Statistical Analysis Plan

Protocol Title:	A MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PLACEBO-CONTROLLED STUDY TO EVALUATE THE SAFETY AND EFFICACY OF ORAL CR845 IN CHRONIC KIDNEY DISEASE PATIENTS WITH MODERATE-TO-SEVERE PRURITUS
Protocol Number:	CR845-210301
Protocol Version, Date	Version 1.2, 14May2018
Document Version, Date:	Final 3.0, 18 October 2019

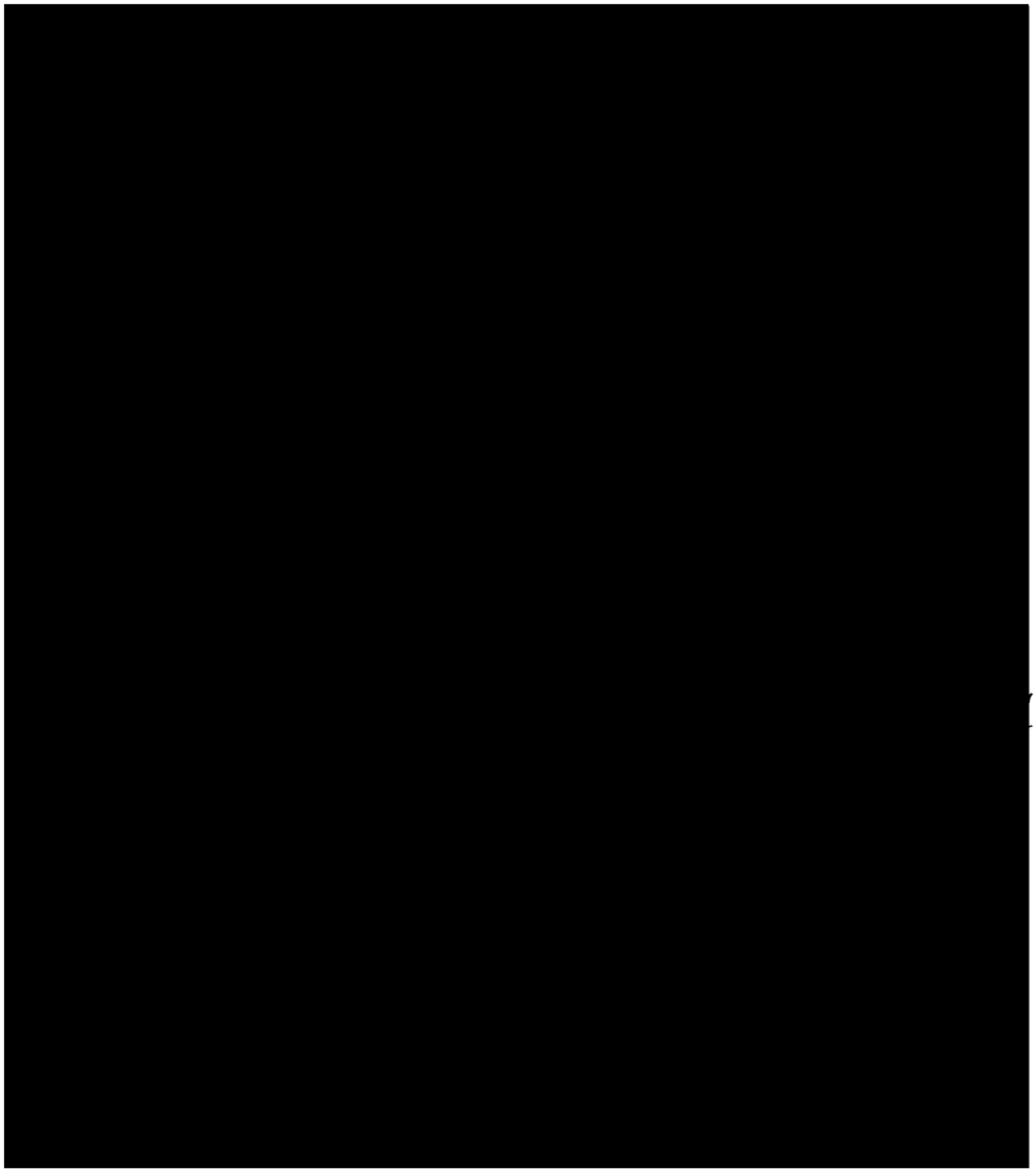
[redacted]

On behalf of:

CARA Therapeutics, Inc.

Confidentiality statement:

- The information provided in this document is strictly confidential.
- The recipients of the SAP must not disclose the confidential information contained within this document or any related information to other persons without the permission of the sponsor.
- In addition, the recipients of the SAP must keep this confidential document in a controlled environment, which prevents unauthorized access to the document.



