STATISTICAL ANALYSIS PLAN

A Multicenter, Multiple-dose, Two-arm, Active-controlled, Double-blind, Double-dummy Study to Compare the Therapeutic Efficacy and Safety of Oral Doses of Cinacalcet HCl With Intravenous Doses of AMG 416 in Hemodialysis Subjects With Secondary Hyperparathyroidism

Protocol Number: AMG 416 20120360
Version: 1.1
Date: 20 October 2014
Authors:

Biostatistics:
Joanne Lin, Yan Sun

Global Health Economics:
Mark Bensink, Vasily Belozeroff
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Abbreviations</td>
<td>4</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>6</td>
</tr>
<tr>
<td>2. Objectives</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Primary</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Secondary</td>
<td>6</td>
</tr>
<tr>
<td>2.3 Safety</td>
<td>6</td>
</tr>
<tr>
<td>2.4 Exploratory</td>
<td>6</td>
</tr>
<tr>
<td>3. Study Overview</td>
<td>6</td>
</tr>
<tr>
<td>3.1 Study Design</td>
<td>6</td>
</tr>
<tr>
<td>3.2 Sample Size</td>
<td>7</td>
</tr>
<tr>
<td>4. Study Endpoints and Covariates</td>
<td>8</td>
</tr>
<tr>
<td>4.1 Study Endpoints</td>
<td>8</td>
</tr>
<tr>
<td>4.2 Planned Covariates</td>
<td>9</td>
</tr>
<tr>
<td>5. Hypotheses</td>
<td>10</td>
</tr>
<tr>
<td>5.1 Study Time Points</td>
<td>10</td>
</tr>
<tr>
<td>5.2 Demographics and Baseline Related Definitions</td>
<td>11</td>
</tr>
<tr>
<td>5.3 Other Study Related Definitions</td>
<td>14</td>
</tr>
<tr>
<td>6. Analysis Subsets</td>
<td>16</td>
</tr>
<tr>
<td>6.1 Full Analysis Set (FAS)</td>
<td>16</td>
</tr>
<tr>
<td>6.2 Completer Analysis Set (CAS)</td>
<td>16</td>
</tr>
<tr>
<td>6.3 Per Protocol Analysis Set (PPAS)</td>
<td>16</td>
</tr>
<tr>
<td>6.4 Safety Analysis Set</td>
<td>17</td>
</tr>
<tr>
<td>6.5 Subgroup Analyses</td>
<td>17</td>
</tr>
<tr>
<td>7. Interim Analysis and Early Stopping Guidelines</td>
<td>17</td>
</tr>
<tr>
<td>8. Data Screening and Acceptance</td>
<td>18</td>
</tr>
<tr>
<td>8.1 General Principles</td>
<td>18</td>
</tr>
<tr>
<td>8.2 Data Handling and Electronic Transfer of Data</td>
<td>18</td>
</tr>
<tr>
<td>8.3 Handling of Missing and Incomplete Data</td>
<td>18</td>
</tr>
<tr>
<td>8.4 Outliers</td>
<td>19</td>
</tr>
<tr>
<td>8.5 Distributional Characteristics</td>
<td>19</td>
</tr>
<tr>
<td>8.6 Validation of Statistical Analyses</td>
<td>19</td>
</tr>
<tr>
<td>9. Statistical Methods of Analysis</td>
<td>20</td>
</tr>
<tr>
<td>9.1 General Principles</td>
<td>20</td>
</tr>
<tr>
<td>9.2 Subject Accountability</td>
<td>20</td>
</tr>
<tr>
<td>9.3 Important Protocol Deviations</td>
<td>21</td>
</tr>
</tbody>
</table>
9.4 Demographic and Baseline Characteristics ............................................ 21
9.5 Efficacy Analyses .................................................................................... 21
  9.5.1 Analyses of Primary Efficacy Endpoint(s) .......................................... 24
     9.5.1.1 Primary Analysis ........................................................................ 24
     9.5.1.2 Sensitivity Analysis .................................................................... 24
  9.5.2 Analyses of Key Secondary Efficacy Endpoint(s) .................................. 25
     9.5.2.1 Primary Analysis ........................................................................ 25
     9.5.2.2 Sensitivity Analysis .................................................................... 26
  9.5.3 Subgroup Analyses ............................................................................ 27
  9.5.4 Analyses of Other Efficacy Endpoints(s) ............................................. 27
     9.5.4.1 Sensitivity Analysis .................................................................... 28
  9.5.5 Analyses of Exploratory Endpoints .................................................... 29
  9.5.6 Variability of PTH values ..................................................................... 30
  9.5.7 Analyses of Health Related Quality of Life Endpoints ....................... 30
9.6 Safety Analyses ....................................................................................... 30
  9.6.1 Adverse Events ............................................................................... 30
  9.6.2 Laboratory Test Results ...................................................................... 31
  9.6.3 Vital Signs and Physical Measurements ............................................. 31
  9.6.4 Antibody Formation .......................................................................... 32
  9.6.5 Exposure to Investigational Product ............................................... 32
  9.6.6 Exposure to Concomitant Medication and Dialysate Calcium ............. 32

10. Changes from Protocol-specified Analyses ............................................. 32
11. Literature Citations / References ............................................................. 33
12. Appendices .............................................................................................. 34

List of Tables
Table 1. Primary and Key Secondary Endpoints Summary Table .................. 22

List of Appendices
Appendix A. Reference Values / Toxicity Grades .......................................... 35
Appendix B. Patient-reported Questionnaires .............................................. 36
## Table of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADPC</td>
<td>Analysis Dataset for Pharmacokinetics Concentrations</td>
</tr>
<tr>
<td>AE</td>
<td>Adverse event</td>
</tr>
<tr>
<td>ALP</td>
<td>Alkaline phosphatase</td>
</tr>
<tr>
<td>ALT</td>
<td>Alanine aminotransferase</td>
</tr>
<tr>
<td>ASP</td>
<td>Aspartate aminotransferase</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>BSAP</td>
<td>Bone specific alkaline phosphatase</td>
</tr>
<tr>
<td>CAS</td>
<td>Completer analysis set</td>
</tr>
<tr>
<td>cCa</td>
<td>Albumin corrected calcium concentration</td>
</tr>
<tr>
<td>CDM</td>
<td>Clinical data management</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CKD</td>
<td>Chronic kidney disease</td>
</tr>
<tr>
<td>CMH</td>
<td>Cochran-Mantel-Haenszel</td>
</tr>
<tr>
<td>CPE</td>
<td>Clinical</td>
</tr>
<tr>
<td>CRF</td>
<td>Case report form</td>
</tr>
<tr>
<td>CTCAE</td>
<td>Common terminology criteria for adverse events</td>
</tr>
<tr>
<td>CTX</td>
<td>Serum collagen type 1 cross-linked C-telopeptide</td>
</tr>
<tr>
<td>DMC</td>
<td>Data monitoring committee</td>
</tr>
<tr>
<td>EAP</td>
<td>Efficacy assessment phase: the period between Week 20 and Week 27, inclusive</td>
</tr>
<tr>
<td>EEAS</td>
<td>Efficacy Evaluable Analysis Set</td>
</tr>
<tr>
<td>EOI</td>
<td>Adverse event of interest</td>
</tr>
<tr>
<td>EOS</td>
<td>End of Study</td>
</tr>
<tr>
<td>ePRO</td>
<td>Electronic patient report outcome</td>
</tr>
<tr>
<td>ESRD</td>
<td>End-stage renal disease</td>
</tr>
<tr>
<td>EVOLVE</td>
<td>EValuation Of Cinacalcet Therapy to Lower CardioVascular Events</td>
</tr>
<tr>
<td>FAS</td>
<td>Full analysis set</td>
</tr>
<tr>
<td>FGF-23</td>
<td>Fibroblast growth factor 23</td>
</tr>
<tr>
<td>FLIE</td>
<td>Functional living index emesis</td>
</tr>
<tr>
<td>HR</td>
<td>Heart rate</td>
</tr>
<tr>
<td>IP</td>
<td>Investigational product</td>
</tr>
<tr>
<td>IPD</td>
<td>Important protocol deviation</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>Definition</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
</tr>
<tr>
<td>iPTH</td>
<td>Intact parathyroid hormone</td>
</tr>
<tr>
<td>LLOQ</td>
<td>Lower limit of quantification</td>
</tr>
<tr>
<td>LVCF</td>
<td>Last value carried forward</td>
</tr>
<tr>
<td>IV</td>
<td>Intravenous</td>
</tr>
<tr>
<td>IXRS</td>
<td>Interactive voice/web response system</td>
</tr>
<tr>
<td></td>
<td>Technology that utilizes either telecommunications or is web based, which is linked to a central computer in real time as an interface to collect and process information.</td>
</tr>
<tr>
<td>KAI</td>
<td>KAI Pharmaceuticals, Inc.</td>
</tr>
<tr>
<td>KDQOL-36</td>
<td>Kidney Disease Quality of Life-36</td>
</tr>
<tr>
<td>NVSA</td>
<td>Nausea and Vomiting Symptom Assessment</td>
</tr>
<tr>
<td>MedDRA</td>
<td>Medical dictionary for regulatory activities</td>
</tr>
<tr>
<td>P</td>
<td>Phosphorus</td>
</tr>
<tr>
<td>PPAS</td>
<td>Per protocol analysis set</td>
</tr>
<tr>
<td>PTH</td>
<td>Parathyroid hormone</td>
</tr>
<tr>
<td>Q1</td>
<td>Quarter 1</td>
</tr>
<tr>
<td>Q3</td>
<td>Quarter 3</td>
</tr>
<tr>
<td>SAP</td>
<td>Statistical analysis plan</td>
</tr>
<tr>
<td>SBP</td>
<td>Systolic blood pressure</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SE</td>
<td>Standard error</td>
</tr>
<tr>
<td>SHPT</td>
<td>Secondary hyperparathyroidism</td>
</tr>
<tr>
<td>TBIL</td>
<td>Total bilirubin</td>
</tr>
<tr>
<td>TEAE</td>
<td>Treatment-emergent adverse event</td>
</tr>
<tr>
<td>ULOQ</td>
<td>Upper limit of qualification</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual analogue scale</td>
</tr>
</tbody>
</table>
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 amendment 3.0 for AMG 416 study 20120360 dated 17 October 2013. The scope of this plan includes the final analyses that are planned and will be executed by the Amgen Biostatistics department unless otherwise specified.

