STATISTICAL ANALYSIS PLAN

A Phase 2, Randomized, Double-blind, Placebo-controlled Study to Evaluate the Efficacy and Safety of AMG 334 in Migraine Prevention

Protocol Number: **20120309**

Version: **Version 1.0**

Date: **February 14, 2016**

Author: PPD

NCT Number: 2630459
This NCT number has been applied to the document
for purposes of posting on clinicaltrials.gov

Table of Contents

Table of Abbreviations 5

1. Introduction 6

2. Objectives 6
   - 2.1 Primary 6
   - 2.2 Secondary 6
   - 2.3 Exploratory 6

3. Study Overview 8
   - 3.1 Study Design 8
     - 3.1.1 Double-blind Treatment Phase 8
     - 3.1.2 Open-Label Treatment Phase 9
     - 3.1.3 Safety Follow-up Phase 9
   - 3.2 Sample Size 9

4. Study Endpoints and Covariates 10
   - 4.1 Study Endpoints 10
     - 4.1.1 Primary Endpoint 10
     - 4.1.2 Secondary Endpoints 10
     - 4.1.3 Exploratory Endpoints 10
   - 4.2 Planned Covariates 12

5. Hypotheses 12

6. Definitions 13
   - 6.1 Definition of Terms Included in Study Endpoints 13
     - 6.1.1 Efficacy Endpoints 13
       - Migraine Physical Function Impact Stand-alone Item 17
     - 6.1.2 Safety Endpoints 18
   - 6.2 Study Dates 19
   - 6.3 Study Points of Reference 20
   - 6.4 Study Time Intervals 21
   - 6.5 Subject Disposition 25
   - 6.6 Arithmetic Calculations 26
   - 6.7 Disease Characteristics 28

7. Analysis Subsets 29
   - 7.1 Full Analysis Set 29
   - 7.2 Efficacy Analysis Set 29
   - 7.3 Safety Analysis Set 29
   - 7.4 Open-Label Treatment Phase Set 29
   - 7.5 Per Protocol Set 29
   - 7.6 Patient Reported Outcomes Analysis Set 31

7.7 Interim Analysis Set 31
7.8 Subgroup Analyses 31

8. Interim Analysis and Early Stopping Guidelines 31

9. Data Screening and Acceptance 31
9.1 General Principles 31
9.2 Data Handling and Electronic Transfer of Data 32
9.3 Handling of Missing and Incomplete Data 32
9.3.1 Missing and Incomplete Dates 33
9.3.2 Missing Baseline Evaluation 33
9.3.3 Missing Post-baseline Evaluation in Double-Blind Treatment Phase 34
9.4 Detection of Bias 35
9.5 Outliers 35
9.6 Distributional Characteristics 35
9.7 Validation of Statistical Analyses 36

10. Statistical Methods of Analysis 36
10.1 General Principles 36
10.2 Subject Accountability 37
10.3 Important Protocol Deviations 37
10.4 Demographic and Baseline Characteristics 37
10.5 Efficacy Analyses 38
10.5.1 Analyses of Primary and Secondary Efficacy Endpoints 42
10.5.2 Analyses of Exploratory Efficacy Endpoints 43
10.5.3 Analyses of Efficacy Endpoints in Open-Label Treatment Phase 44
10.5.4 Analyses of Patient Reported Outcomes (PROs) 44
10.5.5 Pharmacokinetic Endpoints 44
10.6 Safety Analyses 44
10.6.1 Adverse Events 45
10.6.2 Laboratory Test Results 45
10.6.3 Vital Signs 46
10.6.4 Columbia-Suicide Severity Rating Scale (C-SSRS) 46
10.6.5 Electrocardiogram (ECG) 46
10.6.6 Antibody Formation 46
10.6.7 Exposure to Investigational Product 47
10.6.8 Summary Concomitant Medication Use 47
10.7 Pharmacokinetic Analysis 47
10.7.1 Exposure-Response Analysis 47

11. Changes from Protocol-specified Analyses 47

12. Literature Citations / References 48

13. Prioritization of Analyses....48
14. Data not Covered by This Plan....48
15. Appendices....49

List of Tables

Table 1. Summary of Efficacy Endpoints and Analysis Methods....39

List of Appendices

Appendix A. Technical Detail and Supplemental Information Regarding Statistical Procedures and Programs....50
Appendix B. Code Fragments....56
Appendix C. Reference Values/Toxicity Grades....63
Appendix D. Patient-reported Outcome Forms/Instruments....64

Table of Abbreviations

In addition to the study glossaries and abbreviations defined in the protocol, here is a table of abbreviations used in this document:

Abbreviation	Definition
ANCOVA	Analysis of Covariance
CMH	Cochran-Mantel-Haenszel
CPMS	Clinical Pharmacology Modeling and Simulation
CTCAE	Common Toxicity Criteria for Adverse Events
DMP	Data Management Plan
DTP	Data Transfer Plan
EDC	Electronic Data Capture
EOS	End of Study
ET	Early Termination
GEE	Generalized Estimating Equations
GLMM	Generalized Linear Mixed Model
GSO-DM	Global Study Operations-Data Management
IPW	Inverse Probability Weighting
IVRS	Interactive Voice Response System
LOCF	Last Observation Carried Forward
MAR	Missing at Random
MCMC	Markov Chain Monte Carlo
MCP-Mod	Multiple Comparison Procedure – Modelling
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
NRI	Non-responder Imputation
SAP	Statistical Analysis Plan
SSAP	Supplemental Statistical Analysis Plan
WHODRUG	World Health Organization Drug Dictionary

1. Introduction

The purpose of this statistical analysis plan (SAP) is to provide details of the statistical analyses that have been outlined within the **Protocol 1.0** for AMG 334 Study 20120309 dated **June 16, 2015**. The study is being conducted in migraine patients in Japan. The scope of this plan includes the primary analysis and the final analysis, and will be executed by the Biostatistics Department. The analysis plan for the exposure–response analysis and additional exploratory efficacy endpoints from eDiary and bridging analysis will be provided in separate Supplement Statistical Analysis Plans (SSAPs).

2. Objectives

2.1 Primary

To evaluate the effect of AMG 334 compared to placebo on the change from baseline in mean monthly migraine days, in subjects with episodic migraine

2.2 Secondary

Efficacy:

- To evaluate the effect of AMG 334 compared to placebo on the proportion of subjects with at least 50% reduction from baseline in mean monthly migraine days
- To evaluate the effect of AMG 334 compared to placebo on the change from baseline in mean monthly acute migraine-specific medication treatment days

Safety:

To evaluate the safety and tolerability of AMG 334

2.3 Exploratory

- To evaluate the effect of AMG 334 compared to placebo on change from baseline in headache impact scores as measured by the Headache Impact Test (HIT-6)
- To evaluate the month of onset of action of AMG 334 compared to placebo as assessed by monthly migraine days
- To evaluate the effect of AMG 334 compared to placebo on the change from baseline in mean monthly migraine attacks
- To evaluate the effect of AMG 334 compared to placebo on change from baseline in mean monthly headache (migraine and non-migraine headache) days

- To evaluate the effect of AMG 334 compared to placebo on the proportion of subjects with at least 75% reduction from baseline in mean monthly migraine days
- To evaluate the effect of AMG 334 compared to placebo on the proportion of subjects with 100% reduction from baseline in mean monthly migraine days
- To evaluate the effect of AMG 334 compared to placebo on change from baseline in monthly acute headache medication treatment days
- To evaluate the effect of AMG 334 compared to placebo on change from baseline in monthly hours of migraine headache
- To evaluate the effect of AMG 334 compared to placebo on change from baseline in monthly average severity of migraine pain
- To evaluate the effect of AMG 334 compared to placebo on change from baseline in mean monthly migraine days with severe pain
- To evaluate the effect of AMG 334 compared to placebo on change from baseline in mean monthly hours of severe migraine pain
- To evaluate the effect of AMG 334 compared to placebo on migraine pain interference with daily activities as measured by the migraine symptom interference items
- To evaluate the effect of AMG 334 compared to placebo on the change from baseline in the overall impact on everyday activities score as measured by the Migraine Physical Function stand-alone item
- To evaluate AMG 334 pharmacokinetics (PK) in subjects with migraine and characterize the exposure-response (E-R) relationships for efficacy and safety endpoints
- To investigate the dose-response relationship of AMG 334 for efficacy
- CCI [redacted]
- To evaluate the long-term safety, tolerability and maintenance of effect of AMG 334 after 76 weeks of treatment

3. Study Overview

3.1 Study Design

This is a Phase 2, multicenter, randomized, stratified, double-blind, placebo-controlled, parallel-group study of subjects in Japan with episodic migraine. The study is composed of an initial screening phase (up to 3 weeks), a 4-week baseline phase, a 24-week double-blind treatment phase (DBTP), a 52-week open-label treatment phase (OLTP), and a 12-week safety follow-up phase (16 weeks after the last dose of investigational product).

The overall study design is described by a study schema at the end of the protocol synopsis section.

3.1.1 Double-blind Treatment Phase

Approximately 459 eligible subjects will be randomized in a 2:1:2:2 ratio to 1 of 4 treatment groups: placebo, AMG 334 28 mg, AMG 334 70 mg, or AMG 334 140 mg, with approximately 131 subjects assigned to placebo, approximately 66 subjects assigned to AMG 334 28 mg, approximately 131 subjects assigned to AMG 334 70 mg, and approximately 131 subjects assigned to AMG 334 140 mg. The randomization will be stratified by prior/current treatment with migraine prophylactic medication (current migraine prophylactic medication treatment vs prior migraine prophylactic medication treatment only vs no prior or current migraine prophylactic medication treatment). There will be a limit on the percentage of subjects on current migraine prophylactic medication treatment.

Amgen investigational product (ie, AMG 334 28 mg, AMG 334 70 mg, AMG 334 140 mg, or placebo) will be dosed monthly (QM) by subcutaneous (SC) injections. Double-blind AMG 334 140 mg, AMG 334 70 mg, AMG 334 28 mg, or placebo will be administered during the 24-week DBTP (ie, at day 1 and weeks 4, 8, 12, 16, and 20).

Migraine and non-migraine headache outcomes and other efficacy measures will be assessed based on eDiary data. Subjects will also have in-clinic study visits at Day 1 and monthly visits thereafter.

Subjects who permanently discontinue investigational product during the DBTP are to continue to return for all other study procedures until the end of the DBTP and study procedures for the safety follow-up visit 16 weeks after the last dose of investigational product.

3.1.2 Open-Label Treatment Phase

At the week 24 visit, subjects will be entered into the 52-week OLTP and will begin to receive open-label AMG 334 70 mg QM SC.

A subject who discontinues open-label investigational product or the study during OLTP will complete the week 76/early termination (ET) visit and the safety follow-up visit 16 weeks after the last dose of investigational product.

Subjects will use eDiary every day for the first 6 months and last month of the OLTP to report information about their migraine and non-migraine headaches and acute headache medication use. Subjects will participate in monthly in-clinic study visits.

3.1.3 Safety Follow-up Phase

Subjects who complete the OLTP or who discontinue investigational product would complete the safety follow-up visit 16 weeks after the last dose of investigational product.

3.2 Sample Size

The primary endpoint is the change from baseline in mean monthly migraine days over the last three months (months 4, 5, and 6) of the DBTP. Using the treatment effect compared to placebo of -1.12 and -1.30 for the AMG 334 70 mg and 140 mg dose groups, respectively and a common standard deviation of 2.8 based on a placebo-controlled dose-finding study of JNS019 (topiramate) in migraine patients (www.clinicaltrials.gov), the planned sample size of 131 subjects in the placebo, 70 mg and 140 mg dose groups will provide 90% and 96% power for a two-sided test with significance level of 0.05 to show the difference of AMG 334 70 mg and 140 mg compared to placebo, respectively. The proposed number of subjects in the 28 mg (n=66) dose group with the placebo (n=131), the 70 mg (n=131) and the 140 mg (n=131) dose groups is sufficient to demonstrate a dose-response using Multiple Comparison Procedure – Modelling (MCP-Mod) analysis with minimum power of 95% and provide estimates of response in a Japanese population, using an assumed treatment effect of -0.55 for the AMG 334 28 mg dose group.

Power calculations are derived through nQuery 7.0.

4. Study Endpoints and Covariates

4.1 Study Endpoints

4.1.1 Primary Endpoint

Change from baseline in mean monthly migraine days. The mean monthly migraine days will be calculated using the monthly migraine days from each of the last three months (months 4, 5, and 6) of the DBTP.

Change from baseline in mean monthly migraine days is calculated based on the following: mean monthly migraine days – number of migraine days during the 4-week baseline phase.

