STATISTICAL ANALYSIS PLAN
PERIOD: DOUBLE-BLIND
09 March 2020 Final 1.1

A MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PLACEBO-CONTROLLED STUDY TO EVALUATE THE SAFETY AND EFFICACY OF INTRAVENOUS CR845 IN HEMODIALYSIS PATIENTS WITH MODERATE-TO-SEVERE PRURITUS, WITH A 52-WEEK OPEN-LABEL EXTENSION

PROTOCOL NUMBER CR845-CLIN3103

SPONSORED BY

Cara Therapeutics, Inc.
4 Stamford Plaza
Stamford, CT 06902
(203) 406-3700

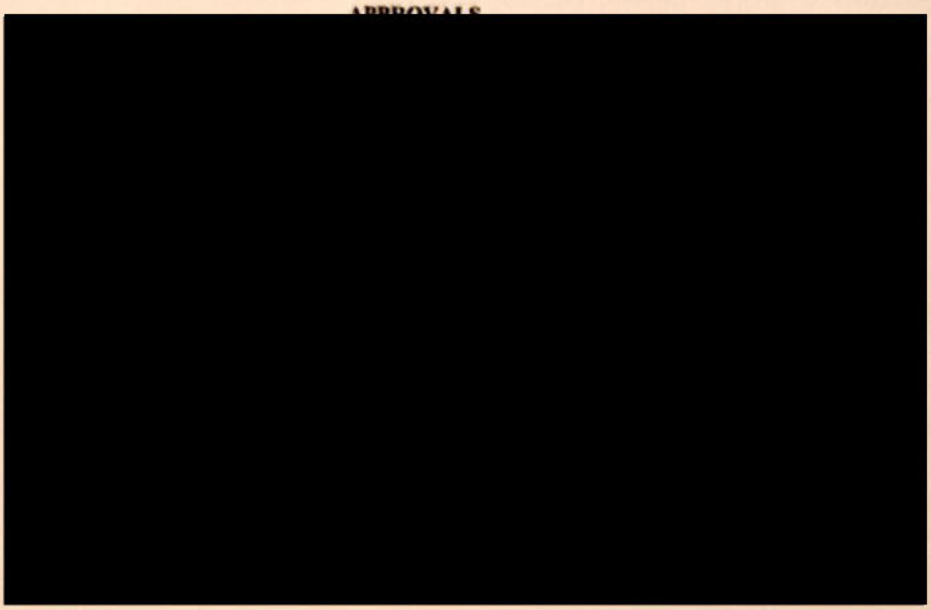
PREPARED BY

Rho, Inc.
6330 Quadrangle Drive, Suite 500
Chapel Hill, NC 27517
Telephone: (919) 408-8000
Fax: (919) 408-0999

DOCUMENT VERSION CONTROL

Version Number	Date	Comments/Changes
Final 1.0	10 September 2019	
Final 1.1	09 March 2020	• Sensitivity analyses added for 5-D and Skindex • K-M plots and SMQ analyses removed • Correction to potassium normal range and hemoglobin units • Added AE analyses by region and dialysis type • Added WI-NRS analyses by region and dialysis type • Minor edits and clarifications around AE analyses

APPROVALS



APPROVALS

Approved:

[redacted]

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Definition
AESI	adverse event of special interest
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
CHW	Cui, Hung, Wang
CI	confidence interval
CKD	chronic kidney disease
CMQ	custom MedDRA query
CRF	case report form
DSMB	Data Safety Monitoring Board
ECG	electrocardiogram
eCRF	electronic case report form
ESA	erythropoiesis-stimulating agents
ESRD	end-stage renal disease
H	above laboratory reference range
ICF	informed consent form
IDMC	Independent Data Monitoring Committee
ITT	intent-to-treat
IV	intravenous or intravenously
L	below laboratory reference range
LS	least squares
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MI	multiple imputation
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed effects model with repeated measures
MNAR	missing not at random
N	within laboratory reference range
NRS	numerical rating scale
PGIC	Patient Global Impression of Change
PP	per protocol
PRO	patient-reported outcome
SAE	serious adverse event
SAP	statistical analysis plan
SOC	system organ class
TEAE	treatment-emergent adverse event
ULN	upper limit of normal

LIST OF IN-TEXT TABLES

1. PURPOSE OF THE ANALYSES

This is a multicenter, randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of intravenous (IV) CR845 for the treatment of chronic kidney disease-associated puritus (CKD-aP) in hemodialysis patients at a dose of 0.5 mcg/kg administered after each dialysis session compared to placebo. The study includes a double-blind phase and an open-label extension phase.

This statistical analysis plan (SAP) provides a detailed description of the strategy and statistical methodology to be used for analysis of data from the double-blind phase of the CR845-CLIN3103 protocol. The open-label analyses will be addressed in a separate document.

The purpose of the SAP is to describe the pre-specified statistical approaches to the analysis of study data prior to database lock. This analysis plan is meant to supplement the study protocol. If differences occur between analyses described in the SAP and the current protocol, those found in this SAP will assume primacy. Any deviations from this plan will be described in the Clinical Study Report.

2. PROTOCOL SUMMARY

2.1 Study Objectives

2.1.1 Primary Objective

The primary objective is to evaluate the efficacy of IV CR845 at a dose of 0.5 mcg/kg compared to placebo, in reducing the intensity of itch in hemodialysis subjects with moderate-to-severe pruritus. This objective will be assessed by comparing the proportion of subjects achieving at least a 3-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity Numerical Rating Scale (NRS) score at Week 12 of the Double-blind Treatment Period between placebo and CR845 0.5 mcg/kg.

2.1.2 Secondary Objectives

The secondary objectives are:

- To evaluate the efficacy of IV CR845 at a dose of 0.5 mcg/kg compared to placebo in improving itch-related quality-of-life measures in hemodialysis subjects with moderate-to-severe pruritus.
- To evaluate the safety of IV CR845 at a dose of 0.5 mcg/kg in hemodialysis subjects with moderate-to-severe pruritus.

2.2 Overall Study Design and Plan

This is a multicenter, randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of IV CR845 at a dose of 0.5 mcg/kg administered after each dialysis session. The study includes a double-blind phase and an open-label extension phase.

2.2.1 Double-blind Phase

The double-blind phase of the study will consist of a Screening Visit, a 7-day Run-in period, and a 12-week Double-blind Treatment Period. Informed consent will be obtained prior to performing any study-specific procedures. The screening visit will occur within 7 to 28 days prior to randomization to assess eligibility. The site has the option to conduct the Screening Visit within the Run-in period at the discretion of the investigator.

Eligible subjects will complete a 7-day Run-in Period during the week prior to randomization to confirm eligibility, preferably starting on the first dialysis session of that week (ie, Monday for subjects on a Monday Wednesday-Friday dialysis schedule or Tuesday for subjects on a Tuesday-Thursday-Saturday dialysis schedule). The purpose of the Run-in Period is to confirm that each subject has moderate to severe pruritus (ie, weekly average worst itch score ≥5), as measured by the patient-daily reported 24-hour Worst Itching Intensity NRS, and to establish a baseline itch intensity.

During the first visit of the Run-in Period, subjects will be trained on completion of the 24-hour Worst Itching Intensity NRS and will start the reporting of their Worst Itching

Intensity NRS daily score. For consistency, subjects will be requested to complete the NRS worksheets (either at home or in the dialysis unit, as required) each day of the run-in period at a similar time of day around the normal start time of their dialysis. Subjects will be trained on other itch-related patient-reported outcome (PRO) worksheets at any time during the run-in period or on Day 1 of the Double-blind Treatment Period.

If subjects continue to meet all inclusion and no exclusion criteria at the end of the 7-day Run-in Period, they will be randomized into the Double-blind Treatment Period in a 1:1 ratio to receive either IV CR845 0.5 mcg/kg or placebo. Subjects will be stratified according to use and no use of concomitant medications to treat itching during the week prior to randomization (Run-in Period) as well as the presence or absence of specific medical conditions. These specific medical conditions include:

- History of fall or fracture (related to fall)
- Confusional state or mental status change or altered mental status or disorientation
- Gait disturbance or movement disorder

Day 1 of the Double-blind Treatment Period will be defined as the day of administration of the first dose of study drug and will occur on the first dialysis session day of the first treatment week (i.e., Monday for subjects on a Monday-Wednesday-Friday dialysis schedule, or Tuesday for subjects on a Tuesday-Thursday-Saturday dialysis schedule). Subjects will be administered CR845 or matched placebo as an IV bolus after the end of each dialysis session during the 12-week Double-Blind Treatment Period. Each subject is to receive CR845 or placebo 3 times weekly for a total of up to 36 doses.

During the Double-blind Treatment Period, subjects will continue to report their daily Worst Itching Intensity NRS score over the previous 24 hours. In addition, during selected study visits (see Appendix 16.1), they will complete other PRO measures [Skindex-10 Scale, 5-D Itch Scale, and Patient Global Impression of Change (PGIC)]. Subjects will be instructed to record PRO measurements, including Worst Itching Intensity NRS scores, at a similar time of day, whether in the dialysis unit (on dialysis days) or at home (on non-dialysis days).

Blood samples for clinical laboratory tests will be collected at Screening and on Days 1 and 85. Blood samples for biomarkers will be collected on Days 1 and 85. Electrocardiograms (ECGs) will be monitored at the Screening Visit and Day 85. Vital signs will be monitored periodically, and adverse events and concomitant medications will be continuously recorded starting at the screening visit until the end of the Double Blind Treatment Period or Early Termination Visit. Use of antipruritic medications, iron, and erythropoiesis stimulating agents (ESAs) and any missed dialysis sessions will be recorded throughout the Double-blind Treatment Period.

A Structured Safety Evaluation will be performed once during the Run-in Period and weekly (preferably on Wednesday/Thursday) during the Double-blind Treatment Period. The Structured Safety Evaluation is performed by study staff using a list of specific signs/symptoms (eg, mental status change, falls, gait disturbance).

2.2.2 Open-label Treatment Period (Extension Phase)

This SAP addresses only analyses for the double-blind phase of the study; however, an open-label extension phase will follow the double-blind phase, and the analyses for the extension phase will be covered in a separate SAP.

2.3 Study Population

Subjects to be included are male and female hemodialysis subjects aged 18 years of age or older with end-stage renal disease (ESRD) who have been on hemodialysis 3 times per week for at least 3 months prior to start of screening, have moderate-to-severe pruritus (mean baseline Worst Itching Intensity NRS score ≥5 with at least 4 out of eight worksheets completed in the run-in period), and meet additional eligibility criteria. A full list of the inclusion and exclusion criteria can be found in the CR845-CLIN3103 protocol (v1.0 10APR2018).

2.4 Treatment Regimens

Subjects will be administered CR845 0.5 mcg/kg or placebo as a single IV bolus 3 times a week after each dialysis session for 12 weeks during the Double-blind Treatment Period.

During the Open-label Treatment Period, subjects will be administered CR845 0.5 mcg/kg as a single IV bolus 3 times a week after each dialysis session for up to 52 weeks.

