Statistical Analysis Plan: I5Q-MC-CGAX (Version 1)

A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study of the Efficacy and Safety of Galcanezumab in Patients with Episodic Migraine – the PERSIST Study

NCT03963232

Approval Date: 24-Aug-2021

1. Statistical Analysis Plan: I5Q-MC-CGAX: A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study of the Efficacy and Safety of Galcanezumab in Patients with Episodic Migraine – the PERSIST Study

Galcanezumab (LY2951742) Episodic Migraine

I5Q-MC-CGAX is a multicenter, randomized, double-blind, parallel, placebo-controlled Phase 3 trial comparing 120-mg galcanezumab with placebo given as subcutaneous injection once monthly over 3 months in patients who meet International Classification of Headache Disorders (ICHD) 3 criteria for a diagnosis of migraine with or without aura (1.1 or 1.2), with 4 to 14 MHDs per month.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I5Q-MC-CGAX
Phase 3

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly on date provided below.

Approval Date: 24-Aug-2021 GMT

2. Table of Contents

Section **Page**

1. Statistical Analysis Plan: I5Q-MC-CGAX: A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study of the Efficacy and Safety of Galcanezumab in Patients with Episodic Migraine – the PERSIST Study 1
2. Table of Contents 2
3. Revision History 6
4. Study Objectives 7
5. Study Design 11
5.1. Summary of Study Design 11
5.2. Determination of Sample Size 11
5.3. Method of Assignment to Treatment 12
5.4. Endpoints 12
5.4.1. Efficacy endpoints 12
5.4.1.1. Primary Measure: The Number of Monthly Migraine Headache Days 15
5.4.1.2. Secondary and Exploratory Efficacy Measures 16
5.4.2. Other efficacy measures 18
5.4.3. Health Outcome Questionnaires 18
5.4.3.1. Migraine Specific Quality of Life (MSQ) v2.1 18
5.4.3.2. Migraine Disability Assessment test (MIDAS) 20
5.4.3.3. Migraine Interictal Burden Scale (MIBS-4) 20
5.4.3.4. Patient Health Questionnaire-9 21
5.4.3.5. 7-Item Generalized Anxiety Disorder Scale 21
5.4.4. Safety endpoints 22
5.4.5. Immunogenicity Endpoints 22
5.4.6. Pharmacokinetic Assessment 22
6. A Priori Statistical Methods 23
6.1. General Considerations 23
6.2. Adjustments for Covariates 30
6.3. Handling of Dropouts or Missing Data 30
6.4. Multiple Comparisons/Multiplicity 31
6.5. Patient Disposition 32
6.6. Patient Characteristics 32
6.7. Treatment Compliance 33
6.8. Electronic Patient-Reported Outcome Diary Compliance 33

6.9. Concomitant Therapy 34
6.10. Efficacy Analyses 34
6.10.1. Primary Outcome and Methodology 34
6.10.2. Sensitivity Analysis for Primary Outcome 35
6.10.3. Secondary and Exploratory Efficacy Analyses 36
6.11. Health Outcomes Analyses 40
6.12. Pharmacokinetic Analyses 41
6.13. Evaluation of Immunogenicity 41
6.14. Safety Analyses 42
6.14.1. Extent of Exposure 42
6.14.2. Adverse Events 42
6.14.2.1. Safety Topics of Interest 43
6.14.2.1.1. Potential Hypersensitivity Events 43
6.14.2.1.2. Adverse Events Related to Injection Sites 43
6.14.2.1.3. Upper Respiratory Tract Infections 44
6.14.3. Clinical Laboratory Evaluation 44
6.14.4. Vital Signs and Weight 45
6.14.5. Electrocardiograms 47
6.15. Subgroup Analyses 49
6.16. Protocol Deviations 49
6.17. Interim Analyses 49
6.18. Clinical Trial Registry Analyses 50
7. Unblinding Plan 51
8. References 52

Table of Contents

Table **Page**

Table CGAX. 5.1 Migraine and Headache Endpoint Definitions 12

Table CGAX. 5.2 Electronic Patient-Reported Outcomes Diary Questions, Responses, and Assignment to Headache Type..... 14

Table CGAX. 5.3 Item Values for Migraine Specific Quality of Life (MSQ) Item Responses 19

Table CGAX. 5.4 Scores for Migraine Interictal Burden Scale (MIBS-4) Item Responses..... 21

Table CGAX. 6.1 Patient Population with Baseline and Postbaseline Definitions by Study Period and Type of Analysis 25

Table CGAX. 6.2 Secondary and Exploratory Efficacy Variables and Analysis Methods 40

Table CGAX. 6.3 Criteria for Categorical Changes of Interest in Vital Signs and Weight..... 47

Table CGAX. 6.4 Criteria for Treatment-Emergent Changes in ECG Intervals and Heart Rate 48

Table CGAX. 6.5 Definition of Subgroup Variables..... 49

Table of Contents

Figure **Page**

Figure CGAX. 5.1 Illustration of study design for Clinical Protocol I5Q-MC-CGAX.....11

Figure CGAX. 6.1 Gatekeeping sequence for testing of primary and key secondary endpoints.....32

3. Revision History

SAP Version 1 was approved prior to unblinding at the primary database lock.

4. Study Objectives

Objectives	Endpoints
Primary Objective To test the hypothesis that galcanezumab (240-mg loading dose followed by 120 mg monthly) is superior to placebo in the prevention of migraine in patients with EM.	The overall mean change from baseline in the number of monthly MHDs during the 3-month double-blind treatment phase
Key Secondary Objectives	
• To compare galcanezumab with placebo with respect to a 50% response rate	• The overall proportion of patients with ≥50% reduction from baseline in monthly MHDs during the 3-month double-blind treatment phase
• To compare galcanezumab with placebo with respect to change in functioning	• The overall mean change of Months 1 to 3 from baseline in the Role Function-Restrictive domain score of the MSQ v2.1
• To compare galcanezumab with placebo with respect to 75% response rate	• The overall proportion of patients with ≥75% reduction from baseline in monthly MHDs during the 3-month double-blind treatment phase
• To compare galcanezumab with placebo with respect to 100% response rate	• The overall proportion of patients with a 100% reduction from baseline in monthly MHDs during the 3-month double-blind treatment phase
Other Secondary Objectives	
• To compare galcanezumab with placebo with respect to a 30% response rate	• The overall proportion of patients with ≥30% reduction from baseline in monthly MHDs during the 3-month double-blind treatment phase
• To compare galcanezumab with placebo with respect to distribution of response rates	• Cumulative distribution of monthly MHDs response rates during the 3-month double-blind treatment phase
• To compare galcanezumab with placebo with respect to change in use of acute headache treatment	• The overall mean change from baseline in the number of monthly MHDs taking medication for the acute treatment of headache during the 3-month double-blind treatment phase
• To compare galcanezumab with placebo with respect to change in headache days	• The overall mean change from baseline in the number of monthly headache days during the 3-month double-blind treatment phase

Objectives (cont.)	Endpoints (cont.)
Other Secondary Objectives (cont.)	
• To compare galcanezumab with placebo with respect to change in moderate to severe headache days	• The overall mean change from baseline in the number of monthly moderate to severe headache days during the 3-month double-blind treatment phase
• To compare galcanezumab with placebo with respect to time to 50% response	• Time to first occurrence of a ≥50% reduction from baseline in the number of monthly MHDs (Kaplan-Meier analysis)
• To compare galcanezumab with placebo with respect to onset of effect	• The initial month at which statistical separation in mean change from baseline in the number of monthly MHDs is demonstrated and maintained at all subsequent months through Month 3
• To compare galcanezumab with placebo with respect to onset of 50% sustained response	• The initial month at which statistical separation in the proportion of patients meeting at least a 50% reduction in monthly MHDs that is maintained at all subsequent months through Month 3
• To compare galcanezumab with placebo with respect to maintenance of 50% response	• The proportion of patients who maintain 50% response criteria for all 3 months of double-blind treatment
• To compare galcanezumab with placebo with respect to changes in other efficacy parameters, specifically: o ICHD MHDs o migraine attacks o migraine headache hours o headache hours o severity of remaining migraines	• Overall mean change from baseline (during the 3-month double-blind treatment phase) on the following monthly measures: o ICHD MHDs o migraine attacks o migraine headache hours o headache hours o severity of remaining migraines
• To compare galcanezumab with placebo with respect to change in patients' global impression of migraine severity	• Mean change from baseline in the PGI-S at Month 3
• To compare galcanezumab with placebo with respect to changes in disability and quality of life	• Mean change from baseline to Month 3 on the MIDAS total score and individual items • Overall mean change from baseline to Months 1 to 3 on MSQ v2.1 total score, and Role Function-Preventive and Emotional Function domain scores

Objectives (cont.)	Endpoints (cont.)
Other Secondary Objectives (cont.)	
• To compare galcanezumab with placebo with respect to safety and tolerability	• Analysis of: o TEAEs o SAEs o discontinuation due to AEs o discontinuation rates o vital signs and weight o ECGs o laboratory measures
• To evaluate the immunogenicity of galcanezumab	• Incidence and consequences of anti-drug antibodies and neutralizing anti-drug antibodies to galcanezumab
• To evaluate the pharmacokinetics of galcanezumab	• Serum concentrations of galcanezumab
Tertiary Objectives	
• To compare galcanezumab with placebo with respect to changes in quality of life	• Percentages of patients with ≥50% improvement in MIDAS total score • The mean change from baseline in the MIBS-4 total score
• To compare galcanezumab with placebo with respect to changes in symptoms that accompany migraine or probable migraine	• Change from baseline in the number of monthly MHDs with: o nausea and/or vomiting o photophobia and phonophobia
• To compare galcanezumab with placebo with respect to changes in symptoms of depression and anxiety	• Changes from baseline to Month 3 on the following measures: o PHQ-9 o GAD-7
• To assess changes in efficacy, safety, and functional outcomes during Study Period IV (open-label treatment)	• In Study Period IV: o Mean changes in all continuous measures of efficacy, safety, and functional outcomes that are also assessed in the double-blind period o Among patients previously treated with galcanezumab who met 50% response criteria at Month 3 in the double-blind treatment period, the proportion of patients who demonstrate response throughout the open-label treatment period

Objectives (cont.)	Endpoints (cont.)
Tertiary Objectives (cont.)	
• To assess changes in efficacy outcomes during Study Period V as collected by ePRO diary data	• In Study Period V: o Mean change in monthly MHDs from baseline to the end of the post-treatment follow-up phase o Time to first loss of response among patients who met the 50% response rate criteria at their last injection interval o Time to initiation of treatment with a migraine prevention medication