REVISION HISTORY

Version/Date	Version name	Section	Changes implemented
Version 1.0/ 26MAR2019	Version 1	N/A	N/A
Version 2.0/ 01JUL2019	Version 2	7.1 / 8.6.2	Update condition to treat the Weekly NRS score as missing based on number of non-missing scores rather then number of days with missing scores.
		8.6.3	Use multiple missing data imputation for secondary endpoints also.
Version 3.0/ 18OCT2019	Version 3	4.3 / 8.6.3 / 8.6.4	New endpoints added: - Proportion of "complete responders" with respect to Worst Itching Intensity score. - Proportion of patients who indicate a severity of 'None' on the Patient Global Impression Worst Itch Severity scale. - Proportion of patients who either rate the degree of itching intensity over the past 2 weeks as 'Not Present' or the direction of itching as 'Completely Resolved' from the 5-D Itch Scale. - Proportion of patients with itching present at each body part and the shift from baseline.
		7.1 / 8.1	Provided further clarification regarding the terms "Baseline", "On-treatment" and "End of Treatment".
		8.4.2	Baseline and screen failure summaries added for Glomerular filtration rate.
		8.4.5	Section for prior/concomitant procedures added.
		8.5.3	Section for CR845 exposure added.
		8.6.3 / 8.6.4	Standardization of the analysis of binary outcomes from non-multiple imputed data. Logistic regression model always performed, followed by Cochran-Mantel-Haenzel test (or equivalent exact test if cell counts < 5). Analyses of "End of Treatment" result were ensured for all endpoints.
		8.7.2	Definitions for "attained" and "sustained" sodium levels added.

TABLE OF CONTENTS

SIGNATURE PAGE 2
REVISION HISTORY 3
TABLE OF CONTENTS 4
LIST OF ABBREVIATIONS 6
1 INTRODUCTION 7
2 STUDY OBJECTIVES 8
2.1 Primary objective 8
2.2 Secondary objective (s) 8
2.3 Exploratory objective 8
3 STUDY DESIGN 9
3.1 General study design 9
3.2 Randomization and blinding 10
3.3 Study treatments and assessments 10
4 STUDY ENDPOINTS 13
4.1 Primary efficacy endpoint 13
4.2 Secondary efficacy endpoint(s) 13
4.3 Additional efficacy endpoint(s) 13
4.4 Safety endpoint(s) 14
5 SAMPLE SIZE AND POWER 15
6 ANALYSIS POPULATIONS 16
6.1 Enrolled Population 16
6.2 Full Analysis Set (FAS) 16
6.3 Safety population (Safety) 16
6.4 Per-Protocol population (PP) 16
6.5 Protocol deviations/violations and exclusions from analysis sets 17
7 Statistical Considerations and analysis 18
7.1 Derived Variables 18
7.2 Handling of missing data and outliers 19
7.2.1 Missing data analysis methods 19
7.2.2 Handling of missing or incomplete dates 19
7.3 Assessment Time Windows 20
8 STATISTICAL METHODS 21
8.1 General statistical conventions 21
8.2 Subject disposition 22
8.3 Protocol deviations 22
8.4 Demographics and baseline characteristics 22
8.4.1 Demographics 22
8.4.2 Baseline and disease characteristics 23
8.4.3 Medical History 23
8.4.4 Prior and concomitant medications 23
8.4.5 Prior and concomitant procedures 24
8.5 Extent of exposure 24
8.5.1 Treatment duration 24

8.5.2 Treatment compliance 25
8.5.3 Exposure to CR845 25
8.6 Efficacy analyses 25
8.6.1 Analysis methods 26
8.6.1.1 Multiplicity 26
8.6.1.2 Treatment by center interaction analysis (multi-center study) 26
8.6.2 Analysis of primary efficacy endpoint 26
8.6.3 Analysis of secondary efficacy endpoint(s) 29
8.6.4 Analysis of additional efficacy endpoint(s) 31
8.7 Safety analyses 34
8.7.1 Adverse events 35
8.7.2 Clinical laboratory evaluations 36
8.7.3 Vital signs 37
8.7.4 Electrocardiograms 38
8.8 Other analysis 39
8.8.1 Subgroup analysis 39
8.9 Interim analysis 39
8.9.1 General Strategy 39
8.9.2 Interim Analysis Statistical Analysis 41
8.9.2.1 Subject disposition including study population 41
8.9.2.2 Demographic and Baseline Characteristic 41
8.9.2.3 Summary and analysis of NRS score 41
8.9.2.4 Summary and analysis of Skindex-10 and 5-D Itch total scale score and Responder Analysis 42
8.9.2.5 Safety Analysis 42
9 CHANGES TO PLANNED ANALYSIS FROM STUDY PROTOCOL 43
10 REFERENCES 44
11 APPENDICES 45