4.1.2 Secondary Endpoints

Efficacy:

- Achievement of at least a 50% reduction from baseline in mean monthly migraine days over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in mean monthly acute migraine-specific medication treatment days over the last three months (months 4, 5, and 6) of the DBTP

Safety:

- Adverse events
- Clinical laboratory values and vital signs
- Anti-AMG 334 antibodies

4.1.3 Exploratory Endpoints

- Change from baseline in mean headache impact scores as measured by the Headache Impact Test (HIT-6) over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in monthly migraine days at assessment timepoints
- Change from baseline in mean monthly migraine attacks over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in monthly migraine attacks at assessment timepoints
- Change from baseline in mean monthly headache (migraine and non-migraine headache) days over the last three months (months 4, 5, and 6) of the DBTP

- Change from baseline in monthly headache (migraine and non-migraine headache) days at assessment timepoints
- Achievement of at least 50% reduction from baseline in monthly migraine days at assessment timepoints
- Achievement of at least 75% reduction from baseline in mean monthly migraine days over the last three months (months 4, 5, and 6) of the DBTP
- Achievement of at least 75% reduction from baseline in monthly migraine days at assessment timepoints
- Achievement of 100% reduction from baseline in mean monthly migraine days over the last three months (months 4, 5, and 6) of the DBTP
- Achievement of 100% reduction from baseline in monthly migraine days at assessment timepoints
- Change from baseline in mean monthly acute headache medication treatment days over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in monthly acute headache medication treatment days at assessment timepoints
- Change from baseline in monthly acute migraine-specific medication treatment days at assessment timepoints
- Change from baseline in mean monthly hours of migraine headache over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in monthly hours of migraine headache at assessment timepoints
- Change from baseline in mean monthly average severity of migraine pain over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in monthly average severity of migraine pain at assessment timepoints
- Change from baseline in mean monthly migraine days with severe pain over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in monthly migraine days with severe pain at assessment timepoints

- Change from baseline in mean monthly hours of severe migraine pain over the last three months (months 4, 5, and 6) of the DBTP
- Change from baseline in monthly hours of severe migraine pain at assessment timepoints
- Change from baseline in migraine pain interference with daily activities as measured by the migraine symptom interference items over the last three months (months 4, 5, and 6) of the DBTP at assessment timepoints
- Change from baseline in the overall impact on everyday activities score as measured by the Migraine Physical Function stand-alone item over the last three months (months 4, 5, and 6) of the DBTP
- AMG 334 exposure and PK-PD relationships

[redacted]

4.2 Planned Covariates

The following covariates will be included in the primary analysis of the efficacy endpoints:

- Stratification factor: prior/current treatment with migraine prophylactic medication (current migraine prophylactic medication treatment vs prior migraine prophylactic medication treatment only vs no prior or current migraine prophylactic medication treatment)
- Baseline value for corresponding endpoint (eg, baseline monthly migraine days will be added in the model as a covariate for the endpoint of change from baseline in monthly migraine days)

Stratification factor will use the values used for randomization unless otherwise noted.

5. Hypotheses

The primary endpoint will be tested for each AMG 334 treatment group compared to the placebo group sequentially at a 2-sided significance level of 0.05 in the order of AMG 334 140 mg vs. placebo, 70 mg vs. placebo and 28 mg vs. placebo. The lower dose group will be tested only when the higher dose group is considered statistically significant.

- Null Hypothesis: In subject with episodic migraine, the AMG 334 treatment group is same as placebo, in terms of the change in the mean monthly migraine days from baseline
- Alternative Hypothesis : In subject with episodic migraine, the AMG 334 treatment group is different from placebo, in terms of the change in the mean monthly migraine days from baseline

6. Definitions

6.1 Definition of Terms Included in Study Endpoints

6.1.1 Efficacy Endpoints

The baseline period for efficacy analysis is defined as the period between week -4 visit (when eDiary device is set up or eDiary device assignment date) and the day prior to study day 1 (study day 1 is not included).

Migraine Day

Any calendar day in which the subject experiences a qualified migraine headache (onset, continuation or recurrence of the migraine headache). A qualified migraine headache is defined as a migraine with or without aura, lasting for ≥ 30 minutes, and meeting at least one of the following criteria (a and/or b):

a) ≥ 2 of the following pain features:
    - Unilateral
    - Throbbing
    - Moderate to severe
    - Exacerbated with exercise/physical activity

b) ≥ 1 of the following associated symptoms:
    - Nausea and/or vomiting
    - Photophobia and phonophobia

If the subject took a migraine-specific medication (ie, triptan or ergotamine) during aura or to treat headache on a calendar day, then it will be counted as a migraine day regardless of the duration and pain features/associated symptoms.

Monthly Migraine Days

Number of migraine days between each monthly IP dose. Monthly migraine days at baseline is the number of migraine days in the baseline period. Days without eDiary data in each monthly interval are handled by proration according to Section 9.3.

Headache Day

Any calendar day in which the subject experiences a qualified migraine or non-migraine headache (initial onset, continuation or recurrence of the headache). A qualified headache is defined as**:**

- a qualified migraine headache (including an aura-only event that is treated with acute migraine-specific medication), or
- a qualified non-migraine headache, which is a headache that lasts ≥ 30 minutes and is not a qualified migraine headache, or
- a headache of any duration for which acute headache treatment is administered.

Monthly Headache Days

Number of headache days between each monthly IP dose. Monthly headache days at baseline is the number of headache days in baseline period. Days without eDiary data in each monthly interval are handled by proration according to Section 9.3.

Migraine Attack

An episode of any qualified migraine headache or migraine specific medication intakes for aura only. The following rules will be used to distinguish an attack of long duration from two attacks, or to distinguish between attacks and relapses:

a) A migraine attack that is interrupted by sleep, or temporarily remits, and then recurs within 48 hours will be considered as one attack and not two.

b) An attack treated successfully with medication but with relapse within 48 hours will be considered as one attack.

A migraine attack lasting more than 48 hours will be counted as one attack.

Monthly Migraine Attacks

Number of migraine attacks between each monthly IP dose. Monthly migraine attacks at baseline is the number of migraine attacks in baseline period. Days without eDiary data are handled by proration according to Section 9.3.

Response

At least 50% (or 75%, 100%) reduction from baseline in monthly migraine days.

Monthly Migraine Days with Severe Pain

Number of migraine days with severe pain between each monthly IP dose. Monthly rate at baseline is the number of migraine days with severe pain in baseline period. Days without eDiary data in each monthly interval are handled by proration according to Section 9.3.

Monthly Cumulative Hours of Severe Migraine Pain

Cumulative duration of any qualified migraine headache with severe pain between each monthly IP dose regardless of acute treatment use. Monthly cumulative hours of migraine pain at baseline is the cumulative duration of any qualified migraine headache with severe pain in baseline period. Days without eDiary data are handled by proration according to Section 9.3.

Monthly Cumulative Hours of Migraine Headache

Cumulative duration of any qualified migraine headache between each monthly IP dose regardless of acute treatment use. Monthly cumulative hours of migraine headache at baseline is the cumulative duration of any qualified migraine headache in baseline period. Days without eDiary data are handled by proration according to Section 9.3.

Monthly Acute Medication Use in Days

Number of days on which acute medications are used as recorded in eDiary between each monthly IP dose. Monthly acute treatment use at baseline is the number of acute treatment days in the baseline period. Days without eDiary data are handled by proration according to Section 9.3.

Monthly Migraine-Specific Medication Use in Days

Number of days on which migraine-specific medications are used between each monthly IP dose. Migraine-Specific Medications including two categories of medications: triptan-based migraine medications and ergotamine-based migraine medications. Monthly migraine-specific medication use at baseline is the number of migraine-specific medication days in the baseline period. Days without eDiary data are handled by proration according to Section 9.3.

Monthly Average Severity of Migraine Pain

Severity of migraine pain is graded as 1=mild, 2=moderate or 3=severe and is rated as its worst or peak intensity per migraine headache. Monthly average severity of migraine pain is defined as the sum of the severity of each observed qualified migraine headache

between each monthly IP dose divided by the total number of observed qualified migraine headaches (as one record in eDiary with start and end time) in that interval. If less than 14 days of eDiary data in each interval is recorded, then the monthly average severity of migraine pain will be set as missing. Monthly average severity of migraine pain at baseline is the average severity of migraine pain in the baseline period.

Migraine Symptom Interference Items

A single item will assess the extent to which the migraine-related symptoms interfered with the subject's daily activities within the past 24 hours. The scale is from 0 to 10 with higher score indicating more severe interference. This question is answered whether or not the subject reported headache. Data will be collected from eDiary.

Monthly average scale of migraine symptom interference is defined as the sum of observed scales divided by the number of days when this question is answered between each monthly IP dose. If less than 14 days of eDiary out of the interval is recorded, then the monthly average scale of migraine symptom interference will be set as missing. For monthly average scale of migraine symptom interference at baseline, data collected from end of initial screening phase (week -4 study visit) through the day that the investigational site runs the eligibility report will be utilized.

Bed Days due to Migraine Symptoms

A single item will assess if the subject spends the day or part of the day in bed due to migraine-related symptoms in the past 24 hours. Data will be collected from eDiary.

Monthly Bed Days is defined as the number of days during which the subject spends the day or part of the day in bed between each monthly IP dose. For monthly bed days at baseline, data collected from end of initial screening phase (week -4 study visit) through the day that the investigational site runs the eligibility report will be utilized. Days without eDiary data are handled by proration according to Section 9.3.

Missed Work/School Days due to Migraine Symptoms

A single item will assess if the subject misses work or school due to migraine-related symptoms in the past 24 hours. Data will be collected from eDiary.

Monthly Missed Work/School Days is defined as the number of Days on which subject missed work/school between each monthly IP dose. For monthly Missed Work/School Days at baseline, data collected from end of initial screening phase (week -4 study visit) through the day that the investigational site runs the eligibility report will be utilized. Data

will be treated as missing if subject answered this question as "I do not work or attend school". Days without eDiary data are handled by proration according to Section 9.3.

Migraine Physical Function Impact Stand-alone Item

The Migraine Physical Function Impact stand-alone item is a global question which provides an assessment of overall impact of migraine on everyday activities. Subjects respond to the item using a 5-point scale, with difficulty items ranging from "Not difficult" to "Extremely difficult." These are assigned scores from 1 to 5, with 5 representing the greatest burden. The score will be rescaled to a 0 - 100 scale, with higher scores representing greater impact of migraine (ie, higher burden).

The recall period is the past 24 hours.

Subjects will complete the Migraine Physical Function Impact stand-alone item using the eDiary.

Headache Impact Test (HIT-6)

The Headache Impact Test (HIT-6) is a short-form self-administered questionnaire based on the Internet-HIT question pool. The HIT-6 was developed as a global measure of adverse headache impact to assess headache severity in the previous month and change in a patient's clinical status over a short period of time. Six questions cover:

- severe pain,
- limitation of daily activity (household, work, school and social),
- wanting to lie down when headache is experienced,
- feeling too tired to work or do daily activities because of headache,
- feeling fed up or irritated because of headache,
- headache limiting ability to concentrate or work on daily activities.

Each of the 6 questions is responded using 1 of 5 response categories: "never," "rarely," "sometimes," "very often," or "always." Besides, 6, 8, 10, 11, or 13 points, respectively, are assigned to the response provided for each HIT-6 item. These points are summed to produce a total HIT-6 score that ranges from 36 to 78. HIT-6 score are categorized into 4 grades, representing little or no impact (49 or less), some impact (50-55), substantial impact (56-59), and severe impact (60-78) due to headache. No recall period is specified for the first 3 items. The recall period is the past 4 weeks for the last 3 items. Please refer to Appendix D for scoring algorithm.

Achievement of at least a 50% reduction from baseline in mean monthly migraine days over the last three months (months 4, 5, and 6) of the double-blind treatment phase

Calculated based on the following: if (mean monthly migraine days over the last three months of the DBTP - baseline monthly migraine days)*100/baseline monthly migraine days is less than or equal to - 50%

Change from baseline in mean monthly acute migraine-specific medication treatment days over the last three months (months 4, 5, and 6) of the double-blind treatment phase

Calculated based on the following: (mean monthly acute migraine-specific medication treatment days over the last three months of the DBTP) – (baseline monthly acute migraine-specific medication treatment days)

6.1.2 Safety Endpoints

Serious Adverse Event (SAE)

SAEs determined by the flag indicating if the adverse event is serious on the Adverse Events eCRF page will include those that occur after signing of the informed consent and up to and including end of study.

Treatment-Emergent Adverse Event

Adverse Events (AEs) recorded on the Adverse Events eCRF page that occurs on or after first dose of investigational product as determined by the flag indicating if the adverse event started prior to the first dose on the Adverse Events eCRF and up to and including 112 days after the end of investigational product (16 weeks after the last dose of IP) or end of study, whichever comes earlier.

Serious Treatment-Emergent Adverse Event

A serious treatment-emergent adverse event is an SAE considered to be treatment-emergent.

Treatment-Related Adverse Event

A treatment-related AE is defined as a treatment-emergent adverse event that is considered by investigators to have reasonable possibility that it may have been caused by IP as determined by the flag indicating if the event is caused by investigational product on the Adverse Events eCRF page.

Serious Treatment-Related Adverse Event

A serious treatment-related adverse event is an SAE considered to be treatment-related.

Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is a clinician rating of suicidal behavior and ideation. Two versions depending on the type of visits will be used in this study: Screening and Since Last Visit. The C-SSRS consists of a maximum of 20 items to evaluate suicidal behavior and suicidal ideation.

Beck Depression Inventory (BDI)-II

The Beck Depression Inventory (BDI)-II is a 21-item questionnaire that assesses severity of depression. Each item is scored from 0 to 3. The total score is categorized into 4 severity grades: minimal depression (0-13), mild depression (14-19), moderate depression (20-28), and severe depression (29-63).

6.2 Study Dates

Informed Consent Date

The date on which subject signs the informed consent form.

eDiary Device Assignment Date

The date on which an eDiary device is assigned to a subject for the first time after completion of initial screening at week -4 visit.

Randomization (Enrollment) Date in DBTP

Randomization (Enrollment) Date in DBTP is the date on which a subject is assigned to one of the treatments through Interactive Voice Response System (IVRS) in DBTP.

