAGI-QOL Subscales	Not applicable
Each item is scored by the subject on a 6-point Likert scale: 0=None of the time, 1=A little of the time, 2=Some of the time, 3 = A good bit of the time, 4=Most of the time, and 5=All of the time.		

The subscales scores will be calculated by taking the mean of the non-missing items in each subscale after reversing the item scores. The subscale scores will range from 0 (lowest QOL) to 5 (highest QOL). Missing data for the subscales will be handled using the half-scale rule, ie, a subscale score will be calculated when ≤50% of the items within a subscale are missing; if >50% of the items within a subscale are missing, the score will be set to missing. The PAGI-QOL total score will be calculated by taking the mean of the corresponding subscales. If any subscale scores are missing, the total score will be set to missing.

Positive changes indicate improved quality of life.

13.5.2 PAGI-SYM

The PAGI-SYM questionnaire includes questions that ask about the severity of symptoms the subject may have related to his gastrointestinal problem. The final version of the questionnaire consists of 20 items, each with response options based on a 6-point Likert scale and with a recall period of the previous 2 weeks. The items are grouped into 6 subscales and a total score, as described in the following table.

PAGI-SYM Subscale	Description	Location on PAGI-SYM eCRF
Nausea/Vomiting	3 items related to severity of nausea, retching, and vomiting	Questions 1-3
Fullness/Early Satiety	4 items related to severity of stomach fullness, ability to finish a meal, feeling full after meals, and loss of appetite	Questions 4-7
Bloating	2 items related to severity of bloating and stomach size	Questions 8-9
Upper Abdominal Pain	2 items related to severity of upper abdominal pain and discomfort	Questions 10-11
Lower Abdominal Pain	2 items related to severity of lower abdominal pain and discomfort	Questions 12-13
Heartburn/Regurgitation	7 items related to severity of heartburn during the day and when lying down, discomfort inside chest during the day and at night, regurgitation during the day and when lying down, and bitter taste in mouth	Questions 14-20
PAGI-SYM Total Score	Mean of PAGI-SYM Subscales	Not applicable
Each item is scored by the subject on a 6-point Likert scale: 0=None, 1=Very Mild, 2=Mild, 3=Moderate, 4=Severe, and 5=Very Severe.		

The subscales scores will be calculated by taking the mean of the non-missing items in each subscale. The items will not be reversed scored prior to calculating the subscale score. The subscale scores will range from 0 (None) to 5 (Very Severe). Missing data for the subscales will be handled using the half-scale rule, i.e., a subscale score will be calculated when ≤50% of the items within a subscale are missing; if >50% of the items within a subscale are missing, the score will be set to missing. The PAGI-SYM total score will be calculated by taking the mean of the corresponding subscales. If any subscale scores are missing, the total score will be set to missing. Negative changes indicate improvement (decreased severity).

13.5.3 EQ-5D-5L

EQ-5D-5L descriptive system, may be converted into a single index value. The index values, presented in country specific value sets, are a major feature of the EQ-5D instrument, facilitating the calculation of quality-adjusted life years (QALYs) that are used to inform economic evaluations of health care interventions. By using the crosswalk link function of US value set for subjects in US and UK value set for subjects in Europe, and the individual responses to the EQ-5D-5L descriptive system, index values for the EQ-5D-5L can be calculated. Documents containing information on the crosswalk project, tables of values for all 3125 health states and the 'EQ-5D-5L Crosswalk Index Value Calculator' can be downloaded from the EuroQol website. EQ-5D-5L Index score is on a scale from 0 (worst imaginable health state) to 1 (best imaginable health state).

13.6 Estimands and Sensitivity Analysis

Estimands and Sensitivity Analysis for the Healing Phase and Maintenance Phase are summarized in the following tables.

13.6.1 Healing Phase

13.6.1.1 Intercurrent Event Types

Table 6 Intercurrent Event Types

Label	Intercurrent Event Type
IcEv1 (Discontinuation)	Discontinuation of treatment prior to completing the Healing Phase[1]
IcEv2 (Prohibited EE Medication[2])	PPI or H2 receptor antagonist taken during the Healing Phase (whether taken directly for EE or a different condition)
IcEv3 (Treatment Non-compliance)	Treatment compliance (%) is defined as (total actual capsules taken / total expected capsules) × 100 in range of 80%-120% during the Healing Phase. Total expected tablets vary according to the timing of the endoscopy confirming the end of Healing Phase.
IcEv4 (COVID-19 Impacts)	Unable to perform endoscopy assessment due to COVID-19 related reasons.