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse Event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
CI	Confidence Interval
CKD	Chronic Kidney Disease
cm	Centimeter
CSR	Clinical Study Report
ECG	Electrocardiogram
EOT	End of Treatment
FAS	Full Analysis Set
GFR	Glomerular filtration rate
IA	Interim Analysis
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
IVRS/IWRS	Interactive Voice or Web Response System
Kg	Kilogram
MAR	Missing At Random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MMRM	Mixed Model for Repeated Measurements
MNAR	Missing Not At Random
NRS	Numerical Rating Scale
PGIC	Patient Global Impression of Change
PP	Per Protocol
PRO	Patient-Reported Outcome
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard error
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TFLs	Tables, Figures and Listings
ULN	Upper Limit of Normal

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide detailed descriptions of the statistical methods, data derivations and data displays for study protocol CR845- 210301, version 1.2, "A multicenter, double-blind, randomized, placebo-controlled study to evaluate the safety and efficacy of oral CR845 in chronic kidney disease patients with moderate-to-severe pruritus" dated 14 May 2018 for Interim and Final analysis. The table of contents and templates for the tables, figures and listings (TFLs) will be produced in a separate document. Any deviations from this SAP will be described and justified in the Clinical Study Report (CSR).

The preparation of this SAP has been based on International Conference on Harmonization (ICH) E9.
Unless otherwise specified, all data analyses and generation of TFLs will be performed using SAS 9.2® or higher.

2 STUDY OBJECTIVES

2.1 Primary objective

To evaluate the efficacy of 3 dose levels of oral CR845 compared to placebo in reducing the intensity of itch in Chronic Kidney Disease (CKD) patients with moderate-to-severe pruritus.

2.2 Secondary objective (s)

- To evaluate the efficacy of 3 dose levels of oral CR845 compared to placebo in improving itch-related quality-of-life measures in CKD patients with moderate-to-severe pruritus
- To evaluate the safety of 3 dose levels of oral CR845 in CKD patients with moderate-to-severe pruritus.

2.3 Exploratory objective

Not Applicable.

3 STUDY DESIGN

3.1 General study design

This is a multicenter, randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of 3 dose levels (0.25, 0.5, and 1 mg) of oral CR845 in moderate-to-severe CKD patients with moderate-to-severe pruritus.

Total study duration for a single patient: up to approximately 18 weeks

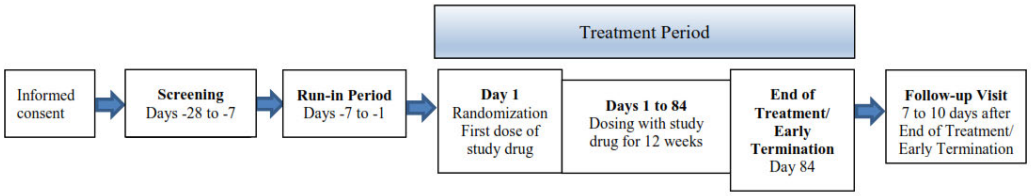
- Screening Period: 7 to 28 days prior to randomization (Day -28 to Day -8)
- Run-in Period:1 week prior to randomization (Day -7 to Day -1)
- Treatment Period: 12 weeks
- Follow-up Visit: 7-10 days after the End-of-Treatment Visit or Early Termination Visit

Eligible patients will enter a 7-day Run-in period during the week prior to randomization when they will be required to complete at least four 24-hr Worst Itching Intensity NRS worksheets. If patients continue to meet all inclusion and no exclusion criteria at the end of the 7-day Run-in period and their mean baseline Worst Itching Intensity NRS score is ≥5, they will be randomized in a 1:1:1:1 ratio to receive orally once daily either placebo or CR845 tablets at doses of 0.25, 0.5 or 1 mg. The mean baseline NRS score is defined as the average of the non-missing scores reported from the start of the 7-day Run-in period (eligible patients cannot have more than 3 missing daily scores during the run-in period). Randomization will be stratified according to the patient's renal disease status: moderate CKD non-dialysis; severe CKD non-dialysis; severe CKD on dialysis. The randomization of severe CKD patients on hemodialysis will be capped at approximately 20% of the total sample size.

The planned sample size is approximately 240 (60 per treatment group) male and female CKD patients with moderate-to-severe pruritus (mean baseline 24-hour Worst Itching Intensity numerical rating scale (NRS) score ≥5). The sample size may be increased up to 480 patients (120 per treatment group) based on the results of a planned interim analysis (IA) to be conducted when approximately 50-60% of the planned 240 patients have been randomized and have either completed the 12-week Treatment period or have discontinued study drug early.

