STATISTICAL ANALYSIS PLAN
22 February 2017 FINAL

A TWO-PART, MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PLACEBO-CONTROLLED STUDY TO EVALUATE THE SAFETY AND EFFICACY OF INTRAVENOUS CR845 IN HEMODIALYSIS PATIENTS WITH MODERATE-TO-SEVERE PRURITUS

PROTOCOL NUMBER CR845-CLIN2101

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

Version: CARA_CR845-CLIN2101_Draft_SAP_PartA_Final.doc

DOCUMENT VERSION CONTROL

Version Number	Date	Comments/Changes
DRAFT	22 NOV 2016	
FINAL	22FEB 2017	Clarifications and updates to model specifications

APPROVALS

Approved: _____ Date: _____

[redacted]

_____ Date: _____

[redacted]

_____ Date: _____

[redacted]

APPROVALS

Approved: [redacted] Date: 22 Feb 2017

[redacted] Date: _____

[redacted] Date: _____

APPROVALS

Approved: [redacted] Date: 22 Feb 2017

[redacted] Date: 27 Feb 2017

[redacted] Date: _____

APPROVALS

Approved: [redacted] Date: 22 Feb 2017

[redacted] Date: _____

Date: 27FEB2017

TABLE OF CONTENTS

LIST OF ABBREVIATIONS 6

LIST OF IN-TEXT TABLES 7

1. PURPOSE OF THE ANALYSES 8

2. PROTOCOL SUMMARY 9

2.1 Study Objectives 9
2.1.1 Part A 9
2.1.2 Part B 9
2.2 Overall Study Design and Plan 9
2.3 Study Population 10
2.4 Treatment Regimens 10
2.5 Treatment Group Assignments or Randomization 11
2.6 Sample Size Determination 11

3. GENERAL ANALYSIS AND REPORTING CONVENTIONS 12

4. ANALYSIS POPULATIONS 14

4.1 Full Analysis (FAP) Population 14
4.2 Per Protocol Population 14
4.3 Safety Analysis Population 15
4.4 Pharmacokinetic Evaluable Population 15

5. STUDY PATIENTS 16

5.1 Disposition of Patients 16
5.2 Protocol Deviations 16

6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS 17

6.1 Demographic Characteristics 17
6.2 Baseline characteristics 17
6.3 et 18
6.4 Prior and Concomitant Medications 18
6.4.1 Prior Medication 18
6.4.2 Concomitant Medication 18
6.4.3 Anti-Itch Medication 19

7. MEASUREMENTS OF TREATMENT COMPLIANCE 20

8. EFFICACY EVALUATION 21

8.1 Overview of Efficacy Analysis Issues 21
8.1.1 Handling of Dropouts or Missing Data 21
8.1.2 Multicenter Studies 21
8.1.3 Assessment Time Windows 21
8.2 Efficacy Variables 21
8.2.1 Efficacy Variables 23
8.2.1.1 Worst Itching Intensity Numerical Rating Scale (NRS) 23
8.2.1.2 Skindex-10 Scale score 24
8.2.1.3 5-D Itch Scale 24
8.2.1.4 Medical Outcomes Study (MOS) Sleep Scale 25
8.2.1.5 Patient Global Impression of Worst Itch Severity 25
8.2.1.6 Patient Global Impression of Change 26
8.2.1.7 Missed Dialysis Visits and Incidence of Infection 26

8.2.1.8 Iron Status and use of Erythropoiesis-stimulating agents (ESA) and IV iron 26
8.2.1.9 Inflammatory biomarkers....26
8.3 Analysis Methods....27
8.3.1 Primary Efficacy Analyses 27
8.3.2 Secondary Efficacy Analyses....30
8.3.3 Other Efficacy Analyses....30
8.3.3.1 Itch Intensity Measures 30
8.3.3.2 Itch-Related Quality-of-Life Measures 31
8.3.3.3 Missed Dialysis Visits and the Incidence of Infection....32
8.3.3.4 Iron Status and use of Erythropoiesis-stimulating agents (ESA) and IV iron 32
8.3.3.5 Change in inflammatory biomarkers....33
8.4 Examination of Subgroups 33

9. SAFETY EVALUATION 34

9.1 Overview of Safety Analysis Methods 34
9.2 Adverse Events....34
9.3 Deaths, Serious Adverse Events, and Other Significant Adverse Events 36
9.4 Clinical Laboratory Evaluation....36
9.5 Vital Signs, Physical Findings, and Other Observations Related to Safety ..37
9.5.1 Vital Signs....37
9.5.2 12-Lead ECGs....39

10. PHARMACOKINETIC EVALUATION 40

10.1 Plasma Concentration Data 40
10.2 Pharmacokinetic Parameters in Plasma.....**Error! Bookmark not defined.**
10.2.1 Pharmacokinetic Endpoints Estimation .**Error! Bookmark not defined.**

11. OTHER ANALYSES 42

12. INTERIM ANALYSES AND DATA MONITORING 43

13. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL....44

14. REFERENCES 46

15. APPENDICES 47

15.1 Schedule of Events 47

LIST OF ABBREVIATIONS

Abbreviation	Definition
AUC	area under the curve
Cmax	maximum concentration
CRF	case report form
DMC	Data Monitoring Committee
ECG	electrocardiogram
eCRF	electronic case report form
H	above laboratory reference range
ICF	Informed consent form
IV	intraveneous
IVRS/IWRS	Interactive voice response system/interactive web response system
L	below laboratory reference range
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed effects model with repeated measures
MOS	Medical Outcomes Study
N	within laboratory reference range
NRS	numerical rating scale
PK	pharmacokinetic(s)
PNS	Peripheral nervous system
PRO	patient reported outcome
Rauc	accumulation ratio for AUC
Rcmax	accumulation ratio for Cmax
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

Table 8-1 Efficacy Variables and Analysis Methods..... 21

1. PURPOSE OF THE ANALYSES

CR845-CLIN2101 is a two-part, multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of CR845 in hemodialysis patients with moderate-to-severe pruritus. This study is sponsored by Cara Therapeutics, Inc.

This statistical analysis plan (SAP) provides a detailed description of the strategy and statistical methodology to be used for analysis of data in Part A of the CR845-CLIN2101 protocol.

The purpose of the SAP is to ensure the credibility of the study findings by pre-specifying the statistical approaches to the analysis of study data prior to data base lock for Part A. 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 (CSR).

2. PROTOCOL SUMMARY

2.1 Study Objectives

2.1.1 Part A

Primary Objective
The primary objective is to evaluate the efficacy of 3 dose levels of IV CR845 administered after each dialysis session, compared to placebo, in reducing the intensity of itch as assessed by the change from baseline to the last week of the 8-week Treatment Period with respect to the Worst Itching Intensity Numerical Rating Scale (NRS) score in hemodialysis patients with moderate-to-severe pruritus.

Secondary Objectives
The secondary objectives are:

- To evaluate the efficacy of 3 dose levels of IV CR845 administered after each dialysis session, compared to placebo, in improving itch-related quality-of-life measures as assessed by the Skindex-10 Scale over an 8-week Treatment Period in hemodialysis patients with moderate-to-severe pruritus.

- To evaluate the safety of 3 dose levels of IV CR845 administered after each dialysis session over an 8-week Treatment Period in hemodialysis patients with moderate-to-severe pruritus.

- To evaluate the Pharmacokinetic (PK) of 3 dose levels of IV CR845 administered after each dialysis session over an 8-week Treatment Period in hemodialysis patients with moderate-to-severe pruritus.

2.1.2 Part B

One dose level will be tested in Part B after selection from the doses tested in part A and which will be the dose which represents best efficacy and safety. Part B analyses are outside the scope of this document.