First IP Dose Date

First IP Dose Date is the date on which a subject is administered the first dose of investigational product following randomization, which may be the same day or after the randomization date. For subjects who are randomized but not dosed with double-blind IP after randomization, First IP Dose Date is considered missing.

First OLTP IP Dose Date

First OLTP IP Dose Date is the date on which a subject is administered the first dose of investigational product in the open-label treatment phase following completion of DBTP.

Last IP Dose Date

Last IP Dose Date for each subject is defined as the latest date IP is administered.

End of IP Admin Date

End of IP Admin for each subject is defined as the date the decision was made to end IP as recorded on the End of IP eCRF page.

Subject-level End of Study (EOS) Date

End of study (EOS) date for each subject is defined as the last date on which the subject participated in the study. The date will be recorded on the End of Study eCRF page.

Primary Completion Date

The Primary Completion Date is the date of the End of Double-Blind Treatment Phase page for the last subject who completes DBTP.

Study Completion Date

The Study Completion Date is the EOS date of the last subject in the study.

6.3 Study Points of Reference

Baseline Assessment

Baseline assessment is defined as the last non-missing measurement for the endpoint of interest taken before the first dose of investigational product. In cases where baseline measurements are taken on the same day as IP, it will be assumed that these measurements are taken prior to IP being administered. For subjects who are randomized but not dosed after the randomization, the baseline of the study is defined as the last non-missing measurement prior to or on the date of randomization.

For C-SSRS, if subject completes C-SSRS form on Day 1 then all individual items from Day 1 visit will be used as baseline. If subject does not complete C-SSRS form on Day 1 then all individual items from Week -4 visit will be used as baseline.

Baseline assessment of data collected from eDiary will be summarized as monthly measurements using data collected during baseline period. See Section 6.1.1 for detail.

Study Day 1

Study Day 1 is defined as the first IP dose date. For subjects who are randomized but not dosed after randomization, the Study Day 1 is defined as the date of randomization.

Study Day

Study Day is defined as the number of days from Study Day 1.

Before Study Day 1:

Study Day = (Date of Interest – Date of Study Day 1)

On or after Study Day 1:

Study Day = (Date of Interest – Date of Study Day 1) + 1

Therefore the day prior to Study Day 1 is -1.

6.4 Study Time Intervals

Monthly Interval for Efficacy Endpoints

Monthly efficacy measurements will be calculated based on the subject's monthly IP dosing schedule defined below using eDiary data collected from beginning of the baseline phase (week -4 visit) up to week 76 (end of eDiary use). Any eDiary data occurring after EOS date will not be included in the analysis.

Study Phase	Assessment Timepoint	Start Date	End Date
Baseline Phase	Baseline	LogPad assignment date (Week -4 study visit)	Day prior to study day 1
Double-blind Treatment Phase	Week 4	• Study Day 1	• Week 4 dose date-1 • Study day 28 if Week 4 dose is not received (either missed or IP discontinued prior to Week 4)
	Week 8	• Week 4 dose date • Study day 29 if Week 4 dose is not received (either missed or IP discontinued prior to Week 4)	• Week 8 dose date-1 • Study day 56 if Week 8 dose is not received (either missed or IP discontinued prior to Week 8)
	Week 12	• Week 8 dose date	• Week 12 IP dose date-1 • Study day 84 if Week 12

		• Study day 57 if Week 8 dose is not received (either missed or IP discontinued prior to Week 8)	dose is not received (either missed or IP discontinued prior to Week 12)
	Week 16	• Week 12 dose date • Study day 85 if Week 12 dose is not received (either missed or IP discontinued prior to Week 12)	• Week 16 dose date-1 • Study day 112 if Week 16 dose is not received (either missed or IP discontinued prior to Week 16)
	Week 20	• Week 16 dose date • Study day 113 if Week 16 dose is not received (either missed or IP discontinued prior to Week 16)	• Week 20 IP dose date-1 • Study day 140 if Week 20 dose is not received (either missed or IP discontinued prior to Week 20)
	Week 24	• Week 20 dose date • Study day 141 if Week 20 dose is not received (either missed or IP discontinued prior to Week 20)	• Week 24 IP dose date-1 • Study day 168 if Week 24 dose is not received (either missed or IP discontinued prior to Week 24)
Open-label Treatment Phase	Week 28	• The 1st OL dose day	• Week 28 dose date-1 • OLTP day 28 if Week 28 dose is not received (either missed or IP discontinued prior to Week 28)

	Week 32 or later when eDiary is used	• Same rule as above	

Study Visit

Since the actual visit for a subject may not exactly coincide with their targeted visit date, the actual visit date is mapped to the study visit as follows.

Study Visit	Target Day	Study Day
Baseline	Please refer to Section 6.3 Baseline Assessment of the Study	
Day 1	1	1
Week 4	29	16-43
Week 8	57	44-71
Week 12	85	72-99
Week 16	113	100-127
Week 20	141	128-155
Week 24	169	• 156 to (week 24 OLTP dose date – 1) for subjects who receive OLTP dose at week 24 • 156-183 for subjects who did not receive OLTP dose at week 24
Week 28	197	• Week 24 OLTP dose date to 211 for subjects who receive OLTP dose at week 24 • 184-211 for subjects who received at least one OLTP dose and did not

		receive OLTP dose at week 24 (e.g., first OLTP dose was postponed)
Week 32	225	212-239
Week 36 to Week 72	Following same monthly windowing as Week 32	
Safety follow-up	Based on visit name from RAVE or other source data	

Note:
If more than one visit (including the unscheduled visits, ie, CPEVENT = 'UNSCHED') fall within the same defined window, scheduled visit will be used regardless of the distance from the target day. Unscheduled visit will only be used when there is no scheduled visit in the defined window. The closest visit to the target day among the same type of visit (scheduled vs. unscheduled) will be considered for analysis. If two assessment dates are equidistant from the target date, the latter visit will be considered for analysis.

For safety analyses which are summarized by study phase, analysis windows will be set up based on study phase:

Study Phase	Start Time Point	End Time Point
Double-Blind Treatment Phase	Study Day 1 after the 1st DB IP Dose	• First OL Dose Date for subjects who receive OL IP dose; • Min(EOS Date, last IP dose date + 112 days) for subjects who do not receive any OL IP dose
Open-Label Treatment Phase	1st OL IP Dose Date	Min(EOS Date, last OL IP dose date + 112 days)
Entire Study	Study Day 1 after the 1st DB IP Dose	Min(EOS Date, last IP dose date + 112 days)

Note: all antibody data will be included in antibody analysis.

6.5 Subject Disposition

Randomized

Individuals are considered randomized if they have been assigned a randomization number. Randomized individuals are referred to as “subjects”.

Completing the Double-Blind Treatment Phase

Subjects are defined as completing the DBTP if they complete week 24 assessment. It will be derived from the End of Double-Blind Treatment Phase CRF page with “Completed” as the reason for ending study phase.

Completing the Open-Label Treatment Phase

Subjects are defined as completing the OLTP if they complete the week 76 assessment**.** It will be derived from the End of Open-Label Treatment Phase CRF page with “Completed” as the reason for ending study phase.

Completing Study

Subjects are defined as completing study if they complete the whole 88 weeks of study evaluation**.** It will be derived from the End of Study CRF page with “Completed” as the primary reason for ending study.

Exposed to Investigational Product

Subjects are defined as exposed if they receive at least one dose of investigational product.

Completing the Double-Blind Investigational Product

Subjects are defined as completing double-blind investigational product if they complete week 20 IP dose or temporarily withhold IP at week 20 and continue to receive IP during OLTP. It will be derived from the End of IP DBTP CRF page with “Completed” as the reason for ending IP.

Completing the Open-Label Investigational Product

Subjects are defined as completing open-label investigational product if they complete week 72 IP dose. It will be derived from the End of IP OLTP CRF page with “Completed” as the reason for ending IP.

On-study

Subjects are considered on-study if they have been randomized and have not yet had their end of study visit.

6.6 Arithmetic Calculations

Duration of Migraine

The number of years from the diagnosis date (DXDT) of migraine (migraine with aura or migraine without aura, whichever is earlier) to the date informed consent is signed.

If month or day of the diagnosis date is missing, follow the formula below to calculate the duration:

Observed portion	Missing portion	Formula to Calculate Duration
Year, Month, Day		(Informed Consent Day – DXDT)/365.25
Year, Month	Day	[Year(Informed Consent Day)-Year(DXDT)]+ [Month(Informed Consent Day)-Month(DXDT)]/12 *if it equals 0, add 1 month or 1/12 years (this is to avoid a disease duration of 0)
Year	Month, Day	[Year(Informed Consent Day)-Year(DXDT)] *if it equals 0, add 1 month or 1/12 years (this is to avoid a disease duration of 0)

Duration of DB IP Exposure

If subject enter into OL treatment phase,

Duration = Minimum (Last DB Dose Date + 27, First OL Dose Date - 1, EOS Date) – First DB Dose Date + 1

Otherwise,

Minimum (Last DB Dose Date + 27, EOS Date) – First DB Dose Date + 1

Duration of OLTP IP Exposure

Minimum (Last OLTP Dose Date + 27, EOS Date, Date of Data Cut-off) – First OLTP Dose Date + 1

Change from Baseline in Monthly Efficacy Measurement

The change from baseline in monthly efficacy measurement is the monthly efficacy measurement prior to the given time point minus the baseline monthly efficacy measurement. Please refer to the Monthly Intervals defined in Section 6.4. For example, change from baseline in monthly migraine days at Week 24 will be calculated based on the following:

> (Monthly migraine days during the week 24 interval) – (monthly migraine days during the baseline phase)

If the baseline or post-baseline value is missing, then the change from baseline is set to missing.

Following Efficacy endpoints will be calculated as above:

- Change from baseline in monthly migraine days
- Change from baseline in monthly migraine attacks
- Change from baseline in monthly headache days
- Change from baseline in monthly migraine days with severe pain
- Change from baseline in monthly cumulative migraine hours
- Change from baseline in monthly cumulative migraine hours with severe pain
- Change from baseline in monthly acute medication use in days
- Change from baseline in monthly migraine-specific medication use in days
- Change from baseline in monthly average severity of migraine pain
- Change from baseline in monthly average migraine symptoms interference scale
- Change from baseline in monthly bed days due to migraine related symptoms
- Change from baseline in monthly missed work/school days due to migraine related symptoms
- Change from baseline in monthly headache impact scores as measured by HIT-6

Mean Change from Baseline in Monthly Efficacy Measurement over multiple months

The Mean Change from Baseline in Monthly Efficacy Measurement is the Arithmetic mean of each change from baseline value for the months considered

Percent Change from Baseline

The change from baseline divided by baseline and multiplied by 100:

(Post-baseline – Baseline) * 100 / Baseline

If the baseline value is 0 and the post-baseline value is also 0, then the percent change from baseline is set to 0. If the baseline value is 0 and the post-baseline value is non-zero, then the percent change from baseline is set to missing.

Percent change from baseline in (mean) monthly migraine days which is used to determine 50%, 75%, and 100% responder will be calculated as above.

Subject Incidence

The subject incidence for a given event in a given period is defined as the number of subjects with at least one reported occurrence of the event divided by the number of subjects who entered that period. For subjects with multiple occurrences of the same event, the event will only be counted once per subject.

Exposure-Adjusted Incidence Rate

The exposure-adjusted incidence rate for a given event in a given period is defined as the number of subjects with at least one reported occurrence of the event in a given time period at risk divided by total exposure time of all subjects who are at risk for the event. For subjects with events, only the time until the first event contributes to the total exposure time. For subjects with no event, the exposure time is the time from the first IP dose to the last follow-up assessment. This rate will be presented per 100 subject years. For subjects with multiple occurrences of the same event, the first occurrence of an event will be counted for each subject.

6.7 Disease Characteristics

Migraine-Specific Medications

Migraine-specific medications include two categories of medications: triptan-based and ergotamine-based migraine medications collected from subject's eDiary.

Treatment Failure of Prior Migraine Prophylactic Medications

Treatment failure of prior migraine prophylactic medications is determined by "Reason for ending medication" as "Lack of efficacy" or "Adverse Reaction" in the Prior Migraine Prophylactic Medication eCRF page.

7. Analysis Subsets

7.1 Full Analysis Set

The Full Analysis Set (FAS) includes all subjects who were randomized in the study. Subjects will be analyzed according to their randomized treatment, regardless of the treatment received. Tabulations of demographic and baseline characteristics, disposition, and important protocol deviations (IPD) will utilize this analysis set.

7.2 Efficacy Analysis Set

The Efficacy Analysis Set (EAS) will be used to carry out the primary analyses of efficacy endpoints, which is a subset of the Full Analysis Set consisting of subjects who received at least one dose of IP and completed at least one post-baseline monthly migraine day measurement in the DBTP**.** Subjects will be analyzed according to their randomized treatment, regardless of the treatment received. Analyses for efficacy endpoints and patient reported outcomes (PROs) will utilize this analysis set.

7.3 Safety Analysis Set

The Safety Analysis Set will consist of all randomized subjects who received at least one dose of IP. Subjects will be analyzed according to the randomized treatment unless a subject has received the incorrect dose during the entire DBTP. Analyses for safety endpoints and summary of IP administration in DBTP will utilize this analysis set.

7.4 Open-Label Treatment Phase Set

The Open-Label Treatment Phase Set (OLTPS) will consist of all subjects receiving at least one dose of AMG 334 in the OLTP. This analysis set will be used when summarizing data collected during the OLTP.