2. Objectives

2.1 Primary

Demonstrate that treatment with AMG 416 is not inferior to treatment with cinacalcet for lowering serum intact parathyroid hormone (PTH) levels by > 30% from baseline among subjects with chronic kidney disease (CKD) and secondary hyperparathyroidism (SHPT) who require management with hemodialysis.

2.2 Secondary

Assess whether treatment with AMG 416 is superior to treatment with cinacalcet as measured by the proportion of subjects with > 50% decrease in serum PTH from baseline, proportion of subjects with > 30% decrease in serum PTH from baseline, mean number of days of vomiting or nausea per week, change from baseline in albumin corrected calcium concentration (cCa), mean pre-dialysis serum phosphorous (P), mean severity of nausea, and mean number of episodes of vomiting. Information about nausea and vomiting will be collected via an electronic patient reported outcome (ePRO).

2.3 Safety

Assess the safety and tolerability of AMG 416 compared with cinacalcet.

2.4 Exploratory

Assess the impact of AMG 416 on P, PTH, on bone turnover parameters such as serum bone specific alkaline phosphatase (BSAP), serum collagen type 1 cross-linked C-telopeptide (CTX), and fibroblast growth factor-23 (FGF-23), and on vomiting and nausea (collected as ePRO).

3. Study Overview

3.1 Study Design

This is a phase 3, multicenter, randomized, active-controlled, double-blind, double-dummy, dose-titration, 26-week treatment period comparison of AMG 416 and cinacalcet. All subjects, regardless of treatment assignment, may receive standard of
care as prescribed by the individual Investigator, with calcium supplements, phosphate
binders, and nutritional vitamin D supplements. If treatment with calcitriol or vitamin D
analogs is ongoing when subjects are enrolled in the study, the doses of these agents
should remain constant for the duration of study, unless treatment with vitamin D is
initiated, interrupted, or adjusted for reasons of safety.

Subjects will be stratified by serum PTH level (< 900 pg/mL, ≥ 900 pg/mL) and region
(North America and non-North America), and be randomized 1:1 to receive AMG 416 or
cinacalcet.

3.2 Sample Size

The planned sample size is 600 subjects (300 subjects per treatment arm).

A non-inferiority margin was determined based on data collected in the Amgen EVOLVE
trial (EValuation Of Cinacalcet Therapy to Lower CardioVascular Events,
Study 20050182). This was a randomized, placebo-controlled trial that enrolled a similar
patient population as intended to be recruited in this study. Rates of 25% and 60% in
the placebo and cinacalcet arms for the subjects who have 30% reduction in PTH from
baseline, respectively, were derived and the two-sided 95% confidence interval for the
treatment difference based on the large sample normal approximation is (31%, 39%).
Half of the lower limit of the confidence interval for the treatment difference (compared to
placebo) is 15.5%. Based on short term variation in serum PTH values, a difference of
12% in the proportion of subjects of achieving a PTH reduction from baseline in
percentage between treatment groups would not be considered a clinically meaningful
difference. Twelve percent, which is smaller than the above margin 15.5% and the loss
of effect that would be clinically acceptable, was selected as the non-inferiority margin
for this study.

Assuming 60% of subjects randomized to cinacalcet and 60% of subjects randomized to
AMG 416 achieve a greater than 30% reduction from baseline in mean pre-dialysis
PTH during the efficacy assessment phase (EAP), 300 subjects per treatment arm will
provide 90% power to demonstrate non-inferiority using a margin of 12% for the upper
bound of the 95% 2-sided confidence interval for the treatment difference between
AMG 416 and cinacalcet (cinacalcet- AMG 416).

For the test of superiority based on the key secondary endpoint of achievement of
>50% reduction from baseline in mean pre-dialysis PTH during the EAP, 300 subjects
per treatment group will provide more than 90% power to detect a statistically significant
difference between the treatment groups at the 5% significance level (two-sided), if we assume a 60% and 45% response rate in subjects randomized to AMG 416 and cinacalcet, respectively. For the test of superiority based on the key secondary endpoint of achievement of >30% reduction from baseline in mean pre-dialysis PTH during the EAP, 300 subjects per treatment group will provide more than 80% power to detect a statistically significant difference between the treatment groups at the 5% significance level (two-sided), if we assume a 68% and 57% response rate in subjects randomized to AMG 416 and cinacalcet, respectively. The assumption of a 45% and 57% response rate at 6 months in the cinacalcet group for >50% and >30% reduction in mean pre-dialysis PTH during the EAP is based on results generated from the Amgen EVOLVE trial in the subset of subjects who had a baseline PTH > 500 pg/mL, and the assumption of a 60% and 68% response rate in the AMG 416 group for >50% and >30% reduction in mean pre-dialysis PTH during the EAP is based on data observed in subjects with baseline PTH > 500 pg/mL in the AMG 416 open-label phase 2 study. Three hundred subjects per treatment group will also have more than 90% power to detect a treatment difference of 0.47 in mean number of days of vomiting or nausea per week in the first 8 weeks (mean of 0.57 for cinacalcet and 0.1 for AMG 416), assuming that the common standard deviation is 1.48 using a two group t-test with a 5% two-sided significance level.

4. Study Endpoints and Covariates

4.1 Study Endpoints

Primary Endpoint

Achievement of a > 30% reduction from baseline in mean pre-dialysis serum PTH level during the EAP (non-inferiority)

Key Secondary Endpoints

- achievement of a > 50% reduction from baseline in mean pre-dialysis serum PTH during the EAP (superiority)
- achievement of a > 30% reduction from baseline in mean pre-dialysis serum PTH during the EAP (superiority)
- mean number of days of vomiting or nausea per week in the first 8 weeks

Other Secondary Endpoints

- percent change from baseline in mean pre-dialysis serum cCa during the EAP
- achievement of mean pre-dialysis serum P ≤ 4.5 mg/dL during the EAP
- mean severity of nausea in the first 8 weeks
- mean number of episodes of vomiting per week in the first 8 weeks

**Safety Endpoints**

- incidence of cCa < 8.3 mg/dL at any time during the study
- incidence of cCa < 8.0 mg/dL at any time during the study
- incidence of cCa < 7.5 mg/dL at any time during the study
- incidence of hyperphosphatemia, defined as serum P > 5.5 mg/dL at any time during the study
- incidence of symptomatic hypocalcemia at any time during the study
- nature, frequency, severity, and relationship to treatment-emergent adverse events

**Exploratory Endpoints**

- percent change from baseline in mean pre-dialysis P during the EAP
- achievement of mean pre-dialysis serum PTH ≤ 300 pg/mL during the EAP
- change in serum BSAP from baseline to Week 27
- change in CTX from baseline to Week 27
- change in FGF-23 from baseline to Week 27
- mean number of episodes of vomiting per week in the first 16 weeks
- mean number of episodes of vomiting per week in the 26 weeks of treatment
- mean severity of nausea in the first 16 weeks
- mean severity of nausea in the 26 weeks of treatment

**4.2 Planned Covariates**

The planned covariates include the baseline stratification factors, including PTH level at baseline (< 900 pg/mL, ≥ 900 pg/mL) and region in 2 strata (North America vs. Non-North America), as well as the following important baseline covariates:

- Race group
- Dialysis vintage (0 -≤1 yr, >1-≤5 yr, >5 yr)
- Vitamin D sterol use (yes/no)
- Calcium containing phosphate binder or calcium supplement use (yes/no).
- Previous cinacalcet use
- Dialysate calcium ( < 3.0 mEq/L vs. ≥ 3.0 mEq/L)
- Sex (male/female)
- Age (< 65 years, ≥ 65 years)
All of these covariates will be used in the Multiple Imputation Method to be implemented in sensitivity analyses of the key secondary endpoint of achievement of > 50% and > 30% reduction in mean pre-dialysis serum PTH from baseline during EAP, respectively.

5. Hypotheses

Treatment of SHPT with AMG 416 is not inferior to treatment with cinacalcet as measured by the proportion of subjects with > 30% reduction from baseline in mean pre-dialysis serum PTH level during the EAP.

Treatment with AMG 416 is superior to treatment with cinacalcet as measured by the proportion of subjects with > 50% decrease in pre-dialysis serum PTH from baseline, by the proportion of subjects with > 30% decrease in pre-dialysis serum PTH from baseline, and by the mean days of vomiting or nausea per week in the first 8 weeks.

Definitions

5.1 Study Time Points

Informed Consent Date

The informed consent date for each subject is the date the subject signs the original informed consent for this study.

Screening Phase

The period after a subject has provided written informed consent and prior to randomization. This period may last up to 8 weeks and it is during this period when eligibility is determined, based on all screening tests and procedures.

Enrollment Date

The date a subject is randomized to a treatment group by the interactive voice/web response system (IXRS) after they have satisfied all enrollment criteria.