2.2 Overall Study Design and Plan

This is a two-part, multicenter, randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of IV CR845 administered after each dialysis session. In Part A, 3 dose levels of CR845 will be evaluated relative to placebo over an 8-week Treatment Period in hemodialysis patients with moderate-to-severe pruritus. A subset of patients (approximately 10 patients per treatment group) will also be consented for the collection of blood samples for evaluating the PK profile of CR845. Based on the overall safety and efficacy profile from Part A, 1 dose of CR845 will be selected for Part B and

evaluated relative to placebo over a 12-week Treatment Period in hemodialysis patients with moderate-to-severe pruritus.

Part A consists of a Screening Period, an 8-week Treatment Period, an End-of-Treatment Visit (approximately 3 days after the last dose), and a Follow-Up Visit (approximately 1 week after the last dose).

Part B consists of a Screening Period, a 12-week Treatment Period, an End-of-Treatment Visit (approximately 3 days after the last dose), and a Follow-Up Visit (approximately 1 week after the last dose). Part B will not be initiated until an unblinded review of the safety and efficacy data from Part A has been completed.

The Screening Period will occur between 8 and 14 days prior to randomization following the signing of the informed consent form.

Eligible patients will then be trained on completion of the 24-hour Worst Itching Intensity NRS scale and required to record their Worst Itching Intensity NRS score each day for at least 7 days prior to randomization. For consistency, patients will be requested to complete the NRS worksheets (either at home or in the dialysis unit, as required) at a similar time of day around the normal start time of their dialysis (+/- 2 hours). During the week prior to randomization up to the first day of study drug (Day 1), patients will be trained on other PRO measures, including quality-of-life measurements. Day 1 of the Treatment Period will be defined as the day of the first dose of administration of study drug.

All study visits (scheduled and unscheduled) will be conducted on dialysis days, including the end of study visit.

A final safety Follow-Up Visit will be conducted 7 days (+3) after the last dose of study drug.

2.3 Study Population

For both Parts A and B of the study, male and female hemodialysis patients aged 18 years of age or older with end stage renal disease who have been on hemodialysis 3 times per week for at least 3 months prior to start of Screening, who have moderate-to-severe pruritus (mean baseline Worst Itching Intensity NRS score > 4), and who meet additional eligibility criteria are included. A full list of the inclusion and exclusion criteria can be found in the CR845 protocol.

2.4 Treatment Regimens

Part A: Four groups with a sample size of approximately 40 patients each will be administered CR845 0.5, 1, or 1.5 mcg/kg, or placebo as a single IV bolus 3 times a week immediately after each dialysis session for 8 weeks.

Part B: Two groups with a sample size of approximately 120 patients each will be administered CR845 or placebo as a single IV bolus 3 times a week immediately after

each dialysis session for 12 weeks. The dose of CR845 and the total sample size will be confirmed based on an unblinded review of the safety and efficacy data collected in Part A.

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 (IVRS/IWRS).

For Part A, patients will be randomized in a 1:1:1:1 ratio to receive CR845 1.5 mcg/kg, CR845 1.0 mcg/kg, CR845 0.5 mcg/kg, or placebo. Patients will be stratified according to their use or non-use (i.e., 2 strata will be defined) of concomitant medications to treat their itch and their inclusion in the PK sample of approximately 10 patients per treatment arm.

For Part B, patients will be randomized in a 1:1 ratio to receive either CR845 or placebo, with the dose of CR845 determined based on safety and efficacy determined in Part A. Patients enrolled in Part B will also be stratified according to their use or non-use of concomitant medications to treat their itch. For both Part A and B, patients will be categorized as "using anti-pruritic medications" if they have used >1 dose of potentially anti-pruritic medications used to treat their uremic pruritus during the week prior to randomization (e.g. antihistamine (H1), pregabalin, gabapentin, nalbuphine, naltrexone, corticosteroids, or buprenorphine).

2.6 Sample Size Determination

The planned sample size for this clinical investigation is approximately 400 male and female hemodialysis patients with persistent moderate-to-severe pruritus (mean baseline Worst Itching Intensity NRS score > 4).

Part A: Approximately 160 male and female hemodialysis patients with moderate-to-severe pruritus will be randomized at approximately 40 clinical sites. A subset of up to 40 patients (approximately 10 per treatment group) will be consented for the collection of blood samples for evaluating the PK profile of CR845. No formal sample size calculation was performed to select this sample size. However, a sample size of 40 patients per group is adequate to provide an appropriate estimate of the magnitude and variability of treatment effect at each dose, and select the appropriate dose to be further evaluated compared to placebo in Part B.

Part B: Approximately 240 male and female hemodialysis patients with moderate-to-severe pruritus will be randomized at approximately 60 clinical sites. Details of the assumptions and the calculations for sample size determination can be found in Section 12.2 of the protocol.

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 may be given in the specific detailed 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 patients in each category. All percentages will be rounded to one decimal point. The number and percentage of patients 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 patients had a response in a particular category. Unless otherwise noted, for all percentages, the number of patients in the analysis population for the treatment group will be the denominator.

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

In general, the baseline value will be considered the last non-missing measurement observed prior to the first dose of study treatment.

For efficacy, patients will be analyzed according to randomized treatment. For safety analyses, patients will be analyzed according to the actual treatment received. Data from all sites will be pooled for the purpose of analysis.

Aside from the primary analysis, P-values are considered descriptive in nature; there are no adjustments for multiple comparisons.

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

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

- CR-845 1.5 mcg/kg
- CR-845 1.0 mcg/kg
- CR-845 0.5 mcg/kg
- CR-845 (All Doses Combined), used for primary and secondary analysis only
- Placebo

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

4. ANALYSIS POPULATIONS

Four analysis populations will be used for this study: the Full Analysis (FAP) population, the Per Protocol (PP) population, the Safety population (same as FAP population), and the Pharmacokinetic (PK) Evaluable population.

The analysis of the primary efficacy and secondary endpoints will be performed using the FAP population. Sensitivity analyses of the primary and secondary efficacy endpoints will also be performed on the PP population (if the criteria are met). For exploratory efficacy endpoints, analyses will be performed on the FAP population only. All safety analyses will be performed using the Safety Population. Pharmacokinetics analyses will be performed using the PK Evaluable population.

4.1 Full Analysis (FAP) Population

The Full Analysis Population is defined as the group of all randomized patients who received at least 1 dose of double-blind study drug. Following the intent-to-treat principle, patients in the Full Analysis Population will be analyzed according to their randomized treatment, regardless of the actual treatment received.

The Full Analysis Population will be used to analyze all efficacy endpoints.

4.2 Per Protocol Population

The Per-Protocol Population is defined as the subset of patients in the Full Analysis Population who do not have any major protocol deviations that could affect the efficacy analyses.

The Per Protocol Population is defined as patients who:

- Received at least 80% of the planned study drug doses
- Had a mean baseline Worst Itching Intensity score > 4.0
- Had a non-missing average 24-hour weekly Worst Itching NRS score available for at least 75% of study weeks (weeks with >3 missing daily values are missing)
- Did not have significant amounts of restricted and prohibited medications listed in protocol Section 10.2 based on medical review.
- Did not have 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 per-protocol population will be determined. Patients will be analyzed in the treatment arm to which they were randomly assigned regardless of which treatment they received. Patients receiving different treatment than the treatment to which they were randomized will be considered major protocol deviations and will not be included in the Per-Protocol Population.

An analysis of the primary and secondary efficacy variables for the Per-Protocol Population may be performed if more than 20% of the patients in the Full Analysis Population are excluded.

4.3 Safety Analysis Population

The Safety Analysis Population is identical to the Full Analysis Population. However, patients in the Safety Population will be analyzed according to their actual treatment.

The Safety Population will be used to analyze all safety endpoints.

4.4 Pharmacokinetic Evaluable Population

The Pharmacokinetic Evaluable Population is defined as all patients who received CR845 and had sufficient plasma concentrations for PK analysis.