2.5 Treatment Group Assignments or Randomization

Before the start of the study, computer-generated randomization schedules will be prepared. Randomization will be performed using an interactive voice or web response system. Subjects will be randomized in a 1:1 ratio to receive either CR845 0.5 mcg/kg IV or matching placebo IV during the Double-blind Treatment Period. Subjects will be stratified according to use and no use of concomitant medications to treat itching during the week prior to randomization (Run-in Period), as well as the presence or absence of specific medical conditions, for a total of 4 strata.

All eligible subjects providing consent for participation in the Open-label Treatment Period will receive CR845 at a dose of 0.5 mcg/kg starting on Day 1 of the Open-label Treatment Period.

2.6 Sample Size Determination

The planned sample size for this study is 350 (175 per treatment group) male and female hemodialysis subjects with chronic moderate-to-severe pruritus (mean baseline 24-hour Worst Itching Intensity NRS score ≥5), randomized at approximately 95clinical sites. The sample size may be increased to 500 subjects (250 per treatment group) based on the results of a planned unblinded interim assessment conducted when approximately 50% of the planned 350 first subjects have been randomized and have either completed the

12-week Double-blind Treatment Period or have discontinued from treatment early. The planned interim assessment will be conducted by an Independent Data Monitoring Committee (IDMC). Details related to the sample size re-estimation are included in Section 11.

The sample size calculation is based on results of a completed phase 2 double-blind, placebo-controlled study (CR845-CLIN2101) of CR845 in hemodialysis subjects with ESRD who had moderate-to-severe pruritus. In this study, 30% of subjects randomized to the placebo group reported ≥3-point improvement from baseline with respect to the 24-hour Worst Itching Intensity NRS at the end of treatment (Week 8). The proportion of subjects who received CR845 and reported a similar improvement in itch scores ranged from approximately 60% to 45% (i.e., 30% to 15% difference from placebo), depending on the dose of active study drug (0.5 mcg/kg, 1.0 mcg/kg, 1.5 mcg/kg).

Given a sample size of 350 subjects (175 per treatment group) and assuming a true response rate of 30% for the placebo group and a true response rate of 50% for the CR845 group (defining response as an improvement from baseline ≥3 points with respect to the Worst Itching Intensity NRS at Week 12), a 2-sided continuity corrected Chi-square will have 96% power to detect a treatment difference. The power of this test statistic would be ≥84% for differences from placebo as low as 0.16 (Table 1).

Table 1 Power as a Function of Odds Ratio (N = 175 per arm)

Placebo Response	0.30	0.30	0.30	0.30
CR845 Response	0.50	0.48	0.46	0.45
Odds Ratio	2.333	2.154	1.988	1.909
Power[a]	96%	92%	84%	79%

a. Power for a 2-sided Chi-square continuity-corrected test and a 5% Type 1 error.

Based on results of a planned interim assessment, the sample size may be increased up to 500 subjects (250 per treatment group). Given this maximum sample size, and assuming a true response rate of 30% in the placebo group, a 2-sided continuity corrected Chi-square would have approximately 90% power to detect a treatment difference when the CR845 response rate is 45% (a 15% difference from placebo) (Table 2).

Table 2 Sample Size as a Function of Odds Ratio (90% Power)

Placebo Response	0.30	0.30	0.30	0.30	0.30
CR845 Response	0.50	0.48	0.46	0.45	0.44
Odds Ratio	2.333	2.154	1.988	1.909	1.833
Sample Size[a]	134	164	204	230	262

a. Sample size for a 2-sided Chi-square continuity-corrected test and a 5% Type 1 error.

3. GENERAL ANALYSIS AND REPORTING CONVENTIONS

This section discusses general policies to be employed in the analysis and reporting of the data from the study. Departures from these general policies will be described, if applicable, in the appropriate sections of this SAP. When this situation occurs, the rules set forth in the specific section take precedence over the general policies.

For categorical variables, summary statistics will consist of the number and percentage of subjects in each category. All percentages will be rounded to one decimal point. The number and percentage of subjects will always be presented in the form XX (XX.X%) where the percentage is in parentheses. To ensure completeness, all summaries for categorical and discrete variables will include all categories, even if none of the subjects had a response in a particular category. Denominators for each analysis will be based on the population of interest (e.g., safety population; subjects with non-missing data).

For continuous variables, summary statistics will consist of the number of subjects with data, mean, median, standard deviation, minimum, and maximum values. The summary statistic n will be the number of subjects with non-missing values. All means and medians will be reported to one more significant digit than the values being analyzed. Standard errors and standard deviations will be reported to two more significant digits than the values being analyzed. The minimum and maximum will be reported to the same number of significant digits as the values being analyzed.

For tests of hypothesis of treatment group differences, the associated p-value will be reported. All p-values will be rounded to three decimal places; p-values that round to 0.000 will be presented as "<0.001". P-values are descriptive for outcomes and analyses not included in the multiplicity algorithm.

In general, the baseline value will be considered the last non-missing measurement observed prior to the first dose of study treatment; the NRS will use the mean of the Run-in Period.

For efficacy, subjects will be analyzed according to randomized treatment. For safety analyses, subjects will be analyzed according to the actual treatment received. Data from all sites will be pooled for the purpose of analysis. For analyses involving randomization strata, subjects will be analyzed according to actual strata recorded. It is possible that some values will be updated for the strata variables between the time of the interim analysis and the final database lock; for the final analyses, the values from the locked database will be used.

Data will be listed by treatment and subject. In general, listings will be sorted in the order that columns are displayed, starting with the first column on the left (treatment). Subject listings of data will be presented for all randomized subjects unless specified otherwise.

Unless otherwise specified, summaries will include the following treatment groups:

- CR845 0.5 mcg/kg
- Placebo

SAS statistical software, version 9.4 or higher, will be used for all analyses.

3.1 Assessment Time Windows

For the primary analysis variable, assessment time windows are not needed since the NRS Itch Intensity Assessments Log collects the individual daily scores; average Worst Itching Intensity NRS score for screening and each post-baseline visit week are derived using the daily scores.

Assessments collected by study week that are collected at early termination visits and unscheduled visits will be assigned to a planned visit window, if the early termination or unscheduled visit day falls between +/- 3 days of the planned visit, with the exception of the double-blind Week 12 visit, which will use a window of +1/-3. Should more than one measurement fall within a visit window, priority is given first to the measurement with a non-missing value in the following order: first, the scheduled assessment; second, an early termination visit; and next, the unscheduled assessment closest to the planned day. In the case that two unscheduled visits are equidistant, the latest will be used. This rule will be applied both to efficacy and safety endpoints.

4. ANALYSIS POPULATIONS

Four analysis populations will be used for this study: the Enrolled Population, the Intent-to-Treat (ITT) Population, the Double-blind Safety Population, the Per Protocol (PP) Population.

4.1 Enrolled Population

The Enrolled Population is defined as the group of subjects who sign informed consent.

4.2 Intent-to-treat Population

The ITT Population is defined as the group of subjects who are randomized to a treatment group. Following the ITT principle, subjects in the ITT Population will be analyzed according to their randomized treatment, regardless of the actual treatment received. The ITT population will be used to analyze all efficacy endpoints collected during the double-blind phase.

4.3 Double-blind Safety Population

The Double-blind Safety Population is defined as the group of randomized subjects who received at least one dose of double-blind study drug during the Double-blind Treatment Period. Subjects in the Double-blind Safety Population will be analyzed according to the actual treatment received. The Double-blind Safety Population will be used to analyze all safety endpoints collected during the double-blind phase.

4.4 Per Protocol Population

The PP Population is defined as the subset of subjects in the ITT Population who do not have any major protocol deviations that could affect the efficacy analyses of the double-blind data. An analysis of the primary and secondary efficacy variables for the PP Population will be performed.

The PP Population is defined as subjects who:

- Received at least 80% of the planned study drug doses while in the study
- Received at least one study dose in each of Week 11 and 12 of the Double-blind period, if present through Week 12
- Did not receive a different treatment than the treatment to which they were randomized
- Had a mean baseline Worst Itching Intensity NRS score ≥5.0
- Had a non-missing average 24-hour weekly Worst Itching Intensity NRS score available for at least 75% of study weeks while in the study (weeks with >3 missing daily values are considered missing)
- Did not have significant amounts of restricted and prohibited medications listed in protocol Section 6.4.9 based on medical review.

- Did not have other major protocol violations that would impact efficacy outcomes

Prior to unblinding, the protocol violations and medications will be reviewed in a blinded manner and the PP Population will be determined. Subjects will be analyzed in the treatment arm to which they were randomly assigned regardless of which treatment they received.

5. STUDY SUBJECTS

5.1 Disposition of Subjects

The number of subjects who enrolled, failed screening, were randomized, received treatment, completed treatment, and discontinued from treatment, along with the reason for discontinuation, and number of subjects who entered the Open-label Treatment Period will be presented by treatment group and overall for the Double-blind Treatment Period. Subjects randomized will also be reported by the randomization stratification factors.

The following provides the definitions of the aforementioned groups:

- Enrolled subjects are all subjects who sign informed consent.
- Randomized subjects consist of all screened subjects who have a randomization record in the interactive voice/web randomization system.
- Treated subjects are all subjects who received at least one dose of double-blind study drug.
- Subjects who completed treatment are those randomized subjects with a "Yes" noted for the question "Did the subject complete the Double-blind Treatment Period?".
- Subjects who discontinue treatment early are randomized subjects with "No" noted for the question "Did the subject complete the Double-blind Treatment Period?". Reasons for treatment discontinuation are also collected on this case report form (CRF).
- Subjects who entered the Open-label Treatment Period are those who received at least one dose of study drug in the Open-label Treatment Period.

For all categories of subjects (except for enrolled subjects and screen failures), percentages will be calculated using the number of randomized subjects as the denominator.

Additionally, the analysis populations will be summarized in a table by subject counts, as well as in a subject listing:

- Intent-to-treat Population;
- Double-blind Safety Population;
- Per Protocol Population

The reasons for exclusion from the PP Population will also be summarized in a table by subject counts. Percentages will be calculated using the ITT population as the denominator.

5.2 Protocol Deviations

Protocol deviations will be identified in several ways: through programmatic checks, through medical reviews, and by clinical research associates during site monitoring. Deviations will be classified as minor or major prior to the database lock. Protocol deviations will be summarized by treatment group. All protocol deviations will be listed.

6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic and baseline characteristics will be collected during the Screening Visit.

Descriptive statistics will be provided for all demographic and baseline characteristics based on the Double-blind Safety Population. For categorical variables, the number and percentage of subjects in each category will be presented. For continuous variables, summaries will include the number of subjects with data, mean, median, standard deviation, minimum, and maximum.

All demographic and other baseline characteristics will be provided in a listing.