Abbreviation: EE, Erosive Esophagitis; IcEv, intercurrent event; PPI, proton pump inhibitor

Notes:

[1] Healing Phase can be completed at Week 2 or Unscheduled Visit (prior to Week 8) if there is absence of EE in endoscopy (confirmed healed status) but otherwise completes at Week 8.

[2] Other major protocol deviations that impact efficacy will be reviewed prior to breaking the blind and will be considered in the same manner as prohibited EE medication.

13.6.1.2 Estimand Specification

Table 7 Objectives, Endpoints, and Estimands

Objective	Estimand Label	Endpoint	Estimand Description
To assess the efficacy of vonoprazan (20 mg QD) compared to lansoprazole (30 mg QD) in healing of EE over 8 weeks in subjects with endoscopically proven EE.	Estimand 1 (Primary)	The percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in rates of EE healing by Week 8 in patients with endoscopically proven EE irrespective of whether they fully adhere to the treatment schedule or take prohibited EE medications (PPIs and H2 receptor antagonists).
As above with further detail given in estimand description.	Estimand 2 (Composite - Supportive)	Percentage of subjects with effective treatment defined as complete healing of EE by Week 8 as assessed by endoscopy and completed treatment with 80% -120% compliance and without intake of prohibited EE medications during the Healing Phase.	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in the effectiveness rates of treatment by Week 8 in patients with endoscopically proven EE where treatment effectiveness for EE healing requires that the patient has • EE healed by Week 8; and • completed treatment with 80%-120% compliance; and • no intake of prohibited EE medications (PPIs and H2 receptor antagonists). Otherwise, the treatment effectiveness for EE healing failed.
As above with further detail given in estimand description.	Estimand 3 (Hypothetical-Supplementary)	Percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in rates of EE healing by Week 8 in patients with endoscopically proven EE assuming the hypothetical situation that they complete treatment during the Healing Phase with 80%-120% compliance and no intake of prohibited EE medications (PPIs and H2 receptor antagonists).

Abbreviations: EE, Erosive Esophagitis; PPI, proton pump inhibitor.

Table 8 Summary of Estimands and Strategies for Managing Intercurrent Events

Estimand Label	Estimand 1 (Primary)	Estimand 2 (Composite - Supportive)	Estimand 3 (Hypothetical-Supplementary)
Estimand Description	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in rates of EE healing by Week 8 in patients with endoscopically proven EE irrespective of whether they fully adhere to the treatment schedule or take prohibited EE medications (PPIs and H2 receptor antagonists).	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in the effectiveness rates of treatment by Week 8 in patients with endoscopically proven EE where treatment effectiveness for EE healing requires that the patient has • EE healed by Week 8; and • completed treatment with 80%-120% compliance; and • no intake of prohibited EE medications (PPIs and H2 receptor antagonists). Otherwise, the treatment effectiveness for EE healing failed.	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in rates of EE healing by Week 8 in patients with endoscopically proven EE assuming the hypothetical situation that they complete treatment during the Healing Phase with 80%-120% compliance and no intake of prohibited EE medications (PPIs and H2 receptor antagonists).
Target Population	Patients with endoscopically proven EE (all Grades A-D) who would meet the inclusion/exclusion criteria of the study.	Patients with endoscopically proven EE (all Grades A-D) who would meet the inclusion/exclusion criteria of the study.	Patients with endoscopically proven EE (all Grades A-D) who would meet the inclusion/exclusion criteria of the study.