7.5 Per Protocol Set

The Per Protocol Set is a subset of the Efficacy Analysis Set which consists of subjects who received IP at Week 12, 16, and 20 and do not satisfy any of the following conditions:

1) Important protocol deviations that will potentially impact primary analysis of efficacy endpoints from week 16 to week 24 or violate GCP at site as specified in the IPD List (Cabinets/AMG 334/TMF/20120309 - Episodic Migraine Prophylaxis

Japan/Study Protocol Compliance/Protocol Deviation Listing) to be applied in per protocol analysis:

- Informed Consent was not provided
- Monthly migraine days at baseline is < 4 or ≥ 15
- Monthly headache days at baseline is ≥ 15
- No therapeutic response with > 2 of the 7 medication categories for prophylactic treatment of migraine after an adequate therapeutic trial
- Received botulinum toxin in the head and/or neck region within 4 months prior to the start of the baseline phase and throughout the study
- Ergotamine-derivatives, steroids, and triptans used for migraine prophylaxis are excluded within 2 months prior to the start of the baseline phase and throughout the study
- Devices and procedures used for migraine prophylaxis are excluded within 2 months prior to the start of the baseline phase and throughout the study
- Using opioid- or butalbital-containing analgesics on ≥ 4 days in any month during the 2 months prior to the start of the baseline phase
- History of chronic pain syndromes (eg, fibromyalgia, chronic back pain, chronic pelvic pain)
- Received current migraine prophylaxis medication prior to the start of the baseline phase after closure of the stratum
- Received less than 75% or more than 125% of planned total dose of AMG 334 at any of the weeks 12, 16, 20 (based on unblinded information) (Example: subject who was randomized to AMG 334 70 mg treatment group received 140 mg, not 70 mg at week 20)
- Important GCP violation resulting in that data can't be used for analysis

2) In addition to above protocol deviations, subjects who did not have an observed monthly migraine day value at any of the weeks 16, 20, 24 will also be excluded from Per Protocol Set

The Per Protocol Set will be used to perform the sensitivity analysis on the primary and secondary efficacy endpoints.

7.6 Patient Reported Outcomes Analysis Set

The Efficacy Analysis Set will be used to carry out analyses for patient reported outcomes.

7.7 Interim Analysis Set

The Interim Analysis Set will include all subjects at the time when 100 subjects have completed 52 weeks of treatment with investigational product. The purpose of this interim analysis is to evaluate the long-term efficacy and safety of AMG 334 in subjects with episodic migraine after the 24-week DBTP.

7.8 Subgroup Analyses

The primary and secondary efficacy will be analyzed in the subgroups defined by the stratification factor, BMI (< median vs ≥ median), baseline monthly migraine days (< 8 vs ≥ 8) and treatment failure of prior migraine prophylactic medications.

8. Interim Analysis and Early Stopping Guidelines

An independent Date Monitoring Committee (DMC) will review and make recommendations regarding the safety of the study participants throughout the DBTP of the study, and until treatment assignment information is available to the study team for the primary analysis. The DMC will be composed of external advisors, including at least 2 clinicians and a biostatistician. Summaries of data at the treatment group level will be prepared and presented by an independent biostatistician at the DMC meeting. Additionally the DMC will review pharmacokinetic data and compare the results against pre-defined criteria for observed and predicted data to potentially recommend discontinuation of enrollment into one or more of the AMG 334 treatment groups, in which case randomization of the remaining subjects would continue for the remaining treatment groups.

During the OLTP of the study, an interim analysis is planned after at least 100 subjects randomized to 70 mg of AMG 334 have completed 52 weeks of treatment with investigational product. The purpose of this interim analysis is to evaluate the long-term efficacy and safety of AMG 334 in subjects with episodic migraine after the 24-week DBTP.

9. Data Screening and Acceptance

9.1 General Principles

The objective of the data screening is to assess the quantity, quality and statistical characteristics of the data relative to the requirements of the planned analyses.

9.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. This study will use the RAVE database. The database will be subjected to edit check outlined in the Data Management Plan (DMP). eDiary data, PK, antibody, biomarkers, and ECG data are outside of RAVE database. All the datasets to be used for planned analyses will be received from GSO-DM department. Additional details will be provided in the DMP and Data Transfer Plan (DTP).

9.3 Handling of Missing and Incomplete Data

Subjects may miss specific data points for a variety of causes. In general, data could be missing due to a subject's early withdrawal from study, a missed visit, or inability to evaluate an endpoint at a particular point in time. For this study, most of the efficacy endpoint will be collected via eDiary and subjects could miss entering several days of data in each monthly interval. The general procedures outlined below describe what will be performed when a data point is missing.

Missing eDiary data in the calculation of monthly measurements about subjects' migraine and non-migraine headaches will be handled using the following method.

1. For monthly intervals with ≥ 14 days of eDiary days (including retrospective eDiary days) in each interval**:**
    a) Monthly frequency measurements (including migraine days, headache days, migraine attacks, cumulative hours of migraine headaches, acute medication use, bed days due to migraine and missed work/school days due to migraine) will be prorated to 28-day equivalents. Prorated result does not need to be rounded.
    b) Monthly average severity of migraine pain, migraine related symptoms and monthly average scale of migraine interference with daily activity will calculated as the average of observed scores
2. For monthly intervals with < 14 days of eDiary use (including retrospective eDiary days), all monthly measurement will be set as missing and will be handled as described in Section 9.3.3

Missing eDiary data in the calculation of cumulative average monthly measurements about subjects' migraine and non-migraine headaches is treated as if subjects do not have headache or medication use to report.

Missing PROs (HIT-6) scheduled to be collected at office visit at certain assessment will not be imputed.

Missing safety endpoints will not be imputed. Missing day portion of AE start time will be imputed based on Section 9.3.1.

Missing pharmacogenetic and antibody data will not be imputed.

Handling of missing or incomplete data for exposure-response analysis will be described in the E-R SSAP or associated documents to support population PK/PD dataset generation and E-R analysis.

9.3.1 Missing and Incomplete Dates

Missing or incomplete dates will be listed as it is in any listings.

Incomplete start date of an adverse event or concomitant medication taken will be handled by following rule:

	Missing	Imputation	Exception
Start date (AE, concomitant medication)	Day	01	Default to Study Day 1 if an adverse event starts the same year and month as Study Day 1 and the flag indicates that the adverse event started on or after the first dose on the Adverse Events eCRF
	Day/Month	01JAN	Default to Study Day 1 if an event started the same year as Study Day 1 and the flag indicates that the adverse event started on or after the first dose on the Adverse Events eCRF
	Day/Month/Year	No imputation	

9.3.2 Missing Baseline Evaluation

Baseline values are defined in Section 6.3 “Baseline Assessment of the Study”. Missing baseline evaluations will not be imputed.

All subjects included in the efficacy analysis set will have baseline monthly rate or monthly average of migraine and non-migraine headaches related measurements after applying proration rule defined in Section 9.3 since only subject with ≥ 80% compliance of eDiary use during baseline will be eligible for randomization.

9.3.3 Missing Post-baseline Evaluation in Double-Blind Treatment Phase

Primary analysis of efficacy endpoints during the 24-week randomized DBTP will be conducted using the repeated measures linear mixed effects model on observed data without imputation.

For the descriptive summary of mean monthly value calculated using the monthly value from each of the last three months (months 4, 5, and 6) of the DBTP, if a subject has at least one month's value in months 4, 5, and 6, then the subject contributes to the summary statistics.

In the sensitivity analysis on primary and secondary efficacy endpoints during the 24-week DBTP, missing continuous efficacy endpoints will be handled using last observation carried forward (LOCF) method, inverse probability weighting (IPW) generalized estimation equations (GEE) method, and multiple imputation (MI) with assumption of missing at random (MAR), respectively.

In LOCF method, post-baseline missing continuous efficacy endpoints during DBTP will be imputed using the last observed value including baseline value. For example, if subject has all of the post-baseline values as missing, then all of the post-baseline values will be imputed using the observed baseline value.

In non-responder imputation (NRI) method, post-baseline missing dichotomous secondary efficacy endpoint (responder (Yes/No) based on ≥50% reduction from baseline in monthly migraine days) during DBTP will be imputed as non-responder at each corresponding time point.

IPW GEE method will be used to handle monotone missing data during DBTP. Intermittent missing data during DBTP will be handled by LOCF or NRI method before applying IPW GEE method.

If the proportion of missing data in primary endpoint is high (eg, > 20% for primary analysis at week 24), further analysis will be performed to

- examine the frequency and reason of missing data
- determine if there are any patterns in the missing data
- distinguish true missing values from other unknown values (eg, due to measurement or sample processing error)

9.4 Detection of Bias

This study has been designed to minimize potential bias by allocating treatment groups randomly, assessing endpoints and handling withdrawals without knowledge of the treatment. Other factors that may bias the results of the study include:

- important protocol deviations likely to impact the analysis and interpretation of the efficacy endpoints
- inadvertent breaking of the blind before formal unblinding
- investigational product dosing non-compliance
- the timing of and reasons for early withdrawal from treatment and from study

The incidence of these factors may be assessed. Important protocol deviations will be listed and/or tabulated in the clinical study report (CSR). If necessary, the incidence of other factors will be tabulated.

Any breaking of the blind for individual subjects prior to formal unblinding of the study will be documented in the CSR.

The timing of and reasons for early withdrawal from treatment and from study will be tabulated and/or listed.

9.5 Outliers

Histograms will be examined to identify outliers in any of the continuous variables used in the analyses. Unexpected and/or unexplained values in categorical data will be identified by utilizing frequency tables.

Outliers due to data entry errors will be corrected by the study team before final database lock. The validity of any questionable values or outliers will be confirmed. Outliers or any questionable values with confirmed validity will be included in the analyses. However, ad-hoc sensitivity analyses may be conducted to evaluate the influence of extreme values in the data.

9.6 Distributional Characteristics

Continuous endpoints of change from baseline value will be analyzed under normality assumption. If they deviate appreciably from normality, appropriate transformations or the non-parametric alternatives such as Quade test (Quade D, 1966) may additionally be considered. For repeated measure analysis, GEE model which is less sensitive to normality violation may be used in addition to the mixed effect model.

9.7 Validation of Statistical Analyses

Programs will be developed and maintained, and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software, for example the SAS System and S-plus.

For the exposure–response analysis, refer to the E-R SSAP for the software used.

10. Statistical Methods of Analysis

10.1 General Principles

The primary objective of this study is to evaluate the effect of AMG 334 compared to placebo on the change from baseline in mean monthly migraine days, in subjects with episodic migraine.

The primary analysis will be performed when the last randomized subject completes the week 24 assessment or is discontinued from the study. The final analysis for the study including DBTP, OLTP and safety follow-up phase will be performed at the end of the study.

Subjects will be analyzed based on their randomized treatment group assignment.

Summary descriptive statistics by each treatment group will be tabulated at each visit. For continuous endpoints, the descriptive statistics include: number of observations, mean, median, standard deviation, first and third quartiles, minimum and maximum. For categorical and ordinal endpoints, frequency and percentage will be given.

Change from baseline for efficacy endpoints will be summarized using both the study baseline and the pre-OLTP baseline for efficacy endpoint, respectively.

Primary analyses for efficacy endpoints are based on a linear mixed effects model including appropriate terms and covariates in the model (See Section 4.2). Nominal p-values will be provided for the comparisons between each AMG 334 treatment group vs. the placebo group for efficacy endpoints. To maintain a family-wise type I error at 0.05, the pair-wise comparison will be tested in a sequential testing procedure in the order of AMG 334 140 mg vs. placebo, 70 mg vs. placebo and 28 mg vs. placebo. The lower dose group will be tested only when the higher dose group is considered statistically significant. For continuous efficacy endpoints, the adjusted mean change from baseline

for each treatment group, and the adjusted treatment difference compared to placebo, associated 95% confidence intervals, and p-values for pairwise comparison will be reported. For dichotomous efficacy endpoints, adjusted odds ratios compared to placebo, associated 95% confidence intervals, and p-values will be reported. Randomized stratum (stratification group) will be included in the statistical models for the efficacy endpoints. However, the actual data collected will be used for the analysis of baseline characteristics and subgroup analyses.

10.2 Subject Accountability

For the primary analysis at week 24, the disposition of all enrolled subjects will be tabulated by the randomized treatment group. The summary will include the number of subjects who are randomized, the number and percent of subjects who receive the double-blind IP, who complete double-blind IP, discontinue double-blind IP and reasons for discontinuing, who complete the 24-week DBTP, and who withdraw prematurely from the study before completion of the 24-week DBTP and their reasons for withdrawal.

For the final analysis, disposition of the OLTP and safety follow-up phases will be added, which include the number and percent of subjects who enter the OLTP, who receive AMG 334, who complete AMG 334, discontinue AMG 334 and reasons for discontinuing, who complete the OLTP, who complete the study, and who withdraw prematurely from the study and their reasons for withdrawal.

A footnote on the subject disposition tables will include the number of subjects screened, date first subject enrolled, date last subject enrolled, the date of study primary completion, the date of study completion (eg, last subject completed the safety follow-up visit or drop out from the study) and corresponding data cutoff date.

10.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's visit and updated during the IPD reviews throughout the study prior to database lock. These definitions of IPD categories, sub-category codes and descriptions will be used during the course of the study. Eligibility deviations are defined in the protocol. IPDs will be summarized by randomized treatment group and listed for DBTP and OLTP, respectively.