The study schematic is shown in Figure 1.

Figure-1 CR845-210301 Study Schematic



3.2 Randomization and blinding

Before the start of the study, a computer-generated randomization schedule will be prepared. Randomization will be performed using an interactive voice or web response system (IVRS/IWRS). Patients will be randomized in a 1:1:1:1 ratio to receive either placebo or CR845 tablets at doses of 0.25, 0.5, or 1 mg. Randomization will be stratified according to the patient's renal disease status: moderate CKD non-dialysis; severe CKD non-dialysis; severe CKD on dialysis (ie, 3 categories). The randomization of severe CKD patients on hemodialysis will be capped at approximately 20% of the total sample size (ie, 48 of 240 patients).

Patients, investigators, study staff, and the sponsor will be blinded to study drug assignment.

For medically urgent or emergent situations that necessitate knowledge of study drug assignment for patient management, the blind may be broken via the IVRS/IWRS. Whenever possible, the medical monitor should be contacted prior to breaking the blind. The sponsor and medical monitor will receive a report whenever a patient blind is broken.

3.3 Study treatments and assessments

As described in Section 3.1, randomized patients will take study drug once daily for 12 weeks and a total of 84 doses of study drug. Each dose should be taken at least 2 hours prior to or after a meal, around the same time of day (patient should not lie down for at least 1 hour after swallowing the tablet).

For severe CKD patients on hemodialysis, the same requirement for meal timing of study drug is required but the study drug must not be taken until after their dialysis procedure on days when they receive dialysis.

A detailed description of procedures and assessments to be conducted during this study is summarized in the scheduled of study assessments in Table 1.

Table 1: Schedule of Events

Study Procedures	Screening Period		Treatment Period							Follow-up[j]
	Screening Period	Run-in Period							EOT or Early Termination	
Study Week			1	2	4	6	8	10	12	
Visit Days	-28 to -7	-7 to -1	1	14	28	42	56	70	84 (or following discontinuation at any time)	7 to 10 days after EOT/Early Termination
				←———		±2 days			———→	
Administrative procedures										
Informed consent	X									
Inclusion/exclusion criteria	X		X							
Medical history	X									
Review drug dosing instructions with patient			X							
Randomization			X							
Safety and efficacy evaluations										
Physical examination	X									
Height	X									
Weight (this is the prescription dry body weight for hemodialysis patients)	X									
12-lead electrocardiogram	X		X[a]						X	
Vital signs[b]	X		X[a]	X	X	X	X	X	X	
Hematology	X		X[a]			X			X	
Serum chemistry	X		X[a]			X			X	
Serum pregnancy (females of childbearing potential only)	X		X[a]						X	
Dispense study drug and drug diary[c]			X	X	X	X	X	X		

Study Procedures	Screening Period		Treatment Period							Follow-up[j]
	Screening Period	Run-in Period							EOT or Early Termination	
Study Week			1	2	4	6	8	10	12	
Visit Days	-28 to -7	-7 to -1	1	14	28	42	56	70	84 (or following discontinuation at any time)	7 to 10 days after EOT/Early Termination
				←——— ±2 days ———→						
Patient training on PRO worksheets		X[d,e]	X[e]							
Worst Itching Intensity NRS (daily)		X	X[f]						X	
Skindex-10, 5-D Itch Scale, Patient Global Impression of Worst Itch Severity[g]			X[a]		X		X	X	X	
Patient Global Impression of Change									X	
Adverse event assessment	←———								———→	
Prior medications	X									
Concomitant medications[h]			←———						———→	
CKD = chronic kidney disease; EOT = end of treatment; NRS = numerical rating scale; PRO = patient-reported outcome										

a) To be collected/performed prior to the first dose of study drug on Day 1.
b) Vital signs, including body temperature, heart rate, and blood pressure, will be obtained while the patient is in a sitting or semi-recumbent position.
c) Study drug will be dispensed to patients in bottles of 15 tablets following randomization on Day 1 and thereafter at each visit during the Treatment Period until the end of treatment/early termination visit. Patients may be contacted to verify or remind them to take their tablet daily per the protocol. Patients must record when they took the study drug in a drug diary every day during the Treatment Period.
d) Training on Worst Itching Intensity NRS on the first day of the Run-in Period.
e) Training on Skindex-10, 5-D Itch, and Patient Global Impression of Worst Itch Severity may be performed at any time during the Run-in Period or on Day 1 of the Treatment Period.
f) Patients will be requested to complete the NRS worksheets at a similar time each day.
g) Skindex-10, 5-D Itch, and Patient Global Impression of Worst Itch Severity are to be completed at the clinical site.
h) Concomitant medications will be updated at each study visit following the first dose of study drug on Day 1.
j) A phone follow-up will be conducted, unless otherwise necessary.