5. STUDY PATIENTS

5.1 Disposition of Patients

The number of patients screened, randomized, treated, completed, or discontinued from the study, along with the reason for discontinuation, will be presented overall and by treatment group. Patient count by analysis population will also be tabulated.

The following provides the definitions of the aforementioned groups:

- Screened patients have a date of informed consent.
- Randomized patients consist of all screened patients who have met the eligibility criteria (based on Inclusion/Exclusion Summary CRF) or were waived and have a date of randomization, regardless of whether treatment was subsequently received.
- Treated patients are all patients were randomized and who received at least one dose of study treatment.
- Completers are randomized patients with 'Yes' noted for the question 'Did the patient complete the 8 week treatment period' on the End of Treatment (Study Completion Status) CRF.
- Patients who discontinued the study are randomized patients with 'No' noted for the question 'Did the patient complete the *8* week treatment period' on the End of Treatment (Study Completion Status) CRF. Reasons for study discontinuation are also collected on this CRF.

For all categories of patients (except for the screened patients and enrolled patients) percentages will be calculated using the number of randomized patients as the denominator.

Additionally, the analysis populations will be summarized in a table by patient counts, as well as in a patient listing, for the randomized population.

- Full Analysis population;
- Per Protocol population;
- Safety population;
- PK Evaluable population

5.2 Protocol Deviations

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

6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic and baseline characteristics were collected during the screening visit.

Descriptive statistics will be provided for all demographic and baseline characteristics based on the Safety population. For categorical variables, the number and percentages of patients in each category will be presented. For continuous variables, summaries will include the number of patients 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)
- Age category at screening (<45, 45-<65, 65-<75, ≥75)
- Gender
- Ethnicity
- Race
- Height (cm)
- Target Dry Body Weight (kg)

Age will be calculated using the following formula:
Age = (Date of the screening visit – the date of birth + 1)/365.25.

6.2 Baseline characteristics

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

- Duration of pruritus (years)
- Patient self-characterization of disease severity
- Localized itch (Palms of Hands) (Yes/No)
- Pruritus during dialysis session only (Yes/No)
- Years since ESRD
- Months on chronic hemodialysis
- Baseline single pool Kt/V (spKt/V)
- Baseline urea reduction ratio (URR)

- Etiology of chronic kidney disease

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

Similarly, years since ESRD will be calculated as:
(Date of the screening visit – first date of ESRD+ 1)/365.25.

Months on chronic hemodialysis will be calculated as:
(Date of the screening visit – date of first chronic hemodialysis + 1)/30.4.

6.3 Medical History

Medical history data will be coded using MedDRA version 19.0 available at the time of reporting. A summary table by treatment group will be presented by MedDRA system organ class and preferred term. The data will also be listed including whether or not the patient reported any medical history (yes/no), if yes, system organ class, preferred term, recorded medical condition, start date, end date, and whether or not the condition is ongoing.

6.4 Prior and Concomitant Medications

All medications, including any itch medications, must be recorded within 14 days prior to the start of Screening through the Follow Up Visit. These will be coded using the March 2016 WHO Drug Dictionary. All prior and concomitant medications will be listed. Additionally, a listing unique medications and their corresponding coding will be presented.

6.4.1 Prior Medication

Prior medications are defined as medications collected on the *Previous or Concomitant Medications CRF page* that the patient has taken any time during the 14 days prior to the start of Screening through to the first dose of study drug on Day 1. Prior medications will be summarized in a table using the Safety Population by treatment group. Medications will be reported by drug class (ATC3) and preferred term; a patient will be counted only once for each of these.

6.4.2 Concomitant Medication

Concomitant medications are defined as medications collected on the *Previous or Concomitant Medications CRF page* that are taken from after the start of the first dose of study drug on Day 1 through the end of the study (i.e., Follow-Up Visit). Concomitant

medications will be summarized for the Safety Population by treatment group. Medications will be reported by drug class (ATC3) and preferred term; a patient will be counted only once for each of these.

6.4.3 Anti-Itch Medication

Prescribed anti-itch medications are identified as medications where 'Yes' is checked on the *Previous or Concomitant Medications CRF page* to the question 'Was this medication used to treat pruritus?'.

The prior and concomitant medication summaries described in Sections 6.4.1 and 6.4.2 will be repeated for the prescribed anti-itch medications

In addition, the number of patients using anti-itch medications and the number of days of anti-itch medication use as recorded on the *Anti-Itch Medications CRF page* will be summarized by week in the study. The mean number of days will also be plotted by treatment group.

7. MEASUREMENTS OF TREATMENT COMPLIANCE

For Part A, the duration of treatment for each individual patient is expected to be 8 weeks, for a total of approximately 24 doses of study drug (3 doses per week) administered immediately following each dialysis session. The overall study duration for each individual patient is expected to be up to 11.5 weeks.

The following variables will be summarized by treatment group.

- Duration of treatment (days)
- Duration of study (days)
- Total number of doses actually received (1-3, 4-6, 7-9, etc)
- Number of missed doses

Duration of treatment (days) = (Date of first dialysis after last dose) – (Date of first dose) + 1.

Duration of study (days) = (End of participation date) – (Date of first dose) + 1.

Note that per protocol patients receiving a 4th dialysis treatment during a given week (e.g., for fluid overload) should receive an extra dose of CR845 following dialysis after this treatment. The number of patients getting such an extra treatment will be summarized.

Missed and extra doses will be determined as follows:
1) Individual weeks for each patient are examined
2) Apart from the last week, each patient should have 3 doses per week. Anything more would count as extra doses; anything less is counted as missed doses
3) Patients that complete at or after day 54 should also have 3 doses for their last week (and anything going beyond week 8 is included in week 8)
4) Patients that do not complete through day 54, 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+=3 doses. This is compared to actual doses for that week to get determine missed/extra
5) The missed and extra doses are then summed across each patient'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

For the primary efficacy analysis (Part A), missing daily worst itching scores will not be imputed. Assuming that the data are missing at random (MAR), the estimates of the treatment differences calculated from the MMRM model described above are unbiased.

Depending on the pattern and the amount of early treatment discontinuations in Part A, sensitivity analyses may be performed for Part A in order to provide estimates of treatment effect under different imputation algorithms and help refine the sample size for Part B. These sensitivity analyses are described in section 8.3.1.

8.1.2 Multicenter Studies

The primary efficacy outcome analysis will be repeated with summary statistics by site, but otherwise data from all sites will be pooled for the purpose of analyses.

8.1.3 Assessment Time Windows

For the primary analysis variable, assessment time windows are not needed since the NRS Itch Intensity Assessments Log collects the individual scores and average NRS score for screening and each post-baseline visit week. Similarly, the assessments for other PROs are collected by visit week on the CRF.

8.2 Efficacy Variables

Table 8-1 Efficacy Variables and Analysis Methods

Efficacy Variables (Part A)	MMRM	Logistic Regression	Fisher's Exact Test	Descriptive Summary Table
Primary				
Change from baseline to Week 8 in weekly mean of daily 24-hour Worst Itching Intensity NRS	X			
Secondary				
Change from baseline to Week 8 in total Skindex-10 Scale score	X			
Other Efficacy Variables				
Itch-intensity Variables				
Change from baseline in weekly mean of daily 24-hour Worst Itching Intensity NRS score, Week 1 to Week 8	X			
Treatment Response at Week 8				X
Proportion of Patient that have a ≥20% Treatment Response		X		
Itch-related quality of life Variables:				
Change from baseline in total Skindex-10 Scale score, Week 2 to Week 8	X			
Change from baseline to Week 8 in each of the 3 Skindex-10 Scale Scores	X			
Change from baseline to Week 8 in the total 5-D Itch Scale	X			
Change from baseline to Week 8 in the total Sleep Disturbance Subscale of the MOS Sleep Scale	X			
Proportion of patients with 'Very Much Improved' or 'Much Improved' on Patient Global Impression of Change at Week 8			X	
Proportion of patients with 1pt. improvement in Patient Global Impression of Worst Itch Severity			X	
Changes in ESA from pre-dose to the end of the Treatment Period.	X			
Missed Dialysis Visits and Incidence of Infection				
Total number of missed dialysis visits during treatment period				X
Proportion of patients who missed 1 or more visits				X
Rate of infections		X		

8.2.1 Efficacy Variables

The baseline value of the primary efficacy endpoint will be calculated using all the 24-hour Worst Itching Intensity NRS scores reported during the last 7 days prior to randomization, including Day 1 scores collected prior to the administration of the first dose of double-blind study drug. The baselines for other PRO measures are the pre-dose assessments collected on Day 1.