Randomization Date

The randomization date is the same as the enrollment date.

Study Week

The 7-day periods beginning with Study Day 1.
**Study Day 1**

The first day that investigational product is administered to the subject. For subjects who did not receive any investigational product during the study, Study Day 1 will be the randomization day.

**Study Day**

For each subject and for a given study visit date, study day is defined as the number of days since Study Day1:

\[
\text{Study day} = (\text{study visit date} - \text{Study Day 1 date}) + 1
\]

If the date of interest is prior to study day 1, study day will be calculated as

\[
\text{Study day} = (\text{study visit date} - \text{Study Day 1 date}).
\]

**Efficacy Assessment Phase (EAP)**

The period between Week 20 and Week 27, inclusive.

**Date of Last Dose IP Received**

For each subject, the last investigational product dose date is defined as the date when the last non-missing dose of investigational product is administered.

**Subject-level End of Study Date**

End of study for each subject is defined as the date of the subject last study assessment in the study. The date subject ended the study is recorded on the End of Study electronic case report form (CRF).

**Follow-up**

The 30 -period that occurs after a subject receives the last dose of investigational product during which safety information will be collected.

**5.2 Demographics and Baseline Related Definitions**

**Age**

Age is calculated as the subject’s (floor) integer age in years at the enrollment date. If the date of birth is not available, collected age will be used in the analyses.

**Body Mass Index (BMI)**

BMI equals a person's weight in kilograms divided by baseline height in meters squared (kg/m²).
Dialysis Vintage (years)

Dialysis vintage is defined as the duration in years from the dialysis start date (recorded on the end-stage renal disease (ESRD) history CRF) to the randomization date.

Baseline Values

Baseline values of PTH, cCa, and serum phosphorus are defined as the average of the last pre-dialysis assessments done during the screening period and the pre-dialysis assessment done on Study Day 1.

All other baseline laboratory values are using the last non-missing assessment taken prior to or on Study Day 1 (Study Day 1 assessment if available, otherwise the last screening assessment prior to Study Day 1).

Baseline vital sign and weight will be using the last post-dialysis assessment values prior to or on Study Day 1.

The use of vitamin D, including nutritional vitamin D (vitamin D supplement) and vitamin D sterol (active vitamin D), is defined as use of vitamin D during the 7 days period prior to Study Day 1, inclusive.

The use of other concomitant medication of interest at baseline is defined as use of each concomitant medication on Study Day 1. This includes medication use that starts prior to Study Day 1 and ends on Study Day 1 or duration of use covers Study Day 1.

The baseline scores of functional living index emesis (FLIE) and Kidney Disease Quality of Life-36 (KDQOL-36™) are defined as the scores collected during screening period.

Change from Baseline

The arithmetic difference between a post-baseline value and baseline value for a given time point:

\[ \text{Change from baseline} = \text{(post-baseline value – baseline value)} \]

Percent Change from Baseline

For each subject, percent change from baseline in a given variable at a given time point is defined as:

\[ 100 \times \left[ \frac{\text{(value at given time point – baseline value)}}{\text{baseline value}} \right] \]
Screening Value of Number of Days of Vomiting or Nausea per Week

The screening value of number of days of vomiting or nausea per week is defined by the number of days with at least one non-missing response to either question 1 or question 2 from the seven consecutive daily NVSA questionnaires provided by subjects in the 2-week period prior to first dose (including the enrollment period if necessary), where a response of 99 in question 2 is considered missing. For subjects providing 7 responses of either question from consecutive daily NVSA questionnaires in the week prior to first dose, these 7 days will be used. Subjects providing less than seven consecutive daily responses in the week period prior to first dose will be handled as below:

- If subject provides more than 7 days (inclusive) of responses in the 2-week period prior to first dose day, responses for the most proximal 7 days before first dose day (exclusive) will be used to calculate the screening value of number of days of vomiting or nausea per week. For example, if a subject provides responses to daily NVSA questionnaires at days 12, 11, 10, 9, 8, 7, 6, 5, 3, and 2 prior to first dose, responses for days of 2, 3, 5 to 9 prior to first dose will be used.

- If a subject provides less than 7 days (inclusive) of responses in the 2-week period prior to the first dose day, then the screening value of number of days of vomiting or nausea per week is considered missing for the analysis. No imputation will be done for defining screening value.

Screening Value of Number of Episodes of Vomiting Per Week

The screening value of number of episodes of vomiting per week is defined by the average response to question 2 from the seven consecutive daily NVSA questionnaires provided by subjects in the 2-week period prior to first dose (including the enrollment period if necessary).

For subjects providing 7 responses to question 2 from consecutive daily NVSA questionnaires in the week prior to first dose, these 7 days will be used. A response equal to 99 in question 2 will be considered as missing. Subjects providing less than 7 consecutive daily responses in the week period prior to first dose will be handled as follows:

- If a subject provides more than 7 days (inclusive) of responses in the 2-week period prior to first dose day, responses for the most proximal 7 days before first dose day (exclusive) will be used to calculate the screening value of number of episodes of vomiting per week. For example, if a subject provides responses to question 2 from daily NVSA questionnaires at days 12, 11, 10, 9, 8, 7, 6, 5, 3, and 2 prior to first dose, days of 2, 3, 5 to 9 prior to first dose will be used.
If a subject provides less than 7 days (inclusive) of responses in the 2-week prior to first dose day, data for screen value is considered missing for the analysis. No imputation will be done for screen value data.

5.3 Other Study Related Definitions

Protocol Scheduled Visits

The clinical planned event (CPE) assigned in the clinical database will be used as protocol scheduled visits. If the value at a scheduled visit is missing, an unscheduled assessment can be used provided the unscheduled assessment occurs within the window for that study week, i.e.

$$\text{ceiling} \left( \frac{\text{study day of the unscheduled visit}}{7} \right) = \text{study week number}$$

All missing scheduled assessments will be replaced using the above method, if possible, prior to deriving any study endpoints.

If the value at a scheduled visit is missing and multiple unscheduled assessments occurs within the window of this visit, or there are multiple values at a scheduled visit, then the first record will be used for the visit.

Mean Value of Efficacy Endpoints during the EAP

The mean of the scheduled (including the missing scheduled assessments replaced by unscheduled assessments as described above) pre-dialysis assessments taken during week 20 to week 27, inclusive.

Efficacy Endpoints Based on the Percent Changes

Percent change from baseline to the mean value during the EAP for the corresponding endpoint.

Efficacy Endpoints Based on Mean Number of Days of Vomiting or Nausea per Week

For each subject and each day, a day of vomiting or nausea will be defined as those where the severity of nausea score from question 1 of NVSA is > 0 or where the episodes of vomiting score from question 2 of NVSA is > 0.

Corrected Total Serum Calcium (cCa)

Corrected calcium (mg/dL) = Total calcium (mg/dL) + (4 – albumin [g/dL]) × 0.8

Total serum calcium will be corrected using above formulae if the serum albumin is < 4 g/dL or 40 g/L, otherwise cCa equals total serum calcium.
Low Calcium Based on Corrected Calcium Values

cCa values will be used to identify cases of low calcium of different severities using the following three categories: < 7.0 mg/dL, 7.0 - < 7.5 mg/dL, and 7.5 - < 8.3 mg/dL.

Exposure Period

For each subject who received at least one dose of IP, the number of weeks of IP exposure is calculated as:

\[
\text{Exposure period (weeks)} = \frac{\text{(date of last dose received – date of first dose) + 1}}{7}
\]

Compliance to Investigational Product (%)

Compliance = number of actual treatment taken / number of prescribed treatment *100

For the IV treatment, the number of actual treatment taken is defined as the actual number of IV doses a subject received, and the number of prescribed treatment is defined as the number of IV doses prescribed.

For the oral treatment, the number of actual treatment taken is defined as the number of days that a subject took the oral dose, and the number of prescribed treatment is defined as the number of days that oral treatment was prescribed.

The treatment compliance calculation only considers the periods when a subject was supposed to receive IP and it will exclude the dose withholding periods due to protocol specified reasons on CRF.

Treatment Emergent Adverse Events (TEAE)

Adverse events (AEs) starting on or after the Study Day 1 as determined by the flag indicating if the adverse event started prior to the first dose of IP on the Adverse Events Summary CRF and up to 30 days after the last dose date.

Serious adverse events (SAEs) starting on or after the Study Day 1 as determined by the flag indicating if the adverse event started prior to the first dose of IP on the Adverse Events Summary CRF and up to 30 days after the last dose date.

Adverse Event Subject Incidence

Defined as the number and percentage of subjects with a reported event(s). For subjects with multiple reports of the same event during the study, the subject will be counted only once. For adverse event tabulations involving severity, the highest severity of the particular adverse events will be used for each subject.
Adverse Events of Interest (EOI)

Adverse events of interest are defined by the most current standardized product-level event of interest list. Unless otherwise specified, the narrow search scope will be used for all EOIs.

Concomitant Medications of Interest

The selected medications of interest are: nutritional vitamin D (vitamin D supplement), vitamin D sterol (active vitamin D), calcium supplements, and phosphate binders, as identified in the concomitant medication CRFs.

6. Analysis Subsets

6.1 Full Analysis Set (FAS)

The FAS includes all randomized subjects. Subjects will be analyzed according to the treatment group assigned by the IXRS. Efficacy analyses will be performed using the FAS.

6.2 Completer Analysis Set (CAS)

The CAS includes all randomized subjects with at least one pre-dialysis PTH value (including the missing scheduled assessments replaced by unscheduled assessments as described previously) during the EAP. Subjects will be analyzed according to randomized treatment group. This analysis set is the same as the Efficacy Evaluable Analysis Set (EEAS) mentioned in the protocol.