4 STUDY ENDPOINTS

4.1 Primary efficacy endpoint

The primary efficacy endpoint is defined as the change from baseline to week 12 of the Treatment period with respect to the weekly mean of the daily 24 hour Worst Itching Intensity NRS score.

4.2 Secondary efficacy endpoint(s)

The secondary efficacy endpoints of this study are:

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

4.3 Additional efficacy endpoint(s)

The additional efficacy endpoints of this study are:

Itch-intensity Measures

- Proportion of patients who have an improvement from baseline at each week of the Treatment period and end of treatment with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS scores ≥1, ≥2, ≥3 (excluding week 12) and ≥4 points.
- Proportion of "complete responders" with respect to Worst Itching Intensity score NRS at each week of the Treatment period and at end of treatment. A patient that has ≥ 80% of the non-missing daily NRS scores in a week equal to 0 or 1 is considered a complete responder at that week.
- Change from baseline in the weekly mean of the 24-hour Worst Itching Intensity NRS score at each week of the Treatment period (Week 1 to Week 11) and at end of treatment.
- Proportion of patients who rate their itch condition as "Very much improved" or "Much improved" at Week 12/end of treatment, as measured by the Patient Global Impression of Change (PGIC).

- Proportion of patients in each of the 5 categories of the Patient Global Impression of worst itch severity at baseline and at each scheduled weekly visit of the Treatment period and end of treatment.
  In addition:
  - Proportion of patients who have a 1-point improvement or more from baseline
  - Proportion of patients who indicate a severity of 'None'.

Itch-related Quality-of-Life Measures

- Change from baseline in itch-related quality of life at each scheduled weekly visit of the Treatment period (excluding Week 12) and end of treatment, as assessed by the total score of the 5-D Itch Scale.
  In addition, at each scheduled weekly visit of the Treatment period and end of treatment:
  - Proportion of patients who either rate the degree of itching intensity over the past 2 weeks as 'Not Present' or the direction of itching as 'Completely Resolved' from the 5-D Itch Scale.
  - Proportion of patients with itching present at each body part and the shift from baseline.

- Change from baseline in itch-related quality-of-life at each scheduled weekly visit (excluding Week 12) of the Treatment period and end of treatment, as assessed by the total score of the Skindex-10 Scale.

- Change from baseline in itch-related quality-of-life at each scheduled weekly visit of the Treatment period and end of treatment with respect to each of the 3 domains of the Skindex-10 Scale.

- Change from baseline in itch-related quality-of-life at each scheduled weekly visit (excluding Week 12) of the Treatment period and end of treatment with respect to each of the domains of 5-D Itch Scale.

4.4 Safety endpoint(s)

The safety assessments to evaluate the overall safety of CR845 will include,

- Adverse events
- Vital signs
- 12-lead Electrocardiograms (ECGs)
- Clinical laboratory evaluations
- Exposure to Study Drug and Compliance

5 SAMPLE SIZE AND POWER

The minimum total sample size for this study is N = 240 patients (60 per treatment group) and the maximum is N = 480 (120 per treatment group). An unblinded interim analysis (IA) will be conducted when approximately 50-60% of the planned 240 patients have been randomized and have either completed the 12-week Treatment period or have discontinued study drug early. The primary goal of the IA is to identify dose(s) of oral CR845 that are both safe and efficacious. Doses that are found to be unsafe or poorly tolerated will be dropped from the study. No dose will be dropped for futility reasons. The details of sample size re-estimation based on conditional power are described in section 8.9.1. A sample size of 60 patients per treatment group (the minimum sample size) will provide adequate power (≥80%) for effect sizes ≥0.52 for a 5% type 1 error and a 2-sided T-test. Assuming a standard deviation of 2.6 for the primary efficacy variable, an effect size of 0.52 corresponds to a difference of approximately 1.4 point on a 0- to 10-point scale.

A sample size of 120 patients per treatment group (the maximum sample size) will provide adequate power (≥80%) for effect sizes ≥0.40 for a 5% type 1 error and a 2-sided T-test. Assuming a standard deviation of 2.6 for the primary efficacy variable, an effect size of 0.40 corresponds to a difference of approximately 1 point on a 0- to 10-point scale. The power of a 2-sided hypothesis test with a 5% Type 1 error as a function of sample size and effect size is presented below.

Power as a Function of Effect Size (N = 120 Per Arm)[a]

Effect Size	0.55	0.52	0.50	0.45	0.40	0.35
Power (N = 120)	98%	97%	97%	93%	86%	77%
Power (N = 60)	84%	80%	77%	68%	58%	47%

[a] Power for a 2-sided T-test with equal variance and a 5% Type 1 error.