Table 8 Summary of Estimands and Strategies for Managing Intercurrent Events (continued)

Estimand Label	Estimand 1 (Primary)	Estimand 2 (Composite - Supportive)	Estimand 3 (Hypothetical-Supplementary)
Variable (measured on individual and defined as endpoint in ICH E9[R1])	EE Healing status ("healed" or "not healed"); taken as "healed" if there is an absence of EE in a centrally adjudicated endoscopy at Week 2, an unscheduled visit between Week 2 and Week 8, or at Week 8.	Treatment effectiveness for EE healing (success/failure) where failure is defined if • EE present at Week 8 as assessed by central adjudication of endoscopy; or • Healing Phase treatment discontinued; or • Healing Phase treatment compliance < 80% or >120%; or • Intake of prohibited EE medications (PPIs and H2 receptor antagonists). Conversely, success requires EE healing status result of healed, completed treatment with 80% -120% compliance and without intake of prohibited EE medications during the Healing Phase.	EE Healing status ("healed" or "not healed"); taken as "healed" if there is an absence of EE in a centrally adjudicated endoscopy at Week 2, an unscheduled visit between Week 2 and Week 8, or at Week 8.
Endpoint	Percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy.	Percentage of subjects with effective treatment defined as complete healing of EE by Week 8 as assessed by endoscopy and completed treatment with 80% - 120% compliance and without intake of prohibited EE medications during the Healing Phase.	Percentage of subjects who have complete healing of EE by Week 8 as assessed by endoscopy.
Treatment Condition(s)	Experimental: Vonoprazan 20 mg Reference: Lansoprazole 30 mg	Experimental: Vonoprazan 20 mg Reference: Lansoprazole 30 mg	Experimental: Vonoprazan 20 mg Reference: Lansoprazole 30 mg
Population-Level Summary	Difference in EE healing rates.	Difference in treatment effectiveness for EE healing rates	Difference in EE healing rates.

Table 8 Summary of Estimands and Strategies for Managing Intercurrent Events (continued)

Estimand Label	Estimand 1 (Primary)	Estimand 2 (Composite - Supportive)	Estimand 3 (Hypothetical-Supplementary)
Intercurrent Event Strategy			
IcEv1 (Discontinuation)	Composite	Composite	Hypothetical
IcEv2 (Prohibited EE Medication)	Treatment policy	Composite	Hypothetical
IcEv3 (Treatment Non-compliance)	Treatment policy	Composite	Hypothetical
IcEv4 (COVID-10 Impacts)	Hypothetical	Hypothetical	Hypothetical
Rationale for Strategies	Treatment policy strategy is used to understand the treatment effect in the Healing Phase irrespective of all intercurrent events in Table 6 as this reflects what might happen in clinical practice. Every attempt is taken to measure the healing status by Week 8 (regardless of prior treatment discontinuation or use of prohibited EE medications). Composite policy is used to understand treatment effectiveness where the decision of the patient to not comply with the treatment schedule or feel the need for additional (prohibited) EE medication is taken as failure. Hypothetical policy is used to estimate a treatment effect that is attributable to the difference between the treatments when taken as indicated without interference from prohibited EE medications. Given the non-inferiority hypothesis this approach has generally higher sensitivity to pick up an inferior treatment.		

Abbreviation: EE, Erosive Esophagitis; IcEv, intercurrent event; PPI, proton pump inhibitor

Non-compliance is defined as <80% or >120% of expected doses administered during the Healing Phase where expected number of doses is dependent on time of progression to the Maintenance Phase.

Note that on completing the Healing Phase (up to Week 8), patients may progress to the Maintenance Phase.

13.6.1.3 Statistical Methods and Sensitivity Analyses

Table 9 Summary of Statistical Methods and Sensitivity Analyses

		Main Estimation			
Estimand Label	**Estimand Description**	**Analysis Set**	**Imputation/Data/Censoring Rules**	**Analysis Model/Method**	**Sensitivity/Supportive Analyses**
Estimand 1 (Primary)	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in rates of EE healing by Week 8 in patients with endoscopically proven EE irrespective of whether they fully adhere to the treatment schedule or take prohibited EE medications (PPIs and H2 receptor antagonists).	MITT[1]	Where subjects have missing EE healing data due to reasons not related to COVID-19 (perhaps due to loss to follow up or withdrawing consent for endoscopy), the healing status will be imputed as "non-responder" (single imputation method known as non-responder imputation [NRI]). Where subjects have missing EE healing data due to reasons related to COVID-19, multiple imputation (MI) under missing at random (MAR) assumption will be utilized. The multiply imputed values of LA classification grade at Week 2 or 8 will be used to derive the missing EE healing status.	Farrington and Manning test with a non-inferiority margin of 10% for the difference in EE healing rates by Week 8 between treatments (vonoprazan 20 mg - lansoprazole 30 mg). The point estimate and 2-sided 95% CI of the difference in EE healing rate between vonoprazan 20 mg and lansoprazole 30 mg will be calculated via the Miettinen and Nurminen method.	1. Tipping point for NRI (Sensitivity). 2. Same as primary analysis on observed data only (Sensitivity). 3. Multiple imputation (MI) under missing at random (MAR) assumption followed by calculation of difference in EE healing rate with 95% CI(Sensitivity) using Rubin's Method. Healing Phase treatment and the LA classification grade at screening (A/B/C/D), Week 2 (Healed/A/B/C/D) will be used as variables in the multiple imputation model. The multiply imputed values of LA classification grade at Week 2 or 8 will be used to derive the missing EE healing status. 4. Per-protocol analysis per the SAP (Supportive).