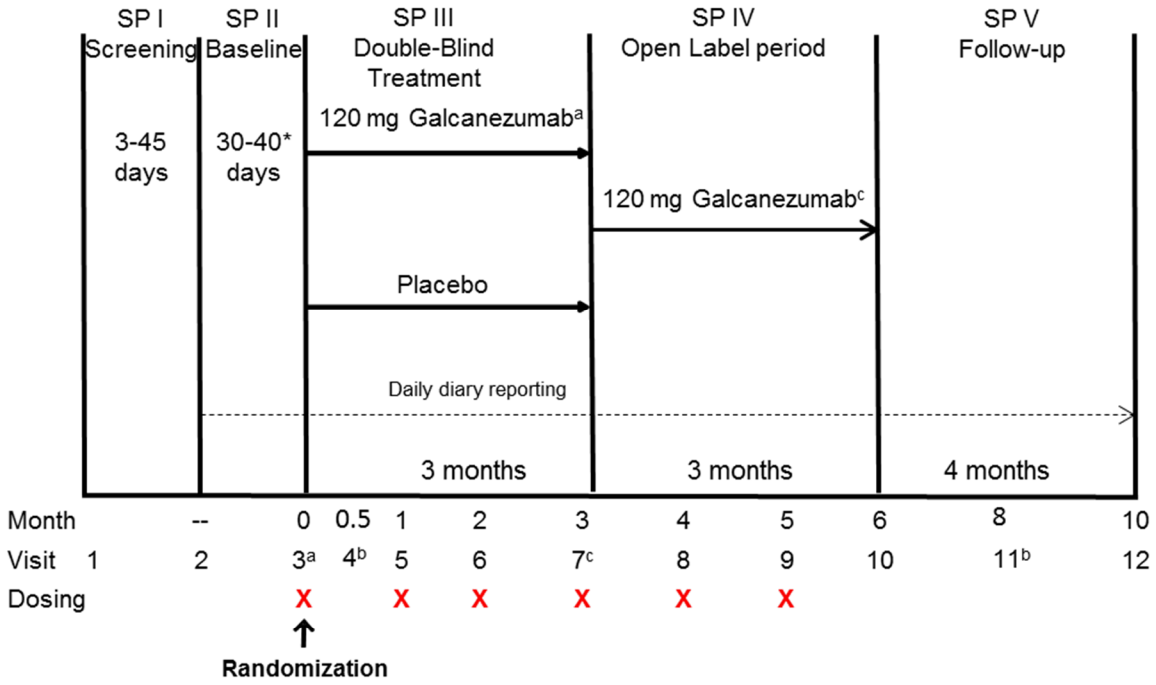
Abbreviations: AE = adverse event; ECG = electrocardiogram; EM = episodic migraine; ePRO = electronic patient reported outcomes; GAD-7 = 7-item Generalized Anxiety Disorder Scale; ICHD = International Classification of Headache Disorders; MHD = migraine headache day; MIBS-4 = Migraine Interictal Burden Scale; MIDAS = Migraine Disability Assessment test; MSQ (v2.1) = Migraine Specific Quality-of-Life Questionnaire; PGI-S = Patient Global Impression of Severity; PHQ-9 = Patient Health Questionnaire-9; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

5. Study Design

5.1. Summary of Study Design

Study CGAX is a Phase 3, multicenter, double-blind, randomized, placebo-controlled study of galcanezumab in patients with EM, who have 4 to 14 MHDs (with or without aura) per month. The study has 5 periods, including a prospective baseline phase to determine patient eligibility.

Figure CGAX. 5.1 illustrates the study design.



*Eligibility period determined between a minimum of 30 days and a maximum of 40 days. Investigators may have up to 5 additional days (beyond the 40 days) if needed to schedule patients' Visit 3 appointment.

[a]Patients randomized to galcanezumab will receive a loading dose of 240 mg at the first injection only (Visit 3).

[b] Visit 4 and Visit 11 will be telephone visits.

[c] At Visit 7, patients randomized to placebo who enter the open-label extension will receive galcanezumab at a dose of 240 mg. Patients randomized to galcanezumab will continue the dose of 120 mg.

Abbreviations: SP = study period.

Figure CGAX. 5.1 Illustration of study design for Clinical Protocol I5Q-MC-CGAX

5.2. Determination of Sample Size

Approximately 486 patients will be randomized in a 1:1 ratio to the galcanezumab 120-mg/month or placebo treatment groups. With 243 patients per treatment group, this study will have approximately 90% power to detect an effect size of 0.33 between galcanezumab 120-mg/month and placebo treatment groups. Within the framework of a mixed model repeated measures (MMRM), the sample size is determined using a between-treatment group t-test, 2-sided with a type I error of 0.05, assuming a discontinuation rate of 20% during the double-blind

phase. Parameters used in the sample size calculations are based on results from 2 double-blind, placebo-controlled Phase 3 studies (CGAG and CGAH) and clinical justification.

5.3. Method of Assignment to Treatment

Following a prospective baseline (30 to 40 days) period, eligible patients will be randomized in a 1:1 ratio to receive placebo or 120 mg/month of galcanezumab (with a 240 mg loading dose at the first injection [administered as 2 injections of 120 mg at Visit 3]), respectively, and will begin a 3-month double-blind treatment phase. Patients who complete the double-blind period may enter a 3 -month open-label extension phase during which all patients will receive galcanezumab 120 mg/month. At Visit 7, patients originally assigned to placebo will receive an initial loading dose of 240 mg galcanezumab; patients originally assigned to galcanezumab will continue the dose of 120 mg but will receive 2 injections (1 injection of 120 mg galcanezumab and 1 injection of placebo) to maintain blinding. At Visit 8 and Visit 9 of the open-label phase, all patients will receive a 120 mg dose of galcanezumab. All patients will be followed for a 4-month, post-treatment phase during which no study medication will be administered.

Patients who meet all criteria for enrollment will be randomized to double-blind treatment at Visit 3. Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS). The IWRS will be used to assign double-blind investigational product to each patient. Site personnel will confirm that they have located the correctly assigned package by entering the confirmation number found on the package into the IWRS.

To achieve between-group comparability, the randomization will be stratified by country and baseline migraine frequency (<8 MHDs versus ≥8 MHDs). To ensure an appropriate balance of patients with low- and high-frequency MHDs, the sponsor will stop enrollment of low-frequency patients if the number exceeds an estimated 292.

5.4. Endpoints

5.4.1. Efficacy endpoints

Migraine and headache endpoints are defined in Table CGAX. 5.1. Each month is defined as a 30-day period with migraine or headache measures normalized from the visit intervals.

Table CGAX. 5.1 Migraine and Headache Endpoint Definitions

Diagnosis	Definition/Criteria
Migraine headache	A headache, with or without aura, of ≥30 minutes duration with both of the following required features (A and B): A. At least 2 of the following headache characteristics: • Unilateral location • Pulsatile quality • Moderate or severe pain intensity • Aggravation by or causing avoidance of routine physical activity

	AND B. During headache at least one of the following: • Nausea and/or vomiting • Photophobia and phonophobia *(Definition adapted from the Standard International Headache Society [IHS] International Classification of Headache Disorders (ICHD)-3)*
Probable migraine	A headache of ⩾30 minutes duration, with or without aura, but missing one of the migraine features in the IHS ICHD-3 definition. To be exact, it meets either at least two A criteria and zero B criteria, or one A criteria and at least one B criteria.
Migraine headache day (MHD) (primary objective)	A calendar day on which a migraine or probable migraine occurred.
ICHD migraine headache day	A calendar day on which a migraine occurs.
Migraine headache attack	Beginning on any day a migraine headache or probable migraine headache is recorded and ends when a migraine-free day occurs.
Non-migraine headache	All headaches of at least 30 minutes duration not fulfilling the definition of migraine or probable migraine are classified as non-migraine headaches.
Non-migraine headache day	A calendar day on which a non-migraine headache occurred.
Headache day	A calendar day on which any type of headache occurs (including migraine, probable migraine, and non-migraine headache).
Episodic migraine	4 to 14 migraine headache days and <15 headache days per 30-day period in the prospective baseline period.

Abbreviations: ICHD = International Classification of Headache Disorders; IHS = International Headache Society; MHD = migraine headache day.

Headache information will be collected via an ePRO diary. Patients will need to enter diary data daily beginning from Visit 2, continuing until Visit 12.

Information recorded in the ePRO diary, the possible responses and the assignment to the type of headache is presented in Table CGAX. 5.2.

Table CGAX. 5.2 Electronic Patient-Reported Outcomes Diary Questions, Responses, and Assignment to Headache Type

QUESTION	RESPONSES	HEADACHE ASSIGNMENT
Q1. Yesterday, did you have a headache that lasted for thirty minutes or more?	Yes	Migraine if at least 2 migraine Criteria As and at least 1 migraine Criterion B were met.
	No[a]	
Q2. Enter the total number of hours you had a headache yesterday.	Range 1 to 24	If ≥1, the headache was counted as a headache day.
Q3. Yesterday, what was the worst headache pain?	Mild	
	Moderate	Migraine Criteria A
	Severe	Migraine Criteria A
Q4. Yesterday, was the headache throbbing or pounding?	Yes	Migraine Criteria A
	No	
Q5. Yesterday, was the headache just on the right or left side of your head?	Yes	Migraine Criteria A
	No	
Q6. Yesterday, was the headache made worse by your usual daily activity (for example, walking or go upstairs)?	Yes	Migraine Criteria A
	No	
Q7. Yesterday, did you want to lay down when you had headache?	Yes	
	No	
Q8. Yesterday, did the headache affect your routine daily activities?	Yes	
	No	
Q9. Yesterday, did the headache come with sensitivity to light and sound?	Yes	Migraine Criteria B
	No	
Q10. Yesterday, did you feel sick to the stomach or throw-up with the headache?	Yes	Migraine Criteria B
	No	
Q11. Yesterday, did you take any medicine for your headache?	Yes	Medication was only counted as headache medication on a day a headache occurred. If yes, patient recorded the name, frequency, and dose of medications taken.
	No	
Q12. Yesterday, do you believe you experienced a migraine with any trigger?	Yes	
	No	
Q13. Yesterday, if you experienced a migraine with any trigger, do you think the trigger was: • Weather change • Environment change • Physical discomfort • Stress • Tiredness • Speical food • Not listed above	Weather change	
	Environment change	
	Physical discomfort	
	Stress	
	Tiredness	
	Speical food	
	Not listed above	
Q14. Yesterday, did you have your menstrual period?	Yes	
	No	

Abbreviations: Q = question.

[a] If "No" was answered for Q1, then the patient skipped Q2-Q8, and only answered Q9-Q14.

Note: Patients who missed a day of data entry had the ability to make-up the missed day by reporting for the day before yesterday and then for yesterday at their next log-in.

5.4.1.1. Primary Measure: The Number of Monthly Migraine Headache Days

The primary measure is the number of monthly migraine headache days (MHD). A migraine headache day is defined as a calendar day on which a migraine or probable migraine occurs.

The primary measure of the number of monthly migraine headache days will be summarized from the daily ePRO data for each patient in that period (including 30-40 days of daily data from the baseline period prior to randomization, 3 months of daily data during the double-blind treatment phase, 3 months of daily data during the open-label treatment phase, and 4 months of daily data during the post-treatment phase).

The daily data will be aggregated, and the number of migraine headache days will be provided for each of the 30-day periods. In calculating the number of migraine headache days for each period, if the period is not equal to 30 days, the number of migraine headache days will be adjusted by multiplying the number of migraine headache days by (30/x) where 'x' is the total number of non-missing diary days in the period.

For the 4 months of the post-treatment period (Study Period V [SP V]), the monthly interval will be derived as below. Firstly, the 2-month visit interval will be split into two one-month periods for efficacy measures. If the number of days between 2 visits (visit x+1 date - visit x date) is even, the first half will be in the first one-month period and the second half will be in the second one-month period. If the number of days between 2 visits is odd, then the days will be split similarly, but the first one-month period will have 1 day more than the second one-month period. Secondly, after the two month visit intervals are split into two one-month periods, the monthly data will be derived in the same way for each one-month period as for treatment phase. For patients who discontinued early during the post-treatment phase, if the date of discontinuation is within 30 days of previous visit date, all data between the previous visit date and the discontinuation date will go to 1 monthly period; if the date of discontinuation is more than 30 days of the previous visit date, then the first 30 days will be the first monthly period, and the rest will be considered as part of the second monthly period.

This approach to missing ePRO diary data assumes that the rate of migraine headaches per day is the same for days with missing and non-missing ePRO diary days.

The same approach will also be applied to secondary and exploratory efficacy measures that are derived from ePRO data.

Additionally, if the compliance rate for each monthly interval is $\leqslant$50%, then all endpoints to be derived from the ePRO diary data for that one-month period will be considered missing. For the post-treatment phase, the derived one-month periods (resulting from splitting two-month visit interval) will be treated similarly.