6 ANALYSIS POPULATIONS

6.1 Enrolled Population

The enrolled population is defined as the group of patients who sign the informed consent form (ICF).

6.2 Full Analysis Set (FAS)

The full analysis set is defined as all randomized patients who received at least 1 dose of study drug. Patients in the full analysis set will be analyzed according to their randomized treatment, regardless of the actual treatment received. The Full Analysis set will be used to analyze all efficacy endpoints.

6.3 Safety population (Safety)

The safety analysis population is defined as all randomized patients who received at least 1 dose of study drug. Patients in the safety analysis population will be analyzed according to their actual treatment. The Safety Population will be used to analyze all safety endpoints.

6.4 Per-Protocol population (PP)

The Per-Protocol Population is defined as the subset of patients in the full analysis set who do not have any major protocol deviations that could affect the efficacy analyses. 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 set are excluded.

The per protocol population include patients who:

- Received ≥ 80% and ≤120% of the planned study drug doses
- Had a mean baseline Worst Itching Intensity NRS score ≥ 5.0
- 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 6.4.7.2.
- Did not have major protocol violations that would impact efficacy outcomes.
- Satisfied all major inclusion and exclusion criteria on study entry and received treatment as planned.

Patients in the per-protocol population will be analyzed according to their actual treatment.

6.5 Protocol deviations/violations and exclusions from analysis sets

All violations and exclusions of patients from analysis sets will be identified through programmatic checks, through medical reviews, and by clinical research associates during site monitoring. Protocol deviations will be classified as minor or major and whether they result in exclusion of the patient from the PP set, prior to the database lock.

7 STATISTICAL CONSIDERATIONS AND ANALYSIS

7.1 Derived Variables

The table below provides the list of derived variables for various duration derivations, drug compliance, baseline derivations and other important derivations applicable for this study.

Variables	Formula
Derivation of Duration	
Duration of treatment (days)	(Date of last dose of drug – date of first dose of study drug) + 1. (dates will be determined based on the patient daily dosing diary)
Duration of study (days)	(Date of study completion/discontinuation (for discontinued patients) – date of first dose) + 1
Study day	• For dates before the date of first dose: Date of interest – date of first dose of study drug. • For dates on or after the date of first dose: Date of interest – date of first dose of study drug + 1
Drug Compliance	
Compliance using accountability records	[(Sum of the total number of tablet dispensed across all visits- Sum of the total number of tablet returned across all visits)/ Duration of treatment (days)]*100
Compliance using daily dosing diary	[(Total number of tablet taken as per daily dosing diary)/ Duration of study drug (days)]*100
Baseline Derivations	
Baseline Worst Itching Intensity NRS score	The baseline score will be calculated as the average of the non-missing daily 24-hour Worst Itching Intensity NRS scores over the last 7 days prior to randomization; at least 4 completed Worst Itching Intensity NRS worksheets will be required prior to randomization.
Baseline (for other variables)	• Time of assessment collected for the variable: baseline is defined as the last assessment prior to date/time of first dose of study drug whether scheduled or unscheduled. • Time of assessment not collected for the variable: The baseline value is defined as the last observation prior to or on the date of the first dose of study drug whether scheduled or unscheduled
Change from baseline	Post baseline value – Baseline value

Variables	Formula
Weekly Mean Worst Itching Intensity NRS score	
Weekly Mean Worst Itching Intensity NRS score	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 during the Treatment period (eg, Days 1 to 7, Days 8 to 14, Days 15 to 21, etc.) divided by the number of non-missing scores for that week. If there are <4 non-missing daily worst itching scores in a specific week, the corresponding weekly mean worst itching score will be set to missing.
Skindex-10	
Total Scores	Patients will be asked to 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 10 questions. The total score is the sum of the numeric value of each answered question. For more details see Appendix B.
5-D Itch	
Total Score	Total 5-D Itch score = duration score (single item) + degree score (single item) + duration score (single item) + maximum disability (4 disability items) + category score based on sum of affected body parts. 5-D Itch scores can potentially range between 5 (no pruritus) and 25 (most severe pruritus). For more details see Appendix C

7.2 Handling of missing data and outliers

7.2.1 Missing data analysis methods

For analyses relating to the Weekly Mean Worst Itching Intensity NRS score, Skindex-10 Total Score and 5-D Itch Total Score, missing data will be imputed using a multiple imputation (MI) approach. Details of imputation methods are described under section 8.6.2.

7.2.2 Handling of missing or incomplete dates

Imputation rules for missing AE start date:

Analyses of adverse events will assume that any AE with a missing start date is treatment emergent. Partial AE start dates are not allowed in the database.