6.3 Per Protocol Analysis Set (PPAS)

The PPAS is defined as all randomized subjects who have no major protocol violations (as defined below) and have at least one post-dose PTH value and had at least 16 weeks exposure of IP. Subjects with the following eligibility deviations will be excluded from the PPAS:

- Less than 18 years at screening
- Not receiving constant maintenance hemodialysis for at least three times a week
- Not meeting either calcium concentration or PTH eligibility criteria at screening
- Received cinacalcet within 3 month prior to screening or underwent a parathyroidectomy within 6 months prior to dosing
- Currently enrolling in another study or less than 30 days since ending another study
- Having known sensitivity to either cinacalcet or AMG 416
Moreover, subjects who have the following deviations during study period will also be excluded in this analysis set:

- Had parathyroidectomy or kidney transplant before EAP
- Received commercialized cinacalcet during the study
- Missed > 14 consecutive days of intravenous (IV) IP or oral IP, excluding the withholding of IP for protocol-specific reasons

For the subjects who undergo parathyroidectomy or kidney transplantation during EAP, only the PTH values before the surgery will be used.

Subjects in the PPAS will be analyzed according to randomized treatment assignment. PPAS will be used in the sensitivity analysis of the primary endpoint.

6.4 Safety Analysis Set

The Safety Analysis Set consists of all randomized subjects who receive at least one dose of investigational product. Subjects will be analyzed by their randomized group unless an incorrect treatment was administered throughout the study, in which case the originally assigned treatment group will be used for analysis. Safety analyses will be performed using the Safety Analysis Set.

6.5 Subgroup Analyses

The primary and secondary efficacy endpoints will be analyzed in the subgroups defined by the stratification factors using RAVE data and the following baseline factors:

- Screening PTH level (< 900 pg/mL, ≥ 900 pg/mL)
- Region (North America, Non-North America)
- Race (black, white, other), if number of subjects is too small, the 'other' group will be combined with 'white' group
- Dialysis vintage (>0 - ≤1 yr, >1 - ≤5 yr, >5 yr)
- Dialysate calcium (< 3.0 mEq/L, ≥ 3.0 mEq/L);
- Baseline vitamin D sterol use (yes, no)
- Baseline calcium containing phosphate binder or calcium supplement use (yes, no)
- Previous cinacalcet use (yes/no)
- Sex (male/female)
- Age (< 65 years, ≥ 65 years)

7. Interim Analysis and Early Stopping Guidelines

No interim analysis is planned for this study. An independent external Data Monitoring Committee (DMC) comprised of experienced clinicians and a statistician external to Amgen will be used to oversee progress of the study and make recommendations.
relating to early closure or alteration of the study based on ongoing monitoring of the study data for the duration of the study.

8. Data Screening and Acceptance

8.1 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.

8.2 Data Handling and Electronic Transfer of Data
Amgen’s Clinical Data Management (CDM) department will provide all data, including eCRF, central lab assessments, antibody and ePRO, to be used in the planned analyses. This study will use the Medidata RAVE database for all data provided by Covance besides ePRO.

8.3 Handling of Missing and Incomplete Data
Subjects may be missing specific data points for various reasons. In general, data may be missing due to a subject’s early withdrawal from study, a missed visit, or non-evaluable of a data point at a particular point in time. Sites will be queried for missing or non-conformant data that are required or considered critical. All efforts will be made to capture complete critical data prior to the database lock.

The following imputation rules will be implemented for critical data if the missing or incomplete data cannot be resolved after the query process:

Only missing or partially missing dates for adverse events will be imputed with the exception of adverse events occurring prior to the first dose date. Stop dates for AEs will not be imputed. Adverse events with a partially missing start date that is conclusively prior to the date of first IP administered (as indicated by ‘Did event start before first dose of IP’ on the AE CRF page) will be considered pre-treatment adverse events and excluded from safety analyses. All other partially missing adverse event start dates will be handled as described below, with the reference date being Study Day 1:

- If the year is available and the day and month are missing, the day and month will be set to the 1st of January of the onset year
- If the year and month are available and the day is missing, the day will be set to the 1st of the onset month
- If the day and month are available and the year is missing, the year will be set to the year of the reference date
• If the year and day are available and the month is missing, the month will be set to January of the onset year

• If the resulting date is prior to the reference date, the date will be reset to the reference date (this applies to AE start date imputation only)

Partial/missing start dates for concomitant medications of interest (Section 6.3) will be imputed using the algorithm above with the reference date being Study Day 1.

Partial/missing stop dates for these concomitant medications will be imputed using the following logic:

• If the medication stop date is completely missing, then the stop date is set as the end of the study date. Otherwise,

• If the stop year is available and stop month and day are missing, the month and day will be reset to 31st of December of the stop year

• If the stop year and month are available and the stop day is missing, the stop day will be set to the last day of the month of the stop year

• If the stop year and day are available and stop month is missing, the stop month will be set to December of the stop year

• If the stop month and day are available and stop year is missing, the stop year will be reset to the year of the start date

• If any of the resulting dates are prior to the start date, the stop date will be reset to the start date

• If any of the resulting dates are after the end of study (EOS) date, the stop date will be reset to the EOS date

8.4 Outliers

Various methods, including univariate summaries, histograms, scatter plots, box plots, and line graphs, may be used to identify outliers in key safety and efficacy variables. Extreme data points will be identified during the blinded review of the data prior to database lock. Such data points will be reviewed with CDM to ensure accuracy. Extreme data points will be included in the analyses.

8.5 Distributional Characteristics

The percent change endpoints will be analyzed using a repeated measures mixed effects model, which assumes that the data are normally distributed. The normality assumption will be assessed using the graphical methods such as the normal quantile-quantile plot (Q-Q plot) along with the Shapiro-Wilk test. If the normality assumption does not hold, data transformation will be considered.

8.6 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 version 9.2 or later.

9. Statistical Methods of Analysis

9.1 General Principles

This is a phase 3 study to evaluate the efficacy of AMG 416 as compared to cinacalcet. The primary objective is to determine if AMG 416 is considered non-inferior to cinacalcet on achievement of a > 30% reduction from baseline in mean pre-dialysis serum PTH during the EAP. AMG 416 will be considered non-inferior if the upper bound of the two-sided 95% confidence interval of the treatment difference (cinacalcet – AMG 416) is smaller than 12%. If this criterion is met, the three key secondary endpoints will be tested for superiority sequentially at the 5% significance level. If the three key secondary endpoints are statistically significant, the rest of the secondary endpoints will be tested at the 5% significance level. The Hochberg procedure will be used to adjust for multiple comparisons among the other secondary endpoints.

For the efficacy endpoints based on laboratory measurements (pre-dialysis PTH, cCa, and phosphorus), mean values within the EAP for each subject will be used in the analyses.

Continuous variables will be summarized using descriptive statistics, including the number of observations (n), mean, standard deviation (SD), standard error (SE), median, the 1st (Q1) and 3rd (Q3) quartiles, minimum, and maximum. Categorical variables will be summarized using the number and percent of subjects.

9.2 Subject Accountability

The number (percentage) of subjects randomized, received IP, completed IP, and completed the study will be summarized by randomized treatment group. Reasons for study discontinuation and IP discontinuation will be summarized separately also by randomized treatment group.

Key study dates for the first subject enrolled, last subject enrolled, last subject’s end of study, and last subject’s end of IP will be presented.
9.3 **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. A table and a list of eligibility deviations will be generated.

9.4 **Demographic and Baseline Characteristics**

Demographics (sex, age, race, and ethnicity) will be summarized using descriptive statistics for all enrolled subjects.

The following baseline characteristic will also be summarized using descriptive statistics for all enrolled subjects:

- Sex, age, weight, height, BMI, race, ethnicity, stratification factors, PTH, corrected total serum calcium, serum phosphorous, cCa x P, vital signs, selected concomitant medications, and selected medical history (history of hemodialysis, parathyroidectomy history, ESRD history, cardiovascular/social/other history).

9.5 **Efficacy Analyses**

All analyses of efficacy endpoints will be performed according to their randomized treatment group. The analyses of primary efficacy endpoint, key secondary efficacy endpoint, and other secondary efficacy endpoints are summarized in the following table and subsequent sections.
### Table 1. Primary and Key Secondary Endpoints Summary Table

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Primary Analysis</th>
<th>Sensitivity Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Endpoint</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 30% reduction from baseline in mean</td>
<td>Mantel-Haenszel (M-H) method with adjustment for the 2 randomization stratification</td>
<td>1. Similar as primary analysis method but using the CAS</td>
</tr>
<tr>
<td>pre-dialysis serum PTH level during the</td>
<td>factors based on the FAS analysis set to determine 95% CI. Multiple imputation</td>
<td>2. Last value carried forward (LVCF): Same method as primary analysis. For the</td>
</tr>
<tr>
<td>efficacy assessment phase (EAP) (non-inferiority)</td>
<td>under the non-inferiority null method will be applied to the subjects who do not have data during the EAP.</td>
<td>subjects without PTH data during the EAP, the mean of the last 2 pre-dialysis PTH values obtained after Study Day 1 will be carried forward. If only one value is available, this single value will be carried forward to the EAP. A similar imputation approach as the primary analysis (non-inferiority null method) will be applied to subjects without a post-baseline PTH value after applying LVCF approach.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Per Protocol: Similar as primary analysis method but using PPAS</td>
</tr>
<tr>
<td><strong>Key Secondary Endpoints</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achievement of &gt; 50% reduction from</td>
<td>A Cochran-Mantel-Haenszel (CMH) test adjusting for the 2 randomization stratification</td>
<td>Multiple Imputation (MI): For subjects without PTH data during the EAP, the multiple imputation method using SAS® PROC MI will be used to impute the PTH value during EAP. Treatment group, stratification factors, baseline covariates (as detailed in Section 4.2), and other study data such as additional laboratory values will be considered for inclusion in the model as auxiliary variables.</td>
</tr>
<tr>
<td>baseline in mean pre-dialysis serum PTH</td>
<td>factors based on FAS. Subjects will be considered as not achieving the primary efficacy endpoint if they do not have data during the EAP (non-responder imputation)</td>
<td></td>
</tr>
<tr>
<td>during the EAP (superiority)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 1 of 2
Table 1. Primary and Key Secondary Endpoints Summary Table