6.1 Demographic Characteristics

Demographic and baseline variables will be summarized by treatment group and include the following:

- Age at screening (years) as recorded on the CRF
- Age category at screening (<45, ≥ 45-<65, 65-<75, ≥75)
- Gender
- Ethnicity
- Race
- Prescription dry body weight (kg)
- Country
- Region: USA, Western (Canada, UK, Germany, Australia, New Zealand), Eastern EU (Poland, Hungary, Romania, Czech Republic), Asia (Taiwan, South Korea)

6.2 Baseline Disease Characteristics

Baseline characteristics of the disease will also be summarized by treatment group and include the following:

- Duration of pruritus (years)
- Years since ESRD
- Years since chronic kidney disease (CKD)
- Years on chronic hemodialysis
- Etiology of CKD
- Baseline Worst Itching Intensity NRS
- Anti-itch medication use during the Run-in Period (stratification factor)
- Presence of specific medical conditions (stratification factor)

- Dialysis type
- Dialysis type by region

Duration of pruritus (years) will be calculated as:
(Date of the Screening Visit – the start date of the pruritus+ 1)/365.25.

Years since ESRD will be calculated as:
(Date of the Screening Visit – first date of ESRD+ 1)/365.25.

Years since CKD will be calculated as:
(Date of the Screening Visit – first date of CKD+ 1)/365.25.

Years on chronic hemodialysis will be calculated as:
(Date of the Screening Visit – date of first chronic hemodialysis + 1)/365.25.

For each of the above, if partial dates are recorded, the first day of the month will be imputed for missing day, and January for missing month.

6.3 Medical History

Medical history data consisted of a fixed list of conditions to be checked off and any additional self-reported medical conditions not contained on the list. Both types of medical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA); however, the pre-specified medical history will use the fixed list of conditions as preferred terms and map to the most appropriate system organ class (SOC) based on MedDRA and blinded medical review. Medical history terms that were not pre-specified will be coded for SOC and preferred term. Both types of conditions (selected from the fixed list or self-reported) will be combined and summarized by MedDRA SOC, preferred term, and treatment group. Both types of conditions will also be reported together for separate summaries of conditions related to dialysis, conditions worsening during dialysis, and conditions not related to dialysis. The data will also be listed, including the verbatim investigator description of the relevant medical condition, the coded terms (SOC, preferred term), start date, end date, and whether or not the condition is ongoing.

A separate coding listing will be created with all the distinct levels of SOC, preferred terms, and the verbatim investigator description reported in the study. Sorting will be alphabetically by SOC, preferred term, and then verbatim description.

6.4 Prior and Concomitant Medications

All medications, including any antipruritic medications, taken during the 3 months prior to the first dose of study drug on Day 1 through the end of the Double-blind Treatment Phase or early termination (i.e., End-of-Treatment/Early Termination Visit) will be recorded.

These will be coded using the March 2018 World Health Organization Drug Dictionary Enhanced plus Herbal Dictionary. All prior and concomitant medications will be listed. Additionally, a listing for unique medications and their corresponding coding will be presented.

6.4.1 Prior Medication

Prior medications (including vitamins and herbal supplements) are defined as medications collected on the *Previous or Concomitant Medications CRF page* that the subject has taken any time during the last 3 months prior to the first dose of study drug on Day 1 of the Double-blind Treatment Period. Prior medications will be summarized in a table by treatment group using the Double-blind Safety Population. Medications will be reported by drug class (Anatomical Therapeutic Chemical [ATC] Level 3) and ingredient; a subject will be counted only once for each medication.

6.4.2 Concomitant Medication

Concomitant medications used during the Double-blind Treatment Period are medications taken from after the first dose of study drug on Day 1 of the Double-blind Treatment Period through the End-of-Treatment or Early Termination Visit. Concomitant medications during the Double-blind Treatment Period will be summarized by treatment group for the Double-blind Safety Population. Medications will be reported by drug class (ATC Level 3) and ingredient; a subject will be counted only once for each medication.

6.4.3 Anti-Itch Medication

Anti-itch medications are identified as medications where "Yes" is checked on the *Previous or Concomitant Medications CRF page* to the question "Medication was given to treat pruritus."

The prior and concomitant medication summaries described in Sections 6.4.1 and 6.4.2 will be repeated for the anti-itch medications, but presented by ingredient (and not ATC level 3).

7. STUDY DRUG EXPOSURE AND TREATMENT COMPLIANCE

For this study, the duration of double-blind treatment for each individual subject is expected to be 12 weeks, for a total of approximately 36 doses of study drug administered immediately following each dialysis session.

The following variables will be summarized by treatment group to describe the duration of exposure and length of participation in the Double-blind Treatment Period:

- Duration of double-blind treatment (days) as: (Date of first dialysis after last dose) – (Date of first double-blind dose) + 1.
- Duration of double-blind phase (days), from first double-blind dose to final assessment. This will be (End of double-blind participation date) – (Date of first double-blind dose) + 1.
- Average dose per administration (mcg/kg)

The following measures will be used to assess compliance:

- Total number of double-blind doses actually received (1-3, 4-6, 7-9, etc.)
- Total number of dialysis visits logged (1-3, 4-6, 7-9, etc.)
- Number of missed doses
- Number of missed dialysis visits
- Number of subjects with extra doses
- Number of subjects with extra dialysis visits

If a subject receives additional dialysis during a given week for any reason, an additional dose of CR845 will be administered following dialysis. A maximum of 4 doses per week is allowed. No additional doses will be given for subjects receiving an additional unscheduled ultrafiltration treatment. The number of subjects getting such an extra treatment will be summarized.

Missed and extra doses/dialysis will be determined as follows:

1. Individual weeks for each subject are examined.
2. Each subject should have 3 doses per week up to the final week of the period. Anything more will be counted as extra doses; anything less will be counted as missed doses.
3. Subjects who do not complete through the final week of the period will be checked for how far they were into the week that they discontinued: 1,2 days means that they should have 1 dose; 3,4 = 2 doses; 5 or more =3 doses. This will be compared to actual doses for that week to determine missed/extra.
4. The missed and extra doses are then summed across each subject's weeks to get the total missed and extra.

8. EFFICACY EVALUATION

8.1 Overview of Efficacy Analysis Issues

8.1.1 Handling of Dropouts or Missing Data

A variety of approaches will be applied for subjects that have missing Week 12 averages of daily Worst Itching Intensity NRS; full details of these for both the primary analysis and sensitivity analyses may be found in Section 8.2.

Note that a subject must report at least 4 values for a week in order for the weekly mean of the 24 hour Worst Itching Intensity NRS to be non-missing.

Handling of missing data for other endpoints is discussed in the specific sections for those endpoints.

8.1.2 Multicenter Studies

The Week 12 change from baseline will be reported in a separate display with summary statistics by site; likewise, the counts and proportions (out of the ITT Population at that site) of subjects achieving ≥3-point improvement from baseline will be reported by site (for sites that have at least two subjects in each treatment with data present at week 12). Otherwise, data from all sites will be pooled for the purpose of analyses. Note that a covariate for the site's region will be included in the key analyses; the country groupings are given in section 6.1.

8.1.3 Multiplicity Handling

The efficacy of CR845 0.5 mcg/kg compared to placebo in pivotal phase 3 study CLIN3103 will be evaluated based on 1 primary and 7 secondary efficacy endpoints.

Testing of the primary efficacy endpoint will be 2-sided and conducted at the 5% error level. The study will be considered positive if the null hypothesis of no treatment difference in the primary efficacy analysis of the primary endpoint (proportion of subjects achieving ≥3-point improvement from baseline with respect to the Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period) is rejected in favor of the alternative that subjects randomized to CR845 experience significantly less itching compared to subjects randomized to placebo.

To protect the Type 1 error, a gate-keeping strategy will be implemented. Although the p-values corresponding to the hypothesis testing of the secondary variables will be reported, they will only be considered inferential if the primary analysis is statistically significant. Testing of the secondary efficacy endpoints will be performed sequentially at a 2-sided 5% error level in the order specified below. If the test of an endpoint in the sequence is not statistically significant, the p-value for the tests corresponding to the remaining endpoints in the sequence will not be considered inferential and the null hypotheses for the subsequent tests will not be rejected.

- The proportion of subjects achieving ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period will be tested first.

- The proportion of subjects achieving ≥3-point improvement from baseline with respect to the Worst Itching Intensity NRS at Week 8 will be tested next, followed by Week 4.

- Testing will continue with the proportion of subjects achieving a ≥4-point improvement at Weeks 8, followed by Week 4, in an identical manner

- Change from baseline in Skindex-10 total score at Week 12 of the Double-blind Treatment Period using the ANCOVA approach in section 8.2.4.3.
- Change from baseline in 5D total scores at Week 12 of the Double-blind Treatment Period using the ANCOVA approach in section 8.2.4.3.

8.1.4 Region Covariate

For the multiple imputation algorithms and statistical models, geographical region (as described in section 6.1) will be included as a covariate. Should convergence issues occur due to small cell size for the covariate corresponding to region, the following approach will be followed: a) region will first be re-defined, grouping all Asian sites with Eastern Europe; b) If convergence issues still occur after step a) has been completed then all US sites will be grouped with sites from other western countries; c) if convergence issues still remain after implementing steps a) and b), the term for region will dropped from the MI algorithm/statistical model.

If the convergence issue occurs in the generation of a MI dataset, all outcomes based on that dataset will use the same region definition. Additionally, should this occur for the NRS outcome, all sensitivity analysis datasets will use the same region definition and the models base off those datasets will use that region definition.

If the convergence issue occurs in modeling of an outcome, all models for that particular outcome will use the same region definition. For example, if the grouping of the Asian sites with Eastern Europe is required for the Week 4 analysis of ≥4-point improvement in worse itching NRS, then all models at all time points of the ≥4-point improvement in worse itching NRS will use that region definition. However, this would not impact the choice of region for ≥3-point improvement.

8.2 Efficacy Variables

Table 3 presents a summary of the study efficacy variables and types of analyses used to evaluate them.

Table 3 Efficacy Variables and Analysis Methods

Efficacy Variables	Analysis Methods			
	ANCOVA	MMRM	Logistic Regression	CMH
Primary				
≥3-point improvement from baseline in Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period			X	
Secondary				
≥4-point improvement from baseline in Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period			X	
≥3 and ≥4 -point improvement from baseline in Worst Itching Intensity NRS at Week 8 and Week 4 of the Double-blind Treatment Period			X	
Change from baseline in Skindex-10 Scale at Week 12 of the Double-blind Treatment Period	X	X		
Change from baseline Week 12 in total 5-D Itch Scale score at Week 12 of the Double-blind Treatment Period	X	X		
Other Efficacy Variables				
Itch-intensity Variables				
>0-, ≥1-, ≥2-, ≥3-, ≥4-, ≥5-, ≥6-point improvement from baseline in Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period			X	
≥3-point improvement from baseline in Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period by the stratification variables			X	
≥3 and ≥4-point improvement from baseline in Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period by region and dialysis type (each individually)			X	
Change from baseline in the weekly mean of the 24 hour Worst Itching Intensity NRS score at each week of the Double-blind Treatment Period		X		
Proportion of subjects with “Very Much Improved” or “Much Improved” on Patient Global Impression of Change at Week 12 of the Double-blind Treatment Period				X
Worst Itching Intensity NRS Complete Responder			X	
Itch-related Quality-of-Life Variables:				
Change from baseline in total Skindex-10 Scale score, at each week		X		
Change from baseline in each of the three Skindex-10 Scale Scores, at each week	X	X		
≥15-point improvement from baseline in total Skindex-10 Scale score, at each week			X	
Change from baseline in the total 5-D Itch Scale, at each week		X		
Change from baseline in the five 5-D Itch Scale domains, at each week	X	X		
≥5-point improvement from baseline in total 5D Itch Scale, at each week			X	

ANCOVA = analysis of covariance; CMH = Cochran-Mantel-Haenszel exact test; MMRM = mixed effects model with repeated measures; NRS = numerical rating scale

8.2.1 Primary Efficacy Endpoint

The primary efficacy endpoint is defined as the proportion of subjects achieving ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period.