Table 9 Summary of Statistical Methods and Sensitivity Analyses (continued)

Estimand Label	Estimand Description	Main Estimation			Sensitivity/Supportive Analyses
		Analysis Set	Imputation/Data/Censoring Rules	Analysis Model/Method	
Estimand 2 (Composite - Supportive)	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in the effectiveness rates of treatment by Week 8 in patients with endoscopically proven EE where treatment effectiveness for EE healing requires that the patient has • EE healed by Week 8; and • completed treatment with 80% -120% compliance; and • no intake of prohibited EE medications (PPIs and H2 receptor antagonists). Otherwise, the treatment effectiveness for EE healing failed. Subjects who do not have endoscopy performed by Week 8 due to COVID-19 related reasons will be imputed.	MITT[1]	The same MI model output generated above for Estimand 1 will be utilized. The multiply imputed values of LA classification grade at Week 2 or 8 will be used to derive the missing EE healing status. LA classification grade combined with treatment discontinuation, compliance and prohibited EE medication data will generate the composite treatment effectiveness endpoint.	The same analysis as for Estimand 1.	Not required as this is itself a supportive analysis and due to the composite definition, there should be only minimal missing data for this treatment effectiveness endpoint with the main reason being subjects withdrawing consent for endoscopy. Thus, the missing at random (MAR) assumption should be reasonable.

Table 9 Summary of Statistical Methods and Sensitivity Analyses (continued)

Estimand Label	Estimand Description	Main Estimation			Sensitivity/Supportive Analyses
		Analysis Set	Imputation/Data/Censoring Rules	Analysis Model/Method	
Estimand 3 (Hypothetical - Supplementary)	Difference between treatments (vonoprazan 20 mg – lansoprazole 30 mg) in rates of EE healing by Week 8 in patients with endoscopically proven EE assuming the hypothetical situation that they complete treatment during the Healing Phase with 80%-120% compliance and no intake of prohibited EE medications (PPIs and H2 receptor antagonists).	MITT[1]	Create a flag to exclude data after intercurrent events. Then, using the same type of MI model as used in Estimand 1 sensitivity analysis, impute missing values of LA classification grade at Week 2 or 8 which will be used to derive the missing EE healing status.	Same as primary analysis on the MI imputed data.	

Abbreviations: ITT, intent-to-treat; NRI, non-responder imputation; MI, multiple imputation; MAR, missing at random; MITT, modified intent-to-treat.

[1] The MITT set for the Healing Phase will be defined as all subjects randomized into the Healing Phase who have documented EE at baseline and receive at least 1 dose of study drug during the Healing Phase. All analyses using the MITT set will group subjects according to the randomized treatment.

13.6.2 Maintenance Phase

13.6.2.1 Intercurrent Event Types

Table 10 Intercurrent Event Types

Label	Intercurrent Event Type
IcEv1 (Discontinuation)	Discontinuation of maintenance treatment prior to completing 24 weeks Maintenance Phase.
IcEv2 (Prohibited EE Medication[1])	PPI or H2 receptor antagonist taken during the Maintenance Phase (whether taken directly for EE or a different condition)
IcEv3 (Treatment Non-compliance)	Treatment compliance (%) is defined as (total actual capsules taken / total expected capsules) × 100 in range of 80%-120% during the Maintenance Phase.
IcEv4 (COVID-19 Impacts)	Unable to perform endoscopy assessment due to COVID-19 related reasons.