For a patient who discontinues early in the double-blind (Study Period III [SP III]) or open-label treatment phase (Study Period IV [SP IV]) or post-treatment phase (SP V), compliance rate for the last month of that study period will be calculated with a denominator of 30 days (or the actual number of days in the interval if greater than 30). For the rest of months and patients, the compliance rate will be calculated as described in Section 6.8.

5.4.1.2. Secondary and Exploratory Efficacy Measures

The same approach to adjusting the number of days within each period to a 30-day period and the same approach to imputing monthly data based on compliance as for the primary measure will be applied to all efficacy measures that are derived from ePRO diary data and need normalization to 30-day period, including:

- **Number of ICHD migraine headache days** (ICHD MHD) is calculated as the number of calendar days in a 30-day period on which a migraine occurs. Probable migraine is excluded.
- **Number of headache days** is calculated as the number of calendar days in a 30-day period on which a headache occurs.
- **Number of moderate to severe headache days** is calculated as the number of calendar days in a 30-day period on which a headache occurs with a moderate or severe severity.
- **Number of headache hours** is calculated as the total number of headache hours in a 30-day period on which a headache occurred.
- **Number of migraine headache hours** is calculated as the total number of headache hours in a 30-day period on days when a migraine or probable migraine occurs.
- **Number of migraine headache days with acute medication use** is calculated as the number of calendar days in a 30-day period on which migraine or probable migraine occurs and acute medication is used.
- **Number of migraine attacks** per 30-day period is calculated as the number of sets of consecutive days with migraine or probable migraine separated by at least one migraine-free day. For example, a migraine or probable migraine starting on 5JAN and ending on 6JAN will result in a migraine/probable migraine-free day on 7JAN (assuming that there is no migraine/probable migraine on 7JAN). This will count as 1 migraine attack that started on 5JAN and ended on 6JAN. For a migraine attack that begins in one 30-day period but ends in another, only 1 migraine attack will be counted in the first of the 2 periods. For example, in the case of 7 days of consecutive migraine/probable migraine headache with 3 days in the baseline period and 4 days in Month 1, only 1 migraine attack will be counted in the baseline period; the 4 days of migraine/probable headache in Month 1 will not be counted as a migraine attack in Month 1.

Additional secondary and exploratory efficacy measures will be derived as follows:

- **Mean severity of remaining migraine** will be calculated at each period (including baseline and any post-baseline periods). For the calculation of mean severity, for days

with migraine or probable migraine, severity varies from 1 to 3 with 1=mild, 2=moderate, and 3=severe. The mean severity for each period will be calculated as:

$$\frac{\text{Sum of Severity of migraine headache days in the period}}{\text{number of migraine headache days in the Period}}$$

For periods with zero migraine headache days, the mean severity is considered not applicable hence missing in the analysis data set.

- **Percent change from baseline in the number of migraine headache days** will be calculated for any post-baseline 30-day period as:

$$\frac{-1 * 100 \times (\#\text{ of MHD in Month Y} - \#\text{ of MHD in baseline period})}{\#\text{ of MHD in baseline period}}$$

# of MHD in Month Y and # of MHD in baseline period here are the number of migraine headache days per 30-day period in MonthY and in baseline period, respectively.

- An **X% responder** is defined as Yes, if any patient who has a $\geqslant$X% reduction in the total number of migraine headache days in a 30-day period relative to baseline period, as No if otherwise. Therefore, if the percent change from baseline in the number of migraine headache days is $\geqslant$X%, the patient will be counted as an X% responder. In other words, if the response rate defined above in a month is $\geqslant$X%, then the patient will be an X% responder in that month. Indicators of X% responders will be derived for X=0, 5, 10, …, 95, and 100.

- **50% responders sustained for all 3 months during double-blind treatment phase in the number of migraine headache days** is defined as meeting 50% responder criterion in migraine headache days for Month 1 to Month 3 in the double-blind treatment phase.

- **Proportion of patients who continue to demonstrate 50% response for all 3 months during the open-label treatment period.** This analysis will include patients previously treated with galcanezumab who met 50% response criteria at Month 3 in the double-blind treatment phase (SP III) and entered the open-label treatment phase (SP IV). If a patient who discontinued early during the open-label phase (SP IV) will be counted as a non-responder.

- **Time to first 50% response (in months)** is defined as the first month when 50% response is met. If a patient has not met 50% response during SP III, the patient will be censored at the last month of SP III where 50% response status is not missing.

- **Time to first loss of 50% response in post-treatment phase (in months)** is defined as the time from the beginning of the post-treatment phase to the first month when 50% response is no longer met during the post-treatment phase (SP V). If a patient has met 50% response during all the months of SP V, the patient will be censored at the last month where 50% response status is not missing. This analysis will be conducted only for patients who met 50% response criteria at their last injection interval and are also entered into SP V.

- **Time to initiation of treatment with a migraine preventive medication (in days) in the post-treatment phase** is defined as the disposition date of the previous study period before entering SP V to the date of start of the migraine preventive medication (based on information collected from concomitant medication electronic case report form [eCRF]) in the post-treatment phase. If a patient did not initiate preventive treatment during SP V, they will be censored at the disposition date of SP V. This analysis will be conducted only for patients who entered SP V and those who are not in concurrent preventive treatment group.
- **Number of migraine headache days with nausea and/or vomiting** is calculated as the total number of migraine headache days with an answer of "yes" to Question 10 "Yesterday, did you feel sick to the stomach or throw-up with the headache?" in a 30-day period.
- **Number of migraine headache days with photophobia and phonophobia** is calculated as the total number of migraine headache days with an answer of "yes" to Question Q9 "Yesterday, did the headache come with sensitivity to light and sound?' in a 30-day period.
- **Number of weekly migraine headache days in Month 1** is calculated as the number of migraine headache days in a 7-day period on which a migraine headache occurs. At month 1, the first 7 calendar days will be counted as week 1, the second 7 calendar days will be counted as week 2, the third 7 calendar days will be counted as week 3, and the rest of days will be counted as week 4.

5.4.2. *Other efficacy measures*

The Patient Global Impression of Severity (PGI-S) will be collected at baseline and at the last visits of the double-blind treatment (SP III), open-label treatment (SP IV), and post-treatment phase (SP V), respectively. In this single-item scale, patients rate the severity of their migraine condition on a scale ranging from "Normal, not at all ill"(coded as 1) to "extremely ill" (coded as 7).

Change from baseline in PGI-S scores will be analyzed.

5.4.3. *Health Outcome Questionnaires*

5.4.3.1. Migraine Specific Quality of Life (MSQ) v2.1

The MSQ v2.1 is a self-administered health status instrument that was developed to address physical and emotional limitations of specific concern to individuals suffering from migraine. The instrument was designed with a 4-week recall period. The instrument consists of 14 items that address 3 domains: 1) Role Function-Restrictive (7 questions), examines the degree to which performance of daily activities is limited by migraine; 2) Role Function-Preventive (4 questions), examines the degree to which performance of daily activities is prevented by migraine; 3) Emotional Function (3 questions), examines feelings of frustration and helplessness due to migraine.

Responses are given using a 6-point Likert-type scale, ranging from "none of the time" to "all of the time." Precoded item values and final item values for each MSQ item response are shown in Table CGAX. 5.3. All item values range from 1 to 6. Final item value will be used as for analysis with higher score reflecting better quality of life.

Table CGAX. 5.3 Item Values for Migraine Specific Quality of Life (MSQ) Item Responses

Response Categories	Precoded Item Value	Final Item Value
None of the time	1	6
A little bit of the time	2	5
Some of the time	3	4
A good bit of the time	4	3
Most of the time	5	2
All of the time	6	1

Questions 1 to 7 of the questionnaire will be grouped together as Role Function-Restrictive domain, questions 8 to 11 as Role Function-Preventive domain and questions 12 to 14 as the Emotional Function domain. In general, no imputation for missing values is necessary because the MSQ was collected using patient direct data entry on an electronic device which did not allow patients to skip items. Patient either completed the scale in its entirety or not at all.

The raw score of each domain will be calculated as the sum of the raw scores of each question in that domain, using imputed scores where applicable. Should it be the case that the number of missing responses was more than half the questions in that domain, meaning that imputation of missing scores will not be done, the raw score for that domain will not be calculated, hence missing.

If any of the Role Function-Restrictive, Role Function-Preventive or Emotional Function domain is missing, then the total score will be missing, otherwise, the total score will be calculated as the sum of Role Function-Restrictive, Role Function-Preventive, and Emotional Function domain scores.

In addition, the raw scores of each domain and the total score will be transformed to a 0 to 100 scale using the following formulae:

- Role Function-Restrictive (range of 7 to 42):

$$\frac{(raw\ score - 7) * 100}{35}$$

- Role Function-Preventive (range of 4 to 24):

$$\frac{(raw\ score - 4) * 100}{20}$$

- Emotional Function (range of 3 to 18):

$$\frac{(raw\ score - 3) * 100}{15}$$

- Total Score (range of 14 to 84):

$$\frac{(raw\ total\ score - 14) * 100}{70}$$

The transformed total score and domain scores will be analyzed.

5.4.3.2. Migraine Disability Assessment test (MIDAS)

The Migraine Disability Assessment questionnaire (MIDAS) is designed to quantify headache-related disability over a 3-month period. This instrument consists of 5 questions (Q1-Q5) and 2 additional questions (A and B). The questionnaire measures the impact that migraine headaches have on migraineurs' life, including days of work or school missed, days with productivity at work or school reduced to half or more, days with household work missed, days with productivity in household work reduced to half or more, and days missed family/social/leisure activities. Each question is answered as a numeric number of days during the past 3 months of assessment, ranging from 0 to 90. The answers to all 5 questions will be added up to a total MIDAS score. A higher value is indicative of more disability. This instrument is considered highly reliable and valid, and is correlated with clinical judgment regarding the need for medical care (Stewart et al. 1999, 2001). For clinical interpretability, 4 categorical grades were developed based on the total score: Grade I (0 to 5) is for little or no disability, Grade II (6 to 10) is for mild disability, Grade III (11 to 20) is for moderate disability, and Grade IV (21+) is for severe disability. The severe disability category has subsequently been subdivided into Grade IV-A (severe [21 to 40]) and Grade IV-B (very severe [41 to 270]) because a high proportion of patients with chronic migraine are in Grade IV (Blumenfeld et al. 2011).

No imputation is needed when calculating the total score as patients are not allowed to send partial data.

The MIDAS responders are defined as patients with $\geqslant$50% improvement in the total MIDAS score.

The total MIDAS score, the raw score of each question, and the indicator of MIDAS responders will be analyzed.

5.4.3.3. Migraine Interictal Burden Scale (MIBS-4)

The MIBS-4 measures the burden related to headache in the time between attacks. The self-administered instrument consists of 4 items that address disruption at work and school, diminished family and social life, difficulty planning, and emotional difficulty. The questionnaire

specifically asks about the effect of the disease over the past 4 weeks on days without a headache attack. Response options include the following: don't know/not applicable, never, rarely, some of the time, much of the time, or most or all of the time.

Table CGAX. 5.4 Scores for Migraine Interictal Burden Scale (MIBS-4) Item Responses

Response Categories	Item Score
Don't know/not applicable	0
Never	0
Rarely	1
Some of the time	2
Much of the time	3
Most or all of the time	3

Each response has an associated numerical score, as shown in Table CGAX. 5.4. With the summation across all 4 items, a total score ranges from 0 to 12, and the level of interictal burden is categorized into the following: 0 for none, 1 or 2 mild, 3 or 4 moderate, and >5 severe (Buse et al. 2007, 2009).