Intensity of itch will be measured using the Worst Itching Intensity NRS (see protocol Appendix 2) on a worksheet in which subjects will be asked to indicate the intensity of the worst itching they experienced over the past 24 hours by marking 1 of 11 numbers, from 0 to 10, that best describes the intensity, where "0" is labeled with the anchor phrase "no itching" and "10" is labeled "worst itching imaginable." Subjects will be provided with these worksheets to record their 24-hour worst itching assessment scores, both at the clinic on dialysis days and at home on non-dialysis days.

The weekly mean of the 24-hour Worst Itching Intensity NRS score will be defined as the sum of the daily Worst Itching Intensity NRS scores reported during a specific week during the Double-blind Treatment Period (e.g., Days 2 to 8, Days 9 to 15, Days 16 to 22) divided by the number of days with non-missing scores for that week. If the daily worst itching score is missing for >3 days during a specific week, the corresponding weekly mean worst itching score will be set to missing. Additionally, subjects who discontinue treatment but continue on study and report NRS scores will have their NRS scores censored following discontinuation of treatment; missing data rules will apply to these censored values as though they had missing values/ dropped out from the study entirely.

The baseline score will be defined as the average of the daily 24-hour Worst Itching Intensity NRS scores collected over the Run-in Period, including pre-randomization assessments collected on Day 1. As defined in the protocol, to be randomized, subjects had to report at least 4 non-missing Worst Itching Intensity scores from the start of the 7-day Run-in Period up to and including the pre-randomization assessment on Day 1. For subjects who deviated from the protocol and were randomized with more than 4 missing Worst Itching Intensity NRS scores during the 7-day Run-in Period, the baseline Itch NRS scores will be calculated using all available non-missing scores collected prior to and including the pre-randomization assessment on Day 1. These subjects will be included in the ITT population and excluded from the Per-Protocol population.

8.2.2 Primary Efficacy Analysis

In the primary efficacy analysis, missing NRS data at the end of Week 12 will be imputed using a multiple imputation (MI) approach, assuming that subjects who discontinue double-blind treatment early would have similar Worst Itching Intensity NRS scores as other subjects in their respective treatment arm who have complete data:

- Intermittent missing NRS scores will first be imputed using the Markov Chain Monte Carlo (MCMC) method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data.
- The monotone missing NRS values will then be multiply imputed with the SAS MI procedure using the monotone regression method.

- For each stage, MI will be performed within treatment group with covariates for baseline NRS score, both randomization stratification factors, region and the non-missing NRS scores for each week. Should convergence issues occur due to small cell size for the categorical covariates corresponding to strata (at either stage), those specific covariates will be removed from the model. Convergence issues with the region covariate will be handled as described in section 8.1.4.
- The proportion of subjects who have an improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score ≥3 points will be calculated for each imputed dataset. Differences between CR845 0.5 mcg/kg and placebo with respect to the primary endpoint will be compared using a logistic regression model containing terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomization, and presence of specific medical conditions. Convergence issues with the region covariate will be handled as described in section 8.1.4.
- Twenty imputations will be performed.
- Results of the logistic regression on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure.

The observed number and proportion of subjects with ≥ 3-point improvement among the non-imputed data will be reported along with the imputed data logistic regression model-based estimates of the proportions of responders, odds ratio, 95% confidence intervals (CIs), and p-value.

Sample size adjustment

The above MI process will be implemented completely independently among subjects contributing the interim results and those following the interim analysis. Likewise, the logistic regression and results described above will be generated independently for both samples, with the samples combined and adjusted as follows:

The final p-value will be calculated using the Cui, Hung, Wang (CHW) procedure where the z-score is a weighted average of the z-score at the interim and the z-score observed for data collected after the interim, following the formula below:

$$Z_{CHW}=Z_{interim} * \sqrt{(n/N)} + Z_{post\text{-}interim} * \sqrt{(1- n/N)}$$

where n is the number of randomized subjects at the interim and N is the initial number of subjects planned (350).

The adjusted odds ratio estimate prior to exponentiating and the 95% CI will be obtained using methodology suggested by Hung and Lawrence:

$$\hat{\delta} = \frac{r_1\hat{\delta}^{(1)} + \sqrt{1-r_1}\sqrt{N^* - r_1}\,\hat{\delta}^{(2)}}{r_1 + \sqrt{1-r_1}\sqrt{N^* - r_1}}$$

where
$r_1 = n/350$

N^* = (Re-estimated total sample size)/350
$\hat{\delta}^{(1)}$ = Least square mean difference from stage 1 analysis
$\hat{\delta}^{(2)}$ = Least square mean difference from stage 2 analysis

95% CI = $\hat{\delta} \pm \left(\frac{\hat{\delta}}{Z_{CHW}} * 1.96 \right)$

Note that it is possible for Z_{CHW} and δ to have opposite directions, particularly when close to zero, in which case the absolute value of the ratio will be used in the above formula.

See Section 11 for further details on the interim analysis.

The primary analysis and each sensitivity analysis will also be performed using the full combined dataset with the missing data imputed separately for interim subjects and post-interim subjects. No CHW adjustment will be conducted. Presentation of results will start with these analyses (additionally including the CHW results) and then follow with the results of the analysis from each individual stage.

8.2.3 Sensitivity Analyses

Sensitivity analyses of the primary efficacy endpoint will be conducted to evaluate the robustness of study results under different assumptions and imputation algorithms. For each of these sensitivity analyses, the final p-value will be calculated based on the CWH procedure, using the formula specified above.

Sensitivity 1: Early Discontinuations as nonresponders

Subjects who discontinue study drug early will be considered nonresponders (including subjects that discontinue study drug, but continue to report NRS scores as described in Section 8.1.1). Subjects who do not discontinue but have missing Week 12 data will be imputed via MI as is done in the primary analysis. The imputed data will be analyzed using a logistic regression model similar to the primary analysis.

Sensitivity 2: Multiple imputation; missing not at random (MNAR)

This sensitivity analysis is an implementation of a pattern mixture model that draws from different populations based on the reason for withdrawal.

- Intermittent missing NRS scores will first be imputed using the MCMC method with the SAS MI procedure, which is appropriate for non-monotonic missing data.
- For subjects who discontinued study drug due to adverse events, NRS scores missing after discontinuation will be imputed using the distribution of the baseline value of all subjects' daily worst itching score assuming a trimmed normal (from 5 to 10).
- For subjects who discontinue due to reasons other than adverse event, missing NRS scores after subjects discontinue study drug early will be multiply

imputed using multiple calls of the SAS MI procedure using data from subjects within the same treatment group who have complete data at that time, including subjects who discontinued due to adverse event. Terms will include baseline values, region, the stratification values, and the weekly data through the time point being imputed. The same MI method will be used for subjects with other monotone missing data.

- The proportion of subjects who have an improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score ≥3 points will be calculated for each imputed dataset
- Similar to the primary analysis, results of the logistic regression on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure.

Sensitivity 3: Tipping point analysis

Multiple imputation with mixed missing data mechanisms (MNAR for a CR845 and MAR [missing at random] for placebo) will be used to assess the robustness of the MAR assumption. This sensitivity analysis is to investigate the departure from MAR assumption by progressively decreasing the treatment differences with respect to the NRS scores over the missing visits in active treatment group until conclusion from the primary analysis is overturned. This will be applied to only Week 12 values. The MI procedure includes the following steps:

- Intermittent missing NRS scores will first be imputed using the MCMC method with the SAS MI procedure, which is appropriate for non-monotonic missing data.
- The monotone missing NRS values will then be multiply imputed with the SAS MI procedure using the monotone regression method.
- For each stage, MI will be performed within treatment group with covariates for baseline NRS score, region, both randomization stratification factors and the non-missing NRS scores for each week. Should convergence issues occur due to small cell size for the categorical covariates at either stage, those covariates will be removed from the model- with the exception of region, as described in the primary.
- For subjects in the active treatment group, a shift parameter running from 0 to 5 points in .25 point increments in PROC MI will be progressively applied to impute the missing data at Week 12, until the p-value is >0.05.
- This sensitivity analysis will not be performed should the initial primary results fail to achieve significance

To evaluate the potential impact of the interim analysis on the properties of statistical inference at the end of the trial, the primary and sensitivity analyses of the primary endpoint will also be presented separately for the sample of subjects enrolled into the study before and after the interim analysis.

8.2.4 Key Secondary Endpoints

The key secondary endpoints are:

- Proportion of subjects achieving ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period;
- Proportion of patients achieving ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 8 of the Double-blind Treatment Period;
- Proportion of patients achieving ≥3-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 4 of the Double-blind Treatment Period;
- Proportion of patients achieving ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 8 of the Double-blind Treatment Period;
- Proportion of patients achieving ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 4 of the Double-blind Treatment Period;
- Change from baseline in itch-related quality of life at the end of Week 12 of the Double-blind Treatment Period, as assessed by the total Skindex-10 Scale score.
- Change from baseline in itch-related quality of life at the end of Week 12 of the Double-blind Treatment Period, as assessed by the 5-D Itch Scale;

8.2.4.1 5-D Itch Scale

The 5-D Itch Scale was developed as a brief but multidimensional questionnaire designed to be useful as an outcome measure in clinical trials. The 5 dimensions of itch being assessed are degree, duration, direction, disability, and distribution (see protocol Appendix 4).

The duration, degree, and direction domains each include 1 item, while the disability domain has 4 items. All items of the first 4 domains were measured on a 5-point Likert scale. The distribution domain included 16 potential locations of itch, including 15 body part items and 1 point of contact with clothing or bandages.

Single-item domain scores (duration, degree, and direction) are equal to the value indicated below the response choice (range 1–5). The disability domain includes 4 items that assess the impact of itching on daily activities: sleep, leisure⁄social activities, housework⁄errands, and work⁄school. The score for the disability domain is achieved by taking the highest score on any of the 4 items. For the distribution domain, the number of affected body parts is tallied (potential sum 0–16), and the sum is sorted into 5 scoring

bins: sum of 0–2 = score of 1, sum of 3–5 = score of 2, sum of 6–10 = score of 3, sum of 11–13 = score of 4, and sum of 14–16 = score of 5.