Abbreviation: EE, Erosive Esophagitis; IcEv, intercurrent event; PPI, proton pump inhibitor

Notes:

[1] Other major protocol deviations that impact efficacy will be reviewed prior to breaking the blind and will be considered in the same manner as prohibited EE medication.

13.6.2.2 Estimand Specification

Table 11 Objectives, Endpoints, and Estimands

Objective	Estimand Label	Endpoint	Estimand Description
To assess the efficacy of vonoprazan (10 or 20 mg QD) compared to lansoprazole (15 mg QD) in maintenance of healing of EE over 24 weeks in subjects with endoscopically proven EE.	Estimand 1 (Primary)	Percentage of subjects who maintain complete healing of EE after24 Weeks as assessed by endoscopy	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in rates of maintained EE healing up to Week 24 in patients healed from an episode of endoscopically proven EE irrespective of whether they fully adhere to the maintenance treatment schedule or take prohibited EE medications (PPIs and H2 receptor antagonists).
As above with further detail given in estimand description.	Estimand 2 (Composite - Supportive)	Percentage of subjects with effective maintenance treatment defined as maintenance of complete healing of EE up to Week 24 as assessed by endoscopy and completed maintenance treatment with 80% -120% compliance and without intake of prohibited EE medications during the Maintenance Phase.	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in the effectiveness rates of maintenance treatment up to Week 24 in patients healed from an episode of endoscopically proven EE where maintenance treatment effectiveness requires that the patient has • Maintained complete EE healing up to Week 24; and • completed 24 weeks of maintenance treatment with 80%-120% compliance; and • no intake of prohibited EE medications (PPIs and H2 receptor antagonists). Otherwise, the maintenance treatment effectiveness failed.
As above with further detail given in estimand description.	Estimand 3 (Hypothetical-Supplementary)	Percentage of subjects who maintain complete healing of EE after 24 Weeks as assessed by endoscopy	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in rates of maintained EE healing up to Week 24 in patients healed from an episode of endoscopically proven EE assuming the hypothetical situation that they complete 24 weeks of maintenance treatment with 80%-120% compliance and no intake of prohibited EE medications (PPIs and H2 receptor antagonists).

Abbreviations: EE, Erosive Esophagitis; PPI, proton pump inhibitor.

Table 12 Summary of Estimands and Strategies for Managing Intercurrent Events

Estimand Label	Estimand 1 (Primary)	Estimand 2 (Composite - Supportive)	Estimand 3 (Hypothetical-Supplementary)
Estimand Description	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in rates of maintained EE healing up to Week 24 in patients who have healed from an episode of endoscopically proven EE irrespective of whether they fully adhere to the maintenance treatment schedule or take prohibited EE medications (PPIs and H2 receptor antagonists).	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in the effectiveness rates of maintenance treatment up to Week 24 in patients healed from an episode of endoscopically proven EE where maintenance treatment effectiveness requires that the patient has • maintained complete EE healing up to Week 24; and • completed 24 weeks of maintenance treatment with 80%-120% compliance; and • no intake of prohibited EE medications (PPIs and H2 receptor antagonists). Otherwise, the maintenance treatment effectiveness failed.	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in rates of maintained EE healing up to Week 24 in patients healed from an episode of endoscopically proven EE assuming the hypothetical situation that they complete 24 weeks of maintenance treatment with 80%-120% compliance and no intake of prohibited EE medications (PPIs and H2 receptor antagonists).
Target Population	Patients who have a confirmed absence of EE having been healed from an episode of endoscopically proven EE (all Grades A-D) and would meet the inclusion/exclusion criteria of the study and be willing to continue directly from healing therapy to maintenance therapy.	Patients who have a confirmed absence of EE having been healed from an episode of endoscopically proven EE (all Grades A-D) who would meet the inclusion/exclusion criteria of the study and be willing to continue directly from healing therapy to maintenance therapy.	Patients who have a confirmed absence of EE having been healed from an episode of endoscopically proven EE (all Grades A-D) who would meet the inclusion/exclusion criteria of the study and be willing to continue directly from healing therapy to maintenance therapy.