The total MIBS-4 score and the level of interictal burden will be analyzed.

5.4.3.4. Patient Health Questionnaire-9

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item patient-completed instrument that was designed for detecting Major Depressive Disorder (MDD) and for measuring the severity of depressive symptoms (Kroenke et al. 2001). The 9 items address anhedonia; depressed mood; trouble sleeping; feeling tired; change in appetite; guilt, self-blame, or worthlessness; trouble concentrating; feeling slowed down or restless; and thoughts of being better off dead or hurting oneself. Each item is rated on a 4-point scale (0 = not at all, 1 = several days, 2 = more than half the days, 3 = nearly every day) based on symptoms over the past 2 weeks. The overall score ranges from 0 to 27, with the levels of depression severity defined as follows: 0 to 4 minimal, 5 to 9 mild, 10 to 14 moderate, 15 to 19 moderately severe, and 20 to 27 severe. Item 10 is rated as "not difficult at all", "somewhat difficult" , "very difficult" , or "extremely difficult". The instrument is considered reliable and valid for use in research and clinical settings (Kroenke et al. 2001), including in patients with migraine (Seo and Park 2015a).

The PHQ-9 total score and the level of depression severity will be analyzed.

5.4.3.5. 7-Item Generalized Anxiety Disorder Scale

The 7-item Generalized Anxiety Disorder Scale (GAD-7) is a 7-item patient-completed questionnaire that measuers the severity of anxiety symptoms. This instrument addresses the

feelings of nervousness, uncontrollable worrying, excessive worrying, trouble relaxing, restlessness, irritability, and fearfulness. The patient identifies how much they have been bothered by these symptoms over the past 2 weeks. Each of the 7 items is rated on a 4-point scale (0 = not at all, 1 = several days, 2 = more than half the days, 3 = nearly every day), with total score ranging from 0 to 21. The levels of anxiety severity are defined as follows: 0 to 4 minimal, 5 to 9 mild, 10 to 14 moderate, and 15 to 21 severe. The instrument is considered reliable and valid for use in research and clinical settings (Spitzer et al. 2006), including in patients with migraine (Seo and Park 2015b).

Binary indicator of anxiety will be derived as "Yes" (1) if the total score is >=10; "No" (0) otherwise (Spitzer et al. 2006).

The GAD-7 total score, the level of anxiety severity, and the binary indicator of anxiety will be analyzed.

5.4.4. *Safety endpoints*

Safety endpoints consist of the incidences of treatment emergent adverse events (TEAEs), serious adverse events (SAEs) and adverse events (AEs) leading to discontinuation, vital signs (blood pressure, pulse, and body temperature), weight, electrocardiograms (ECGs), and laboratory measures (chemistry, hematology, and urinalysis).

5.4.5. *Immunogenicity Endpoints*

Immunogenicity endpoints consist of the incidences of anti-drug antibodies (ADAs) in all trial patients at baseline (pre-existing ADAs), and in all trial patients at post-baseline (treatment emergent ADAs). An additional endpoint is the incidence of neutralizing antibodies (NAbs) present in those trial patients with ADAs.

5.4.6. *Pharmacokinetic Assessment*

Pharmacokinetic assessments will be summarized in the PK/PD analysis plan.

6. A Priori Statistical Methods

6.1. General Considerations

Unless otherwise specified, efficacy analyses will be conducted on an intent-to-treat (ITT) population, which will include all patients who are randomized and received at least 1 dose of study drug. Patients in the ITT population will be analyzed according to the treatment group to which they are randomized. Safety analyses will be conducted on safety population, which is the same as the ITT population. Patients in the safety populaiton will be analyzed according to the actual treatment the patient received.

There are 5 analysis populations defined:

Intent-to-treat (ITT) population: All patients who are randomized and receive at least 1 dose of study drug.

Safety population: This population is the same as the ITT population defined above.

Open-label treatment population: All patients who enter the open-label treatment phase as indicated by receiving any injections starting from Visit 7. Patients in the open-label treatment population will be analyzed as a single group as they start with the same treatment in the study period. Analyses for open-label phase only will be based on the open-label treatment population.

Post-treatment population: All patients who enter the post-treatment phase (Study Period V) as indicated by entering any post-treatment visit (telephone or office visit). Analyses for post-treatment phase only (that is, excluding earlier study periods) will be based on post-treatment population.

GMB-treated population: All patients who have exposure to galcanezumab (GMB). The GMB-treated population will be used for summaries for GMB-treated time.

GMB-treated time is the period during which a patient received GMB.

- For patients who receive GMB in the double-blind treatment period (SP III) only, GMB-treated time begins and ends in SP III.
- For patients who receive GMB in both treatment periods, GMB-treated time begins in SP III and extends through SP IV.
- For patients who receive placebo in the double-blind treatment period, GMB-treated time begins in SP IV.

Treatment effects will be evaluated based on an overall 2-sided significance level of 0.05 for all efficacy and safety analyses. The 95% confidence intervals (CIs) for the difference in least-square means (LSMeans) between treatment groups will be provided.

For continuous variables with repeated measures, change from baseline will be analyzed using a mixed model repeated measures (MMRM) analysis. An MMRM analysis refers to a restricted maximum likelihood (REML)-based, mixed-effects repeated measures analysis using all the longitudinal observations at each post-baseline visit.

For other continuous variables, the change from baseline to last-observation-carry-forward (LOCF) endpoint will be analyzed using an analysis of covariance (ANCOVA) model.

Unless otherwise specified, when ANCOVA model is used to analyze a continuous efficacy variable, type III sum-of-squares for the LSMeans will be used for the statistical comparisons.

Binary variables with repeated measures will be analyzed in a generalized linear mixed models (GLIMMIX) as pseudo-likelihood-based mixed effects repeated measures analysis.

For categorical efficacy variables without repeated measures, comparisons between treatment groups will be performed using logistic regression model. For other categorical variables without repeated measures, comparisons between treatment groups will be performed using the Fisher's exact test. Unless specified otherwise, Fisher's exact test will be used for comparisons of categorical safety measures; logistic regression will be used for comparisons of post-baseline efficacy measures.

Table CGAX. 6.1 describes the rules for determining the patient population and baseline and postbaseline observations for each study phase and type of analysis. When "last of Visit x-x" is used in the table, the last non-missing observation obtained in the visit interval will be used.

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the clinical study report or SAP. Changes may only be made in the SAP prior to unblinding.

Additional exploratory analyses of the data will be conducted as deemed appropriate.

Statistical analysis of this study will be the responsibility of Eli Lilly and Company or designee. SAS® software will be used to perform most or all statistical analyses.

Table CGAX. 6.1 Patient Population with Baseline and Postbaseline Definitions by Study Period and Type of Analysis

Study Period/Analysis	Patient Population	Baseline Observation	Postbaseline Observation(s)
Study Period III			
Efficacy analyses (repeated measures) or average of observed monthly values	ITT population with a baseline and at least 1 postbaseline observation	Visit 3	All scheduled visits 3 < Visits ≤7
Efficacy analyses at LOCF endpoint		Visit 3	Last of Visit 3.01–7
Health outcome analyses (repeated measures)		Visit 3	All scheduled visits 3 < Visits ≤7
Health outcome analyses at LOCF endpoint or for average of observed monthly values		Visit 3	Last of Visit 3.01-7
Treatment-emergent adverse events	Safety population	All Visits 1–3*a	All Visits 3.01–7
Serious adverse events, discontinuations due to adverse events	Safety population	NA	All Visits 3.01–7
Treatment-emergent abnormal laboratory values	Safety population with normal (with respect to direction being analyzed) laboratory values at all non-missing baseline visits and who have at least 1 postbaseline observation	Low: Min of Visits 1–3 High: Max of Visits 1–3 Abnormal: All Visits 1–3	Low: Min of Visits 3.01–7 High: Max of Visits 3.01–7 Abnormal: All Visits 3.01–7
Treatment-emergent changes in temperature and weight	Safety population with a baseline and at least 1 postbaseline observation	Low: Min of Visits 1–3 High: Max of Visits 1–3	Low: Min of Visits 3.01–7 High: Max of Visits 3.01–7
Treatment-emergent changes in blood pressures, pulse, and ECGs	Safety population with a baseline and at least 1 postbaseline observation	Low: Last of Visits 1–3 High: Last of Visits 1–3	Low: Min of Visits 3.01–7 High: Max of Visits 3.01–7
Treatment-emergent immunogenicity	Safety population with a baseline and at least 1 postbaseline observation	Visit 3	All Visits 3.01–7
Continuous safety analyses (repeated measures)	Safety population with a baseline and at least 1 postbaseline observation	Last of Visits 1–3	All scheduled visits 3 < Visits ≤7
Continuous safety analyses – change from baseline to LOCF endpoint (ANCOVA)	Safety population with a baseline and at least 1 postbaseline observation	Last of Visits 1–3	Last of Visits 3.01–7

Patient Population with Baseline and Postbaseline Definitions by Study Period and Type of Analysis

Study Period/Analysis	Patient Population	Baseline Observation	Postbaseline Observation(s)
Study Period IV			
Maintenance of 50% Response	Open-label treatment population – GMB-treated patients with 50% response at month 3	Visit 3	All scheduled visits 7 < Visits ≤10
Treatment-emergent adverse events	Open-label treatment population	All Visits 1–7*a	All Visits 7.01–10
Serious adverse events, discontinuations due to adverse events	Open-label treatment population	NA	All Visits 7.01–10
Treatment-emergent abnormal laboratory values	Open-label treatment population with normal (with respect to direction being analyzed) laboratory values at all non-missing baseline visits and who have at least 1 postbaseline observation	Low: Min of Visits 1–7 High: Max of Visits 1–7 Abnormal: All Visits 1–7	Low: Min of Visits 7.01–10 High: Max of Visits 7.01–10 Abnormal: All Visits 7.01–10
Treatment-emergent changes in temperature and weight	Open-label treatment population with a baseline and at least 1 postbaseline observation	Low: Min of Visits 1–7 High: Max of Visits 1–7	Low: Min of Visits 7.01–10 High: Max of Visits 7.01–10
Treatment-emergent changes in blood pressures, pulse, and ECGs	Open-label treatment population with a baseline and at least 1 postbaseline observation	Low: Last of Visits 1–7 High: Last of Visits 1–7	Low: Min of Visits 7.01–10 High: Max of Visits 7.01–10
Continuous safety analyses – change from baseline to LOCF endpoint (ANCOVA)	Open-label treatment population with a baseline and at least 1 postbaseline observation	Last of Visits 1–7	Last of Visits 7.01–10
Study Periods III / IV Combined			
Efficacy/Health outcome analyses	ITT population with a baseline and at least 1 postbaseline observation	Visit 3	All scheduled visits 3 < Visits ≤10
Continuous safety analyses (repeated measures)	Safety population with a baseline and at least 1 postbaseline observation	Last of Visits 1–3	All scheduled visits 3 < Visits ≤10

Patient Population with Baseline and Postbaseline Definitions by Study Period and Type of Analysis