Imputation rules for missing or partial medication start/stop dates are defined below:

Missing or partial medication start date:

- If the start date is completely missing, then the medication will be classified as 'concomitant'.
- If the start date is partially missing then it will be compared to the partial portion of the date of first dose. If the partial medication start date is greater than the

partial portion of the first dose date and the medication is ongoing, the particular medication will be classified as 'concomitant'.

Missing or partial medication stop date:

- If the stop date is completely missing then it will be assumed to be ongoing
- If the stop date is partially missing then it will be compared to the partial portion of the date of first dose. If the partial medication stop date is greater than the partial portion of the first dose date, the particular medication will be classified as 'concomitant'.

7.3 Assessment Time Windows

For the primary analysis variable, assessment time windows are not needed since the NRS Itch Intensity Assessments Log collects the daily individual scores and average NRS score for the run-in and each post-baseline visit week are calculated based on these daily scores.

Assessments collected by study week that are collected at early termination visits and unscheduled visits will be assigned to a post-baseline planned visit window, if the early termination or unscheduled visit day falls between +/- 2 days of the planned visit. 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 to an early termination visit, then the unscheduled assessment closest to the planned day will be used. 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.

8 STATISTICAL METHODS

8.1 General statistical conventions

All statistical procedures will be completed using SAS version 9.2 or higher.

Unless otherwise noted, continuous variables will be summarized using descriptive statistics, including number of patients with non-missing observation (n), mean, median, standard deviation (SD), minimum (min) and maximum (max). Means and medians will be reported to one decimal place more than the raw data. Standard deviation will be reported to two more decimal places than the raw value. The minimum and maximum will be reported to the same decimal as the raw data.

Categorical variables will be summarized using the frequency count and the 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 in a particular category. Counts of zero in any category will be presented as "0". Unless otherwise noted, for all percentages, the number of patients in the analysis population for the treatment group will be the denominator.

"Baseline" will be defined as per section 7.1. For efficacy endpoints, "on-treatment" is defined as any assessment following the first dose of study drug and within 2 days from the last dose of study drug recorded in the patient diary. For safety endpoints (excluding adverse events), any assessment within 7 days after last dose of study drug will be deemed "on-treatment". For all endpoints "End of treatment" (EOT), in summaries, will be represented by the last non-missing on-treatment value recorded up to and including the Week 12/Early Termination visit. Treatment emergence for adverse events is described in section 8.7.1. Summaries will be presented across all scheduled visits and EOT unless stated otherwise. Only on-treatment data will be used in summaries and statistical analyses (with the exception of the Patient Global Impression of Change).

Unless otherwise stated, all statistical testing will be two-sided and will be performed using a significance (alpha) level of 0.05. Two-sided 95% confidence intervals (CI) will be provided appropriately. 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".

All patients data, including those derived, will be presented in individual patient data listings. Unless otherwise stated, unscheduled visit results will be included in date/time chronological order, within patient listings only. All listings will be sorted by treatment, patient number, date/time and visit. Patient listings of efficacy and safety data will be presented for all full analysis set and safety set patients respectively unless specified otherwise.

8.2 Subject disposition

The number of enrolments, screen failures, randomizations and treated patients will be summarized over all patients. In addition, the number of each reason for screen failure will be provided. In this summary table, percentages will be calculated using the number of enrolments as the denominator.

The number and percent of screen failure patients failing each inclusion and exclusion criteria at least one once will be summarized. The percentages will be calculated using the number of patients in the enrolled population.

A separate display will present the number and percent of patients who receive study treatment, complete the 12-week study treatment, who discontinue from double-blind treatment, (and the reason for study treatment discontinuation), and who complete the follow-up. For this table, percentages will be calculated using the number in patients in the safety population.

Additionally, the above table will summarize the number and percent of randomized patients in the following populations:

- Full Analysis set;
- Safety population;
- Per Protocol population;

8.3 Protocol deviations

Protocol deviations will be classified as minor or major and whether they led to exclusion of the patient from the Per Protocol Population prior to the database lock. Major protocol deviations overall and major protocol deviations leading to exclusion of the the patient from the Per Protocol Population will be summarized by category and overall. All protocol deviations will be listed.

8.4 Demographics and baseline characteristics

Demographic and baseline patient characteristics will be summarized overall and by treatment group based on the safety population.