10.4 Demographic and Baseline Characteristics

Subject demographic and baseline characteristics will be summarized using descriptive statistics by randomized treatment group and overall study population using FAS. If

multiple races have been reported for a subject, the subject will be categorized as multiple race.

The following demographic and baseline characteristics will be summarized:

- Age
- Sex
- Ethnicity
- Race
- Height (cm)
- Weight (kg)
- Body Mass Index (BMI, kg/m^2)
- Targeted neurological disease diagnosis at baseline
- Disease duration of migraine with or without aura
- Age at onset of migraine
- Migraine specific-medications subjects used during baseline phase:
    a. Triptan-based
    b. Ergotamine-based
- Monthly acute migraine-specific medication use in days during baseline phase
- Prior/current treatment with migraine prophylactic medication (current migraine prophylactic medication treatment vs prior migraine prophylactic medication treatment only vs no prior or current migraine prophylactic medication treatment)
- Monthly migraine days during baseline phase
- Monthly migraine attacks during baseline phase
- Monthly headache days during baseline phase
- Prior prophylactic treatment failure (Yes vs. No)

10.5 Efficacy Analyses

Primary analysis of efficacy endpoints will utilize the efficacy analysis set. Subjects will be analyzed according to their randomized treatment group regardless of the actual

treatment received during the study. Per protocol analysis of the primary and secondary endpoints will utilize the Per Protocol Set.

For primary analysis at week 24, the continuous change from baseline efficacy endpoints as specified in Section 9.3 will be analyzed using linear mixed effect models adjusted by stratification factor and baseline value on observed data as the primary analysis method. The dichotomous efficacy endpoints will be analyzed using the stratified Cochran-Mantel-Haenszel (CMH) test after the missing data are imputed as non-response.

Detailed primary analysis methods, sensitivity analyses, and covariates included in the models are summarized in the table below.

Table 1. Summary of Efficacy Endpoints and Analysis Methods

Endpoint	Primary Summary and Analysis Method	Sensitivity Analysis
Primary Endpoint		
Change from baseline in monthly migraine days (*Note: Results from time points during DB treatment phase will also be generated in the same model.*)	1. Summary statistics by visit using observed data, and using mean monthly migraine days at month 4,5,6 2. Least squares mean at each timepoint from a linear mixed effect model adjusted by stratification variable and baseline value using observed data 3. Test pairwise treatment difference between active and placebo for the primary endpoint – change from baseline in mean monthly migraine days, using a contrast from the model in #2.	1. LOCF: Summary statistics by visit and analyze using an ANCOVA model for the mean monthly migraine days. 2. Per-Protocol subset: Same as primary summary and analysis method. 3. IPW GEE model for change from baseline in monthly migraine days 4. MI with assumption of MAR
Secondary Endpoints		
Response defined as at least a 50% reduction from baseline in mean monthly migraine days (*Note: Results*	1. Summary statistics by visit using observed data, and responder rate calculated using mean monthly migraine days at month 4,5,6 2. A stratified Cochran-Mantel-Haenszel (CMH)	1. Without imputation, adjusted odds ratios from a generalized linear mixed model adjusted by stratification variable and baseline migraine days by visit using observed data, and using responder rate calculated using mean monthly migraine days at month 4,5,6 2. NRI: Summary statistics by visit and

Endpoint	Primary Summary and Analysis Method	Sensitivity Analysis
from time points during DB treatment phase will also be generated in the same model.)	test will be used after the missing data are imputed as non-response	analyze using a logistic regression model, and using responder rate calculated using mean monthly migraine days at month 4,5,6 3. Per-Protocol subset: Same as primary summary and analysis method.

Endpoint	Primary Summary and Analysis Method	Sensitivity Analysis
Change from baseline in monthly acute migraine-specific medication treatment days (*Note: Results from time points during DB treatment phase will also be generated in the same model.*)	1. Summary statistics by visit using observed data and using mean monthly acute migraine-specific medication treatment days 2. Least squares mean at each timepoint from a linear mixed effect model adjusted by stratification variable and baseline value using observed data 3. Test pairwise treatment difference between active and placebo for the endpoint of change from baseline in mean monthly acute migraine-specific medication treatment days, using a contrast from the model in #2.	1. LOCF: Summary statistics by visit and analyze using an ANCOVA model. 2. Per-Protocol subset: Same as primary summary and analysis method. 3. IPW GEE model for change from baseline in monthly acute migraine-specific medication treatment days 4. MI with assumption of MAR
Exploratory Endpoints		
Change from baseline in monthly headache (migraine and non-migraine headache) days at assessment time points [a]	1. Summary statistics by visit using observed data and mean monthly headache 2. Least squares mean at each timepoint from a linear mixed effect model adjusted by stratification variable and baseline value using observed data 3. Test pairwise treatment difference between active and placebo for the corresponding endpoint of change from baseline in mean monthly values, using a contrast from the model in #2.	
Response defined as at least a 75% reduction from baseline in mean monthly	1. Summary statistics by visit using observed data and responder rate calculated using mean monthly migraine days at month 4,5,6	

Endpoint	Primary Summary and Analysis Method	Sensitivity Analysis
migraine days (*Note: Results from time points during DB treatment phase will also be generated in the same model.*) [b]	2. Adjusted odds ratios from a generalized linear mixed model adjusted by stratification variable and baseline migraine days by visit using observed data, and using responder rate calculated using mean monthly migraine days at month 4,5,6	

a: The same analyses methods will be applied to the endpoints:

Change from baseline in mean headache impact scores as measured by HIT-6 (HIT-6 total score),

Change from baseline in monthly migraine attack,

Change from baseline in monthly cumulative migraine hours,

Change from baseline in monthly average severity of migraine pain,

Change from baseline in monthly migraine days with severe pain,

Change from baseline in monthly cumulative hours of migraine with severe pain,

Change from baseline in monthly acute medication use in days,

Change from baseline in monthly average migraine symptoms interference scale,

Change from baseline in monthly bed days due to migraine related symptoms,

Change from baseline in monthly missed work/school days due to migraine related symptoms.

b: The same analyses methods will be applied to the endpoints:

Response defined as at least a 100% reduction from baseline in mean monthly migraine days.

10.5.1 Analyses of Primary and Secondary Efficacy Endpoints

For the primary analyses at the end of DBTP, the continuous primary and secondary endpoints will be tested using a linear mixed model based on observed monthly data from 24-week DBTP with appropriate contrasts provided in Appendix B for pairwise comparisons for the corresponding efficacy endpoints: change from baseline in mean monthly values.

The model will include treatment, visit, treatment by visit interaction, stratification variable and baseline value as covariates. If applicable, the first-order autoregressive covariance structure is assumed. Least squares means (LSMs) for each treatment group, standard errors, associated 95% confidence intervals, difference of LSMs compared to placebo group, associated 95% confidence intervals and nominal two-sided p-values will be tabulated by visit and treatment, as well as for the mean monthly values over the last 3 months in DBTP.

For binary endpoints, adjusted odds ratios compared to placebo group, associated 95% confidence intervals and nominal two-sided p-values will be tabulated by visit and treatment.

Sensitivity analysis described below will be performed for the primary and secondary endpoints:

1. Summary statistics and ANCOVA model for continuous endpoints (logistic regression model for binary endpoints) by visit using imputed data by LOCF (NRI for binary endpoint). Factor of treatment, baseline covariate and stratification variable will be included in the model.
2. Inverse probability weighted generalized estimation equation (IPW GEE) model on the primary and secondary endpoints.
3. MI with assumption of MAR
4. Primary summary and analysis method based on per protocol analysis set.
5. Primary summary and analysis by the subgroup of
    - Prior/current treatment with migraine prophylactic medication (current migraine prophylactic medication treatment vs prior migraine prophylactic medication treatment only vs no prior or current migraine prophylactic medication treatment)
    - BMI (< median vs. ≥ median)
    - Baseline monthly migraine days (< 8 vs ≥ 8)
    - Treatment failure of prior migraine prophylactic medications
6. Primary summary and analysis method with interaction: If the inclusion criterion at the 0.15 level is met, the interaction of treatment group by stratification variable will be included in the model as sensitivity analysis for the primary analysis method.

The purpose of the subgroup analyses is to explore if the treatment effect varies across subgroups of interest. Subgroup analyses are performed for primary and secondary efficacy endpoints using the same method as Primary Summary and Analysis Method but performed within each described subgroup.

10.5.2 Analyses of Exploratory Efficacy Endpoints

For exploratory efficacy endpoints as specified in Section 9.3, summary statistics and primary analysis method will be conduct in the same way as that for the primary endpoint.

The exploratory efficacy endpoints at each assessment time will be analyzed using the repeated measures linear mixed effects model (a generalized linear mixed model for dichotomized variables) that includes treatment group, baseline values, stratification factor, scheduled visit, and the interaction of treatment and scheduled visit without any imputation for missing data. The least squares mean (or odds ratio) of treatment group and placebo with associated 95% confidence intervals and p-values will be reported.

10.5.3 Analyses of Efficacy Endpoints in Open-Label Treatment Phase

For the OLTP of the study, descriptive summaries of efficacy endpoints will be tabulated by randomized treatment group and visit based on observed data without imputation.

10.5.4 Analyses of Patient Reported Outcomes (PROs)

PROs include migraine symptom interference items and HIT-6. Change from baseline in total score (or subscale if applicable) of each PRO will be analyzed similarly as the primary analysis of primary and secondary efficacy endpoints described in the Section 10.5.1 above during DBTP**.**

For HIT-6, in addition to the analysis of continuous change from baseline value, proportions of subject with a ≥ 5 point reduction from baseline and proportions of subject with HIT-6 ≥ 60 (severe impact) will be analyzed similarly as the primary analysis method for dichotomous efficacy endpoints described in the in the Section 10.5.1 above.

No sensitivity analysis will be conducted for PROs.

For the OLTP of the study, descriptive summaries of PRO endpoints will be tabulated by randomized treatment group and visit based on observed data without imputation.

10.5.5 Pharmacokinetic Endpoints

The pharmacokinetic concentration of all subjects will be summarized with descriptive statistics by treatment groups and visits using PK analyses set.

For the population PK analysis, please refer to the separate population PK/PD analysis plan for details.

10.6 Safety Analyses

For safety endpoints, all randomized subjects who received at least one dose of investigational product (ie, Safety Analysis Set) will be analyzed based on the randomized treatment unless a subject has received the incorrect dose the entire period of interest (phase or study).

No statistical testing comparing treatment groups will be performed in the safety analyses.

10.6.1 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 18 or later will be used to code all adverse events (AE) to a system organ class and a preferred term. All adverse events will be graded using the Common Terminology Criteria for Adverse Events (CTCAE) Version 4. All adverse event tables will be summarized by treatment group. Refer to Section 6.1.2 for the definition of treatment-emergent adverse event.

The subject incidence of AEs will be summarized for all treatment-emergent AEs (TEAEs), serious AEs, AEs leading to withdrawal of investigational product, fatal AEs, and device-related AEs for DBTP and OLTP, respectively.

Subject incidence of all treatment-emergent AEs, serious AEs, AEs leading to withdrawal of investigational product, serious AEs leading to withdrawal of investigational product, fatal AEs and device-related AEs will be tabulated by system organ class in alphabetical order and preferred term in descending order of frequency, for DBTP and OLTP, respectively.

In addition, subject incidence of all treatment-emergent AEs and serious AEs will be tabulated by system organ class in alphabetical order, preferred term and CTCAE grade in descending order of frequency, for DBTP and OLTP, respectively. Subject incidence of all treatment-emergent AEs and serious AEs will also be tabulated by preferred term in descending order of frequency.

Treatment-related treatment-emergent AEs will be summarized for OLTP by system organ class, preferred term and CTCAE grade.

In addition, exposure-adjusted incidence rate will be produced for DBTP and OLTP, respectively. All treatment-emergent AEs and serious AEs will be tabulated by preferred term in descending order of frequency.

10.6.2 Laboratory Test Results

Shifts tables of the laboratory toxicity for absolute neutrophil count (ANC) based on CTCAE grade relative to baseline will be tabulated by treatment group for each study phase.

Summary of change from baseline for ANC, alanine transaminase (ALT) and aspartate aminotransferase (AST) will also be provided by visit for each study phase.

Subject incidence of liver function test abnormalities (including AST, ALT, Total Bilirubin (TBL) and Alkaline Phosphatase (ALP)) will also be summarized by treatment group and study phase.

10.6.3 Vital Signs

The analyses of vital signs (systolic/diastolic blood pressure and weight) will include summary statistics of change from baseline over time by treatment group.

Systolic blood pressure and diastolic blood pressure may be analyzed by change from baseline in categories: 10-20 mmHg and ≥ 20 mmHg and analyzed for SBP >140 mmHg (Yes vs. No) and DBP > 90 mmHg (Yes vs. No) at each time point.

10.6.4 Columbia-Suicide Severity Rating Scale (C-SSRS)

No statistical testing will be performed on C-SSRS. The number and percentage of subjects reporting any suicidal ideation and any suicidal behavior will be summarized descriptively by treatment group separately for the DBTP and OLTP. Shift table of C-SSRS maximum severity of suicidal ideation/behavior compared to baseline will be provided by treatment group separately for the DBTP and OLTP.

10.6.5 Electrocardiogram (ECG)

The ECG measurements from this clinical study were performed as per standard of care for routine safety monitoring, rather than for purposes of assessment of potential QTc effect. Since these evaluations may not necessarily be performed under the rigorous conditions expected in a thorough QT study, neither summaries nor statistical analyses for QTc will be provided, and these data are not expected to be useful for meta-analysis with data from other trials.