Study Period/Analysis	Patient Population	Baseline Observation	Postbaseline Observation(s)
Study Period V			
Post-treatment-emergent adverse events	Post-treatment population	All Visits 1–10*a	All Visits 10.01–12
Serious adverse events, discontinuations due to adverse events	Post-treatment population	NA	All Visits 10.01–12
Post-treatment-emergent abnormal laboratory values	Post-treatment population with normal (with respect to direction being analyzed) laboratory values at all non-missing baseline visits and who have at least 1 postbaseline observation	Low: Min of Visits 1–10 High: Max of Visits 1–10 Abnormal: All Visits 1–10	Low: Min of Visits 10.01–12 High: Max of Visits 10.01–12 Abnormal: All Visits 10.01–12
Post-treatment-emergent changes in temperature and weight	Post-treatment population with a baseline and at least 1 postbaseline observation	Low: Min of Visits 1–10 High: Max of Visits 1–10	Low: Min of Visits 10.01–12 High: Max of Visits 10.01–12
Post-treatment-emergent changes in blood pressures, pulse, and ECGs	Post-treatment population with a baseline and at least 1 postbaseline observation	Low: Last of Visits 1–10 High: Last of Visits 1–10	Low: Min of Visits 10.01–12 High: Max of Visits 10.01–12
Continuous safety analyses – change from baseline to LOCF endpoint (ANCOVA)	Post-treatment population with a baseline and at least 1 postbaseline observation	Last of Visits 1–10	Last of Visits 10.01–12
Study Periods III / IV / V Combined			
Efficacy/Health outcome analyses	ITT population with a baseline and at least 1 postbaseline observation	Visit 3	All scheduled visits 3 < Visits ≤12
Continuous safety analyses (repeated measures)	Safety population with a baseline and at least 1 postbaseline observation	Last of Visits 1–3	All scheduled visits 3 < Visits ≤12
Treatment-emergent immunogenicity	Safety population - GMB-treated patients with a baseline and at least 1 postbaseline observation	Visit 3	All Visits 3.01–12

Patient Population with Baseline and Postbaseline Definitions by Study Period and Type of Analysis

Study Period/Analysis	Patient Population	Baseline Observation	Postbaseline Observation(s)
GMB-Treated Time			
Treatment-emergent adverse events	GMB-treated population	All Visits 1–3 before double-blind phase dosing for patients treated with GMB during SP III; All Visits 1–7 before open-label phase dosing for patients treated with placebo during SP III*a	All visits after first dosing of GMB through Visit 10
Serious adverse events, discontinuations due to adverse events	GMB-treated population	NA	All visits after first dosing of GMB through Visit 10
Treatment-emergent abnormal laboratory values	GMB-treated population with normal (with respect to direction being analyzed) laboratory values at all non-missing baseline visits and who have at least 1 postbaseline observation	Visits 1–3 before double-blind phase dosing for patients treated with GMB during SP III; Visits 1–7 before open-label phase dosing for patients treated with placebo during SP III	All visits after first dosing of GMB through Visit 10: Visits 3.01–10 for GMB-treated patients during SP III; Visits 7.01–10 for placebo-treated patients during SP III
Treatment-emergent changes in temperature and weight	GMB-treated population with a baseline and at least 1 postbaseline observation	Min/max of Visits 1–3 before double-blind phase dosing for patients treated with GMB during SP III for low/high, respectively; Min/max of Visits 1–7 before open-label phase dosing for patients treated with placebo during SP III for low/high, respectively	Min/max of all visits after first dose of GMB through Visit 10 for low/high, respectively: Visits 3.01–10 for GMB-treated patients during SP III; Visits 7.01–10 for placebo-treated patients during SP III.

Patient Population with Baseline and Postbaseline Definitions by Study Period and Type of Analysis

Study Period/Analysis	Patient Population	Baseline Observation	Postbaseline Observation(s)
GMB-Treated Time			
Treatment-emergent changes in blood pressure, pulse, and ECGs	GMB-treated population with a baseline and at least 1 postbaseline observation	Last non-missing value from Visits 1–3 before double-blind phase dosing for patients treated with GMB during SP III; Last non-missing value from Visits 1–7 before open-label phase dosing for patients treated with placebo during SP III.	Min/max of all visits after first dosing of GMB through Visit 10 for low/high, respectively: Visits 3.01–10 for GMB-treated patients during SP III; Visits 7.01–10 for placebo-treated patients during SP III.
Treatment-emergent immunogenicity	GMB-treated population with a baseline and at least 1 postbaseline observation	Visit 3	All Visits 3.01–10

Abbreviations: AE = adverse event; ANCOVA = analysis of covariance; ECG = electrocardiogram; GMB = galcanezumab; ITT = intent-to-treat; LOCF = last observation carried forward; max = maximum; min = minimum; NA = not applicable; SP = study period.

*a, AEs that occurred on the visit date of Visit 3 or Visit 7 will be determined to be predose or after dose based on AE start time and the injection time: predose if AE start time is before the injection time and post dose if AE start time is after the injection time. Note: Visit 3.01 indicates the first unscheduled visit occurring after Visit 3 and prior to Visit 4. Visit 7.01 indicates the first unscheduled visit occurring after Visit 7 and prior to Visit 8.

6.2. Adjustments for Covariates

The MMRM models will include the fixed, categorical effects of treatment, country, month, and treatment-by-month interaction, as well as the continuous, fixed covariates of baseline value and baseline-by-month interaction. The baseline value and baseline-by-month interaction are included to account for the differential influence over time that the baseline value has on the post-baseline values. Country will be excluded from MMRM models for safety measures.

When an ANCOVA model is used to analyze a continuous efficacy at the LOCF endpoint, the model will contain the main effects of treatment, country, and appropriate baseline value as a covariate. When an ANCOVA model is used to analyze a continuous safety variable, the country will be removed from the model.

The GLIMMIX models for the repeated binary outcomes will include the fixed, categorical effects of treatment, month, and treatment-by-month interaction, as well as the continuous, fixed covariate of baseline value. Country and the baseline value-by-month interaction will be excluded from the model in order to increase the likelihood of convergence.

When a logistic regression is used to analyze a binary variable, the model will include the main effect of treatment, country (if appropriate), and appropriate baseline value as a covariate. Country may be excluded from the model in case of non-convergence.

With the exception of efficacy analyses on migraine headache days or categorical analysis of response rate (such as 50% response rate) derived from migraine headache days where the continuous value of baseline migraine headache days will be used as covariate, all other efficacy analyses will include baseline number of migraine headache days category (<8 vs ≥8) as a covariate in the MMRM, ANCOVA, GLIMMIX and logistic regression model. Specifically, for time to event analysis of 50% response (such as Time to first 50% response in double-blind treatment phase and Time to first loss of 50% response in post-treatment phase), stratified log rank test will be used with the baseline number of migraine headache days category (<8 vs ≥8) as one of the covariates.

6.3. Handling of Dropouts or Missing Data

Two statistical approaches to handling missing data will be used as appropriate: repeated measures analyses and ANCOVA model using change from baseline to LOCF endpoint.

For the repeated measures analyses, the model parameters are simultaneously estimated using restricted likelihood estimation incorporating all of the observed data. Estimates have been shown to be unbiased when the missing data are missing at random.

Please refer to Section 5.4.1 for approach of handling missing diary data for derivation of migraine headache days and other efficacy measures (with the exception of migraine attacks) derived from ePRO data per 30-day period.

Approaches of handling missing diary data for the derivation of migraine attack

For the derivation of the number of monthly migraine attacks, the LOCF method will be used to impute the missing ePRO diary days. If the patient was migraine headache-free on the day before the missing ePRO diary day, this would be carried forward as no migraine headache day until the actual next non-missing diary day. On the other hand, if the day before the missing ePRO diary day is a migraine headache day, then it would be carried forward as migraine headache day until the next non-missing diary day. The imputation will be carried out for all the missing diary days between the first non-missing to the last day.

If the diary compliance rate for a monthly interval is ≤50%, the number of monthly migraine attacks during that month will be considered missing. Please refer to Section 6.8 for diary compliance rate calculation.

6.4. Multiple Comparisons/Multiplicity

In order to provide strong control of the Type I error, the key secondary analyses will be tested using a gated testing approach at a 2-sided alpha level of 0.05. If the null hypothesis is rejected for the primary endpoint, key secondary endpoints will be sequentially tested following the gatekeeping hierarchy as shown in Figure CGAX. 6.1.

Following the primary objective, the sequential procedure for key secondary objectives starts with the comparison between treatment groups in the 50% response rate. If that null hypothesis is rejected, then the next comparison in the sequence will be tested, following this same pattern until all hypotheses are tested or until the null hypothesis is accepted for an endpoint, at which point, any further testing would stop for the key secondary objectives.

Abbreviations: MHD = the number of monthly migraine headache days; MSQ RR = Migraine Specific Quality of Life Questionnaire Role Function-Restrictive domain; response = response rate. See Section 5.4 for full definitions of endpoints. Note: All testing will be conducted at a 2-sided alpha of 0.05.

Figure CGAX. 6.1 Gatekeeping sequence for testing of primary and key secondary endpoints.

6.5. Patient Disposition

The number and percentage of ITT patients who complete the study or discontinue early will be tabulated overall for both treatment groups for double-blind treatment (Study Period III), open-label treatment (Study Period IV), and post-treatment follow-up (Study Period V) separately. Reasons for discontinuation will be presented for each period. In addition, subcategories of discontinuation due to subject decision will be summarized.

Patient allocation by investigator will be summarized for all ITT patients.

6.6. Patient Characteristics

The following patient characteristics at baseline will be summarized by treatment group for all ITT patients:

- Demographic (age, gender, race, height, weight, body mass index)
- Duration of migraine illess
- Migraine and/or headache measures per 30-day baseline period
    - number of migraine headache days
    - number of ICHD migraine headache days
    - number of headache days
    - number of moderate to severe headache days
    - number of migraine headache hours
    - number of headache hours

  - number of monthly days with acute medication use
  - number of migraine attacks
  - mean severity of migraine headaches
  - number of migraine headache days with nausea and/or vomiting
  - number of migraine headache days with photophobia and phonophobia
  - number of migraine headache days during menstrual period
  - number of migraine headache days with any trigger during menstrual period
  - baseline number of migraine headache day category (<8 versus ⩾8)

- Patient Global Impression - Severity
- MIDAS total score
- MSQ total score and each domain score
- Medical history and pre-existing condition

Duration of migraine illess in years will be derived as (inform consent date – migraine start date + 1)/365.25. If multiple migraines were reported in CRF, the earliest start date will be used.

Medical history and pre-existing conditions will be summarized by preferred term (PT) within system organ class (SOC). Medical history is defined as illness(es) that ended prior to the signing of informed consent. Pre-existing conditions are those events that either have an ongoing status of Yes or have an end date on or after informed consent.

6.7. Treatment Compliance

Dosing will occur at monthly study visits. Treatment compliance for each patient will be calculated as the number of completed scheduled dosing visits in which the patient receives an injection, divided by the number of scheduled dosing visits, including any skipped dosing visits at or before the last dosing visit (Visit 9) or early discontinuation visit. For this analysis, partial dose (for example, a patient only received 1 injection instead of 2 at Visit 3 or Visit 7) will be considered as no dose received.

Treatment compliance for SP III and SP IV will be summarized overall and by treatment group.

6.8. Electronic Patient-Reported Outcome Diary Compliance

Electronic patient-reported outcome diary compliance at each 1-month period (including baseline, Month 1, 2, 3, … till Month 10) as well as for each treatment phase (SP III, SP IV, SP V) will be calculated. Diary compliance at each period is calculated as follows:

$$\frac{\text{Actual number of diary days in the period}}{\text{Expected number of diary days in the period}} * 100$$

Actual number of diary days is calculated as the total number of days with nonmissing answers. Expected number of ePRO diary days is calculated as date of injection (or date of visit for post-treatment follow-up phase) at the end of the interval minus date of injection (or date of visit for post-treatment follow-up phase) at the beginning of the interval +1.

For a patient who discontinues early in the double-blind (SP III) or open-label treatment phase (SP IV) or post-treatment phase (SP V), compliance rate for the last month of that patient will be calculated with a denominator of 30 days (or the actual number of days in the interval if it is greater than 30).

Diary compliance for each period will be summarized overall and by treatment group.

6.9. Concomitant Therapy

The proportion of patients who received concomitant medication collected from eCRFs will be summarized for ITT population for each period separately.