Table 12 Summary of Estimands and Strategies for Managing Intercurrent Events (continued)

Estimand Label	Estimand 1 (Primary)	Estimand 2 (Composite - Supportive)	Estimand 3 (Hypothetical-Supplementary)
Variable (measured on individual and defined as endpoint in ICH E9[R1])	Maintenance of EE healing up to Week 24 status ("healing maintained" or "recurrence"); taken as "healing maintained" if there is an absence of EE in a centrally adjudicated endoscopy at Week 24.	Maintenance of EE healing treatment effectiveness (success/failure) where failure is defined if • EE recurrence is present by Week 24 as assessed by endoscopy; or • Maintenance Phase treatment discontinued; or • Maintenance Phase treatment compliance < 80% or >120%; or • Intake of prohibited EE medications (PPIs and H2 receptor antagonists). Conversely, success requires EE healing maintained up to Week 24, completed 24 weeks of maintenance treatment with 80% -120% compliance and without intake of prohibited EE medications during the Maintenance Phase.	Maintenance of EE healing up to Week 24 status ("healing maintained" or "recurrence"); taken as "healing maintained" if there is an absence of EE in a centrally adjudicated endoscopy at Week 24.
Endpoint	Percentage of subjects who maintain complete healing of EE after 24 weeks as assessed by endoscopy	Percentage of subjects with effective maintenance treatment defined as maintenance of complete healing of EE up to Week 24 as assessed by endoscopy and completed maintenance treatment with 80% - 120% compliance and without intake of prohibited EE medications during the Maintenance Phase.	Percentage of subjects who maintain complete healing of EE after 24 weeks as assessed by endoscopy

Table 12 Summary of Estimands and Strategies for Managing Intercurrent Events (continued)

Estimand Label	Estimand 1 (Primary)	Estimand 2 (Composite - Supportive)	Estimand 3 (Hypothetical-Supplementary)
Treatment Condition(s)	Experimental: Vonoprazan 10 mg, Vonoprazan 20 mg Reference: Lansoprazole 15 mg	Experimental: Vonoprazan 10, Vonoprazan 20 mg Reference: Lansoprazole 15 mg	Experimental: Vonoprazan 10 mg, Vonoprazan 20 mg Reference: Lansoprazole 15 mg
Population-Level Summary	Difference in maintained EE healing rates.	Difference in maintenance treatment effectiveness rates	Difference in maintained EE healing rates.
Intercurrent Event Strategy			
IcEv1 (Discontinuation)	Composite	Composite	Hypothetical
IcEv2 (Prohibited EE Medication)	Treatment policy	Composite	Hypothetical
IcEv3 (Treatment Non-compliance)	Treatment policy	Composite	Hypothetical
IcEv4 (COVID-10 Impacts)	Hypothetical	Hypothetical	Hypothetical
Rationale for Strategies	Treatment policy strategy is used to understand the treatment effect in the Maintenance Phase irrespective of all intercurrent events in Table 10 as this reflects what might happen in clinical practice. Every attempt is taken to measure the healing status at Week 24 (regardless of maintenance treatment discontinuation or use of prohibited EE medications). Composite policy is used to understand maintenance treatment effectiveness where the decision of the patient to not comply with the maintenance treatment schedule or feel the need for additional (prohibited) EE medication is taken as failure. Hypothetical policy is used to estimate a treatment effect that is attributable to the difference between the maintenance treatments when taken as indicated without interference from prohibited EE medications. Given the non-inferiority hypothesis this approach has generally higher sensitivity to pick up an inferior treatment.		

Abbreviation: EE, Erosive Esophagitis; IcEv, intercurrent event; PPI, proton pump inhibitor
Non-compliance is defined as <80% or >120% of expected doses administered during the maintenance phase.