Subject incidence of ECG diagnosis will be summarized by treatment group separately for the DBTP and OLTP.

10.6.6 Antibody Formation

The number and percentage of subjects who develop anti-AMG 334 antibodies (binding and, if positive, neutralizing) will be tabulated by treatment group for the entire study. The list of subjects with positive antibodies at any time will be provided.

In addition, a listing of treatment-emergent AEs will be provided for subjects with positive binding or neutralizing antibodies.

10.6.7 Exposure to Investigational Product

Descriptive statistics will be produced to describe the exposure to investigational product by treatment group and by phases. The number and percentage of subjects with dose change, reason for dose change and duration of exposure to investigational product in days will be summarized by treatment group.

10.6.8 Summary Concomitant Medication Use

The number and proportion of subjects receiving headache-related medications will be summarized by acute medication category for each treatment group.

10.7 Pharmacokinetic Analysis

All PK-related tables, figures, listings and other deliverables will be generated by Clinical Pharmacology Modeling and Simulation (CPMS).

10.7.1 Exposure-Response Analysis

Exposure-Response (E-R) analysis will be performed. Please refer to the E-R SSAP for additional details.

11. Changes from Protocol-specified Analyses

N/A

12. Literature Citations / References

- Fitzmaurice GM, Laird NM and Ware JH. (2004). Applied Longitudinal Analysis. New York: John Wiley and Sons.
- Breslow, N. E. and Clayton, D. G. (1993), Approximate Inference in Generalized Linear Mixed Models, Journal of the American Statistical Association, 88, 9–25.
- Wolfinger, R. and O'Connell, M. (1993), Generalized Linear Mixed Models: A Pseudo Likelihood Approach, Journal of Statistical Computation and Simulation, 4, 233–243.
- Rubin, D. B. (1987), Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc.
- Heinze, G. and Schemper, M. (2002), A Solution to the Problem of Separation in Logistic Regression, Statistics in Medicine, 2002; 21:2409–2419.
- Quade, D. (1966), Rank analysis of covariance, University of North Carolina, Institute of Statistics Mimeo Series. No. 483
- Gould AL, Shih WJ. Sample size re-estimation without unblinding for normally distributed outcomes with unknown variance. *Commun Stats-theory Meth. 1992;21(10):2833-2853.*
- Fairclough DL1, Cella DF. Functional Assessment of Cancer Therapy (FACT-G): non-response to individual questions. Qual Life Res. 1996 Jun;5(3):321-9.

13. Prioritization of Analyses

The tables, listings and figures for Flash Memo of primary analysis at week 24 will be prioritized.

14. Data not Covered by This Plan

There are no plans to specifically analyze or summarize the following data points.

- ECG interval data
- PK
- CCI [redacted]
- CCI [redacted]

15. Appendices

Appendix A. Technical Detail and Supplemental Information Regarding Statistical Procedures and Programs

A1 Generalized Linear Mixed Model

Generalized mixed effects repeated measure model fits data with correlations or nonconstant variability and where the response is not necessarily normally distributed, which is also known as generalized linear mixed models (GLMM). The correlations can arise from repeated observation of the same sampling units, shared random effects in an experimental design, spatial (temporal) proximity, multivariate observations, and so on.

GLMMs, like linear mixed models, assume normal (Gaussian) random effects. Conditional on these random effects, data can have any distribution in the exponential family. The exponential family comprises many of the elementary discrete and continuous distributions. The binary, binomial, Poisson, and negative binomial distributions, for example, are discrete members of this family. The normal, beta, gamma, and chi-square distributions are representatives of the continuous distributions in this family.

Suppose Y represents the $(n \times 1)$ vector of observed data and γ is a $(r \times 1)$ vector of random effects. GLMM Models assume that

$$E[Y|\gamma] = g^{-1}(X\beta + Z\gamma)$$

where $g(.)$ is a differentiable monotonic link function and $g^{-1}(.)$ is its inverse. The matrix X is an $(n \times p)$ matrix of rank k, and Z is an $(n \times r)$ design matrix for the random effects. The random effects are assumed to be normally distributed with mean 0 and variance matrix G.

The GLMM contains a linear mixed model inside the inverse link function. This model component is referred to as the linear predictor,

$$\eta = X\beta + Z\gamma$$

The variance of the observations, conditional on the random effects, is

$$Var[Y|\gamma] = A^{1/2}RA^{1/2}$$

The matrix A is a diagonal matrix and contains the variance functions of the model. The variance function expresses the variance of a response as a function of the mean.

The GLIMMIX procedure in SAS fits Generalized linear mixed model. Sample SAS code is provided in Appendix B.

The GLIMMIX procedure distinguishes two types of random effects. Depending on whether the parameters of the covariance structure for random components in your model are contained in G or in R, the procedure distinguishes between “G-side” and “R-side” random effects. The associated covariance structures of G and R are similarly termed the G-side and R-side

covariance structure, respectively. R-side effects are also called "residual" effects. Simply put, if a random effect is an element of γ , it is a G-side effect and you are modeling the G-side covariance structure; otherwise, you are modeling the R-side covariance structure of the model. Models without G-side effects are also known as marginal (or population-averaged) models. Models fit with the GLIMMIX procedure can have none, one, or more of each type of effect.

For a model containing random effects, the GLIMMIX procedure, by default, estimates the parameters by applying pseudo-likelihood techniques as in Wolfinger and O'Connell (1993) and Breslow and Clayton (1993). You can also fit generalized linear mixed models by maximum likelihood where the marginal distribution is numerically approximated by the Laplace method (METHOD=LAPLACE) or by adaptive Gaussian quadrature (METHOD=QUAD).

Once the parameters have been estimated, you can perform statistical inferences for the fixed effects and covariance parameters of the model. Tests of hypotheses for the fixed effects are based on Wald type tests and the estimated variance-covariance matrix.

A2 Marginal Model and Inverse Probability Weighted Generalized Estimating Equation (IPW GEE)

Marginal model (Fitzmaurice et.al. 2004) refers to a method for extending generalized linear models to longitudinal data. The term "*marginal*" in this context indicates that the model for mean responses depends only on the covariates of interest, and not on any random effects or previous responses. That is, the term marginal is used to emphasize that the model for the mean response at each occasion does not incorporate dependence on any random effects or previous responses. This is in contrast to *mixed effect models*, where the mean response depends not only on covariates but also on a vector of random effects. Marginal models do not require distributional assumptions for the observations, only a regression model for the mean response. That is, marginal models provide a unified method for analyzing diverse types of longitudinal responses, which avoids making assumptions about the distribution of the vector of responses; the method replies solely on assumptions about the mean responses. Marginal models are primarily used to make inferences about population means. As a result, marginal models for longitudinal data separately model the mean response and within-subject association among the repeated responses. In a marginal model, the goal is to make inferences about the former, whereas the latter is regarded as a nuisance characteristic of the data that must be accounted for to make correct inferences about the changes in the population mean responses.

Assume that there are n_i repeated measurements of the response on the i^{th} subject and that each Y_{ij} is observed at time t_{ij}, with associated covaraites X_{ij}, which can be time-stationary or time-varying. A marginal model for longitudinal data has the following three-part specification:

1. The conditional expectation or mean of each response, $E\left(Y_{ij}\middle|X_{ij}\right) = \mu_{ij}$, is assumed to depend on the covariates through a known link function

$$g(\mu_{ij}) = \eta_{ij} = X_{ij}^{'}\beta.$$

2. The conditional variance of Y_{ij}, given the covariates, is assumed to depend on the mean according to

$$\mathrm{Var}(Y_{ij}) = \phi v(\mu_{ij}),$$

where $v(\mu_{ij})$ is a known "variance function" (ie a known function of the mean, μ_{ij}) and ϕ is a scale parameter that may be known or may need to be estimated.

3. The conditional within-subject association among the vector of repeated responses, given the covariates, is assumed to be a function of an additional set of association parameters, α (and also depends upon the means, μ_{ij}). For example, the components of α might represent the pairwise correlations or log odds ratios among the repeated responses.

The avidance of distributional assumptions leads to a method of estimation known as *generalized estimating equations (GEE).* The GEE estimator of β for marginal models can be thought of as arising from minimizing the following objective function:

$$\sum_{i=1}^{N} \{\mathrm{y_i} - \mu_\mathrm{i}(\beta)\}^{'} \mathbf{V}_\mathrm{i}^{-1} \{\mathrm{y_i} - \mu_\mathrm{i}(\beta)\} \ ,$$

With respect to , where V_i is treated as known (by ignoring its dependence on β through μ_i) and μ_i is the vector of mean responses, with elements

$$\mu_{ij} = \mu_{ij}(\beta) = g^{-1}(\boldsymbol{X}_{ij}^{'}\beta)$$

Using calculus, it can be shown that if a minimum of the above function exists, it must solve the following *generalized estimating equations:*

$$\sum_{i=1}^{N} \mathrm{D}^{'} \mathbf{V}_\mathrm{i}^{-1} \{\mathrm{y_i} - \mu_\mathrm{i}\} = 0 \ ,$$

where $D_i = \partial\mu_i / \partial\beta$ is the derivative matrix and V_i is the so called working covariance matrix, which approximates the true underlying covariance matric for Y_i, that is, $V_i \cong \mathrm{Cov}(Y_i)$, recognizing that $V_i \neq \mathrm{Cov}(Y_i)$ unless the models for the variance and he within-subject associations are correct. Parameter estimates from the GEE are precise or efficient as the MLE and are consistent even when the variance structure is misspecified, under mild regularity conditions.

The solution to standard Generalized Estimating Equation yields consistent estimator of coefficients provided the data are MCAR (missing completely at random) or provided that missingness depends only on the covariates included in the model for the mean response. However, when dropout is MAR (missing at random), the standard GEE can yield badly biased estimates of coefficients. The inverse probability weighted GEE (IPW-GEE) approach was

developed to circumvent this specific problem. In the IPW-GEE the idea is to base the estimation on the observed responses but weight them to account for the probability of remaining on the study. Assume random variable D_i is recorded for all individuals and $D_i = k$ if an individual drops out between $(k-1)^{th}$ and k^{th} occasion, that is, only the first $D_i - 1$ responses are observed. Then we estimate $w_{ik} = \Pr(D_i = k+1)$, the probability that the i^{th} subject is still in the study at k^{th} occasion, then it is representated by

$$w_{ik} = (1-\pi_{i1}) \times (1-\pi_{i2}) \times \cdots \times (1-\pi_{ik})$$

where $\pi_{ik} = \Pr(D_i = k \mid D_i \geq k)$, which can be estimated from those still remaining at the $(k-1)^{th}$ occasion, given the record history of all available data up to $(k-1)^{th}$ occasion. Then the available data at the k^{th} occasion are weighted by $\frac{1}{w_{ik}}$ in the analysis. The intuition behind the inverse probability weighted methods is that each subject's contribution to the weighted complete case analysis is replicated $\frac{1}{w_{ik}}$ times, in order to count once for herself, and $\left(\frac{1}{w_{ik}} - 1\right)$ times for those who do not complete the study. In general, the weight methods are valid provided that the model that produces the estimated w_{ik} is correctly specified.

A3 Multiple Imputation (MI) and MCMC Method

The multiple imputation assume that the missing data are missing at random (MAR), that is, the probability that an observation is missing may depend on the observed values but not the missing values. It also assumes that the parameters *q* of the data model and the parameters *f* of the missing data indicators are distinct. That is, knowing the values of *q* does not provide any additional information about *f*, and vice versa. If both MAR and the distinctness assumptions are satisfied, the missing data mechanism is said to be ignorable. The MI procedure provides three methods for imputing missing values and the method of choice depends on the type of missing data pattern. For monotone missing data patterns, either a parametric regression method that assumes multivariate normality or a nonparametric method that uses propensity scores is appropriate. For an arbitrary missing data pattern, a Markov chain Monte Carlo (MCMC) method that assumes multivariate normality can be used.

In MCMC, one constructs a Markov chain long enough for the distribution of the elements to stabilize to a common, stationary distribution. By repeatedly simulating steps of the chain, it simulates draws from the distribution of interest.

In Bayesian inference, information about unknown parameters is expressed in the form of a posterior distribution. MCMC has been applied as a method for exploring posterior distributions in Bayesian inference. That is, through MCMC, one can simulate the entire joint distribution of the unknown quantities and obtain simulation-based estimates of posterior parameters that are of interest.

Assuming that the data are from a multivariate normal distribution, data augmentation is applied to Bayesian inference with missing data by repeating a series of imputation and posterior steps. These two steps are iterated long enough for the results to be reliable for a multiply imputed data set (Schafer 1997). The goal is to have the iterations converge to their stationary distribution and then to simulate an approximately independent draw of the missing values.

Sample SAS code for MI using MCMC method is provided in Appendix B.

A4 Quade Test for Continuous Response Variable with Covariates

When assumption of normality in analysis of covariance model is violated, nonparametric methods, eg rank test, can then be used. Quade (1966) proposed the use of rank analysis of covariance. This technique can be combined with the extended Mantel-Haenszel statistics to carry out nonparametric comparisons between treatment groups, after adjusting for the effects of one or more covariates.