The scores of each of the 5 domains are achieved separately and then summed together to obtain a total 5-D score. 5-D scores can potentially range between 5 (no pruritus) and 25 (most severe pruritus).

Total 5-D Itch score = duration score (single item) + degree score (single item) + duration score (single item) + maximum (4 disability items) + category score based on sum of affected body parts.

The scoring manual does not give specific direction regarding scoring when some questions are missing; therefore, each domain and the total score will be set to missing when any of their individual components are missing, with the exception of the disability domain. The maximum of any items is present for disability will be used for that domain. Missing data will be handled implicitly in the MMRM model or explicitly via multiple imputation (see below).

8.2.4.2 Skindex-10 Scale

Developed specifically for uremic pruritus, the Skindex-10 Scale (see protocol Appendix 3) is an instrument for measurement of quality of life. Subjects are asked the question "During the past week, how often have you been bothered by" and respond by filling in 1 of 7 circles numbered from 0 (labeled with the anchor phrase "never bothered") to 6 (labeled as "always bothered") for each of the 10 questions.

The total score is the sum of the numeric value of each answered question.

Additionally, the total score is subdivided into 3 domain scores, which are sums of the scores of the following questions: disease domain (questions 1 to 3), mood/emotional distress domain (questions 4 to 6), and social functioning domain (questions 7 to 10).

The scoring manual does not give specific direction regarding scoring when some questions are missing; therefore, the three domains and the total score will be set to missing when any of their individual components are missing. Missing data will be handled implicitly in the MMRM model or explicitly via multiple imputation (see below).

8.2.4.3 Key Secondary Analyses

Worst Itching Intensity ≥4-Point Improvement at Week 12

The proportion of subjects achieving ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Week 12 of the Double-blind Treatment Period will be analyzed following a methodology identical to the one employed for the primary analysis of the primary endpoint (see section 8.2.2). The same dataset will be utilized; separate flags for various cut points of improvement will be

utilized to facilitate these analyses. The analysis will be completed for both the ITT and PP Populations.

Worst Itching Intensity ≥3-Point Improvement and ≥4-Point Improvement at Week 8 and Week 4

The proportion of subjects achieving ≥3-point improvement and ≥4-point improvement from baseline with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS at Weeks 8 and 4 of the Double-blind Treatment Period will be analyzed following a methodology identical to the one employed for the primary analysis of the primary endpoint (see section 8.2.2). The same dataset will be utilized; separate flags for various cut points of improvement will be utilized to facilitate these analyses. The analysis will be completed for both the ITT and PP Populations. These will be tested in the order described in section 8.1.3.

5-D Itch Scale and the Skindex-10 Scale

The 5-D Itch Scale and the Skindex-10 Scale scores will be analyzed only at Week 12 using an analysis of covariance (ANCOVA). The model will contain treatment as fixed effects, with baseline score, region, and the randomization stratification variables as covariates. The baseline 5-D total score and the Skindex-10 total score will be defined as the value collected on Day 1, prior to randomization. Convergence issues with the region covariate will be handled as described in section 8.1.4.

Additionally, subjects who discontinue treatment but continue on study and have 5-D and Skindex-10 scores recorded will have these scores censored following discontinuation of treatment; missing data rules will apply to these censored values as though they had missing values/ dropped out from the study entirely.

For each domain in each questionnaire, missing values at Week 12 will be imputed using an MI approach, assuming that subjects who discontinue double-blind treatment early would have similar 5-D Itch and Skindex-10 scores as other subjects in their respective treatment arm that have complete data. All available visits will be included in the MI to better inform the Week 12 imputed results. The MI approach will proceed as follows, using all ITT subjects (that is, it will not be split into IA subjects and post- IA subjects):

- Intermittent missing scores will first be imputed using the Markov Chain Monte Carlo (MCMC) method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data.
- The monotone missing values will then be multiply imputed with the SAS MI procedure using the monotone regression method.
- For each stage, MI will be performed within treatment group with covariates for baseline score, region, both randomization stratification factors, and all non-missing visit scores for each domain of the questionnaire. Should convergence issues occur due to small cell size for the categorical covariates at either stage, they will be removed from the model- with the exception of region as described

for the primary analysis. Convergence issues with the region covariate will be handled as described in section 8.1.4.

- Twenty imputations will be performed for each domain.
- For each questionnaire, the total score at each visit will be computed from the domain scores for each imputed dataset.
- The ANCOVA analysis described above will be implemented for each imputed dataset.
- Results of the ANCOVA on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure.

Since the observed change from baseline at Week 12 among the non-imputed data will be reported with the MMRM analyses described below, only the LS means, standard errors, 95% CIs, and differences between treatment groups from this ANCOVA will be reported. The analysis will be completed for both the ITT and PP Populations.

Sensitivity Analyses of 5-D Itch Scale and the Skindex-10 Scale

Three additional analyses will be performed for these outcomes. These will include MMRM, missing imputed using control distribution, and a missing imputed using baseline distribution.

- **Mixed effects model with repeated measures**

The 5-D Itch Scale and the Skindex-10 Scale scores will be analyzed using a mixed effects model with repeated measures (MMRM). The model will contain treatment, week, and treatment-by-week interaction as fixed effects, and baseline score, region, and the randomization stratification variables as covariates. The baseline 5-D total score and the Skindex-10 total score will be defined as the value collected on Day 1, prior to randomization. Repeated measures will include values assigned at the end of Weeks 4, 8, 10, and 12 (end of treatment); see Section 3.1 for visit windowing.

It is important to note that, in HD subjects, the study drug administered during the last dialysis of a particular week is not cleared until the first dialysis of the next week. Therefore, measurements that would reflect treatment effect at the end of a specific week (e.g., Week 4) and labelled “end of Week xx” will actually be collected during the first day of the next week (e.g., Week 5).

An unstructured covariance matrix will be used to model the within-subject errors. Should the model fail to converge, a compound symmetric covariance matrix will be used instead. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. Missing scores will not be imputed. Assuming that the data are MAR, the estimates of the treatment differences calculated from the MMRM described above are unbiased.

Standard descriptive statistics will be reported for each time point on the values and changes from baseline along with the least squares (LS) means, standard errors, 95% CIs

and differences between treatment groups reported with LS means, standard errors, and 95% CIs.

- **Missing imputed using control distribution**

For this analysis, the 5-D Itch Scale and the Skindex-10 Scale scores will be analyzed only at Week 12 using an analysis of covariance (ANCOVA). The model will contain treatment as fixed effects, with baseline score, region, and the randomization stratification variables as covariates. The baseline 5-D total score and the Skindex-10 total score will be defined as the value collected on Day 1, prior to randomization. Convergence issues with the region covariate will be handled as described in section 8.1.4.

As with the main MI analysis subjects who discontinue treatment but continue on study and have 5-D and Skindex-10 scores recorded will have these scores censored following discontinuation of treatment; missing data rules will apply to these censored values as though they had missing values/ dropped out from the study entirely.

For each domain in each questionnaire, missing values at Week 12 will be imputed using values from the control arm. This assumes that subjects (regardless of randomized arm) who discontinue double-blind treatment early would have similar 5-D Itch and Skindex-10 scores as subjects in the ***placebo*** arm that have complete data.

All available visits will be included in the MI to better inform the Week 12 imputed results. The MI approach will proceed as follows, using all ITT subjects (that is, it will not be split into IA subjects and post- IA subjects).

- Intermittent missing scores will first be imputed within treatment group using the Markov Chain Monte Carlo (MCMC) method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data. Covariates for baseline score, region, both randomization stratification factors, and all non-missing visit scores for each domain of the questionnaire will be included.
- The monotone missing values will then be multiply imputed with the SAS MI procedure using the monotone regression method. Covariates for baseline score, region, both randomization stratification factors, and all non-missing visit scores for each domain of the questionnaire will be included.
- The MNAR option in Proc MI will be used with the visit values and will reference the placebo group for informing the imputation.
- For each stage, should convergence issues occur due to small cell size for the categorical covariates at either stage, they will be removed from the model- with the exception of region as described for the primary analysis. Convergence issues with the region covariate will be handled as described in section 8.1.4.
- Twenty imputations will be performed for each domain.
- For each questionnaire, the total score at each visit will be computed from the domain scores for each imputed dataset.

- The ANCOVA analysis described above will be implemented for each imputed dataset.
- Results of the ANCOVA on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure.

The LS means, standard errors, 95% CIs, and differences between treatment groups from this ANCOVA will be reported. The analysis will be completed for both the ITT and PP Populations.

- **Missing imputed using baseline distribution**

This imputation approach will largely be identical to the Missing imputed using control distribution approach, but will utilize the baseline distribution to impute missing values; this assumes that the unobserved values for subjects that discontinue will revert to scores similar to the population at baseline.

- Intermittent missing scores will first be imputed within treatment group using the Markov Chain Monte Carlo (MCMC) method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data. Covariates for baseline score, region, both randomization stratification factors, and all non-missing visit scores for each domain of the questionnaire will be included.
- For the monotonic missing scores following discontinuation, each will be imputed using the distribution of the baseline value of all subjects' daily worst itching score assuming a trimmed normal (from 5 to 10).
- Twenty imputations will be performed for each domain.

An identical approach to above will be applied for analysis and combination of the MI repetitions. The LS means, standard errors, 95% CIs, and differences between treatment groups from this ANCOVA will be reported. The analysis will be completed for both the ITT and PP Populations.

8.2.5 Other Efficacy Endpoints

The remaining efficacy analyses will present further presentations of the itch intensity measures, itch-related quality-of-life measures, plus the PGIC.

8.2.5.1 Itch Intensity Measures

- Proportion of subjects who have an improvement from baseline at Week 12 of the Double-Blind Treatment Period with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS scores >0, ≥1, ≥2, ≥3, ≥4, ≥5, and ≥6 will be analyzed and reported in a manner identical to that described in section 8.2.4.3 above. The same dataset from the primary analysis will be used, utilizing additional flags for these cut points of improvement. However, no adjustment for interim analysis will be implemented.