Concomitant therapies for SP III are those which stopped during SP III or continued in SP III. If medication started and stopped on the same day of injection, it will still be considered as concomitant medication for SP III. If a medication started before the first day of injection but stopped on the same day of injection, then it will not be counted as concomitant medication for SP III. Concomitant therapies for SP IV are those which either started, stopped or continued in SP IV.

6.10. Efficacy Analyses

6.10.1. Primary Outcome and Methodology

The primary efficacy measure is the overall mean change from the baseline period in the number of monthly MHDs during the 3-month double-blind treatment phase, and the primary analysis will evaluate the efficacy of galcanezumab (120 mg/month) compared with placebo.

The primary analysis will be performed using a restricted maximum likelihood-based MMRM technique. The analysis will include the fixed categorical effects of treatment, country, month, and treatment-by-month interaction, as well as the continuous fixed covariates of baseline number of MHDs and baseline number of MHDs-by-month interaction.

An unstructured covariance structure will be used to model within-patient errors. The Kenward-Roger (Kenward and Roger 1997) approximation will be used to estimate denominator degrees of freedom. If the model does not converge with both the Hessian and the G matrix being positive definite under the default fitting algorithm used by PROC MIXED, the Fisher scoring algorithm will be implemented by specifying the SCORING option in SAS®. If the model still fails to converge, the model will be fit using covariance matrices of the following order specified by a decreasing number of covariance parameters until convergence is met:

- heterogeneous Toeplitz
- heterogeneous first-order autoregressive
- Toeplitz
- first-order autoregressive

When the unstructured covariance matrix is not utilized, the sandwich estimator (Diggle and Kenward 1994) will be used to estimate the standard errors of the fixed effects parameters. The sandwich estimator is implemented by specifying the EMPIRICAL option in SAS®. When the

sandwich estimator is utilized, the Kenward-Roger approximation for denominator degrees of freedom cannot be used. Instead, the denominator degrees of freedom will be partitioned into between-subject and within-subject portions by the DDFM=BETWITHIN option in SAS®. SAS® PROC MIXED will be used to perform the analysis.

The primary endpoint of this study for galcanezumab compared with placebo will be estimated as the treatment main effect from the MMRM analysis model. This provides the average treatment effect across Month 1, Month 2, and Month 3. The Type I error rate for the study will be controlled at a 2-sided 0.05 level (equivalently, 1-sided 0.025 level). Specific details of the testing procedure for the primary outcome and the secondary gatekeeper objectives are provided in Section 6.4.

The results of the statistical tests at each month in the double-blinded treatment phase will be used to assess the onset of effect for galcanezumab compared with placebo. In particular, if the primary efficacy analysis is statistically significant, then the earliest month where the statistically significant improvement is observed and maintained for all the subsequent months during the double-blinded treatment phase will be considered as the period that demonstrated the onset of effect. If the onset of effect occurs in Month 1, the number of weekly migraine headache days in Month 1 will be analyzed to identify the earliest week where the statistically significant improvement is observed and maintained for all the subsequent weeks in Month 1. Details about the analysis for the number of weekly migraine headache days are described in Section 6.10.3.

6.10.2. Sensitivity Analysis for Primary Outcome

Two types of sensitivity analyses are planned to assess the robustness of deviations from the assumptions of primary analysis including normality assumption and missing data assumption.

Missing Data Assumption

Sensitivity analyses will be performed to assess the robustness of the primary analysis conclusions to deviations from missing at random (MAR) assumption. The approach for these analyses is to vary the assumptions of missing data for the primary analysis in a systematic way. Basically, the method will predict the missing outcomes and then add values (Δ_{GMB}, Δ_P) to the predictions in the galcanezumab and placebo treatment groups respectively, regardless of the reason the data are missing. This approach is consistent with the sensitivity approach suggested in Permutt (2015). This procedure will be repeated multiple times for different values of (Δ_{GMB}, Δ_P) using the following steps:

1) Predict the missing outcomes for each treatment via multiple imputation based on observed primary endpoint and baseline values. Such imputation will be carried out using a Markov Chain Monte Carlo method with a Jeffreys prior via SAS® PROC MI. Thirty (30) such imputations will be created.
2) Add the corresponding Δ value (that is, Δ_{GMB}, Δ_P) to the imputed values based on the patient treatment group.
3) Conduct the primary analysis separately for each of the 30 imputations.

4) Combine the results of these analyses using Rubin's combining rules, as implemented in SAS® PROC MI ANALYZE.

The above steps will be repeated multiple times for different values of (Δ_{GMB}, Δ_P) with Δ_P ranging from 0 to twice the absolute value of the mean value seen for placebo in the primary analysis, Δ_{GMB} ranging from Δ_P to Δ_P + absolute value of the mean treatment difference seen within the primary analysis. For example, if the overall mean change from baseline for placebo is -3.6 and the overall treatment difference is -1.5, then Δ_P would range from 0 to7.2 and Δ_{GMB} would range from Δ_P to Δ_P + 1.5.

Normality Assumption

To assess the robustness of the MMRM results to deviations from normality assumption, a sensitivity analysis for raw number of migraine headache days (total number of migraine headache days for each interval without normalization to 30-day period) will be conducted with a repeated measures negative binomial regression model fitted with SAS PROC GLIMMIX. The model will include treatment, country, month, and treatment-by-month interaction, as well as the continuous fixed covariates of baseline number of monthly migraine headache days and baseline-by-month interaction, log (number of compliant days within each month/30) as the offset in the model. In case of non-convergence, country and/or baseline-by-month interaction may be excluded from the model. Directional consistency of treatment effects from this model and the primary analysis MMRM model as specified in Section 6.10.1 will be examined.

6.10.3. Secondary and Exploratory Efficacy Analyses

Key secondary efficacy measures will be tested in the order as specified in Section 6.4.

Table CGAX. 6.2 summarizes all the planned secondary and exploratory efficacy analyses for SP III, SP III/IV, and SP III/IV/V, not including analyses for health outcome measures. Additional statistical analysis details are also provided below.

Continuous Efficacy Measures

For the continuous secondary efficacy measures, the change from baseline will be estimated from MMRM as described for analysis for primary outcome or ANCOVA as described in Section 6.1.

For continuous secondary efficacy measures where the objective is to assess overall mean change during 3-month double-blind treatment phase, the endpoint for galcanezumab compared with placebo will be estimated as the treatment main effect from the MMRM analysis assessing the average treatment effect across Months 1, 2, and 3.

Binary Efficacy Measures

For the repeated binary efficacy measures such as responder indicators based on the number of migraine headache days, the visit-wise binary outcomes indicating whether patients meet X% response criteria will be analyzed using a categorical, pseudo-likelihood-based repeated measures analysis. This analysis will be implemented using the GLIMMIX procedure in SAS to compare treatments with covariates adjustment described in Section 6.2.

An unstructured covariance structure will be used to model the within-patient errors (denoted by TYPE=CHOL in the RANDOM statement). The Newton-Raphson method with ridging will be used for nonlinear optimization (denoted by including NLOPTIONS TECH=NRRIDG). The Kenward-Roger approximation will be used to estimate denominator degrees of freedom. If the model does not converge, the Fisher's scoring algorithm will be utilized by the SCORING option in SAS.

If the model still fails to converge, the model will be fit using covariance matrices in the following order specified by a decreasing number of covariance parameters until convergence is met:

- Heterogeneous Toeplitz,
- Heterogeneous autoregressive,
- Toeplitz, and
- Autoregressive.

If necessary, both fitting algorithms will be used in the pre-specified order before proceeding to the next covariance structure in the sequence.

For models where the unstructured covariance matrix is not utilized, the sandwich estimator (Diggle et al. 1994) will be used to estimate the standard errors of the fixed effects parameters. The sandwich estimator is utilized by the EMPIRICAL option in SAS. When the sandwich estimator is utilized, the Kenward-Roger approximation for denominator degrees of freedom cannot be used. Instead, the denominator degrees of freedom will be partitioned into between-subject and within-subject portions by the DDFM=BETWITHIN option in SAS.

For the binary secondary efficacy measures where the objective is to assess proportion of patients with X% response during the 3-month double-blind treatment phase, the endpoint for galcanezumab compared with placebo will be estimated as the treatment main effect from the categorical MMRM analysis assessing the average response rate across Month 1, Month 2, and Month 3.

For visit wise indicators of 50% responders, the results of the statistical tests at each month in the double-blinded treatment phase from categorical MMRM analysis will be used to assess the onset of 50% sustained response for galcanezumab compared with placebo. In particular, if the gated secondary measure of 50% response rate is statistically significant, then the earliest month where the statistically significant improvement is observed and maintained for all the subsequent months during double-blinded treatment phase will be considered as the period that demonstrated the onset of 50% sustained response.

For binary secondary efficacy measures such as the proportion of patients who maintain 50% response criteria for all 3 months of double-blind treatment, as well as the proportion of patients who demonstrate 50% response for all 3 months in the open-label treatment period among patients previously treated with galcanezumab who met 50% response criteria at Month 3 in the

double-blind treatment period, treatment differences will be determined using logistic regression with covariate of treatment, country and continuous baseline value.

Measures Conditional on the Postbaseline Number of Migraine Headache Days >0

The following measures are conditional on the number of migraine headache days >0:

- Mean severity of remaining migraine headaches

Those measures will be analyzed using MMRM model, assuming that data are missing for months without migraine headache.

The number of weekly migraine headache days in Month 1

If statistically significant improvement is observed at Month 1 and maintained for all the subsequent months during the double-blind treatment phase for the number of monthly migraine headache days, the number of weekly migraine headache days in Month 1 will be analyzed to identify the earliest week where the statistically significant improvement is observed and maintained for all the subsequent weeks in Month 1 for patients who have a monthly diary compliance rate greater than 50% for Month 1.

The number of weekly migraine headache days in Month 1 can be considered as ordinal data with possible values of 0, 1, 2, 3, 4, etc. and will be analyzed using an ordinal repeated measures model using the GLIMMIX procedure in SAS. In this model, a proportional odds model with cumulative logit link will be used, and a random intercept will be applied to the observations for each patient to account for repeated measures. The model will include the fixed, categorical effects of treatment, country, week, and treatment-by-week interaction, as well as the continuous, fixed covariate of weekly baseline number of migraine headache days and weekly baseline number of migraine headache days-by-week interaction. Log (number of compliant calendar days within each week/7) will be included as the offset in the model. For weekly diary compliance rate ≤50%, the number of weekly migraine headache days during that week will be considered missing.

The weekly baseline number of migraine headache days will be derived as (the monthly number of migraine headache days)/30*7.

Time to Event Measures

For the following time to event measures, a Kaplan-Meier curve of the time to event and treatment group comparison using stratified log-rank test stratifying for country, baseline MHD frequency category will be provided.

- Time to first 50% response in SP III for ITT population.
- Time to first loss of 50% response in SP V for patients who met 50% response criteria at their last injection interval and also entered SP V.

- Time to initiation of preventative treatment for migraine or probable migraine in SP V for patients who entered SP V and those who are not on concurrent preventive treatment group.

Distribution of Response Rate

Overall x% response rate during the double-blind treatment phase will be estimated for X=0, 5, 10, …, 95, and 100, using GLIMMIX model as described earlier in this section. These estimated response rates will be plotted and points within each treatment group will be connected to show a curve of response rates. No statistical comparisons will be conducted among different treatment groups.

Analysis for PGI-I

When analyzing PGI-I, baseline PGI-S score will be included as a covariate. Specifically, it will be analyzed with an MMRM analysis while the covariates include the fixed categorical effects of treatment, month, country, baseline MHD frequency category and treatment-by-month interaction, as well as the continuous fixed covariates of baseline PGI-S score and baseline PGI-S score-by-month interaction.