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Primary Analysis</th>
<th>Sensitivity Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Secondary Endpoints (continued)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achievement of &gt; 30% reduction from baseline in mean pre-dialysis serum PTH during the EAP (superiority)</td>
<td>A Cochran-Mantel-Haenszel (CMH) test adjusting for the 2 randomization stratification factors based on FAS. Subjects will be considered as not achieving the primary efficacy endpoint if they do not have data during the EAP (i.e., non-responder imputation)</td>
<td>Multiple Imputation (MI): For subjects without PTH data during the EAP, the multiple imputation method using SAS® PROC MI will be used to impute the PTH value during EAP. Treatment group, stratification factors, baseline covariates (as detailed in Section 4.2), and other study data such as additional laboratory values will be considered for inclusion in the model as auxiliary variables.</td>
</tr>
<tr>
<td>Mean number of days of vomiting or nausea per week in the first 8 weeks</td>
<td>The mean number of days of vomiting or nausea per week will be analyzed using either a repeated measures mixed effects model or generalized linear mixed model with Poisson regression, depending on the normality assumption achievement, including the randomization stratification factors and other pre-specified covariates.</td>
<td>1. The negative binomial regression with GEE, adjusting for the same covariates as in primary analysis and time in weeks for the first 8 weeks, will be used. 2. Repeat the primary analysis with imputed data, by including the subjects who provide at least 6 days of responses in a week, for each week. The days of nausea or vomiting per week for the first 8 weeks are imputed by the available valid responses in the week, which is from the total number of days of vomiting or nausea that are equivalent to a 7-day week. 3. Repeat the primary analysis with imputation method at numbering item 2 above by changing the threshold to include the subjects who provide at least 5 days of responses per week. 4. Repeat the primary analysis with imputation method at numbering item 2 above by changing the threshold to include the subjects who provide at least 4 days of responses per week.</td>
</tr>
</tbody>
</table>
9.5.1 Analyses of Primary Efficacy Endpoint(s)

The analysis of the primary efficacy endpoint is to assess if the treatment of SHPT with AMG 416 is not inferior to treatment with cinacalcet as measured by the proportion of subjects with > 30% reduction from baseline in mean pre-dialysis serum PTH level during the EAP. AMG 416 will be considered non-inferior to cinacalcet if the upper bound of the two-sided 95% confidence interval of the treatment difference (cinacalcet – AMG 416) is smaller than 12%. The superiority of the primary endpoint will not be formally tested. However, an odds ratio and 95% CI along with analyses described in the subgroup analysis section with non-responder imputation method will be completed.

9.5.1.1 Primary Analysis

The primary endpoint will be analyzed using Mantel-Haenszel (M-H) method [Agresti and Hartzel, 2000] stratified by the enrollment serum PTH level (< 900 pg/mL, ≥ 900 pg/mL) and region (North America and Non-North America). The M-H estimator will be used to estimate the difference in the proportion of subjects who achieve the primary efficacy endpoint between two treatment groups and the standard error of the M-H estimator will also be estimated. The 95% two-sided confidence interval will be provided.

For the subjects who have at least one PTH value during EAP, the available data will be used to calculate the mean. For the subjects who have no PTH values during EAP, including subjects who discontinue the study due to rising PTH (recorded as primary reason for ending the study on the End of Study CRF), the missing primary endpoint will be imputed using non-inferiority null method [Koch, 2008]. The presumed responder rates are 60% and 48% for the cinacalcet and AMG 416 groups, respectively, for the imputation, and the imputation will be performed 5 times. The combination of estimating result for the M-H estimator and its standard error from each imputed data set will follow the method used in SAS® PROC MIANALYZE.

9.5.1.2 Sensitivity Analysis

The following sensitivity analyses will be conducted using the same M-H method as in primary analysis:

- CAS: Similar as primary analysis method but using the CAS.
- LVCF: The subjects without PTH data during the EAP, the mean of the last 2 pre-dialysis PTH values obtained after Study Day 1 will be carried forward. If only one value is available, this single value will be carried forward to the EAP. The same
imputation strategy for primary analysis will be employed for subjects without a post baseline PTH value.

- Per Protocol: Only the subjects who meet the criteria for inclusion in the PPAS are included.

9.5.2 Analyses of Key Secondary Efficacy Endpoint(s)
Sequential testing of the following three key secondary efficacy endpoints will be performed in this order if non-inferiority is demonstrated on the primary endpoint.

- achievement of > 50% reduction in mean pre-dialysis serum PTH from baseline during the EAP (superiority)
- achievement of > 30% reduction in mean pre-dialysis serum PTH from baseline during the EAP (superiority)
- mean number of days of vomiting or nausea per week in the first 8 weeks

9.5.2.1 Primary Analysis
The endpoint of achievement of > 50% reduction in mean pre-dialysis serum PTH from baseline during the EAP will be analyzed using the CMH test performed using FAS and stratified by the randomization stratification factors. The mean pre-dialysis serum PTH will be derived based on all available PTH values during EAP. The non-responder imputation strategy will be used for subjects with no PTH value during the EAP. The endpoint of achievement of > 30% reduction in mean pre-dialysis serum PTH from baseline during the EAP for superiority will be analyzed in the same fashion using CMH test with non-responder imputation strategy for subjects who have no PTH value during EAP. The endpoint of achievement of > 30% reduction in mean pre-dialysis serum PTH from baseline during the EAP for superiority will be analyzed in the same fashion using CMH test with non-responder imputation strategy for subjects who have no PTH value during EAP.

The endpoint of mean number of days of vomiting or nausea per week will be analyzed on FAS using a repeated measure mixed effects model, including screening value of number of days of vomiting or nausea per week, randomization stratification factors, study week, treatment group and treatment by study week as fixed effects and subject as random effect to account for within subject correlation. If the normality assumption does not hold, the data will be analyzed by generalized linear mixed model using Poisson regression with link function, including the same fixed effects and random effect as described for the repeated measure mixed effects model. The difference (AMG 416 -cinacalcet) or ratio (AMG 416 / cinacalcet) of the mean number of days of vomiting or nausea per week between the treatment groups will be estimated by contrast of least square means (LSMESTIMATE). For the primary analysis, no imputation will be completed for subjects providing less than 7 days of responses to NVSA questions in any given week.
9.5.2.2  Sensitivity Analysis

The sensitivity analysis for the key secondary endpoint of proportion of subjects who have achievement of > 50% reduction in mean pre-dialysis serum PTH from baseline during EAP will be performed. The multiple imputation method will be used to carry out the multiple imputation using SAS® PROC MI. The Fully Conditional Specification (FCS) method will be used to carry out the multiple imputation using SAS® PROC MI. PTH values from all post Study Day 1 scheduled visits will be included in the imputation model. The mean PTH value over the EAP is treated as a single time point. Treatment group, stratification factors, baseline covariates (as detailed in Section 4.2), and other study data such as additional laboratory values will be considered for inclusion in the model as auxiliary variables. With FCS method, the missing PTH values are imputed sequentially in the order specified in the VAR statement (chronological order). Each imputed dataset will be analyzed using stratified CMH method (as described in section 10.5.1). Final result of CMH estimate and odds ratio are based on combined inference from each of the imputed datasets.

The same sensitivity analyses will be repeated for the key secondary endpoint of proportion of subjects who have achievement of > 30% reduction in mean pre-dialysis serum PTH from baseline during EAP for superiority in the same manner with the same group of auxiliary variables when imputing the mean of PTH value over the EAP. Each imputed dataset will be analyzed using stratified CMH method (as described in section 10.5.1). Final result of CMH estimate and odds ratio are based on combined inference from each of the imputed datasets.

The sensitivity analysis for the key secondary endpoint of mean number of days of vomiting or nausea per week will be performed using negative binomial regression with generalized estimation equation (GEE). The model includes the same response as in primary analysis with log link function and the randomization stratification factors, study week, treatment group and treatment by study week as covariates.

Additionally, three alternative analyses with missing data imputation will be implemented for use with the same model, adjusted for the same covariates, as that used in the primary analysis. The first analysis will include subjects who provide at least 6 days of NVSA questionnaire responses in a week. For subjects providing less than 7 days of responses in a given week, the number of days of vomiting or nausea for that week will be imputed using the 6 available valid responses in the week such that the subsequent total number of days of vomiting or nausea are equivalent to a 7-day week. The other
two analyses change the threshold of at least 6 days to at least 5 days and at least 4 days, respectively with available valid responses in the week used to calculate the total number of days of vomiting or nausea equivalent to a 7-day week. See Appendix B for details about handling missing data from the NVSA questionnaire.

9.5.3 Subgroup Analyses
The key secondary efficacy endpoints will be analyzed within each subgroup (defined in section 7.5) provided there are adequate number of subjects within the subgroups.

For binary endpoints, the point estimate and 95% confidence interval of the odds ratio for each subgroup will be provided to evaluate the consistency of the treatment effect across subgroups. The CMH test will also be implemented to compare the treatment groups within each subgroup. For subjects who have no PTH values during EAP, the missing endpoints will be imputed using non-responder imputation method.