13.6.2.3 Statistical Methods and Sensitivity Analyses

Table 13 Summary of Statistical Methods and Sensitivity Analyses

		Main Estimation			
Estimand Label	**Estimand Description**	**Analysis Set**	**Imputation/Data/Censoring Rules**	**Analysis Model/Method**	**Sensitivity/Supportive Analyses**
Estimand 1 (Primary)	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in rates of maintained EE healing up to Week 24 in patients healed from an episode of endoscopically proven EE irrespective of whether they fully adhere to the treatment schedule or take prohibited EE medications (PPIs and H2 receptor antagonists).	MITT[1]	Where subjects have missing EE healing data due to reasons not related to COVID-19 (perhaps due to loss to follow up or withdrawing consent for endoscopy), the EE healing maintenance will be imputed as "non-responder/recurrence" (single imputation method known as non-responder imputation [NRI]). Where subjects have missing EE healing data due to reasons related to COVID-19, multiple imputation (MI) under missing at random (MAR) assumption will be utilized. The multiply imputed values of LA classification grade at Week 24 will be used to derive the missing EE maintenance status.	Farrington and Manning test with a non-inferiority margin of 10% for the difference in maintenance of EE healing rates up to Week 24 between treatments [each dose of vonoprazan (10 or 20 mg) - lansoprazole 15 mg)]. The point estimates and 2-sided 95% CI of the differences in maintenance of EE healing rates between each dose of vonoprazan (10 or 20 mg) and lansoprazole 15 mg will be calculated via the Miettinen and Nurminen method.	1. Tipping point for NRI (Sensitivity). 2. Same as primary analysis on observed data only (Sensitivity). 3. Multiple imputation (MI) under missing at random (MAR) assumption followed by calculation of difference in maintenance of EE healing rate with 95% CI (Sensitivity) using Rubin's Method. Maintenance Phase treatment and the LA classification grade at screening (A/B/C/D) will be used as variables in the multiple imputation model. The multiply imputed values of LA classification grade at Week 24 will be used to derive the missing maintenance of EE healing status. 4. Per-protocol analysis per the SAP (Supportive).

Table 13 Summary of Statistical Methods and Sensitivity Analyses (continued)

Estimand Label	Estimand Description	Main Estimation			Sensitivity/Supportive Analyses
		Analysis Set	Imputation/Data/Censoring Rules	Analysis Model/Method	
Estimand 2 (Composite - Supportive)	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in the effectiveness rates of maintenance treatment up to Week 24 in patients healed from an episode of endoscopically proven EE where maintenance treatment effectiveness requires that the patient has • Maintained EE healing up to Week 24; and • completed 24 weeks of maintenance treatment with 80% -120% compliance; and • no intake of prohibited EE medications (PPIs and H2 receptor antagonists). Otherwise, the maintenance treatment effectiveness failed. Subjects who do not have endoscopy performed by Week 24 due to COVID-19 related reasons will be imputed.	MITT[1]	The same MI model output generated above for Estimand 1 will be utilized. The multiply imputed values of LA classification grade at Week 24 will be used to derive the missing EE healing maintenance. LA classification grade combined with maintenance treatment discontinuation, compliance and prohibited EE medication data will generate the composite maintenance treatment effectiveness endpoint.	The same analysis as for Estimand 1.	Not required as this is itself a supportive analysis and due to the composite definition, there should be only minimal missing data for this treatment effectiveness endpoint with the main reason being subjects withdrawing consent for endoscopy. Thus, the missing at random (MAR) assumption should be reasonable.

Table 13 Summary of Statistical Methods and Sensitivity Analyses (continued)

Estimand Label	Estimand Description	Main Estimation			Sensitivity/Supportive Analyses
		Analysis Set	Imputation/Data/Censoring Rules	Analysis Model/Method	
Estimand 3 (Hypothetical - Supplementary)	Difference between maintenance treatments (vonoprazan 10 or 20 mg – lansoprazole 15 mg) in rates of maintained EE healing up to Week 24 in patients healed from an episode of endoscopically proven EE assuming the hypothetical situation that they complete 24 weeks of maintenance treatment with 80%-120% compliance and no intake of prohibited EE medications (PPIs and H2 receptor antagonists).	MITT[1]	Create a flag to exclude data after intercurrent events. Then, using the same type of MI model as used in Estimand 1 sensitivity analysis, impute missing values of LA classification grade at Week 24 which will be used to derive the missing maintenance of EE healing status.	Same as primary analysis on the MI imputed data.	

Abbreviations: ITT, intent-to-treat; NRI, non-responder imputation; MI, multiple imputation; MAR, missing at random; MITT, modified intent-to-treat.

[1] The MITT set for the Maintenance Phase will be defined as all subjects randomized into the Maintenance Phase who have healed EE at the end of the Healing Phase and receive at least 1 dose of study drug during the Maintenance Phase. All analyses using the MITT set will group subjects according to the randomized treatment.