- Proportion of subjects who have an improvement from baseline at Week 12 of the double-blind treatment period with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS scores ≥3 will be reported overall and by each of the randomization stratification variables using the primary analysis and model removing the stratification factor that's being reported from the model. The same dataset from the primary analysis will be used, utilizing additional flags for these cut points of improvement. However, no adjustment for interim analysis will be implemented.
- Proportion of subjects who have an improvement from baseline at Week 12 of the double-blind treatment period with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS scores ≥3 and ≥4 will be reported by region using the primary analysis and model removing region from the model. The same dataset from the primary analysis will be used, utilizing additional flags for these cut points of improvement. However, no adjustment for interim analysis will be implemented.
- Proportion of subjects who have an improvement from baseline at Week 12 of the double-blind treatment period with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS scores ≥3 and ≥4 will be reported by dialysis type using the primary analysis. The same dataset from the primary analysis will be used, utilizing additional flags for these cut points of improvement. However, no adjustment for interim analysis will be implemented. Given the possibility of low counts, model results may not be reported for some types of hemodialysis.
- Proportion of subjects that are complete responders will be reported. A subject that has >= 80% of the non-missing 24-hour Worst Itching Intensity NRS scores equal to 0 or 1 on Week 12 is considered a complete responder; subjects that have less than 4 NRS scores reported or drop out prior to week 12 are also considered non-responders. Differences between CR845 0.5 mcg/kg and placebo will be compared using a logistic regression model containing terms for treatment group, baseline NRS score, region, use of anti-itch medication during the week prior to randomization, and presence of specific medical conditions.
- Change from baseline in the weekly mean of the 24-hour Worst Itching Intensity NRS score at each week of the Double-blind Treatment Period (Week 1 to Week 12). Treatment differences between CR845 and placebo at each post-baseline time point will be analyzed using an MMRM. The model will contain treatment, week, and treatment-by-week interaction as fixed effects, and baseline score, region, and the randomization stratification variables as covariates.

  An unstructured covariance matrix will be used to model the within-subject errors. Should the model fail to converge, a compound symmetric covariance matrix will be used instead. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. Two analyses will be conducted: in the first analysis; missing scores will not be imputed, and in the second analysis, the same dataset from the primary analysis will be used to address missing data. Standard descriptive statistics will be reported for each time point on the values and changes from baseline along with the LS means, standard errors, 95% CIs,

and differences between treatment groups reported with LS means, standard errors, and 95% CIs.

8.2.5.2 Itch-related Quality-of-Life Measures

In addition to the Week 12 results and treatment comparisons, all time points for the 5-D Itch Scale and Skindex-10 will be reported as described in Section 8.2.4.3 , with the exception of the ANCOVA analysis, and utilized as further efficacy endpoints.

Additionally, the ANCOVA and MMRM analyses described in Section 8.2.4.3 for the Skindex-10 will be repeated for all 3 of the Skindex-10 subdomains and the 5 Itch 5-D domains.

The proportion of subjects who have an improvement from baseline in Skindex-10 score of at least 15 points by week will be analyzed and reported along with a logistic regression model-based estimates of the proportions of responders, odds ratio, 95% CIs, and p-value. No imputation of data will be implemented.

The proportion of subjects who have an improvement from baseline in the 5-D Itch Scale total score of at least 5 points by week will be analyzed and reported along with a logistic regression model-based estimates of the proportions of responders, odds ratio, 95% CIs, and p-value. No imputation of data will be implemented.

8.2.5.3 Patient Global Impression of Change

The PGIC is a global PRO measure that assesses the change (improvement or worsening) in overall status of itch relative to the start of the study. The scale has only 1 item, with values ranging from 1 (Very Much Improved) to 7 (Very Much Worse) (see protocol Appendix 5, Section 14.5).

Counts and percentages of subjects for each response will be reported as well as a count of subjects with missing values. All subjects should complete the PGIC Week 12 or early termination, so missing data should be minimal.

The counts and percentages of subjects who rate their itch condition as “Very much improved” or “Much Improved” at the end of Week 12 of the Double-blind Treatment Period/end of double-blind treatment, will be reported. Treatment difference will be tested using the Cochran-Mantel-Haenszel exact test, adjusting for the randomization stratification variables. The Mantel-Haenszel estimate of common odds ratio, exact 95% CI for the common odds ratio, and CMH exact test p-value will be reported for this treatment comparison. Additionally, the exact Clopper Pearson 95% CIs for the proportion of subjects who rate their itch condition as “Very much improved” or “Much Improved” will be reported.

9. SAFETY EVALUATION

9.1 Overview of Safety Analysis Methods

The following assessments will be used to evaluate the safety of CR845 in hemodialysis subjects with moderate-to-severe pruritus:

- Adverse events
- Clinical laboratory parameters
- Vital signs
- 12-lead ECG

All safety analysis will be performed based upon the treatment the subject actually received after randomization. All safety endpoints will be summarized by treatment group (and visit as appropriate) for the Double-blind Safety Population.

9.2 Adverse Events

The period of adverse event reporting will start after the signing of the informed consent form (ICF) through the study Follow-up Visit or Early Termination Visit (or 7 days after the last dose if no Early Termination Visit was conducted). All adverse events that occur during this reporting period will be collected for all subjects, including subjects who are deemed to be screen failures.

Treatment-emergent adverse events (TEAEs) relative to the Double-blind Treatment Period are identified as any adverse event with an onset date after the first dose of the study drug up to the study End of Treatment/Early Termination Visit or 7 days after the last dose if no End of Treatment/Early Termination Visit was conducted, whichever is later.

All tabular adverse event summaries will be for TEAEs.

For events with missing start dates, the following criteria will be used:

- If the start date for a particular event is missing, then the event is considered treatment-emergent for the Double-blind Treatment Period, unless the end date is reported and prior to the Double-blind Treatment Period start date.
- If a partial date is consistent with being in more than 1 period, then it will be assumed to have occurred in the Double-blind Treatment Period,
- If the start time is missing and the start date is the same as the start of a given period, the adverse event will be considered to have occurred in that period.

All adverse events will be coded to SOC and preferred term using the MedDRA version 20.1. The MedDRA treatment dictionary will be used to map adverse events verbatim to SOC and preferred term for standardization and summary purposes.

The incidence of TEAEs will be summarized by treatment group and overall. If a subject experienced more than one episode of an adverse event, the subject is counted once for that preferred term. If a subject had more than 1 adverse event in a SOC, the subject is counted only once in that SOC. The summary tables will include incidence estimates for overall SOC, as well as for preferred terms within each SOC. Incidence for SOC will be presented by decreasing frequency overall and then alphabetically; for preferred terms, incidence will be presented by decreasing frequency overall within each SOC and then alphabetically.

The investigator is to record the severity of each adverse event as mild, moderate, or severe. If the same TEAE occurs for a subject on multiple occasions, the TEAE will be categorized according to the highest severity rating for that TEAE in that subject. If the severity of the TEAE is not reported, then the severity of the TEAE will be counted as severe. For each treatment group and period, the incidence within each category will be presented.

The investigator is to record their opinion on the relationship of each adverse event to study drug (not related, related). If a subject experiences the same adverse event multiple times, the event with the strongest relationship to study drug will be counted. For each treatment group, the incidence within each category will be presented. For the summary of TEAEs by relationship to study drug, if the relationship is missing, it will be counted as related. The incidence of drug-related events will be summarized for each period by treatment group and overall.

Separate tables summarizing the incidence of TEAEs of special interest (AESIs), will be presented. Selected preferred terms shown in Appendix 16.3 are combined into the following categories:

- Gait disturbance
- Fall
- Dizziness
- Somnolence
- Seizure
- Syncope
- Mental status changes
- Mood changes
- Unusual feeling, sensation
- Tachycardia
- Palpitation

Incidence rate will be calculated for serious TEAEs, and TEAEs leading to death, and will be summarized by SOC and preferred term; and by preferred term only for AESIs. Incidence rate will be calculated as (1000 * the number of events)/total person-years, where total person-years is the sum across all subjects in the treatment group of the individual subject risk times, and individual subject risk time is the number of days from

first dose to the last day of the period where an event would be deemed to be treatment emergent.

The summary statistics for the time to first onset among subjects with a given AESI will be presented for each AESI. Additionally, the durations will be summarized with descriptive statistics; if a subject has more than one AE within and AESI category, the longest duration will be summarized. This will also be presented for a subset of common AEs.

To explore the duration and temporal relationship of AESIs, a graphical representation of occurrences and durations of AEs within each AESI will be presented with subjects on the y-axis and time on the x-axis (Swimmer plot).

The following summary tables will be presented for the Double-blind Treatment Period:

- An overall summary showing for each treatment group, the number and percentage of subjects with a TEAE, serious TEAE, related TEAE, severe TEAE, TEAE leading to dose interruption, TEAE leading to study drug discontinuation, TEAE leading to study discontinuation, TEAE of special interest. This table will also include number of events. This display will be repeated for subgroups by randomization stratification variables, region and dialysis type.
- TEAEs by SOC and preferred term
- Serious TEAEs by SOC and preferred term
- TEAEs by SOC, preferred term, and maximum severity
- Related TEAEs by SOC and preferred term
- TEAEs leading to study drug discontinuation by SOC and preferred term
- SAEs by stratification factors, SOC, and preferred term
- TEAEs by region, SOC, and preferred term
- TEAEs by dialysis type, SOC, and preferred term
- Most common TEAEs (≥2% or more of subjects in any treatment group) by preferred term
- AESIs by preferred term
- Related AESIs by preferred term
- AESIs by stratification factors and preferred term
- AESI incidence rate by preferred term
- AESI time to onset and duration summary statistics
- Serious TEAE incidence rate by SOC and preferred term
- TEAEs leading to death incidence rate by SOC and preferred term

- Custom MedDRA query (CMQ) events (see Section 16.2).

In addition, all adverse events will be listed in chronological order including subject identifier, age, race, gender, and all related event status information (start and stop dates, whether the event was ongoing, study day of onset, severity, seriousness, relationship to study drug, action taken with study treatment, and outcome). Note: For the all adverse event listing only, any screen failure subject who has an adverse event after signing ICF will be included for completeness. Separate listings will be generated for serious adverse events (SAEs), deaths, and adverse events leading to treatment discontinuation. Additionally, a coding list of preferred terms and the verbatim text associated with them will be produced.

The incidence of TEAEs in each CMQ category (see Section 16.2) will be summarized. If a subject experienced more than 1 episode of an adverse event, the subject is counted once for that preferred term; the total number of subjects reporting an event in the category will also be reported.

No statistical tests will be performed on adverse events.

9.3 Deaths, Serious Adverse Events, and Other Significant Adverse Events

An SAE is defined as any adverse event occurring at any dose and regardless of causality that results in death, is life-threatening, requires in-patient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, is a congenital anomaly/birth defect, or is an important medical event.

Serious adverse events will be collected on the electronic CRF (eCRF) from the date the ICF is signed up to the Follow-up Visit or 7 days following the last dose of study drug, whichever is later. Serious adverse events that occur after the Follow-Up Visit and up to 30 days thereafter should also be documented on an SAE form if they are deemed by the investigator to be "related" to the study drug. Serious adverse events that occur after the Follow-up Visit and up to 30 days thereafter do not need to be documented on an SAE form if they are deemed by the investigator to be not related to study drug. A more detailed definition of SAEs is provided in Protocol Section 6.5.4. The analysis of SAEs is similar to that of adverse events described in Section 9.2.

Subject deaths are captured on the *Adverse Events* eCRF page. Subject death listings will include all death data available, including date of death and cause of death. Additionally, SAEs and adverse events resulting in discontinuation will be listed as discussed in Section 9.2.