For the continuous endpoint of mean number of days of vomiting or nausea per week for the first 8 weeks, the mean and 95% confidence interval of the endpoint for each subgroup will be provided by treatment group. The treatment group differences in each subgroup will be compared by mixed effects model as described in section 10.5.2.

In addition, the effect of the interaction between the treatment and each of the subgroup factors will be evaluated for the primary and key secondary endpoints. Each factor and its interaction with treatment will be evaluated one at a time. For endpoints based on proportions, logistic regression including all stratification factors will be used. For endpoints based on the mean number of days of vomiting or nausea, a repeated measures mixed effects model or generalized linear mixed model, depending on the model used for the primary analysis of the same endpoint in section 10.5.2, adjusting for all stratification factors, screening value of number of days of vomiting or nausea, study week and study week by treatment as fixed effect and subject as random effect will be used.

9.5.4 Analyses of Other Efficacy Endpoints(s)
The following secondary endpoints will be evaluated if the above key secondary endpoints are considered statistically significant. To control the experiment-wise error rate, the Hochberg procedure of adjustment for multiple comparisons will be used for the following:

- percent change from baseline in mean pre-dialysis serum cCa during the EAP
- achievement of mean pre-dialysis serum P ≤ 4.5 mg/dL during the EAP
- mean severity of nausea in the first 8 weeks
- mean number of episodes of vomiting per week in the first 8 weeks

A repeated measures mixed effects model will be used to compare treatment groups with respect to the percent change from baseline in serum cCa levels during the EAP, and will include the treatment group, randomization stratification factors, study week and study week by treatment as fixed effects. The FAS will be used for these analyses.

The secondary endpoint of proportion of subjects with mean pre-dialysis P ≤ 4.5 mg/dL during the EAP will be analyzed using the CMH test stratified by the randomization stratification factors. These analyses will be performed on the FAS using the non-responder method of imputation for those subjects without laboratory values during the EAP. The mean pre-dialysis P will be derived based on all available values of P during EAP.

The analysis of covariance (ANCOVA) stratified by the randomization stratification factors will be used to compare the mean severity of nausea in the first 8 weeks between two treatment groups. For each subject, the mean severity of nausea will be calculated by averaging all available daily severities (including zeroes) reported in the first 8 weeks. FAS will be used for this analysis.

For the mean number of episodes of vomiting per week in the first 8 weeks, a repeated measures mixed effects model will be used to compare the difference between the two treatment groups with the treatment group, including screening value of number of episodes of vomiting per week, study week, treatment by study week and the randomization stratification factors as fixed effects in the model and subject as a random effect. The number of episodes in a week is the sum of all reported daily episodes in the week. If the normality assumption of repeated measures mixed effects model does not hold, the data will be analyzed by generalized linear mixed model using Poisson regression with link function, including the same fixed effect and random effect as in repeated measure mixed model. No imputation will be completed for subjects providing less than 7 days of responses to NVSA questions in any given week. See Appendix B for details about handling missing data from the NVSA questionnaire.

9.5.4.1 Sensitivity Analysis

The sensitivity analyses will be conducted for the endpoint of mean number of episodes of vomiting per week in the first 8 weeks. Three alternative analyses of missing data imputation will be implemented for use with the same model, adjusted for the same
covariates, as that used in the primary analysis. The first analysis will include subjects who provide at least 6 days of responses to question 2 of the NVSA in a week. For subjects providing less than 7 days of responses in a given week, the number of episodes of vomiting for that week will be imputed using the 6 available valid responses in the week such that the subsequent total number of episodes of vomiting are equivalent to a 7-day week. The other two analyses change the threshold of at least 6 days to at least 5 days and at least 4 days, respectively with available valid responses in the week used to calculate the total number of episodes of vomiting equivalent to a 7-day week. See Appendix B for details about handling missing data from the NVSA questionnaire.

9.5.5 Analyses of Exploratory Endpoints

The following exploratory endpoints will be analyzed:

- percent change from baseline in mean pre-dialysis P during the EAP
- achievement of mean pre-dialysis serum PTH $\leq 300$ pg/mL during the EAP
- change in serum BSAP from baseline to Week 27
- change in CTX from baseline to Week 27
- change in FGF-23 from baseline to Week 27
- mean number of episodes of vomiting per week in the first 16 weeks
- mean number of episodes of vomiting per week in the 26 weeks of treatment
- mean severity of nausea in the first 16 weeks
- mean severity of nausea in the 26 weeks of treatment

All exploratory efficacy endpoints will be summarized by treatment group using descriptive statistics. Additionally, the treatment groups will be compared for each exploratory efficacy endpoint using the approaches described below correspondingly. All these analyses are performed on FAS.

The mean and 95% confidence interval of percent change from baseline in mean pre-dialysis P during the EAP and change in BSAP, CTX and FGF-23 from baseline to Week 27 will be provided for the two treatment groups.

The proportion of subjects with mean pre-dialysis serum PTH $\leq 300$ pg/mL during EAP will be calculated in the two treatment groups and will be compared using CMH test using FAS. Missing data will be imputed by non-responder method. Nominal p-value will be provided.
The mean number of episodes of vomiting per week in the first 16 weeks and in the 26 weeks of treatment will be analyzed, using descriptive statistics. No imputation will be completed for subjects providing less than 7 days of responses to question 2 of the NVSA in any given week. Similarly, the mean severity of nausea in the first 16 weeks and in the 26 weeks of treatment will be analyzed, using descriptive statistics.

9.5.6 Variability of PTH values
Summary statistics for various measures of PTH variability during the EAP will be provided for cinacalcet and AMG 416 group. The variability measures include intrasubject standard deviation (SD), coefficient of variation (CV), and range. Only PTH values collected while subjects are receiving IP will be included.

9.5.7 Analyses of Health Related Quality of Life Endpoints
Subjects will complete the KDQOL-36™ questionnaires and FLIE questionnaires at the protocol specified visits during the study. The questionnaires and scoring algorithms of KDQOL-36™ and FLIE can be found in Appendix B.

Summaries of the score and changes from screening will be provided by treatment group using descriptive statistics using FAS.

9.6 Safety Analyses
9.6.1 Adverse Events
The Medical Dictionary for Regulatory Activities (MedDRA) version 15.1 or later will be used to code all adverse events (AE). All AE tables will be summarized by treatment group and only TEAEs will be included. Subject incidence of all AEs, serious AEs, AEs leading to withdrawal of IP and fatal AEs will be tabulated by system organ class and preferred term in descending order of frequency. Subject incidence of EOIs will be summarized according to the EOI search strategy categories defined by the EOI steering committee.

Summaries of AEs and serious AEs will also be provided by preferred term in descending order of frequency.

Summaries of AEs and serious AEs will be tabulated by system organ class, preferred term, and severity (mild, moderate, severe, maximal/life-threatening) in descending order of frequency.
Proportion of subjects (with 95% confidence interval) meeting each of the following criteria at any time (post-baseline) in the study will be presented by treatment group. No formal statistical testing will be performed.

- cCa < 8.3 mg/dL
- cCa < 8.0 mg/dL
- cCa < 7.5 mg/dL
- hyperphosphatemia, defined as serum P > 5.5 mg/dL
- symptomatic hypocalcaemia (from AE preferred term of Hypocalcaemia)

### 9.6.2 Laboratory Test Results

Laboratory values below the parameter-specific the lower limit of quantification (LLOQ) or above the parameter-specific upper limit of quantification (ULOQ) will be imputed as the LLOQ or ULOQ value, respectively.

Laboratory parameters including corrected calcium, PTH, P, creatinine and Ca x P will be summarized by treatment group and the protocol-specified scheduled visit. Summaries of the absolute value change from baseline, and/or percent change from baseline will also be provided.

Shift tables will be provided for albumin, corrected calcium, potassium, phosphorus, magnesium, hemoglobin, alkaline phosphatase (ALP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin (TBIL) by summarizing the most extreme changes in toxicity grade from baseline (increase and/or decrease as appropriate) to any visit during the study while the subjects are receiving IP. The toxicity grading will be based on common terminology criteria for adverse events (CTCAE) version 4.0.

Number and percentage of subjects with low calcium as defined by cCa < 7.0 mg/dL, 7.0 - < 7.5 mg/dL, and 7.5 - < 8.3 mg/dL during the study will be summarized by treatment group. Each subject’s lowest corrected calcium value during the study (post-baseline) will be used in this analysis.

Number and percentage of subjects with 2 consecutively low corrected calcium, defined by 2 consecutive post-baseline corrected calcium < 7.5 mg/dL, < 8.0 mg/dL and < 8.3 mg/dL during the study will be summarized by treatment group.

### 9.6.3 Vital Signs and Physical Measurements

Vital signs and physical measurements will be summarized by treatment group at each protocol-specified time point while subjects are receiving IP and at follow up.
9.6.4 Antibody Formation
The incidence and percentage of subjects who develop anti-AMG 416 antibodies at any time will be tabulated by treatment group.

9.6.5 Exposure to Investigational Product
Descriptive statistics will be calculated to describe the exposure to IP by treatment group.

Summary statistics will be provided for: exposure period defined by number of days on IP, the minimum and maximum weekly dose, the cumulative total dose, average weekly dose during EAP, and the number and percentage of subjects receiving each dose level (5 mg, 10 mg, 15 mg, etc. for AMG 416 and 30 mg, 60 mg, 90 mg etc. for cinacalcet) of IP at each visit during the study.

Compliance to the investigational product will be summarized.