Table CGAX. 6.2 Secondary and Exploratory Efficacy Variables and Analysis Methods

Efficacy Variables	SP III	SP IV, SP III/IV, SP III/IV/V, SP V
Number of migraine headache hours	MMRM	MMRM
Number of migraine attacks		
Number of days with acute medication use		
Number of headache days		
Number of moderate-severity headache days		
Number of headache hours		
Number of ICHD migraine headache days		
Mean severity of migraine headache		
Number of migraine headache days with nausea and/or vomiting		
Number of migraine headache days with photophobia and phonophobia		
Number of weekly migraine headache days in Month 1	GLIMMIX for ordinal outcome	N/A
X% response rate (X=30, 50, 75, or 100)	GLIMMIX	GLIMMIX
50% response sustained from month 1 to month 3	Logistic Regression	N/A
Proportion of patients maintained 50% response in SP IV among all galcanezumab-treated 50% responders at Month 3	N/A	Logistic regression
Time to first 50% response in SP III	Kaplan-Meier curve and stratified log-rank test	N/A
Time from the end of SP IV to no longer meeting 50% response criterion	N/A	Kaplan-Meier curve and stratified log-rank test for SP V
Time from the end of SP IV to start use of preventative treatment for migraine		
Distribution of response rate in SP III	GLIMMIX	N/A
Global Patient Impression – Severity	ANCOVA	MMRM

Abbreviations: ANCOVA = analysis of covariance; GLIMMIX = Generalized linear mixed model (for categorical variables); MMRM = Mixed models repeated measures; N/A= Not applicable.

6.11. Health Outcomes Analyses

All health outcome measures will be analyzed or summarized for SP III, SP III/IV, and SP III/IV/V, if appropriate.

For MSQ (total score and 3 domain scores), MIDAS (total score and item scores), MIBS-4 (total score), PHQ-9 (overall score), and GAD-7 (total score),

- when there are repeated measures, they will be evaluated using MMRM as described in Section 6.1.
- when there is single post-baseline measure, they will be evaluated using ANCOVA model as described in Section 6.1.

PHQ-9 binary indicator of MDD and GAD-7 binary indicator of anxiety will be analyzed using GLIMMIX method as described in Section 6.1.

For PHQ-9 overall level of depression severity and GAD-7 overall level of anxiety severity, as the severity categories do not follow normal distribution, the change from baseline score will be analyzed using nonparametric tests: Kruskal–Wallis test (Kruskal and Wallis 1952) for treatment comparison, and Wilcoxon Signed-rank test (Wilcoxon 1945) to compare the change from baseline within each treatment group.

Individual item scores of these health outcome questionnaire may be summarized if deemed necessary.

6.12. Pharmacokinetic Analyses

Pharmacokinetic assessment will be summarized in the pharmacokinetic/pharmacodynamic (PK/PD) SAP.

6.13. Evaluation of Immunogenicity

To evaluate the changes in immunogenicity data (Anti-LY2951742 [ADA and NAb]) after treatment, the following statistical analyses are planned for comparison between treatment groups:

- The incidence of ADA positive during baseline will be summarized.
- The incidence of treatment-emergent ADA (TE ADA) between treatment groups will be summarized and compared for SP III. The incidence of TE ADA for galcanezumab-treated patients during galcanezumab treatment and the incidence of TE ADA for galcanezumab-treated patients during galcanezumab treatment and post-treatment phase combined will be summarized. This analysis will be done for each immunogenicity analyte (ADA and NAb). The baseline and postbaseline definitions for each Study Period is shown in Table CGAX. 6.1. Treatment-emergent ADA will be defined as any of the following:
    - a negative baseline result and a positive postbaseline ADA result with a titer ≥20. This is also called treatment-induced ADA.
    - a positive baseline result and a positive postbaseline ADA result with a ≥4-fold increase in titers (for example, baseline titer of 10 increasing to ≥40 postbaseline). This is called treatment-boosted ADA.
- The incidence of TE ADA and NAb Positive combined between treatment groups will be summarized for the same time periods as planned for the incidence of TE ADA.

The following will also be provided:

- Listing of subjects with TE ADA at any time during study, NAb Status will also be displayed.

- Listing of subjects with TE hypersensitivity reactions or TEAEs related to injection sites for subjects with ADA present at any time.

6.14. Safety Analyses

6.14.1. Extent of Exposure

Patients will receive the investigational medicinal product (IMP) at the following planned time points:

- Beginning of Month 1 (Visit 3)
- Beginning of each month from Month 2 to Month 6 (Visit 5 to Visit 9)

The following information will be recorded on the eCRF for each dose:

- Confirmation that the patient received the IMP (including reason if the IMP was not given)
- Date and time of administration

The following will be derived from the information recorded on the eCRF:

- For treatment phase (SP III), duration of exposure in days is calculated as treatment phase disposition date – first date IMP administered +1.
- For open label phase (SP IV), duration of exposure in days is calculated as open label disposition date - first date IMP administered during SP IV +1.
- For SP III and SP IV separately, number and percentage of patients with 1 dose, 2 doses, and 3 doses injected.
- For galcanezumab-treated time, number and percentage of patients with 1 dose, 2 doses, 3 doses, 4 doses, 5 doses, and 6 doses injected.

6.14.2. Adverse Events

Treatment-emergent adverse events are defined as the reported AEs that first occurred or worsened during the post-baseline phase compared with baseline phase. For each TEAE, the severity level of the event (mild, moderate, or severe) will be determined by patient or physician opinion. The Medical Dictionary for Regulatory Activities (MedDRA) Lowest Level Term will be used in the treatment-emergent computation. For each Lowest Level Term, the maximum severity at baseline will be used as the baseline severity. If the maximum severity during post-baseline is greater than the maximum baseline severity, the event is considered to be treatment-emergent for the specific post-baseline period.

Unless specified otherwise, adverse events will be summarized as follows:

- Treatment-emergent adverse events (TEAEs)

  - By preferred term (PT)
  - By PT nested within system organ class (SOC)
  - By maximum severity
  - By considered to be related to study treatment by investigator
- Serious adverse events (SAEs) by PT
- AEs leading to treatment discontinuation by PT

For each patient and TEAE, the maximum severity for the MedDRA level being displayed is the maximum post-baseline severity observed from all associated Lowest Level Terms mapping to the MedDRA level.

For events that are gender specific, the denominator and computation of the percentage will include only patients from the given gender.

6.14.2.1. Safety Topics of Interest

6.14.2.1.1. Potential Hypersensitivity Events

Potential hypersensitivity events (immediate and non-immediate) will be identified from a search of preferred terms generated from the following queries:

- Anaphylactic reaction (SMQ 20000021)
- Broad and narrow terms in the Angioedema (SMQ 20000024)
- Broad and narrow terms in the Hypersensitivity (SMQ 20000214)

A listing of patients having an event identified from these analyses will be reviewed by medical to determine if the terms identified likely represent hypersensitivity events. Listings for the medical review should include information on timing of event relative to latest dose of study drug administration, the event term from this query, other AEs for the patient and timing, any abnormal laboratory findings, and medical history. Events that are judged not hypersensitivity-related will be excluded from the final tables, unless otherwise stated.

The number and percentage of patients with TEAEs, SAEs, and AEs resulting in study drug discontinuation will be summarized by treatment groups using MedDRA PT nested within the SMQ. Events will be ordered by decreasing frequency within the SMQ.

The number and percentage of patients with treatment-emergent (TE) hypersensitivity events by maximum severity will be summarized by treatment groups using MedDRA PT.

6.14.2.1.2. Adverse Events Related to Injection Sites

Adverse events related to injection sites will be defined using terms from the MedDRA High Level Term “Injection Site Reactions”.

The number and percentage of patients with TEAEs, SAEs, and AEs resulting in study drug discontinuation will be summarized by treatment group using MedDRA PTs. Events will be ordered by decreasing frequency. Events will be ordered by decreasing frequency.

The number and percentage of patients with TEAEs related to injection sites by maximum severity will be summarized by treatment groups using MedDRA PT.

The number and percentage of patients with TEAEs related to injection sites by timing will be summarized using MedDRA preferred terms ordered by decreasing frequency.

6.14.2.1.3. Upper Respiratory Tract Infections

Upper respiratory tract infections will be defined using all the PTs from the 2 high level terms of "upper respiratory tract infections" and "upper respiratory tract infections NEC" as defined in MedDRA.

The number and percentage of patients with TEAEs, SAEs, and AEs resulting in study drug discontinuation will be summarized by treatment group using MedDRA PTs. Events will be ordered by decreasing frequency.

The number and percentage of patients with TEAEs of upper respiratory tract infections by maximum severity will be summarized by treatment groups using MedDRA PT.

6.14.3. Clinical Laboratory Evaluation

The incidence rates of patients with treatment-emergent abnormal high or low laboratory values based on Covance reference ranges at any time postbaseline will be assessed using Fisher's exact test for each laboratory test for SP III.

Patients will be defined as having a treatment-emergent low value if they have all normal or high values at baseline, followed by a value below the lower reference limit at any postbaseline visit. Patients with all normal or high values at baseline (no low values) will be included in the analysis of treatment-emergent low laboratory values. Patients will be defined as having a treatment-emergent high value if they have all normal or low values at baseline, followed by a value above the upper reference limit at any postbaseline visit. Patients with all normal or low values at baseline (no high values) will be included in the analysis of treatment-emergent high laboratory values.

For analytes simply classified as normal or abnormal, patients will be defined as having a treatment-emergent abnormal value if they have all normal values at baseline, followed by an abnormal value at any postbaseline visit. Patients with all normal values at baseline will be included in the analysis of treatment-emergent abnormal laboratory values.

The incidence of patients with the following elevations in hepatic laboratory tests at any time postbaseline will also be summarized. Comparison between treatment groups will be done for SP III using Fisher's exact test.

- The percentages of patients with an alanine aminotransferase (ALT) or aspartate aminotransferase (AST) measurement greater than or equal to 3 times (3×), 5 times (5×), and 10 times (10×) the Covance upper limit of normal (ULN) during the treatment period will be summarized for all patients with a postbaseline value.
- The percentages of patients with an Alkaline phosphatase (ALP) greater than or equal to 2 times (2×) the Covance ULN during the treatment period will be summarized for all patients with a postbaseline value.
- The percentages of patients with a total bilirubin (TBIL) measurement greater than or equal to 2 times (2×) ULN during the treatment period will be summarized for all patients with a postbaseline value.

Hy's law is defined as the combination of drug related elevation of ALT≥3× ULN and TBIL≥2× ULN, in the absence of significant cholestasis (that is ALP< 2× ULN), and in the absence of other causes of liver injury.

The analysis of elevation in ALT, AST, ALP, and TBIL will contain 3 subsets:

- patients whose non-missing maximum baseline value is less than or equal to 1× ULN for ALT, AST, ALP, and TBIL.
- patients whose non-missing maximum baseline value is greater than 1× ULN for ALT, AST, ALP, and TBIL , at the same time less than or equal to 2× ULN for ALT and AST, 1.5× ULN for ALP and TBIL.
- patients whose non-missing maximum baseline value is greater than 2× ULN for ALT and AST, 1.5× ULN for ALP and TBIL.

6.14.4. Vital Signs and Weight

Vital signs collected during the study include systolic and diastolic blood pressure (SBP and DBP), pulse, and temperature. Blood pressure and pulse measurements will be taken when the patient is in a sitting position.

The number and percent of patients meeting criteria for categorical changes of interest in vital signs and weight at any time during study will be summarized. Treatment group comparisons will be performed using Fisher's exact test.

Table CGAX. 6.3 displays the criteria for categorical changes of interest in vital signs and weight. The last column of the table displays the patient populations defined by baseline categories.