8.2.1.1 Worst Itching Intensity Numerical Rating Scale (NRS)

Intensity of itch will be measured using the Worst Itching Intensity NRS (see protocol Appendix III, Section 18.3) on a worksheet in which patients will be asked to indicate the intensity of the worst itching they experienced over the past 24 hours by marking one of 11 numbers, from 0 to 10, that best describes it, where "0" is labeled with the anchor phrase "no itching" and "10" is labeled "worst itching imaginable". Patients 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 Worst Itching Intensity NRS is collected daily during Screening and during the entire treatment period.

8.2.1.1.1 Weekly mean of daily 24-hour Worst Itching Intensity NRS

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 score reported during a specific week (Week 1: Days 2-8; Week 2: Days 9-15; Week 3: Days 16–22; etc.) 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. The baseline score will be defined as the mean of the 24-hour worst itching score over the last 7 days prior to randomization (Day -7 to Day -1), including pre-treatment assessments collected on Day 1 (Day 1).

Weekly averages are calculated within the EDC system on all available data; these will be retained in the analysis datasets, but analyses will be performed on the weekly averages with weekly mean values set to missing when more than 3 days in a week are missing.

8.2.1.1.2 Treatment Response

Treatment Response is defined as the percent improvement from baseline with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS score during the last week of the Treatment Period (Week 8 for Part A).

If a patient's mean daily itch score during the last week of the Treatment Period is greater than their baseline score (i.e., the patient has an increase in itch compared to baseline), his/her response to treatment will be assigned a value of 0. The treatment response for patients who discontinue treatment early will be estimated depending on the reason for discontinuation. For patients who discontinue due to their underlying condition (i.e., non-treatment related adverse event), the treatment response will be

estimated by carrying forward the last non-missing weekly mean worst itch score. Patients who discontinue due to any other reason will be considered non-responders and will be assigned a treatment response value of 0.

8.2.1.1.3 Proportion of patients who have a ≥20% treatment response

This is defined as the number of patients with a treatment response ≥20% based on the percent improvement in weekly mean of the 24-hour Worst Itching Intensity NRS. The denominator is the number of patients in the Full Analysis Population.

8.2.1.2 Skindex-10 Scale score

Developed specifically for uremic pruritus, the Skindex-10 Scale (see protocol Appendix V, Section 18.5) is an instrument for measurement of quality-of-life. Patients are asked the question 'During the past week, how often have you been bothered by' and fill 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 Skindex-10 scale is collected prior to dosing on Week 3, 5, 7 and 8 (Day 15, 29, 43, and 57).

8.2.1.2.1 Average Total Skindex-10 Scale Score

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

8.2.1.2.2 Skindex-10 Scale Domain Scores

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).

8.2.1.3 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 VII, Section 18.7).

The duration, degree and direction domains each include one item, while the disability domain has four items. All items of the first four domains were measured on a five-point Likert scale. The distribution domain included 16 potential locations of itch, including 15 body part items and one 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 four 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 four items. For the distribution domain, the number of affected body parts is tallied (potential sum 0–16) and the sum is sorted into five 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 five 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 5-D Itch Scale is collected prior to dosing on Weeks 3, 5, 7 and 8 (Day 15, 29 and 43, 50).

8.2.1.4 Medical Outcomes Study (MOS) Sleep Scale

The MOS Sleep Scale (see protocol Appendix VI, Section 18.6) was developed by the Medical Outcomes Study sleep survey in order to measure sleep disturbance. For most questions, the scale instructs patients to circle 1 of 6 numbers, ranging from 1 (labeled with the anchor phrase “all of the time”) to 6 (labeled “none of the time”), indicating the frequency of various aspects of sleep disruption over the preceding week. Patients are also provided instructions to estimate the average amount of sleep per night during the past week.

Ten of the scale’s 12 items are scored using a 6-point response scale, 1 using a 5-point scale, and sleep quantity is an open-ended question recording the actual number of hours slept. High score of sleep disturbance, somnolence, sleep indices or low score of sleep quantity, adequacy indicates relatively worse sleep problem, whereas lower scores for sleep quantity and sleep adequacy indicate worse sleep problems. The recall period for the MOS Sleep Scale is over the past week.

The original survey items except item 2 will be converted to a 0 to 100 range. For the 5-point scale (1 to 0, 2 to 25, …, and 5 to 100). For the 6-point scales (1 to 0, 2 to 20, …, and 6 to 100). Items in each sleep scale are averaged together to create the score for the scale. The scoring method follows MOS Sleep Scale User’s Manual.

8.2.1.5 Patient Global Impression of Worst Itch Severity

The Patient Global Impression of Worst Itch Severity is a PRO measure that assesses itch severity. The scale has only 1 item, with 5 possible values ranging from none to very severe (see Protocol Appendix VIII, Section 18.8).

8.2.1.6 Patient Global Impression of Change

The Patient Global Impression of Change is a global PRO measure that assesses the change (improvement or worsening) in overall status 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 IX, Section 18.9).

8.2.1.7 Missed Dialysis Visits and Incidence of Infection

Missed dialysis visits will be examined by the percentage of patients who missed 1 or more visits at the dialysis unit and the total number of missed dialysis visits during the Treatment Period; this will be derived in a manner identical to the missed doses in section 7, but counting dialysis instead of dialysis and dosing.

Incidence of infection is defined based on adverse events, hospitalizations, and/or use of antibiotics for treatment of infection related to uremic pruritus.

8.2.1.8 Iron Status and use of Erythropoiesis-stimulating agents (ESA) and IV iron

The dose and type of ESAs and IV iron administered at each dialysis visit and at the End of Treatment will be recorded on the *ESA log CRF* and *Parenteral (IV or dialysate) Iron Log CRF* respectively.

The baseline ESA (mcg) will be defined as the average ESA dose for each patient during the 4 weeks prior to randomization. The average ESA dose across dialysis visits during each week and each month will also be calculated. Change from baseline will be defined as the difference between the average ESA dose during month 1 and month 2 and the baseline average ESA. A similar definition applies to change from baseline in IV iron dose (mcg).

Serum chemistry will include ferritin and transferrin saturation at baseline and on Days 1, 15, 29, 43, and 57 only. Change from baseline in ferritin and similarly in transferrin saturation will be defined as the difference between the value at a specific visit and the baseline value.

8.2.1.9 Inflammatory biomarkers

Inflammatory biomarkers will include hepcidin, IL-1, IL-2, IL-6, IL-8, IL-31, TNF, IFN, MCP and C-reactive protein. A blood sample of sufficient volume to provide for replicate assays of each biomarker, as specified by the laboratory performing the analysis, will be collected prior to dialysis on Day 1 and at the end of the Treatment Period (Week 8 for Part A).