9.4 Clinical Laboratory Evaluation

Summaries of actual values and the changes from baseline to each time point (when applicable) will be presented for quantitative laboratory parameters (e.g., white blood cell count, lymphocyte count). Only data from the central laboratory will be used.

Baseline is defined as the last measurement taken on or prior to the first day of dosing. Note that the Day 1 assessment can be included in the evaluation of baseline if the assessment is performed prior to dosing.

All laboratory evaluation summaries will include the subjects in the Double-blind Safety Population who have at least 1 post-baseline time point (for criteria based on post-baseline assessments) and with both a baseline and at least 1 post-baseline time point (for criteria evaluating changes from baseline).

Laboratory values will be reported in Système International units.

Laboratory test results will be classified according to whether the value was below (L), within (N), or above (H) the laboratory parameter reference range. A summary of treatment-emergent shifts will compare the baseline L/N/H classification for each laboratory test to the highest and/or lowest L/N/H classification during the treatment period. Clinically important laboratory values based on pre-specified criteria appropriate for the study population will also be summarized.

Additionally, alanine aminotransferase, aspartate aminotransferase, bilirubin, and alkaline phosphatase will be presented in a separate table, with 3× and 5× upper limit of normal (ULN) flagged for alanine aminotransferase and aspartate aminotransferase; 2×ULN flagged for bilirubin, and 1.5×ULN flagged for alkaline phosphatase.

Additionally, clinically significant laboratory findings will be reported for hemoglobin, calcium, serum albumin, and phosphate. Ranges are given below.

Test	Low	High
Hemoglobin (g/dL)	<7	>15
Calcium (mg/dL)	<7	>10.5
Serum albumin (mg/dL)	<3	>5.5
Phosphate (mg/dL)	<2.5	>8
Potassium (mmol/L)	<2.5	>7

9.5 Vital Signs, and ECG

9.5.1 Vital Signs

Summary tables for vital signs will include descriptive statistics (number of subjects, mean, standard deviation, median, minimum, and maximum) for baseline and each post-baseline assessment. Descriptive statistics will be calculated on both the actual values and the change from baseline. Baseline is defined as the last measurement taken on or prior to the first day of dosing. Note that the Day 1 assessment can be included in the evaluation of baseline given that the assessment is performed prior to dosing.

All double-blind vital sign summaries will include the subjects in the Double-blind Safety Population who have at least 1 post-baseline assessment (for criteria based on post-baseline assessments) and with both a baseline and at least 1 post-baseline assessment (for criteria evaluating changes from baseline).

Clinically notable vital signs will be identified based on the criteria below. For each vital sign parameter, the number and percentage of subjects with at least 1 notable value will be tabulated by week and overall for the Double-blind Treatment period.

Vital Sign Parameter	Value
Systolic blood pressure	≥180 mm Hg ≤90 mm Hg
Diastolic blood pressure	≥100 mm Hg ≤60 mm Hg
Heart rate	>130 bpm <55 bpm

bpm = beats per minute

All vital signs will be listed in by-subject listings, including visit and collection date/time, and will be sorted by subject identifier and date/time of assessment.

9.5.2 12-Lead ECGs

Standard 12-lead ECG readings with the subject in a supine position will be performed. Electrocardiogram results include an overall interpretation of "normal," "abnormal but not clinically significant," or "abnormal and clinically significant." These results will be tabulated by treatment group and overall at each time point. Electrocardiogram values will also be listed by subject.

Clinically significant abnormalities at screening will be recorded as medical history. Clinically significant abnormalities or worsening of ECG findings observed after the first dose of study drug should be reported as adverse events.

Electrocardiogram results will be listed for each visit, including whether ECG was performed (yes/no), explanation (if not performed), assessment date/time, study date, overall interpretation, and relevant medical history identifier or adverse event identifier if deemed a clinically significant abnormality.

10. OTHER ANALYSES

10.1 Inflammatory Biomarkers

The observed value and the change in inflammatory biomarkers (e.g., interleukin-6, interleukin-8, and granulocyte macrophage-colony stimulating factor) from pre-dose to the end of the Double-blind Treatment Period (Week 12) will be presented by treatment group. Univariate and multivariate analyses of the change in inflammatory biomarkers will be described in an analysis plan, separate from this document.

10.2 Incidence of Infections Related to Uremic Pruritus

Incidence of infections related to uremic pruritus occurring during the Double-blind Treatment Period based on adverse events, hospitalizations, and/or use of antibiotics for treatment of infection related to uremic pruritus will be reported with counts and percentages for each treatment group.

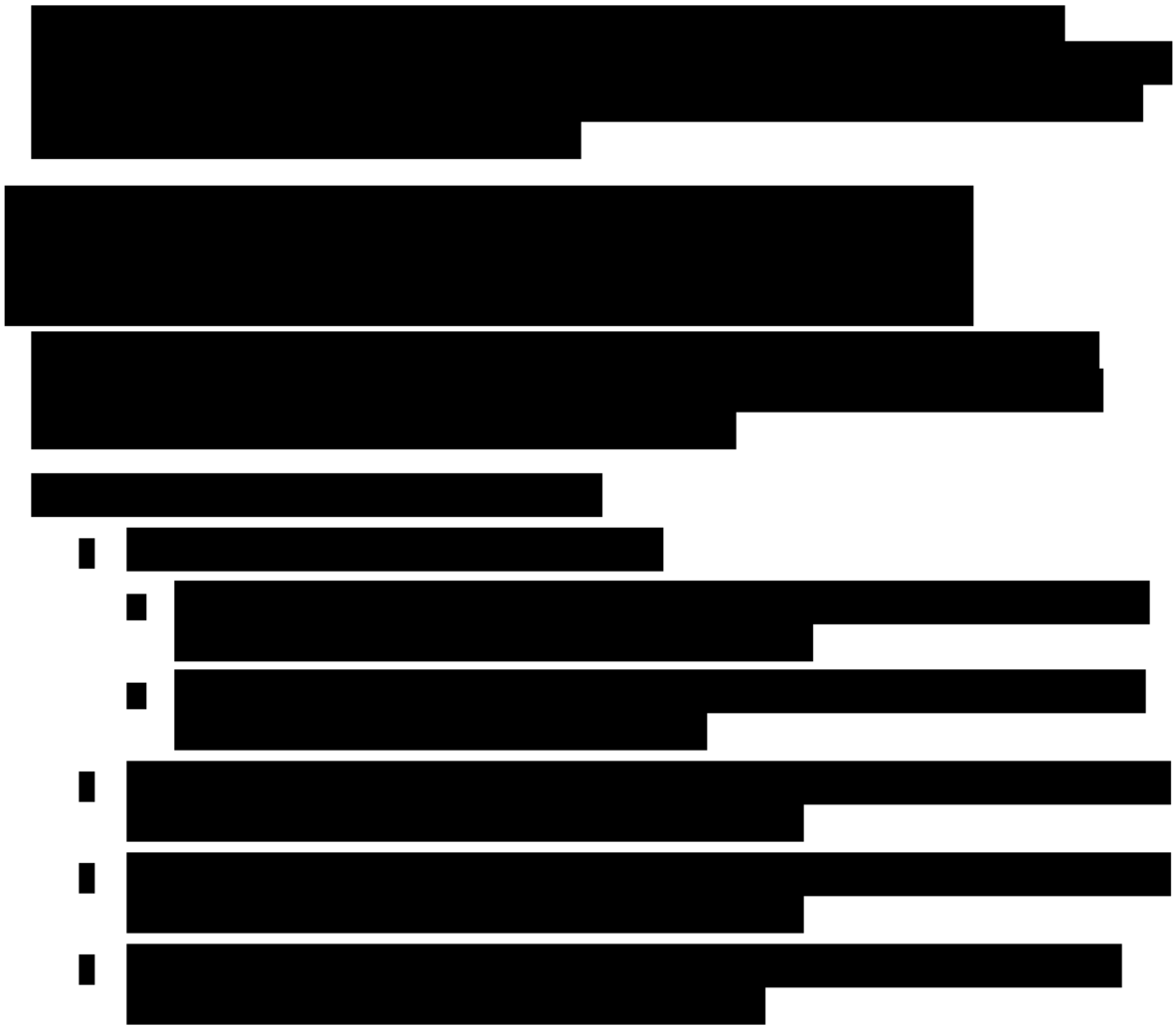
10.3 Hospitalizations, Emergency Department Encounters

The proportion of subjects who had in-patient hospitalization, and the count and percentage for whether dialysis was performed during hospitalization, and summary statistics for duration of stay will be reported by treatment for each study period.

11. INTERIM ANALYSES AND DATA MONITORING

11.1 Sample Size Re-estimation

An unblinded interim analysis for sample size re-estimation will be conducted when approximately 50% of the first 350 subjects have been randomized and have either completed the 12-week treatment period or have discontinued from treatment early. The planned interim assessment will be conducted by an IDMC. An unblinded statistician who will not be part of the study team will provide results of the interim analysis to the IDMC members. Members of the IDMC will not participate in the Data Safety Monitoring Board (DSMB) and will not be members of the study team. During the interim assessment, the study team will remain blinded to the data; however, the IDMC will receive unblinded summary results to implement the decision rule for sample size re-estimation. The IDMC will only communicate the decision either to keep the original sample size or to increase it; no other results will be provided to blinded staff. The DSMB will be made aware of the decision, but not given the results that were the basis of the decision.



11.2 DSMB

The operation of the DSMB will be governed by a charter that will describe the group's meeting frequency, procedures, and requirements for reporting its observations to the sponsor.

Safety data will be reviewed on an ongoing basis by the sponsor and a DSMB (for details see Safety Surveillance Plan and Safety Management Plan). In order to actively monitor subject safety, a prospective approach will be taken to collect and analyze certain adverse events designated as adverse events of special interest.

These analyses are outside the scope of this document.

12. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

There is no change to the analysis planned in the protocol.

13. RANDOM SEEDS FOR MULTIPLE IMPUTATION

For MIs and other instances where random seeds are required, the values used (in order) are:
9392857
3985729
2843255
1086284
6547484
9017456
1597295
3252256
5821871
1852467
6126715
4232132
1841654
4645284
2587345

Note that not all seeds may be required in programming the MIs. If additional seeds are needed, they will be chosen by adding 1 to each of the values above, again using them in order.

14. SAMPLE SAS CODE

Below is the SAS code used to create the dataset to be used in the primary analysis. Weekly NRS scores will be subset on the ITT Population and the dataset will be transposed to create one observation per subject with separate variables for baseline (v1) through week 12 (v13) values. These values as well as a numeric variable for the planned treatment group (trt01pn) and numeric flags for region, anti-itch medication use at baseline (antiIn) and specific medical conditions (specmedn) will be used during the imputation procedures.

The first stage of imputation will create monotone missing data using MCMC and the following code:

```
proc mi data=INDATA seed=9392857 nimpute=20 MAXIMUM=10 MINIMUM=0
out=OUTDATA1 minmaxiter=1000000 ;
mcmc chain=multiple initial = EM(CONVERGE=0.001 maxiter=100000) NBITER=500
NITER=100 impute = monotone displayinit ;
var antiIn specmedn region v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11 v12 v13;
by trt01pn;
run;
```

Using the output data, imputed data following the last non-missing will be reset to missing so that only intermittent values are imputed in this first pass.