9.6.6 Exposure to Concomitant Medication and Dialysate Calcium
The number and proportion of subjects receiving the following selected medications: nutritional vitamin D (vitamin D supplement), vitamin D sterol (active vitamin D), calcium supplements, and phosphate binder, will be summarized by treatment group at baseline and during the study.

Weekly dose of vitamin D sterol will be summarized by treatment group during the study. The dose will be presented as IV paricalcitol equivalent.

Summary of changes in dialysate calcium concentration over the course of the study will be provided.

10. Changes from Protocol-specified Analyses
For the exploratory endpoint of “achievement of mean pre-dialysis serum PTH < 300 pg/mL during the EAP” in the protocol, the threshold is changed from “less than” (<) to “less or equal to (≤)” in this SAP in order to align with the endpoints in AMG 416 pivotal studies 20120229 and 20120230.
11. Literature Citations / References


12. Appendices
Appendix A. Reference Values / Toxicity Grades

For CTCAE grading system V4, please refer to:

Appendix B. Patient-reported Questionnaires

Nausea/Vomiting Symptom Assessment (NVSA)

There are two questions on the NVSA questionnaire:

1. Please circle the one number that best describes how severe your nausea (feel like throwing up) was in the past 24 hours, at its worst?

   No Nausea
   0 1 2 3 4 5 6 7 8 9 10
   Nausea As Severe As you Can Imagine

   2. In the past 24 hours, how many times did you vomit (throw up)? Please enter a number in the space: _____________

Key Secondary Endpoint

The mean number of days of vomiting or nausea per week in the first 8 weeks is one of the two key secondary endpoints.

To prepare data for the primary analysis of this endpoint, we will define whether the subject had vomiting or nausea for each subject and each day. If missing data occurs in either question on certain days of a week, then data will be handled as below:

- If a subject provides a valid response to the severity of nausea question that is > 0 but does not provide a valid response to the episodes of vomiting question, this day will be counted as a day of vomiting or nausea for the purposes of the mean number of days of vomiting or nausea per week endpoint. Similarly, if a subject provides a valid response to the episodes of vomiting question that is > 0 but does not provide a valid response to the severity of nausea question, this day will be counted as a day of vomiting or nausea for the purposes of the mean number of days of vomiting or nausea per week endpoint.

- If the answer to the number of episodes of vomiting question of the NVSA is 99, then the data from this question are considered as missing.

- If, after completing the procedures described in the first and second bullet points above, a subject is missing valid responses to any number of days in that week, then the data for this week are considered missing.
For missing data sensitivity analyses, data will be handled as below:

- If, after completing the procedures described in the first and second bullet points above, a subject is missing valid responses on ≤ 3 days in a week, the available valid responses in the week will be used to estimate the total number of days of vomiting or nausea that are equivalent to a 7-day week using different missing data thresholds. For example, using the ‘at least 6 days of NVSA questionnaire responses in a week’ threshold, a subjects who provides 6 days of NVSA questionnaire responses at week 5, in which the subject had nausea/vomiting on 3 out of 6 days, then the total number of days of vomiting or nausea for this week will be estimated by rounding (3/6)*7 to 1 decimal place, which is 3.5 days. Similarly, using the ‘at least 4 days of NVSA questionnaire responses in a week’ threshold, a subjects who provides 4 days of NVSA questionnaire responses at week 3, in which the subject had vomiting or nausea on 3 out of 4 days, then the total number of days of nausea or vomiting for this week will be estimated by rounding (3/4)*7 to 1 decimal place, which is 5.3 days.

- If, after completing the procedures described in the bullet point above, a subject is missing responses from > 3 days for a particular week, data from this week is considered missing for the analysis. No imputation will be done for data at this particular week.

Other and Exploratory Secondary Endpoints

Analysis of mean severity of nausea in the first 8 weeks, first 16 weeks or in the 26 weeks period will use data from questions 1 on the NVSA questionnaire. For each subject, mean severity of nausea will be calculated by averaging all available daily severity (including zeroes) reported in each of the time periods (i.e. first 8 weeks, first 16 weeks, and during 26 weeks). No imputation will be performed for missing severity of nausea responses.

The analysis of mean number of episodes of vomiting per week in the first 8 weeks, first 16 weeks or in the 26 weeks period will use data from the number of episodes of vomiting question in the NVSA. The number of episodes in a week is the sum of all reported daily episodes in the week. If missing data occurs in the episodes of vomiting question on certain days of a week, then data will be handled in the same fashion as in key secondary endpoint:

- If the answer to the number of episodes of vomiting question of the NVSA is 99, then the data from this questionnaire are considered as missing.

- If, after completing the procedures described in the first bullet point above, a subject is missing valid responses to any number of days in that week, then the data for this week are considered missing in the primary analysis of this endpoint.
For missing data sensitivity analyses, data will be handled as below:

- If, after completing the procedures described in the first bullet points above, a subject is missing valid responses on $\leq 3$ days in a week, the available valid responses in the week will be used to estimate the total number of episodes of vomiting that is equivalent to a 7-day week using different missing data thresholds. For example, using the ‘at least 6 days of NVSA questionnaire responses in a week’ threshold, a subject who provides 6 days of NVSA question 2 responses at week 5, in which the subject has 4 episodes of vomiting in total during the 6 days, then the total number of episodes of vomiting for this week will be estimated by rounding $(4/6)*7$ to 1 decimal place, which is 4.7 episodes. Similarly, using the ‘at least 4 days of NVSA questionnaire responses in a week’ threshold, a subject who provides 4 days of NVSA questionnaire responses at week 3, in which the subject had 5 episodes of vomiting in total during the 4 days, then the total number of episodes of vomiting for this week will be estimated by rounding $(5/4)*7$ to 1 decimal place, which is 8.8 episodes.

- If, after completing the procedures described in the bullet point above, a subject is missing responses from $> 3$ days for a particular week, data from this week is considered missing for the analysis. No imputation will be done for data at this particular week.
Kidney Disease Quality of Life-36 (KDQOL-36™) questionnaire

The Kidney Disease Quality of Life-36 (KDQOL-36™) questionnaires will help keep track of how subjects feel and how well they are able to do usual activities.

**Your Health**

This survey includes a wide variety of questions about your health and your life. We are interested in how you feel about each of these issues.

1. In general, would you say your health is: [Mark an X in the one box that best describes your answer.]

<table>
<thead>
<tr>
<th>Excellent</th>
<th>Very good</th>
<th>Good</th>
<th>Fair</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
</tr>
<tr>
<td>□ 1</td>
<td>□ 2</td>
<td>□ 3</td>
<td>□ 4</td>
<td>□ 5</td>
</tr>
</tbody>
</table>

   The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much? [Mark an X in a box on each line.]

<table>
<thead>
<tr>
<th>Yes, limited a lot</th>
<th>Yes, limited a little</th>
<th>No, not limited at all</th>
</tr>
</thead>
</table>

2. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf: ...................................................... □ 1 ....... □ 2 ....... □ 3

3. Climbing several flights of stairs: ........................................... □ 1 ....... □ 2 ....... □ 3
During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

4. Accomplished less than you would like..............

5. Were limited in the kind of work or other activities

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

6. Accomplished less than you would like..............

7. Didn’t do work or other activities as carefully as usual

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

<table>
<thead>
<tr>
<th>Not at all</th>
<th>A little bit</th>
<th>Moderately</th>
<th>Quite a bit</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ 1</td>
<td>□ 2</td>
<td>□ 3</td>
<td>□ 4</td>
<td>□ 5</td>
</tr>
</tbody>
</table>
These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the past 4 weeks...

<table>
<thead>
<tr>
<th>All of the time</th>
<th>Most of the time</th>
<th>Some of the time</th>
<th>A little of the time</th>
<th>None of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
</tr>
</tbody>
</table>

9. Have you felt calm and peaceful? .................................. □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

10. Did you have a lot of energy? ................................. □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

11. Have you felt downhearted and blue? .......... □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

12. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

<table>
<thead>
<tr>
<th>All of the time</th>
<th>Most of the time</th>
<th>Some of the time</th>
<th>A little of the time</th>
<th>None of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
</tr>
<tr>
<td>□ 1</td>
<td>□ 2</td>
<td>□ 3</td>
<td>□ 4</td>
<td>□ 5</td>
</tr>
</tbody>
</table>
## Your Kidney Disease

How true or false is each of the following statements for you?

<table>
<thead>
<tr>
<th>Statement</th>
<th>Definitely True</th>
<th>Mostly True</th>
<th>Don’t Know</th>
<th>Mostly False</th>
<th>Definitely False</th>
</tr>
</thead>
<tbody>
<tr>
<td>13. My kidney disease interferes too much with my life</td>
<td>![Box] 1</td>
<td>![Box] 2</td>
<td>![Box] 3</td>
<td>![Box] 4</td>
<td>![Box] 5</td>
</tr>
<tr>
<td>14. Too much of my time is spent dealing with my kidney disease</td>
<td>![Box] 1</td>
<td>![Box] 2</td>
<td>![Box] 3</td>
<td>![Box] 4</td>
<td>![Box] 5</td>
</tr>
<tr>
<td>15. I feel frustrated dealing with my kidney disease</td>
<td>![Box] 1</td>
<td>![Box] 2</td>
<td>![Box] 3</td>
<td>![Box] 4</td>
<td>![Box] 5</td>
</tr>
<tr>
<td>16. I feel like a burden on my family</td>
<td>![Box] 1</td>
<td>![Box] 2</td>
<td>![Box] 3</td>
<td>![Box] 4</td>
<td>![Box] 5</td>
</tr>
</tbody>
</table>
During the past 4 weeks, to what extent were you bothered by each of the following?