The criteria to identify patients with treatment-emergent abnormal changes generally consist of a combination of two parts, an absolute threshold and a change from baseline amount.

- The absolute threshold in the criteria is based on 1) minimum postbaseline when the direction is low; 2) maximum postbaseline when the direction is high.
- The change from baseline amount in the criteria is 1) decrease from baseline (defined below and in Table CGAX. 6.1) to minimum postbaseline when the direction is low;

2) increase from baseline (defined below and in Table CGAX. 6.1) to maximum postbaseline when the direction is high.

The baseline for SBP, DBP, and pulse is defined as the last non-missing baseline value during the baseline period (See Table CGAX. 6.1). To be exact,

- For analyses including double-blind treatment phase, the baseline for SBP, DBP, and pulse is defined as the last non-missing value before randomization.
  The rationale for using the last available value in the baseline period is to minimize the potential confound of discontinuing or dose stabilization of medications that modulate BP and pulse during the screening phase (which is early in the baseline period).
- Similarly, for GMB-treated time, the baseline is defined as the last non-missing value before patient's first dose of GMB. This baseline definition was chosen to be consistent with the analysis approach for double-blind treatment phase as described above.

This baseline definition for SBP, DBP, and pulse applies to all analyses (both continuous and categorical) for SBP, DBP, and pulse.

The baseline and postbaseline values for temperature and weight are defined in Table CGAX. 6.1.

- For continuous analyses of temperature and weight, last non-missing baseline during the baseline period will be used as the baseline.
- For the analyses of categorical changes of interest in temperature and weight,
  - the baseline is defined as the minimum value during baseline period when the direction is low.
  - the baseline is defined as the maximum value during the baseline period when the direction is high.

Table CGAX. 6.3 Criteria for Categorical Changes of Interest in Vital Signs and Weight

Parameter	Direction	Criteria	Patients Population defined by Baseline Categories
Systolic BP (mm Hg) (sitting)	Low	≤90 and decrease ≥20	All patients; >90; ≤90
	High	≥140 and increase ≥20	All patients; <140, ≥140
	PCS High	≥180 and increase ≥20	All Patients; <180; ≥180
	Sustained Elevation	≥140 and increase ≥20 at 2 consecutive visits	All patients; < 140; ≥140
Diastolic BP (mm Hg) (sitting)	Low	≤50 and decrease ≥10	All patients; >50; ≤50
	High	≥90 and increase ≥10	All patients; <90; ≥90
	PCS High	≥105 and increase ≥15	All Patients; <105; ≥105
	Sustained Elevation	≥90 and increase ≥10 at 2 consecutive visits	All patients; < 90; ≥90
Systolic BP or Diastolic BP (mm Hg) (sitting)	Sustained Elevation	Meeting criteria for systolic BP for 2 consecutive visits or meeting criteria for diastolic BP for 2 consecutive visits or both	All Patients
Pulse (bpm) (sitting)	Low	<50 and decrease ≥15	All patients; ≥50; <50
	High	>100 and increase ≥15	All patients; ≤100; >100
	Sustained Elevation	>100 and increase ≥15 at 2 consecutive visits	All patients; ≤100; >100
Weight (kg)	Low	(Loss) decrease ≥7%	All patients
	High	(Gain) increase ≥7%	All patients
Temperature (° F)	Low	<96° F and decrease ≥2° F	≥96°F
	High	≥101° F and increase ≥2° F	<101°F

Abbreviations: BP = blood pressure; bpm = beats per minute; F = degrees Fahrenheit; kg = kilograms; mm Hg = millimeters of mercury; PCS= Potentially Clinically Significant.

6.14.5. Electrocardiograms

Analyses of corrected QT (QTc) interval will be calculated using Fridericia's correction formula, as $QT/RR^{1/3}$ (QTcF (msec)). For the QTcF calculation, the unit for QT is millisecond and the unit for RR is second. For patients with QRS ≥120 msec at any time during the study, the QT and QTc interval will be excluded from the analyses. A listing of ECG data for patients with QRS ≥120 msec at any time during the study will be provided.

The baseline for ECG is defined as the last non-missing baseline value during the baseline period (See Table CGAX. 6.1). To be exact,

- For analyses including double-blind treatment phase, the baseline for ECG is defined as the last non-missing value before randomization.
  The rationale for using the last available value in the baseline period is to minimize

the potential confound of discontinuing or dose stabilization of medications that modulate ECG during the screening phase (which is early in the baseline period).

- Similarly, for GMB-treated time, the baseline is defined as the last non-missing value before the patient's first dose of GMB. This baseline definition was chosen to be consistent with the analysis approach for the double-blind treatment phase as described above.

This baseline definition for ECG applies to all analyses (both continuous and categorical, quantitative and qualitative) for ECG.

The baseline and postbaseline values are summarized in Table CGAX. 6.1.

The number and percent of patients meeting criteria for treatment-emergent abnormalities in ECG intervals (PR, QRS, and QTcF) and heart rate at any time during study will be summarized. Treatment group comparisons will be performed using Fisher's exact test.

Table CGAX. 6.4 displays the criteria for treatment-emergent changes in ECG intervals and heart rate. The absolute threshold in the criteria is based on 1) minimum postbaseline when the direction is low; 2) maximum postbaseline when the direction is high.

- For treatment-emergent low analyses: Patients with all normal or high values at baseline (no low values) will be included.
- For treatment-emergent high analyses: Patients with all normal or low values at baseline (no high values) will be included.
- For treatment-emergent increase analyses: Patients with a baseline and at least 1 postbaseline result will be included.

Table CGAX. 6.4 Criteria for Treatment-Emergent Changes in ECG Intervals and Heart Rate

Parameter	Direction	Criteria	
Heart Rate (bpm)	Low	<50 and decrease ≥15	
	High	>100 and increase ≥15	
PR Interval (msec)	Low	<120	
	High	≥220	
QRS Interval (msec)	Low	<60	
	High	≥120	
QTcF (msec)	Low	Males: <330	Females: <340
	High	Males: >450	Females: >470
	PCS High	>500 msec	
	Increase	Increase >30 msec	
		Increase >60 msec	
		Increase >75 msec	

Abbreviations: bpm = beats per minute; ECG = electrocardiogram; msec = milliseconds; PCS = potentially clinically significant; QTcF = Fridericia's corrected QT interval.

6.15. Subgroup Analyses

Subgroup analyses will be performed for the primary efficacy measure (change from baseline in the number of MHDs) separately for each of the subgroup populations listed in Table CGAX. 6.5. Subgroup analyses will be conducted only for the ITT patients in Study Period III.

Each of the subgroup analyses for the primary measure of change from baseline in the number of MHDs will be conducted using MMRM. The same MMRM model described in Section 6.10.1 will be used, with terms of subgroup, subgroup-by-treatment, subgroup-by-month, and subgroup-by-treatment-by-month interactions added as additional covariates.

Table CGAX. 6.5 Definition of Subgroup Variables

Subgroup Variable	Categories
Sex	Male, female
Country	China, India, Russia
Baseline number of migraine headache days (MHDs)	2 levels of baseline migraine frequency : • <8 MHDs • ≥8 MHDs

6.16. Protocol Deviations

Protocol deviations will be identified throughout the study. Important protocol deviations (IPDs) are defined as those deviations from the protocol that would potentially compromise patients' safety, data integrity, or study outcome. The details of the categories and subcategories of IPDs and how the IPDs would be identified are contained in a separate document (Trial Issues Management Plan for study CGAX).

A table and listing of important protocol deviations for Intent-to-Treat patients during baseline, double-blind treatment phase, open-label treatment phase, or post-treatment phase, will be provided by each treatment groups and overall.

6.17. Interim Analyses

No interim analyses are planned for this study. If an unplanned interim analysis is deemed necessary, the appropriate Lilly medical director, or designee, will be consulted to determine whether it is necessary to amend the protocol.

This study will include two database locks. The first is to occur after all patients have had the opportunity to complete the 3-month double-blind treatment phase (Study Period III). The purpose of this analysis is the final analysis of the primary efficacy endpoint, as well as efficacy and safety analyses of the double-blind phase. Study sites, patients, and all Lilly personnel directly involved in the continuing trial will remain blinded to patients' double-blind treatment assignment. Unblinding details are specified in a separate unblinding plan document. This

analysis will not be considered an interim analysis because the double-blind phase is of primary interest.

The second database lock will occur at the post-treatment follow-up phase once all patients have had the opportunity to complete the entire study. This reporting database will include all data for all patients enrolled into the study, including all open-label and post-treatment phase data.

6.18. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

Summary of adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Adverse Events are summarized: by treatment group, by MedDRA preferred term.

- An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (TEAE).
- An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of patients at risk of an event
    - the number of patients who experienced each event term
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

7. Unblinding Plan

After the reporting database is locked for statistical analysis of the double-blind treatment phase, a limited number of Lilly personnel will be unblinded to complete the study report and prepare for submission. However, study sites, patients and all Lilly personnel directly involved in the ongoing open-label and post-treatment follow-up phases will remain blinded to patients' double-blind treatment assignment until final database lock. Unblinding details are specified in a separate unblinding plan document.

A final database lock is planned at the end of the study when all patients complete the study. The final analysis will include all data that are collected during the study, including all open-label and post-treatment phase data.

8. References

Blumenfeld AM, Varon SF, Wilcox TK, Buse DC, Kawata AK, Manack A, Goadsby PJ, Lipton RB. Disability, HRQoL and resource use among chronic and episodic migraineurs: results from the International Burden of Migraine Study (IBMS). *Cephalalgia*. 2011;31(3):301-315.

Buse DC, Bigal MB, Rupnow M, Reed M, Serrano D, Lipton R. Development and validation of the Migraine Interictal Burden Scale (MIBS): A self-administered instrument for measuring the burden of migraine between attacks. *Neurology*. 2007;68:A89.

Buse DC, Rupnow MF, Lipton, RB. Assessing and managing all aspects of migraine: migraine attacks, migraine-related functional impairment, common comorbidities, and quality of life. *Mayo Clin Proc*. 2009;84(5):422-435.

Diggle P, Kenward MG. Informative dropout in longitudinal data analysis (with discussion). *Appl Stat*. 1994;43(1):49-93.

Diggle P, Liang K, Zeger S. Analysis of Longitudinal Data. Oxford: Clarendon Press; 1994.

[IHS] International Headache Society, Headache Classification Subcommittee. The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018;38(1):1–211.

Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics.* 1997;53(3):983-997.

Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-613.

Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc*. 1952;47(260):583-621. Addendum. 1953;48:907-911.

Permutt T. Sensitivity analysis for missing data in regulatory submissions. *Stat Med*. 2015;35(17):2876-2879.

Seo JG, Park SP. Validation of the Patient Health Questionnaire-9 (PHQ-9) and PHQ-2 in patients with migraine. *J Headache Pain*. 2015a;16:65.

Seo JG, Park SP. Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. *J Headache Pain*. 2015b;16:97.

Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-1097.

Stewart WF, Lipton RB, Kolodner K, Liberman J, Sawyer J. Reliability of the migraine disability assessment score in a population-based sample of headache sufferers. *Cephalalgia*. 1999;19(2):107-114.

Stewart WF, Lipton RB, Dowson AJ, Sawyer J. Development and testing of the Migraine Disability Assessment (MIDAS) Questionnaire to assess headache-related disability. *Neurology*. 2001;56(6 Suppl 1):S20-S28.

Wilcoxon F. Individual comparisons by ranking methods. *Biometr Bull*. 1945;1(6):80-83.

Leo Document ID = f8e2cf98-0749-43b9-b35f-bac93baa0a0f

Approver: PPD

Approval Date & Time: 24-Aug-2021 01:17:53 GMT

Signature meaning: Approved