8.3 Analysis Methods

The primary goal for Part A of this clinical investigation is to evaluate the efficacy of different dose levels of IV CR845 administered after each dialysis session in reducing the intensity of itch over an 8-week Treatment Period in hemodialysis patients with moderate-to-severe pruritus. Three doses will be studied (1.5 mcg/kg, 1.0 mcg/kg, and 0.5 mcg/kg), and each will be compared against placebo. One of these doses will be selected for further evaluation compared to placebo in Part B. Although a hypothesis test of each CR845 dose against placebo will be performed for each of the primary, secondary and exploratory variables, these hypothesis tests are not expected to be statistically significant based on the planned sample size of approximately 40 patients per treatment group and effect size data from previous studies of CR845 in the same population. Nevertheless, the estimates of treatment effect and p-values resulting from these hypothesis tests will be used, in addition to a review of safety data, to select the most appropriate dose for Part B. A sample size of 40 patients per group is adequate to provide an appropriate estimate of the magnitude and variability of treatment effect at each dose.

For each analysis variable (primary, secondary, exploratory), all pairwise comparisons against placebo will be evaluated. There will be no adjustments for multiplicity.

In addition, an analysis combining the two highest doses of CR845 and analysis of all CR845 doses combined against placebo will be performed with respect to the primary and secondary variables. An analysis of all CR845 doses combined versus placebo with respect to the exploratory variables may be considered. A dose response analysis will be performed with respect to the primary and secondary variables.

8.3.1 Primary Efficacy Analyses

Part A

The primary efficacy variable, change from baseline to Week 8 in weekly mean of daily 24-hour Worst Itching Intensity NRS score, will be analyzed using a mixed effects model with repeated measures (MMRM)[1-3] using all visit weeks in the model. The model will contain treatment, week, and treatment-by-week interaction as fixed effects; baseline score and prior anti-itch medication usage as covariates, and patient as a random effect.

The baseline score will be defined as the mean of the 24-hour worst itching score over the last 7 days prior to randomization (Day -7 to Day -1), including pre-treatment assessments collected on Day 1 (Day 1). 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 score reported during a specific week (Week 1: Days 2-8; Week 2: Days 9-15; Week 3: Days 16–22; etc.) 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. The average NRS score at each week is also collected on the CRF.

The mixed model can be written as:

$$dNRS_{ij} = \beta_0 + \beta_1 treat_i + \beta_2 week_j + \beta_3 treat \times week_{ij} + \beta_4 Anti\text{-}Itch_k + b_m patient_m + dNRS_0 + \varepsilon_{ij}$$

where $b_m \sim N(0, \sigma^2)$,

dNRS – $dNRS_0$ - $dNRS_j$, change from baseline to Week 8 in weekly mean of daily 24 hr NRS score,
$dNRS_0$: baseline Worst Itching NRS score
treat : treatment, CR845 1.5 mcg/kg, CR845 1.0 mcg/kg, CR845 0.5 mcg/kg, Placebo
week : visit week (Week 1, Week 2, Week 3,….. Week 8)
patient : patient

An appropriate covariance matrix will be used to model the within-patient errors. The use of an unstructured covariance matrix structure as well as other structures, such as spatial patterns, that require fewer parameters (Toeplitz, autoregressive autoregressive, or compound symmetry) will be examined. The Akaike information criterion (AIC) will be used to determine the appropriate covariance matrix for the MMRM model. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom.

In the primary efficacy analysis, missing daily worst itching scores will not be imputed. Assuming that the data are missing at random (MAR), the estimates of the treatment differences calculated from the MMRM model described above are unbiased.

For Week 8, the adjusted Least Square Means and standard error from the model will be presented for each treatment group. For each CR845 treatment group, the treatment group difference versus placebo will be estimated as the simple contrast in the treatment effect on the last week of treatment (Week 8). For each week, the Least Square Means estimate for the difference, standard error, 95% CI, and p-value will be presented. The primary analysis will be conducted using the Full Analysis Population.

An additional supporting table will present descriptive statistics (n, mean, SD, median, min, max) for the observed value and change from baseline at each timepoint by treatment group.

As an exploratory analysis, the primary analysis may be repeated separately using dialysis days only. Additionally, the largest sites will be presented individually, with identical statistics as sample size allows (if the mixed models do not converge, summary statistics only will be presented).

For the analysis of the primary and secondary variables, the same MMRM model will be used to compare all CR845 doses combined vs. placebo by specifying appropriate contrasts in the model. The aforementioned statistics will be presented for this comparison.

For the primary analysis, the two higher doses will be combined and compared against placebo by specifying appropriate contrasts in the model. A linear trend test across the doses

will also be presented for week 8 to explore the dose response curve. These may also be added to other secondary analyses as needed to better understand the response curve.

Depending on the pattern and the amount of early treatment discontinuations in Part A, sensitivity analyses may be performed in order to provide estimates of treatment effect under different imputation algorithms and help refine the sample size for Part B. The following sensitivity analyses will be considered:

1. **Sensitivity analysis 1 (Multiple Imputation; MAR):**

This sensitivity analysis assumes that patients with missing data follow the same model as other patients in their respective treatment arm that have complete data.

- Intermittent missing data 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.
- Data missing after patients discontinue treatment early will then be multiply imputed with the SAS MI procedure using a method appropriate for monotone missingness (e.g., regression statement).
- Results of the MMRM on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure.

2. **Sensitivity analysis 2 (Trimmed mean):**

This sensitivity analysis uses a trimmed mean approach that assumes that all dropouts are "bad" outcomes.

- The percent change from baseline to week 8 for each outcome and treatment group will be determined and the worst 30% percent (or more) of patients will be trimmed from each group. Patients with missing weekly data at Week 8 or a percent change from baseline at Week 8 that represents a worsening in symptoms are automatically considered to be in the set to be trimmed. The choice of trimmed cut-off may be revised based on a blinded review of the data prior to DBL.
- If this exceeds 30% in any treatment group, the percent trimmed from each group will be increased for both groups to cover the larger percentage.
- Once the population of patients to be trimmed has been established, the mean change from baseline to week 8 for each group among the non-trimmed patients and the difference between groups will be calculated.
- Confidence intervals and p-values will be calculated using an approximate permutation test: treatments will be randomly assigned to the observed data for the full analysis population (including records for discontinued patients and those with missing values) and the above process repeated 10,000 times to establish the distribution of the outcome under the null, then the resulting CI and p-value will be reported. If the resulting p-value is in close proximity to 0.05, the number of samples will be increased accordingly to firmly establish on which side of 0.05 the p-value falls.

3. **Sensitivity Analysis 3 (Multiple Imputation; Missing Not at Random):**

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

- Intermittent missing data will first be imputed using the MCMC method implemented with the SAS MI procedure, which is appropriate for non-monotonic missing data
- For patients who discontinued study drug due to adverse events, data missing after discontinuation will be imputed using the distribution of the baseline value of all patients' daily worst itching score assuming a trimmed normal (from 0 to 10).
- For patients who discontinue due to reasons other than adverse event, missing data will after patients discontinue treatment early will be multiply imputed using multiple calls of the SAS MI procedure. At each time point, missing data will be imputed using data from patients in with in each group that have complete data at that time.

8.3.2 Secondary Efficacy Analyses

The secondary efficacy variable, change from baseline to Week 8 in the average total Skindex-10 Scale score, will be analyzed using a MMRM similar to that used for the primary efficacy analysis described in Section 8.3.1.

The baseline value will be defined as the value of the Skindex-10 Scale total score collected on Day 1, prior to randomization. Timepoints will include: Week 3, Week 5, and Week 7, and Week 8 (Day 57). For each dose group, the treatment group difference versus placebo will be estimated as the simple contrast in the treatment effect at Week 8 (Part A).

The total score is the sum of the numeric value of each of 10 questions. Missing Skindex-10 Scale scores will not be imputed. Assuming that the data are MAR, the estimates of the treatment differences calculated from the MMRM models above are unbiased.