Suppose that from each of *m* treatment groups, we have observations (Y_{ia}, X_{ia}), where Y_{ia} is the univariate response of the i^{th} observation in the a^{th} group $(1 \le i \le n_a, 1 \le a \le m)$, and X_{ia} is the corresponding value of covariates, possibly multivariate, whose marginal probability distribution is the same in each treatment group. Assume, in order to make ranking possible, that each variate has been measured on at least an ordinal scale; continuity is not required, however, and even a dichotomy is permitted as extreme case. So let the rank of Y_{ia} among all the $N = \sum n_a$ observed values of Y be $R_{ia} - (N+1)/2$, where the term $(N+1)/2$ has been inserted for the convenience so that $\sum\sum R_{ia} = 0$, thus correcting the ranks for their mean; use “average ranks” in case of ties, and (for definiteness) rank from the smallest first. Similarly, if X is actually a p-variate variable$(X^{(1)}, X^{(2)}, \dots, X^{(p)})$, let $C_{ia}^{(k)} - (N+1)/2$ be the rank of $X_{ia}^{(k)}$ among the N observed values of $X^{(k)}$. Then characterize the relationship between Y and X by performing an ordinary multiple linear regression of R on $C^{(1)}, C^{(2)}, \dots, C^{(p)}$; calculate fitted values $\hat{R}_{ia}$, and assign as scores the residuals from this regression of ranks: ie let

$$Z_{ia} = R_{ia} - \hat{R}_{ia}.$$

Finally, to test the hypothesis of identical conditional distributions of Y on X among treatment groups, use the following the variance ratio:

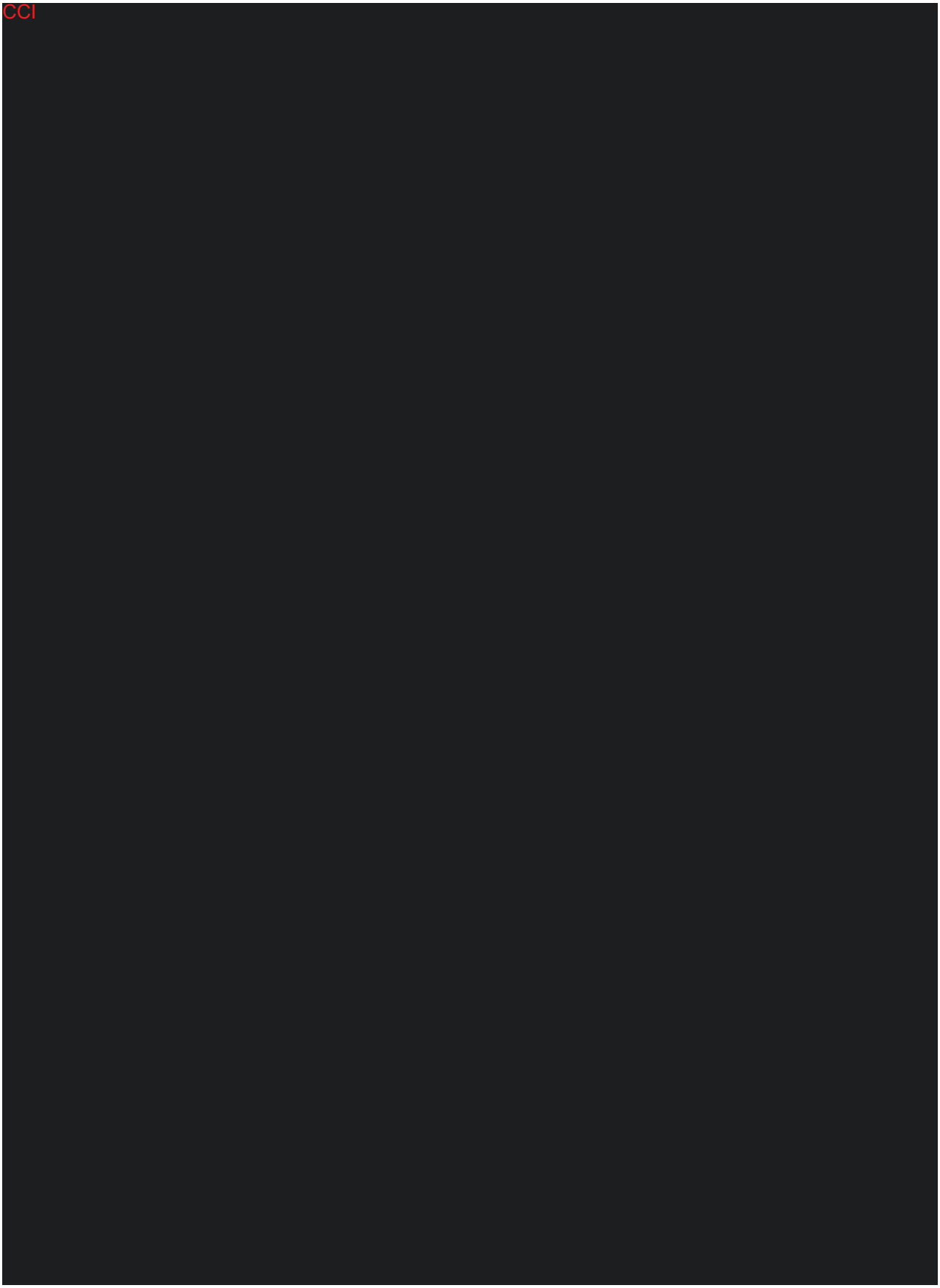
$$VR = \frac{(N-m)\sum_a (\sum_i Z_{ia})^2 / n_a}{(m-1)[\sum_a \sum_i Z_{ia}^2 - \sum_a (\sum_i Z_{ia})^2 / n_a]}$$

as comparing it with the critical value of an F with (m-1,N-m) degrees of freedom. (Note that no explicit correction for the mean is required in VR since $\sum_a \sum_i Z_{ia} = \bar{\bar{Z}} = 0$)

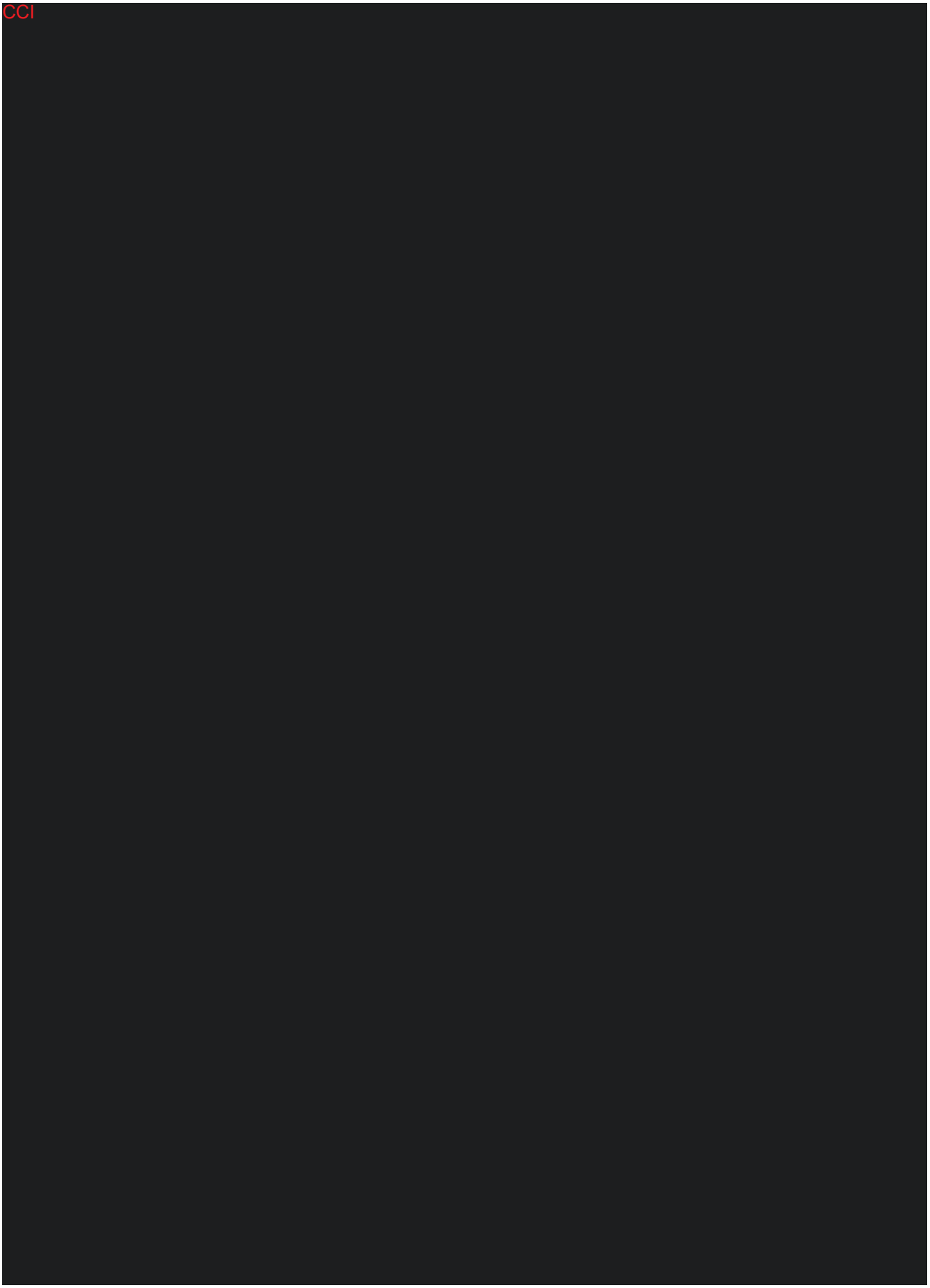
The methodology, which has also been described by Koch et al. (1982, 1990), can easily be implemented using the SAS system. First, ranks of response variable and covariate in the combined group of treatments are computed using PROC RANK. Then perform linear regression of response variable ranks on covariate ranks using PROC REG. The residuals are then used to compare treatment difference by the Mantel-Haenszel mean score statistic, which can be done in PROC FREQ using TABLE scores. See Appendix B for codes.