The second stage will use this output, then fill in the monotone missing values with the following MI code:

```
proc mi data=OUTDATA1 seed=3985729 nimpute=1 MAXIMUM=10 MINIMUM=0
out=OUTDATA2 minmaxiter=1000000 ;
monotone reg;
var antiIn specmedn region v1 v2 v3 v4 v5 v6 v7 v8 v9 v10 v11 v12 v13;
by trt01pn;
run;
```

15. REFERENCES

Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics. 1999;55(3):853-7.

Lawrence J, Hung H.M. (2003), "Estimation and Confidence after Adjusting the Maximum Information". Biometrical Journal 45, 143–152

Mehta C and Pocock S. Adaptive increase in sample size when interim results are promising: A practical guide with examples. Statistics in Medicine. 2010; 30: 3267–3284.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Arlington, VA: American Psychiatric Association; 2013.

16. APPENDICES

16.1 Schedules of Events

Table 4 Schedule of Events: Double-blind Treatment Phase

Study Procedures	Screening Period		Double-blind Treatment Period[a]						Double-blind End of Treatment[b]/ Early Termination	Follow-Up Period (for patients not participating in Open-label Extension Phase ONLY)
	Screening Visit	Run-in Period	Week 1			Weeks 2 to 12			Week 13	FU Days 1[b]-10
	Day -28 to Day -7	Day -7 to Day 1	M/Tu	W/Th	F/Sa	M/Tu	W/Th	F/Sa		
Visit Days →	-28 to -7	-7 to 1	1	3	5	8	10	12	85	85 to 95
						15	17	19		
						22	24	26		
						29[k]	31	33		
						36	38	40		
						43	45	47		
						50	52	54		
						57[k]	59	61		
						64	66	68		
						71[k]	73	75		
						78	80	82		
Administrative procedures										
Informed consent	X									
Inclusion/exclusion criteria	X		X[c]							
Medical history/Prior Medications (including antipruritic medications)/Demographics	X	X[c]	X[c]							
Randomization			X							
Safety and efficacy evaluations										
Physical examination	X									
Prescription dry body weight	X									
Pre-dialysis 12-lead electrocardiogram	X[d]								X[d]	
Pre-dialysis vital signs	X		X[e]			X[e]			X[e]	X[f]
Hematology, serum chemistry (pre-dialysis)[g]	X		X						X	
Serum pregnancy (females of childbearing potential only)	X								X	

Error! Reference source not found. **(Continued)**

Study Procedures	Screening Period		Double-blind Treatment Period[a]						Double-blind End of Treatment[b]/ Early Termination	Follow-Up Period (for patients not participating in Open-label Extension Phase ONLY)
	Screening Visit	Run-in Period	Week 1			Weeks 2 to 12				
	Day -28 to Day -7	Day -7 to Day 1							First Day of Week 13	FU Days 1[b] -10
			M/Tu	W/Th	F/Sa	M/Tu	W/Th	F/Sa		
Visit Days →	-28 to -7	-7 to 1	1	3	5	8	10	12	85	85 to 95
						15	17	19		
						22	24	26		
						29[k]	31	33		
						36	38	40		
						43	45	47		
						50	52	54		
						57[k]	59	61		
						64	66	68		
						71[k]	73	75		
						78	80	82		
Safety and efficacy evaluations										
Patient training on PRO worksheets	X[h,i]	X[i]	X						X	
Worst Itching Intensity NRS (daily)[j]		X	Record on an ongoing basis						X	
Skindex-10 Scale, 5-D Itch Scale[k]			X			X[k]			X[k]	
Patient Global Impression of Change									X	
Record number of missed dialysis visits and reason(s)			Record on an ongoing basis							
IV administration of study drug			Record on an ongoing basis							
Inflammatory biomarker samples[l]			X						X	
Adverse event monitoring	X	X	Record on an ongoing basis						X	X
Concomitant medications (including antipruritic medications)[m]			X	Record on an ongoing basis					X	X
Structured Safety Evaluation[n]		X		X			X			

EoT = End of Treatment; ET = Early Termination; F = Friday; FU = follow-up; IV = intravenous; M = Monday; NRS = numerical rating scale; PRO = patient-reported outcome; Sa = Saturday; Th = Thursday; Tu = Tuesday; W = Wednesday

a. Each visit during the Double-blind Treatment Period will coincide with the patient's normal dialysis treatments.
b. The End-of-Treatment Visit in the Double-blind Phase will be the first dialysis visit following the last dose of study drug (ie, first dialysis on Week 13 [Day 85]), which also corresponds to Day 1 of the Follow-up Period (FU Day 1). Only patients not participating in the Open-label Extension Phase are required to complete the Follow-up Period.
c. Medical history will be updated on Day 1 with any changes since the Screening Visit, and inclusion/exclusion criteria will be confirmed prior to randomization. Antipruritic medication will be updated at each dialysis visit during the Run-in Period.
d Electrocardiograms must be performed prior to the start of dialysis at Screening, Day 85 (End of Treatment), or at Early Termination visit.
e Pre-dialysis vital signs, including body temperature, heart rate, and blood pressure, will be recorded on Days 1, 15, 29, 43, 57, 71 and 85 (End of Treatment), or at Early Termination visit only when the patient is in a sitting or semi-recumbent position. Heart rate will be measured at each dialysis; if heart rate is clinically significant and outside the prespecified visits per the Schedule of Events, the heart rate will be recorded on the relevant CRF page.
f Pre-dialysis vital signs, including body temperature, heart rate, and blood pressure, will be recorded at the Follow-up Visit (7-10 days after EoT/ET visit). Heart rate will be measured at each dialysis.
g. Blood samples for clinical laboratory evaluation will be taken at Screening, and on Days 1 and 85 (End of Treatment), or at Early Termination visit only.
h. Training on Worst Itching Intensity NRS will be conducted prior to the first day of the Run-in Period (Day -7).
i. Training on Skindex-10 Scale and 5-D Itch Scale may be performed at any time during Screening prior to randomization on Day 1 of the Double-blind Treatment Period.
j. Patients will be requested to complete their Worst Itching Intensity NRS worksheets each day at a similar time (either at home on non-dialysis days around the normal start time of their dialysis or in the dialysis unit). On dialysis days, the worksheets will be completed prior to or during dialysis, but must be completed prior to dosing.
k. 5-D Itch Scale and Skindex-10 Scale completed on Day 1 and the first visit of Weeks 5, 9 and 11 (on Days 29, 57 and 71) and Week 13 (Day 85). The 5-D Itch Scale will preferably be completed first. If the first visit of the week is missed, the patient may complete the worksheets at their next visit for the same week. The worksheets will be completed prior to or during dialysis (preferably within 1 hour of the dialysis), but must be completed prior to dosing.
l. Biomarker samples must be collected prior to the start of dialysis on Day 1 and Day 85.
m. Concomitant medications including antipruritic medication will be updated at each dialysis visit during the Double-blind Treatment Period, and until the end of the Follow-up Period.
n. A list of specific signs/symptoms will be verified with the patient by qualified site staff, preferably to be completed on Wednesday/Thursday each week during the Run-in Period, the Double-blind Treatment Period and the Follow-up Period. It is not to be completed on Monday/Tuesday.

16.2 Custom MedDRA Query (CMQ) Preferred Terms

CMQs

Gait Disturbance
- Ataxia
- Balance disorder
- Coordination abnormal
- Gait disturbance
- Gait inability
- Tandem gait test abnormal

Dizziness
- Dizziness
- Dizziness Postural
- Vertigo

Syncope
- Presyncope
- Syncope

Fall & Potentially Drug Related Injury
- Accident
- Back injury
- Contusion
- Fall
- Fracture (all preferred terms containing “fracture”)
- Head injury
- Injury
- Limb injury
- Muscle contusion
- Muscle injury
- Road traffic accident
- Skeletal injury
- Haematoma
- Subdural hematoma
- Epidural haematoma
- Traumatic haematoma
- Periorbital haematoma
- Ecchymosis
- Subdural hematoma
- Subdural haemorrhage

Mood Changes & Behavioral Changes
- Abnormal behavior

Affect lability
Aggression
Agitation
Anger
Anxiety
Apathy
Blunted affect
Crying
Depressed mood
Disinhibition
Dysphoria
Emotional distress
Emotional poverty
Euphoric mood
Flat affect

Grandiosity
Hostility
Inappropriate affect

Listless
Mood altered
Mood swings
Morose
Nervousness
Patient uncooperative
Restlessness
Social avoidant behavior

Seizures

Autonomic seizure
Clonic convulsion
Drug withdrawal convulsions
Epilepsy
Epileptic aura
Focal dyscognitive seizures
Generalised tonic-clonic seizure
Partial seizures
Partial seizures with secondary generalisation

Seizure
Seizure cluster
Simple partial seizures
Status epilepticus
Tonic convulsion

Mental Status & Cognitive Changes
- Acute psychosis
- Delirium
- Altered state of consciousness
- Hallucination (any preferred term containing "Hallucination")
- Bradyphrenia
- Change in sustained attention
- Cognitive disorder
- Confusional state
- Delusion (any preferred term containing "Delusion")
- Depressed level of consciousness
- Disorientation
- Encephalopathy
- Illusion
- Judgment impaired
- Lethargy
- Mental impairment
- Mental status changes
- Stupor
- Thinking abnormal

Unusual Feeling, Sensation
- Asthenia
- Depersonalisation/derealisation disorder
- Derealisation
- Dissociation
- Feeling abnormal
- Feeling cold
- Feeling despair
- Feeling drunk
- Feeling guilty
- Feeling hot
- Feeling jittery
- Feeling of body temperature change
- Feeling of despair
- Feeling of relaxation
- Feelings of worthlessness
- Malaise
- Psychiatric symptom
- Sensation of foreign body
- Suffocation feeling

Palpitations & Tachychardia
- Palpitations
- Heart rate irregular
- Heart rate increased

Tachycardia
Atrial tachycardia
Junctional ectopic tachycardia
Supraventricular tachycardia
Ventricular tachycardia
Sinus tachycardia
Tachycardia paroxysmal
Tachyarrhythmia
Ventricular tachyarrhythmia

Somnolence
Somnolence
Sleep disorder
Abnormal sleep related event
Microsleep
Sleep attacks
Sedation

16.3 Adverse Events of Special Interest (AESIs) and Preferred Terms

AE of Special Interest Terms	Preferred Terms
Gait disturbance	Gait disturbance
Falls	Fall
Dizziness	Dizziness
Somnolence	Somnolence
Seizures	Seizure
Syncope	Syncope
Mental status changes	Mental status changes
Mood changes	Mood altered
Unusual feeling/sensation	Feeling abnormal
Tachycardia	Sinus tachycardia, Tachycardia, Tachyarrhythmia
Palpitation	Palpitations