8.3.3 Other Efficacy Analyses

8.3.3.1 Itch Intensity Measures

Change from baseline in weekly mean of the 24-hour Worst Itching Intensity NRS score from Week 2 through the end of the Treatment Period (Week 8)

This analysis will focus on looking at treatment differences at each post-baseline timepoint with respect to this variable and will be evaluated using the same MMRM model fitted for the primary efficacy analysis.

For each time point, the adjusted Least Square Means and standard error from the model will be presented for each treatment group. For each CR845 treatment group, the

treatment group difference versus placebo will be estimated as the simple contrast in the treatment effect at each time point. The Least Square Means estimate for the difference, standard error, 95% CI, and p-value will be presented.

Treatment Response
Treatment Response is defined as the percent improvement from baseline with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS score during the last week of the Treatment Period (Week 8 for Part A).

If a patient's mean daily itch score during the last week of the Treatment Period is greater than their baseline score (i.e., the patient has an increase in itch compared to baseline), his/her response to treatment will be assigned a value of 0. The treatment response for patients who discontinue treatment early will be estimated depending on the reason for discontinuation. For patients who discontinue due to their underlying condition (i.e., non-treatment related adverse event), the treatment response will be estimated by carrying forward the last non-missing weekly mean worst itch score. Patients who discontinue due to any other reason will be considered non-responders and will be assigned a treatment response value of 0.

The number and percentage of patients in each category of treatment response will be tabulated. A continuous responder graph of percentage of patients who have a treatment response ≥ 10%, ≥ 20%, ≥ 30%, ≥ 40%, ≥ 50%, ≥ 60%, ≥ 70%, ≥ 80%, ≥ 90%, and ≥ 100% will be presented.

Proportion of patients who have a ≥20% treatment response
The proportion of patients who have a ≥20% treatment response will be tabulated and analyzed with a logistic regression model containing terms for treatment group, and additional covariates chosen using forward selection techniques. Possible covariates for inclusion are age, gender, and Patient Self-Categorization of Pruritus Disease Severity at screening (B or C). The cutoff of 20% was shown to be clinically significant in a longitudinal study of uremic pruritus in hemodialysis patients (Mathur, 2010). This analysis may be repeated for other cutoffs.

8.3.3.2 Itch-Related Quality-of-Life Measures

The following endpoints will be analyzed using a MMRM similar to that used for the primary efficacy analysis described in Section 8.3.1:

- Change from baseline in each of the 3 domains of the Skindex-10 score
- Change from baseline in the total 5-D Itch Scale Score
- Change from baseline in the total sleep disturbance subscale of the MOS Sleep Scale

Patient Global Impression of Change
The number and percentage of patients in each of the 7 categories of the Patient Global Impression of Change at the end of the Treatment Period (Week 8 for Part A) will be tabulated by treatment group.

The proportion of patients who rate their itch condition as "Very Much Improved" or "Much Improved" will be tabulated. Pairwise treatment differences with each CR845 treatment group vs. placebo with respect to this variable will be tested using the Fisher's exact test.

Patient Global Impression of Worst Itch Severity
The number and percentage of patients in each of the 5 categories of the Patient Global Impression of Worst Itch Severity at baseline, Weeks 2, 4, and 8 (Part A) will be tabulated. Pairwise treatment differences with each CR845 treatment group vs. placebo with respect to this variable will be tested using a CMH row means score differ test; the number and percentage of patients with a 1 unit improvement will also be tabulated and tested using the Fisher's exact test.

8.3.3.3 Missed Dialysis Visits and the Incidence of Infection

The total number of missed dialysis days and the number of patients with at least 1 missed day will be summarized descriptively.

The incidence of infection will be tabulated and compared with a Fisher's exact test.

8.3.3.4 Iron Status and use of Erythropoiesis-stimulating agents (ESA) and IV iron

The dose and type of ESAs and IV iron administered at each dialysis visit and at the End of Treatment will be recorded on the *ESA log CRF* and *Parenteral (IV or dialysate) Iron Log CRF* respectively.

The baseline ESA (mcg) will be defined as the average ESA dose for each patient during the 4 weeks prior to randomization. The average ESA dose across dialysis visits during each week and each month will also be calculated. Descriptive statistics will be presented for average ESA dose at each timepoint and change from baseline in average ESA. ESA over time will also be plotted by treatment group. Similar analysis will be conducted for IV iron dose (mcg). LS means and 95%CIs for each treatment groups and pairwise comparisons will be reported using a MMRM approach identical to that for the primary analysis (but with monthly values analyzed rather than weekly).

Serum chemistry will include ferritin and transferrin saturation at baseline and on Days 1, 15, 29, 43, and 57 only. Observed values and change from baseline in ferritin and transferrin saturation from pre-dose to the end of the Treatment Period (Week 8 for Part A) will be summarized.

8.3.3.5 Change in inflammatory biomarkers

Changes in blood levels of inflammatory biomarkers, including but not limited to hepcidin, interleukin [IL]-2, IL-6, IL-31, pre-albumin, C-reactive protein from pre-dose to the end of the Treatment Period (Week 8 for Part A) will be summarized descriptively.

8.4 Examination of Subgroups

Given the small number of patients per treatment group in Part A, the MMRM employed as part of the primary and secondary efficacy analyses will not be repeated for subgroups. However, the descriptive summaries for the primary and secondary outcome measures will be provided for the following subgroups:

1. Use of concomitant medication 'anti-itch medication' during the week prior to randomization (use or non-use). This is also a stratification factor. Patients will be categorized as "using anti-pruritic medications" if they have used >1 dose of potentially anti-pruritic medications used to treat their uremic pruritus during the week prior to randomization (e.g. antihistamine (H1), pregabalin, gabapentin, nalbuphine, naltrexone, corticosteroids, or buprenorphine).

2. Patient Self-Categorization of Pruritus Disease Severity at screening (B or C)
   - B profile
     - I sometimes have scratch marks on my skin.
     - I sometimes have problems sleeping because of itching.
     - My itching can sometimes make me feel agitated or sad.
   - C profile
     - I often have scratch marks on my skin that may or may not bleed or get infected.
     - I often have a problem sleeping because of itching.
     - My itching often makes me feel agitated or sad.

9. SAFETY EVALUATION

9.1 Overview of Safety Analysis Methods

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

- Adverse Events (AEs)
- Clinical Laboratory Parameters
- Vital Signs
- 12-lead ECG

All safety analysis will be performed using the Safety Population based upon the treatment the patient actually received after randomization. All safety endpoints will be summarized by treatment group (and visit as appropriate).

9.2 Adverse Events

The period of adverse event reporting will start after the signing of the 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 patients, including patients who are deemed to be a screen failure.

All tabular AE summaries will be for treatment-emergent AEs (TEAEs). Treatment-emergent adverse events are defined as those events which:

- Start any time after the first dose of study drug up to the follow-up visit or early termination visit (or 7 days after the last dose if no early termination visit was conducted), whichever is later.
- Increase in severity any time after the first dose of study drug up to the follow-up visit or early termination visit (or 7 days after the last dose if no early termination visit was conducted), whichever is later.

Given that worsening of a pre-existing condition should be captured as an adverse event if the frequency, intensity, or character of the condition worsens after the use of study drug as judged by the Investigator, for purposes of analysis, TEAEs are identified as any AE starting after the first dose of the study drug up to the study follow-up visit or early termination visit (or 7 days after the last dose if no early termination visit was conducted).

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;
- if the start time is missing and the start date is the same as the first dosing date, then the event is considered treatment-emergent.

If it cannot be determined whether or not an event is treatment emergent due to a missing or partial date, then the event will be assumed to be treatment emergent.

All AEs will be coded to a system organ class (SOC) and preferred term (PT) using the *Medical Dictionary for Regulatory Activities* (MedDRA) version 18.0. The MedDRA treatment dictionary will be used to map adverse events verbatim to system organ class (SOC) and MedDRA preferred term for standardization and summary purposes.