<table>
<thead>
<tr>
<th></th>
<th>Not at all bothered</th>
<th>Somewhat bothered</th>
<th>Moderately bothered</th>
<th>Very much bothered</th>
<th>Extremely bothered</th>
</tr>
</thead>
<tbody>
<tr>
<td>17. Soreness in your muscles?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Chest pain?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Cramps?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Itchy skin?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Dry skin?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Shortness of breath?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Fainting or dizziness?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Lack of appetite?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Washed out or drained?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Numbness in hands or feet?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Nausea or upset stomach?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28a. (Hemodialysis patient only)</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Problems with your access site?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28b. (Peritoneal dialysis patient only)</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Problems with your catheter site?</td>
<td>□ 1 □ 2 □ 3 □ 4 □ 5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Effects of Kidney Disease on Your Daily Life

Some people are bothered by the effects of kidney disease on their daily life, while others are not. How much does kidney disease bother you in each of the following areas?

<table>
<thead>
<tr>
<th>Not at all bothered</th>
<th>Somewhat bothered</th>
<th>Moderately bothered</th>
<th>Very much bothered</th>
<th>Extremely bothered</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
</tr>
</tbody>
</table>

29. Fluid restriction?...

30. Dietary restriction?...

31. Your ability to work around the house?...

32. Your ability to travel?...

33. Being dependent on doctors and other medical staff?...

34. Stress or worries caused by kidney disease?...

35. Your sex life?...

36. Your personal appearance?...

Thank you for completing these questions!
The scoring procedure for the KDQOL-36™ first transforms the raw precoded numeric values of items to a 0 to 100 range, with higher scores always reflecting better quality of life. Each item is put on a 0 to 100 range so that lowest and highest possible scores are set at 0 and 100, respectively. Scores represent the percentage of total possible score achieved. Table below provide the recoding necessary for each of the KDQOL-36™ questions:

<table>
<thead>
<tr>
<th>KDQOL-36™ Question Number</th>
<th>Raw Value to Coded Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF-12 Questions</td>
<td></td>
</tr>
<tr>
<td>1 1</td>
<td>1→ 100; 2 → 75; 3→50; 4 →25; 5 → 0</td>
</tr>
<tr>
<td>2 1</td>
<td>1→ 0; 2 → 50; 3→100</td>
</tr>
<tr>
<td>3 1</td>
<td>1→ 0; 2 → 50; 3→100</td>
</tr>
<tr>
<td>4 1</td>
<td>1→ 0; 2 →100</td>
</tr>
<tr>
<td>5 1</td>
<td>1→ 0; 2 →100</td>
</tr>
<tr>
<td>6 1</td>
<td>1→ 0; 2 →100</td>
</tr>
<tr>
<td>7 1</td>
<td>1→ 0; 2 →100</td>
</tr>
<tr>
<td>8 1</td>
<td>1→ 100; 2 → 75; 3→50; 4 →25; 5 → 0</td>
</tr>
<tr>
<td>9 1</td>
<td>1→ 100; 2 → 80; 3→60; 4 →40; 5 → 20; 6 → 0</td>
</tr>
<tr>
<td>10 1</td>
<td>1→ 100; 2 → 80; 3→60; 4 →40; 5 → 20; 6 → 0</td>
</tr>
<tr>
<td>11 1</td>
<td>1→ 0; 2 →20; 3→40; 4 →60; 5 → 80; 6 →100</td>
</tr>
<tr>
<td>12 1</td>
<td>1→ 0; 2 →25; 3→50; 4 →75; 5 → 100</td>
</tr>
<tr>
<td>Burden of Kidney Disease</td>
<td>13</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----</td>
</tr>
<tr>
<td></td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Symptoms of Kidney Disease</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>28a</td>
</tr>
<tr>
<td>Effects of Kidney Disease</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>29</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
<tr>
<td>30</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
<tr>
<td>31</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
<tr>
<td>32</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
<tr>
<td>33</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
<tr>
<td>34</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
<tr>
<td>35</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
<tr>
<td>36</td>
<td>$1 \rightarrow 100; 2 \rightarrow 75; 3 \rightarrow 50; 4 \rightarrow 25; 5 \rightarrow 0$</td>
</tr>
</tbody>
</table>

In the second or final step of the scoring process, the coded score in the same scale (0-100) are averaged together to create the scale scores. Questions that are missing answers are not taken into account when calculating the scale scores. Hence the final scale scores represent the average for all questions that the respondent answered.
Functional Living Index-Emesis (FLIE) questionnaire

The Functional Living Index--Emesis (FLIE) has 18 questions. These questions are divided into two domains: Nausea (questions 1-9) and Vomiting (questions 10-18).
**FUNCTIONAL LIVING INDEX - EMESIS (Continued)**

**NAUSEA**

Think about the times you felt **nauseated** during the **past 5 days**.

(Please mark your answers with a vertical mark (I) so that it intersects the horizontal line.)

<table>
<thead>
<tr>
<th>Study Coordinator Use Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q6</td>
</tr>
<tr>
<td>Q7</td>
</tr>
</tbody>
</table>

6. How much has nausea affected your willingness to see and spend time with family and friends in the past 5 days?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>A Great Deal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. Has nausea affected your daily functioning in the past 5 days?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>A Great Deal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Rate the degree to which your nausea has imposed a hardship on you (personally) in the past 5 days.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>A Great Deal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

9. Rate the degree to which your nausea has imposed a hardship on those closest to you in the past 5 days.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>A Great Deal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FUNCTIONAL LIVING INDEX - EMESIS

VOMITING

Think about the times you *vomited* during the past 5 days.
(Please mark your answers with a vertical mark (I) so that it intersects the horizontal line.)

10. How much vomiting have you had in the past 5 days?
   1  2  3  4  5  6  7
   Not at all  A Great Deal

11. Has vomiting affected your ability to maintain usual recreation or leisure activities during the past 5 days?
   1  2  3  4  5  6  7
   A Great Deal  Not at all

12. Has vomiting affected your ability to make a meal or do minor household repairs during the past 5 days?
   1  2  3  4  5  6  7
   Not at all  A Great Deal

13. How much has vomiting affected your ability to enjoy a meal in the past 5 days?
   1  2  3  4  5  6  7
   Not at all  A Great Deal

14. How much has vomiting affected your ability to enjoy drinking liquids in the past 5 days?
   1  2  3  4  5  6  7
   Not at all  A Great Deal
Scoring the FLIE

1. The Visual Analogue Scale (VAS) for a Question

When each question is completed using eDiary device, the response is automatically transformed to a Visual Analogue Scale (VAS) scale which has 100 virtual (behind the scenes) tick marks and corresponds to a value from 0 to 100 (from left to right). The VAS from 0 to 100 is used as raw outcome value of FLIE.

2. Calculating the Score (in mm) for a Question

- For questions 1, 2, 4, 5, 7, 8, 9, 10, 12, 13, 14, 16, and 17:
  Subtract the raw outcome value (in mm) from 100 (Score=100-raw outcome value)

- For questions 3, 6, 11, 15, and 18:
  The score (in mm) for the question is the raw outcome value (in mm) (Score=raw outcome value).
The minimum score for any question is 0 and the maximum score is 100.

3. Handling Invalid Item Responses

In some cases, a respondent may answer one or more items incorrectly because the scale anchors are in the opposite direction on items 3, 6, 11, 15, and 18 compared to the other items. In general, these errors should be relatively rare (<10%). A response on one of these items can be considered invalid if the item score is ±50 or more from the mean score of the items from the same domain (excluding items 3, 6, 11, 15 and 18). In these cases, it is recommended that the item, domain, and total scores should be calculated and reported “as is” (i.e., using the value marked by the respondent). A sensitivity analysis may be performed to assess the impact of these errors by setting the invalid item(s) to missing before calculating the domain and total scores.

4. Calculating a Score (in mm) for a Domain

There are two domains: Nausea (questions 1-9) and Vomiting (questions 10-18). When there are missing data, the domain scores is calculated by multiplying the average item score for the nonmissing items by 9.

- Compute the score (in mm) for each of the 18 questions.
- The score (in mm) for the nausea domain is the sum of the scores for questions 1-9 divided by the number actually answered and multiplied by 9 (i.e., nausea domain score = \([\text{sum of nausea item scores}] / \text{[N of items answered]}\) x 9). The minimum domain score is 0 and the maximum domain score is 900. At least 5 of the 9 FLIE nausea items (i.e., >50% overall item response rate) must be nonmissing to calculate a FLIE nausea domain score.
- The score (in mm) for the vomiting domain is the sum of the scores for questions 10-18 divided by the number actually answered and multiplied by 9 (i.e., vomiting domain score = \([\text{sum of vomiting item scores}] / \text{[N of items answered]}\) x 9). The minimum domain score is 0 and the maximum domain score is 900. At least 5 of the 9 FLIE vomiting items (i.e., >50% overall item response rate) must be non-missing to calculate a FLIE vomiting domain score.

5. Calculating the Total Score

The total score (in mm) is the sum of the nausea and vomiting domain scores. At least 12 of the 18 FLIE items (i.e., ≥66% overall item response rate) must be non-missing, and both the vomiting and nausea domains must be non-missing, to calculate a FLIE total score.

6. Interpretation of Scores

Higher scores indicate less impairment on daily life as a result of nausea or vomiting. No (or minimal) impact on patient’s Daily Life (NIDL) is generally considered a domain score > 54 FLIE points (or > 750 in mm) (Functional Living Index for Emesis, 2010).

Reference


URL: http://www.springerreference.com/docs/html/chapterdbid/74252.html 2011-01-31 23:00:00 UTC