Appendix B. Code Fragments

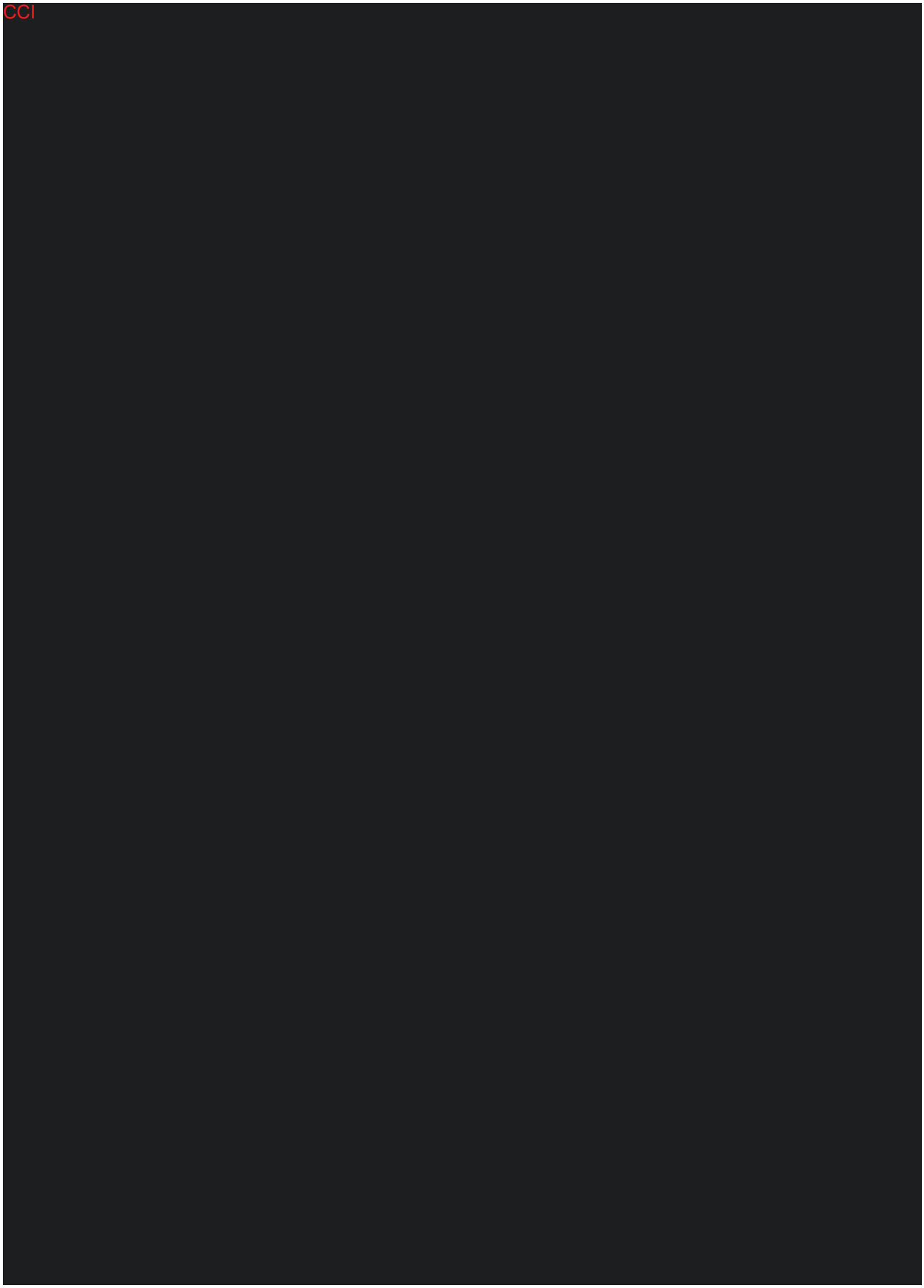
CCI



CCI



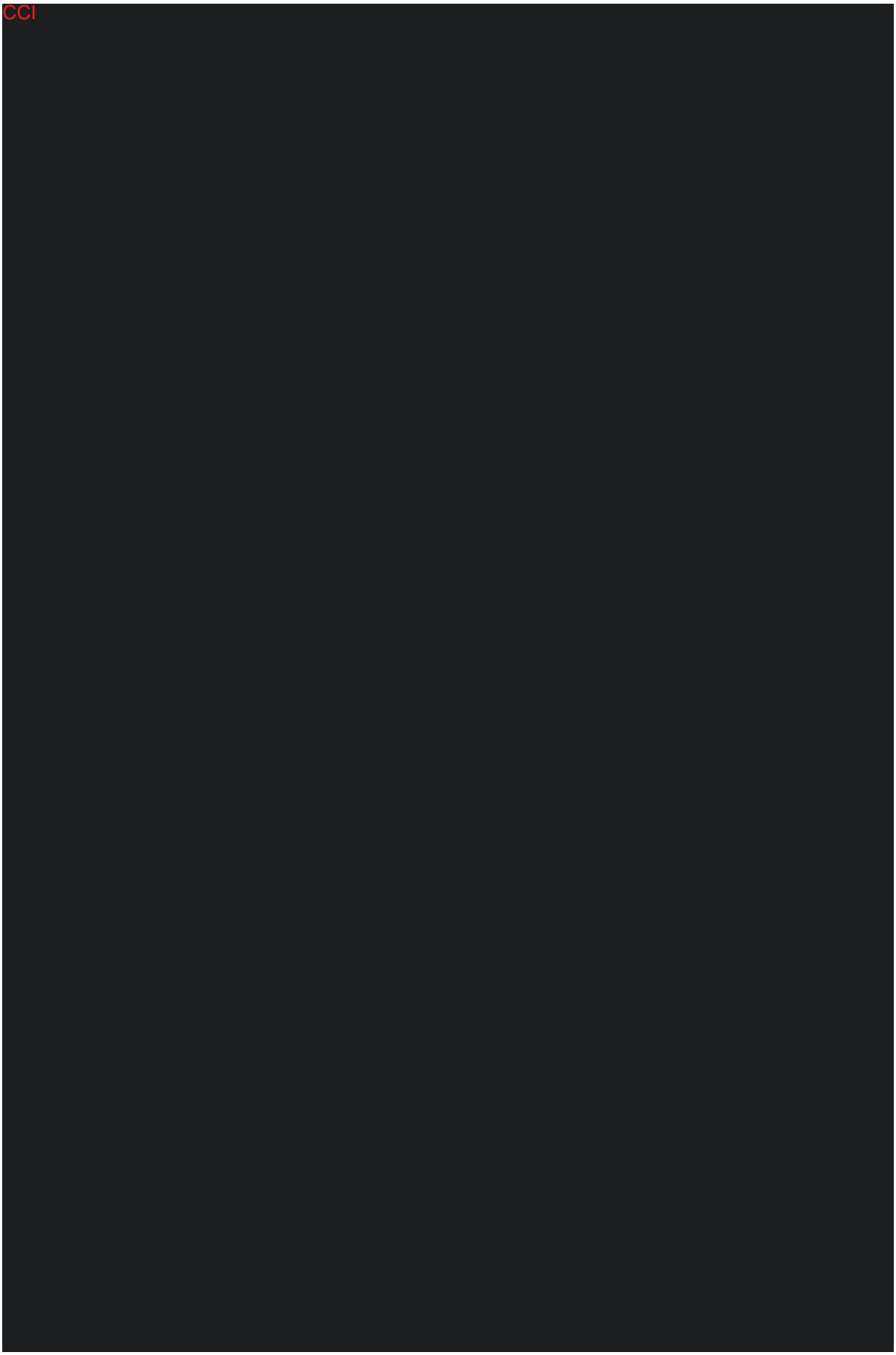
CCI



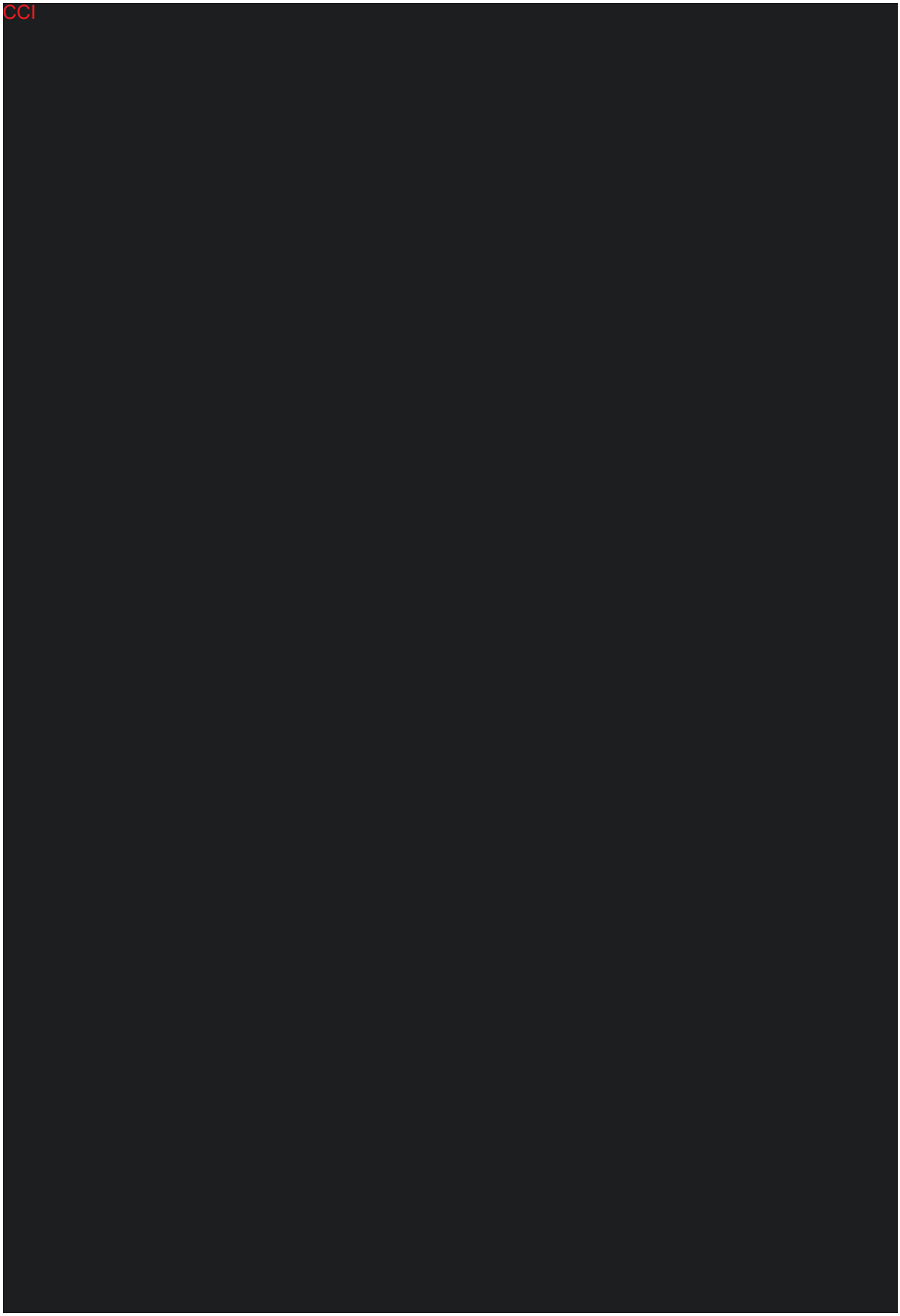
CCI

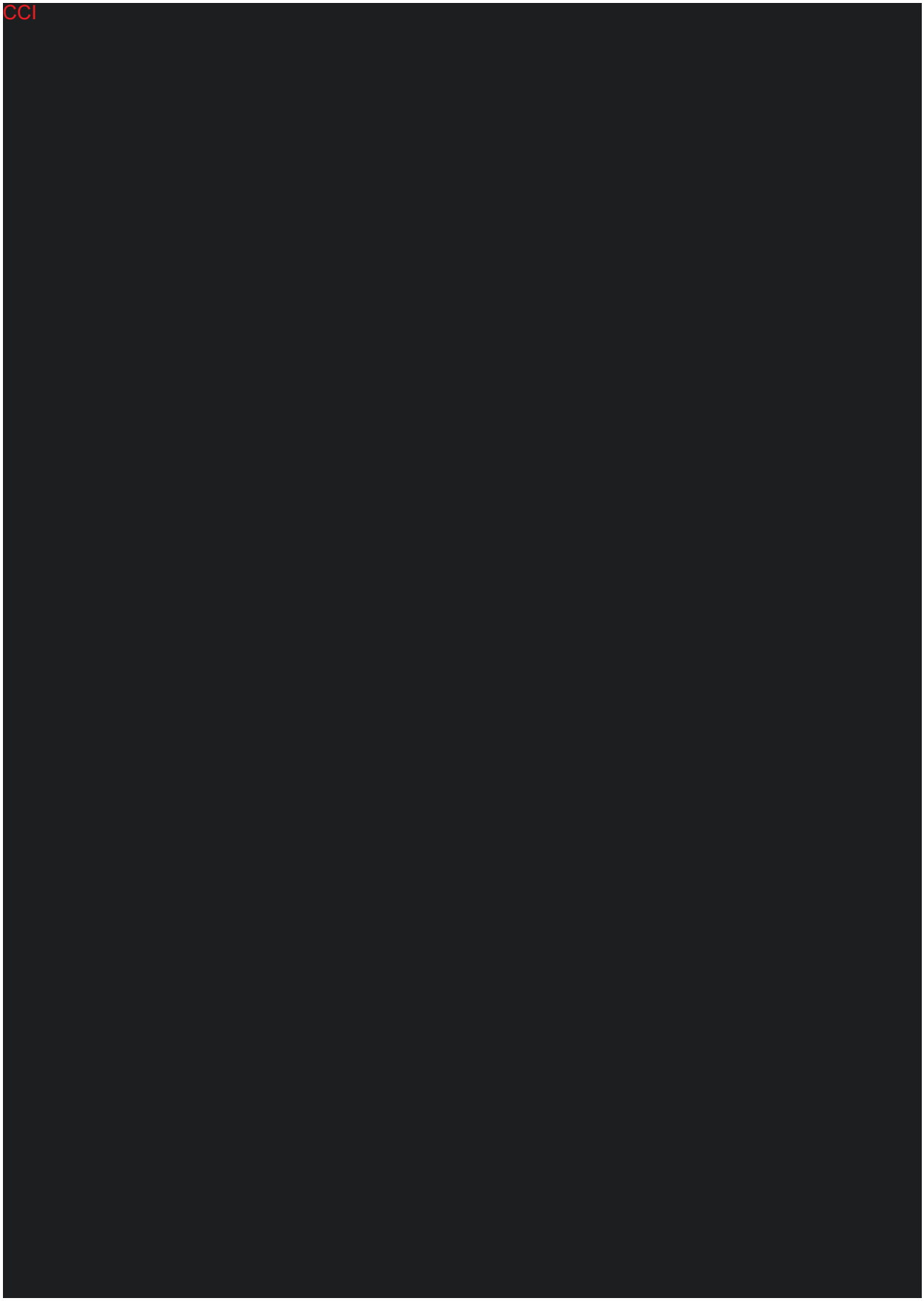
[redacted]

CCI



CCI




CCI

Appendix C. Reference Values/Toxicity Grades

Adverse event severity and laboratory toxicity are graded based on NCI Common Toxicity Criteria version 4 or higher, which is available at the following:

http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf

Appendix D. Patient-reported Outcome Forms/Instruments

D1. Headache Impact Test (HIT-6)

This questionnaire was designed to help you describe and communicate the way you feel and what you cannot do because of headaches.

To complete, please check one box for each question.

1. When you have headaches, how often is the pain severe?

☐ Never ☐ Rarely ☐ Sometimes ☐ Very Often ☐ Always

2. How often do headaches limit your ability to do usual daily activities including household work, work, school, or social activities?

☐ Never ☐ Rarely ☐ Sometimes ☐ Very Often ☐ Always

3. When you have a headache, how often do you wish you could lie down?

☐ Never ☐ Rarely ☐ Sometimes ☐ Very Often ☐ Always

4. In the past 4 weeks, how often have you felt too tired to do work or daily activities because of your headaches?

☐ Never ☐ Rarely ☐ Sometimes ☐ Very Often ☐ Always

5. In the past 4 weeks, how often have you felt fed up or irritated because of your headaches?

☐ Never ☐ Rarely ☐ Sometimes ☐ Very Often ☐ Always

6. In the past 4 weeks, how often did headaches limit your ability to concentrate on work or daily activities?

☐ Never ☐ Rarely ☐ Sometimes ☐ Very Often ☐ Always

COLUMN 1 (6 points each)	COLUMN 2 (8 points each)	COLUMN 3 (10 points each)	COLUMN 4 (11 points each)	COLUMN 5 (13 points each)

Scoring:

Scoring of HIT-6 total score is process by QualityMetrics Software with total score ranging from 36 to 78. HIT-6 scores are categorized into 4 grades, representing little or no impact (49 or less), some impact (50-55), substantial impact (56-59), and severe impact (60-78) due to headache.