The incidence of TEAEs will be summarized by treatment group and overall. If a patient experienced more than one episode of an adverse event, the patient is counted once for that Preferred Term. If a patient had more than one adverse event in a System Organ Class, the patient is counted only once in that System Organ Class. The summary tables will include incidence estimates for overall system organ class as well as for preferred terms within each system organ class. Incidence will be presented for system organ class by decreasing frequency overall and then alphabetically and by preferred term within each system organ class by decreasing frequency overall 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 patient on multiple occasions, the TEAE will be categorized according to the highest severity rating for that TEAE in that patient. If the severity of the TEAE is not reported, then the severity of the AE will be counted as severe. For each treatment group, 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, unlikely related, possibly related, probably related and definitely related). If a patient 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 (those categorized as possibly related or related) will be summarized by treatment group and overall.

The following summary tables will be presented for the Safety population:

- An overall summary showing for each treatment group, the number and percentage of patients with a TEAE, serious TEAE, related TEAE, severe TEAE, TEAE leading to dose interruption, TEAE leading to dose reduction, TEAE leading to study drug discontinuation, TEAE leading to study discontinuation. This table will also include number of events.
- 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
- Most common TEAEs (5% or more of patients in any treatment group) by preferred term

In addition, all AEs will be listed in chronological order including patient 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 medication, action taken with study treatment, and outcome). Note: For the all AE listing only, any screen failure patient that has an adverse event after signing ICF will be included for completeness. Separate listings will be generated for SAEs, deaths, and AEs leading to treatment discontinuation. Additionally, a coding list of preferred terms and the verbatim text associated with them will be produced.

No statistical tests will be performed on adverse events.

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

A serious adverse event is defined as any AE 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.

SAEs will be collected from date the informed consent form is signed up to the Follow-up visit or 7 days following the last dose of study drug, whichever is later. SAEs 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 "definitely related", "probably related", or "possibly related" to the study drug. SAEs 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 "unlikely related" or "not related" to study drug. A more detailed definition of SAEs is provided in Protocol sections 11.3.5 and 11.3.6. The analysis of SAEs is similar to that of AEs described in Section 9.2.

Patient deaths are captured on the *Adverse Events* eCRF page. Patient death listings will include all death data available including date of death and cause of death. Additionally, SAEs and AEs resulting in discontinuation will be listed.

9.4 Clinical Laboratory Evaluation

Blood and urinalysis samples will be collected during screening and day 1 before study drug administration, on days 15, 29 and 43 and day 57 (end of treatment).

All clinical laboratory data will be reported in SI units. Baseline will be defined as the last lab value obtained prior to treatment. Note that the Day 1 assessment can be

included in the evaluation of baseline given that the assessment is performed prior to dosing.

Summaries of actual values and change from baseline values by visit will be presented for quantitative laboratory parameters (e.g., WBC, lymphocytes). Only data from the central laboratory will be used.

If two or more evaluations occur in the same visit window, the evaluation closest to the target visit day will be selected for inclusion in the analysis. If multiple evaluations are equally close to the target visit day, then the latest evaluation will be selected for inclusion in the analysis.

All clinical laboratory values will be presented in a listing; additionally, ALT, AST, bilirubin and ALP will be presented in a separate listing with 3x and 5x ULN flagged for ALT and AST; 2x ULN flagged for bilirubin; and 1.5x ULN flagged for ALP.

Figures will be presented for a subset of laboratory values including Calcium, Ferritin, Transferritin Saturation, Intact Parathyroid Hormone, Serum Albumin and Phosphate.

9.5 Vital Signs, Physical Findings, and Other Observations Related to Safety

9.5.1 Vital Signs

Vital signs (including body temperature, heart rate, systolic and diastolic blood pressure) will be collected during screening and on days 1, 15, 29, 57 (end of treatment visit), and 64 (follow-up visit). 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.

Summary tables will include descriptive statistics (number of patients, mean, std, median, min, and max) for baseline and each post-baseline assessment. Descriptive statistics will be calculated on both the actual score and the change from baseline score. If two or more evaluations occur in the same visit window, the evaluation closest to the target visit day will be selected for inclusion in the analysis. If multiple evaluations are equally close to the target visit day, then the latest evaluation will be selected for inclusion in the analysis.

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

A second summary will focus on the individual patient measures with the intent of identifying post-baseline abnormalities defined as clinically significant. Using the

clinically significant ranges specified below, descriptive statistics will summarize the post-baseline patient frequencies (percentages) that are outside the range defined for each vital sign measure. All vital sign values will be considered when evaluating the criteria.

The number and percent of patients meeting the criteria for each of the following categories during the study will also be presented:

Body Temperature	>38 degrees centigrade or <36 degrees centigrade
Heart Rate	>130 bpm or <55 bpm
	Change from baseline (increase or decrease) of ≥30 bpm
Blood Pressure	BP > 180/100 or BP < 90/60
	Change from baseline (increase or decrease) in SBP of ≥40 mmHg
	Change from baseline (decrease) in SBP of ≥60 mmHg
	Change from baseline in DBP (increase or decrease) of ≥20 mmHg
	Change from baseline in DBP (decrease) of ≥40 mmHg

All vital signs will be listed in by patient listings including visit, collection date/time, and will be sorted by treatment group, patient ID, and date/time of assessment.

9.5.2 12-Lead ECGs

Standard 12-lead ECG readings with the patient in a supine position will be performed in triplicate during screening and day 1 prior to study drug administration, and on days 29 and 57 (end of treatment).

ECG 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. ECG values will also be listed by patient. Summaries of actual values and change from baseline by visit will be presented.

If two or more evaluations occur in the same visit window, the evaluation closest to the target visit day will be selected for inclusion in the analysis. If multiple evaluations are equally close to the target visit day, then the latest evaluation will be selected for inclusion in the analysis. The worst post-baseline result across all visits will be flagged and reported separately; this may include measures that were not in the individual visit summaries.

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.

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

10. PHARMACOKINETIC EVALUATION

Plasma samples will be analyzed for CR845 using liquid chromatography with tandem mass spectrometric detection according to validated analytical methods. All PK analyses will be performed using the PK Evaluable population.
Missing data will not be imputed. If any patients are noncompliant with the dosing schedule or have incomplete data, a decision will be made on a case-by-case basis as to their inclusion in the analysis prior to database lock.

10.1 Plasma Concentration Data

Blood samples for PK analysis will be collected from the predialyzer (arterial) line at Week 1, Day 1, Day 2, and Day 3; Week 2, Day 8; Week 4, Day 22; Week 6, Day 36; Week 8, Day 50, Day 51 and 52; at time points listed in the Pharmacokinetic Sampling Schedule (Protocol Appendix II, Section 18.2).

Treatment		Pharmacokinetic Samples							
				Time After Study Drug Administration (Post-dialysis)					
Week	Day	Pre-dialysis	Post-dialysis	5 minutes	30 minutes	1 hour	2 hours	4 hours	24 hours
1	1	X	X	X	X	X	X	X	
	2								X
	3	X		X					
2	8	X		X					
4	22	X		X					
6	36	X		X					
8	50	X	X	X	X	X	X	X	
	51								X
	52	X							

Plasma concentrations of CR845 will be summarized by dosing group and overall at each nominal sampling time point. Descriptive statistics will include n, arithmetic mean, SD, median, minimum, maximum, and geometric mean with 95% CI.

- Geometric mean = exp (mean of log-transformed concentration data).
- Below quantitation limit (BQL) concentrations will be treated as 0.5 x the LOQ for descriptive statistics. Mean concentrations that are BQL will be presented as BQL, and the SD will be reported as not applicable (N/A).

In plasma concentration tables and plots of mean profiles, statistics will be calculated having set concentrations to missing if one of the following cases is true:

1. A concentration has been collected as ND (i.e., not done) or NS (i.e., no sample),

2. A deviation in sampling time is of sufficient concern or a concentration has been flagged anomalous by the pharmacokineticist.

Mean and individual patient plasma concentration-time profiles will be plotted by dose on linear and semi-logarithmic scales for CR845.

All individual patient concentration data will be listed, including those being excluded from PK analyses.

11. OTHER ANALYSES

Not applicable at this time.

12. INTERIM ANALYSES AND DATA MONITORING

No formal interim analysis of efficacy data is planned for Part A or Part B.

Following completion of Part A, there will be an unblinded analysis to determine the dose of CR845 to be used in Part B. This analysis will not affect the Type I error for Part B, as it will represent the final analysis for Part A.

An ongoing review of the cumulative safety data for this study will be conducted by an external DMC for both Part A and Part B of this study. The operation of the DMC will be governed by a charter that will describe the groups' frequency of meeting, procedures (including, but not limited to, periodic safety monitoring) and requirements for reporting its observations to the Sponsor.

13. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

Sensitivity analyses have been updated from those described in the protocol to reflect more current methods being applied by FDA.

14. STATISTICAL APPENDIX

For the mixed model analyses, the code will be of the following form:

The MMRM with maximum likelihood estimation will utilize the following SAS Proc Mixed code:

proc mixed data = adweek;
class trt01p avisitn usubjid antiitch;
model aval = trt01p avisitn trt01p*avisitn treat*week base antiitch /ddfm=KR;
repeated avisitn / type = <*selected covariance matrix*> patient = usubjid;
run;

Note that it is at the programmer's discretion to use numeric vs character variables for treatment and use other parameterizations of an equivalent model.

For multiple imputations and other instances where random seeds are required, the values used (in order) are:
8392857
2985729
1843255
9086284
5547484
8017456
9597295
2252256
4821871
9852467
5126715
3232132
9841654
3645284
1587345

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

15. REFERENCES

1. Brown H., Prescott R. Applied Mixed Models in Medicine. John Wiley & Sons, 2002.
2. Verbeke G., Molenberghs G. Linear Mixed Models for Longitudinal Data. Springer-Verlag, New York, 2002.
3. Mallinckrodt C.H. et al. Assessing Response Profiles from Incomplete Longitudinal Clinical Trial Data Under Regulatory Considerations. *Journal of Biopharmaceutical Statistics*; Vol. 13, No. 2, pp. 179-190, 2003.
4. Mathur VS, Lindberg J, Germain M, Block G, Tumlin J, Smith M, et al. A longitudinal study of uremic pruritus in hemodialysis patients. Clin J Am Soc Nephrol. 2010; 5:1410-9.

16. APPENDICES

16.1 Schedule of Events

PART A									
Study Procedures	**Screen**	**Treatment Period[a]**						**End of Treatment/ Early Termination**	**Follow-Up**
	Day -14 to Day -1	**Week 1**			**Week 2 to 8**				**+7 days after EOT**
		M/Tu	**W/Th**	**F/Sa**	**M/Tu**	**W/Th**	**F/Sa**		
Visit Days →	**-14 to -1**	**1**	**3**	**5**	**8**	**10**	**12**	**57 (+3)**	**64 (+3)**
					15	**17**	**19**		
					22	**24**	**26**		
					29	**31**	**33**		
					36	**38**	**40**		
					43	**45**	**47**		
					50	**52**	**54**		
Administrative Procedures									
Informed Consent	X								
Inclusion/Exclusion Criteria	X	X[b]							
Medical History	X	X[b]							
Randomization		X							
Safety and Efficacy Evaluations									
Physical Examination	X							X	
Height	X								
Weight (estimated dry body weight)	X	X						X	
Post-dialysis weight		X						X	
12-lead Electrocardiogram	X[c]	X[c]			X[c]			X[c]	
Pre-dialysis Vital Signs[d]	X	X			X[e]			X	X
Hematology, Serum Chemistry (pre-dialysis)	X	X[f]			X[f]			X[f]	
Serum Pregnancy (females of childbearing potential only)	X							X	
FSH	X[g]								

PART A - Continued

Study Procedures	Screen	Treatment Period[a]						End of Treatment/ Early Termination	Follow-Up
	Day -14 to Day -1	Week 1			Week 2 to 8				+7 days after EOT
		M/Tu	W/Th	F/Sa	M/Tu	W/Th	F/Sa		
Visit Days →	-14 to -1	1	3	5	8	10	12	57 (+3)	64 (+3)
					15	17	19		
					22	24	26		
					29	31	33		
					36	38	40		
					43	45	47		
					50	52	54		
Safety and Efficacy Evaluations									
Patient Self-categorization of Pruritus Disease Severity	X								
Patient training on PRO Worksheets	X[h,i]	X							
Worst Itching Intensity NRS (daily)	X	X						X	
Skindex-10 Scale, MOS Sleep Scale, 5-D Itch Scale, Patient Global Impression of Worst Itch Severity		X[j]			X[j,k]			X	
Patient Global Impression of Change								X	
Record procedural data for dialysis (including ESA and IV iron)	X	X	X	X	X	X	X	X	
IV administration of study drug after dialysis		X	X	X	X	X	X		
Inflammatory biomarker samples[l]		X						X	
Adverse Event Monitoring	X	X						X	X
Prior Medications	X								
Concomitant Medications		X[m]			X[m]			X	X

EOT=end of treatment; ESA= erythropoiesis-stimulating agent; F=Friday; FSH=follicle-stimulating hormone; IV=intravenous; M=Monday; MOS=Medical Outcomes Study; NRS=numerical rating scale; PRO=patient reported outcome; Sa=Saturday; Th=Thursday; Tu=Tuesday; W=Wednesday;

See footnotes as well as Section 8 and 11 for additional procedural details.

[a] Each visit during the Treatment Period will coincide with the patient's normal dialysis treatments.

[b] Medical history will be updated on Day 1 with any changes since the Screening Visit, and inclusion/exclusion criteria will be confirmed prior t randomization.

[c] Electrocardiogram must be performed prior to dialysis at Screening, Day 1, Day 29, and Day 57 (or early termination).

[d] Vital signs will be obtained in a sitting or semi-recumbent position (for at least 3 min) and will include body temperature, heart rate, and blood pressure.

[e] On Day 15 and 29 only, pre-dialysis vital signs should be recorded

[f] On Day 1, 15, 29, 43, and 57 only. Serum chemistry should include parathyroid hormone, pre-albumin, ferritin and transferrin saturation.

[g] Obtain on women who have been amenorrheic for at least 1 year and are between 45 and 55 years of age.

[h] Training on Worst Itching Intensity NRS only

[i] Training on Skindex-10 Scale, MOS Sleep Scale, 5-D Itch Scale and Patient Global Impression of Worst Itch Severity scales may be performe at any time during the week prior to randomization.

[j] To be performed prior to or during dialysis, but must be completed prior to dosing.

[k] Skindex-10 Scale and 5-D Itch Scale completed on the first visit of Week 3, 5, and 7 (Day 15, 29, and 43, respectively). MOS Sleep Scale completed on the first visit of Week 4 and 8 (Day 22 and 50, respectively). Patient Global Impression of Worst Itch Severity completed on the first visit of Week 2, 4, and 8 (Day 8, 22, and 50, respectively). If the first visit of the week is missed, the patient may complete the procedure at their next visit for the same week. Questionnaires will be completed in strict adherence to the Patient Reported Outcomes Instructions (as pe Study Reference Manual).

[l] Biomarker samples must be collected prior to the start of dialysis.

[m] Concomitant medications will be updated on a weekly basis at the beginning of each week. Prior medications will be recorded until the time of first dosing.