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

8.4.1 Demographics

Following demographic characteristics will be summarized

- Age (years)at screening,
- Age category (<45, 45-64, 65-74, ≥75 years, Not Reported)
- Gender (Male, Female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not reported)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Other, Not Reported)

- Height (cm)
- Weight (kg) (for hemodialysis patients- prescription dry body weight)
- Body mass index (kg/m^2)
- Female Reproductive status (Pre-menarche, Sterile, Post-menopausal, Potentially able to bear children, Other and Not applicable)

8.4.2 Baseline and disease characteristics

Baseline characteristics of the disease will also be summarized. Baseline disease characteristics include variables such as

- Glomerular filtration rate (GFR) (mL/min/1.73m²)
- Grouped GFR (<30 mL/min/1.73 m^2 or ≥30 and <60 mL/min/1.73 m^2)
- Renal disease status (Moderate CKD Non-Dialysis, Severe CKD Non-Dialysis, Severe CKD On Dialysis)
- Baseline NRS score

As a separate summary, the GFR (mL/min/1.73m²) and grouped GFR (<15 ml/min/1.73m^2, 15-29 ml/min/1.73m^2, 30-59 ml/min/1.73m^2, 60-89 ml/min/1.73m^2, >90 ml/min/1.73m^2) will be summarized across patients who did not meet inclusion criteria #4 (renal status). If multiple GFR results were recorded during the screening/run-in period then the maximum will be used. This data will be listed in the screen failure listing (see section 8.2).

8.4.3 Medical History

Medical history data will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 21.0. A summary table by treatment group will be presented by MedDRA system organ class (SOC) and preferred term.

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 the condition is ongoing.

A separate coding listing will be created with all the distinct levels of SOC, High Level Terms, Preferred Terms, Lowest Level Term and the verbatim Investigator description reported in the study. Sorting will be alphabetically by SOC, Preferred Term, and then verbatim description.

8.4.4 Prior and concomitant medications

Prior medications (including vitamins and herbal supplements) are defined as those that the patient has taken any time during the 30 days prior to the screening up until the first dose of study drug on Day 1.

Concomitant medications are medications that the patient has taken after the start of the first dose of study drug on Day 1 of the Treatment period through the end-of-treatment (i.e., 12-week Treatment period)/early termination visit.

All medications will be coded using the World Health Organization Drug Dictionary (WHODD) version March 2018. All prior and concomitant medications will be listed and summarized separately by Anatomical Therapeutic Chemical (ATC) class 3 and preferred term.

Additionally, a coding listing will be created which will include all distinct ATC class 1, 2 and 3 codes and preferred terms along with the corresponding verbatim description of the medications; sorting will be alphabetically by ATC class 3 and preferred term. Prior and concomitant medications will be listed separately.

8.4.5 Prior and concomitant procedures

Prior procedures are defined as those performed on the patient during the 30 days prior to the screening up until the first dose of study drug on Day 1.

Concomitant procedures are defined as those performed on the patient after the start of the first dose of study drug on Day 1 of the Treatment period through the end-of-treatment (i.e., 12-week Treatment period)/early termination visit.

All procedures will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 21.0. All prior and concomitant procedures will be listed and summarized separately by SOC and preferred term.

Additionally, a coding listing will be created which will include all distinct SOC, High Level Terms, preferred terms and Lowest Level Terms along with the corresponding verbatim description of the procedure; sorting will be alphabetically by SOC and preferred term. Prior and concomitant procedures will be listed separately.

8.5 Extent of exposure

Study drug summary and compliance summaries will be based on the safety population. Drug administration details including derived variables will be provided in the form listings in addition to the below summaries.

8.5.1 Treatment duration

The duration of treatment will be derived using the dates of first/last study drug administration from the patient diary as date of last dose – date of first dose + 1.

Duration of treatment (days) and duration of study (days) will be summarised based on the safety population using descriptive statistics. Refer to section 7.1 for the derivation of variables. In addition, descriptive summary will be provided for total number of doses taken (based on the patient daily dosing diary). This will also be presented in categories “1-7, 8-14, 15-21, 22-28, 29-35, 36-42, 43-49, 50-56, 57-63, 64-70, 71-77, 78-84” using count and

percentage.

Total number of missed doses collected from patient daily dosing diary presented in categories 0, 1, ≥2 using count and percentage.

8.5.2 Treatment compliance

Compliance will be calculated in 2 different ways: based on the patient daily dosing diary and using drug accountability records.

Study drug compliance based on the number of tablets taken will be calculated as: [(Total number of tablet dispensed across all visits-Total number of tablet returned across all visits)/ Duration of study drug (days)] ×100.

Study drug compliance based on the patient daily dosing diary will be calculated as: [(Total number of tablet taken as per daily dosing dairy)/ Duration of treatment] ×100.

Study drug compliance (based on the patient daily dosing diary and using drug accountability records) will be summarized. Compliance will also be summarized in categories <80%, 80% - 120%, >120% using frequency tables.

Study drug compliance summaries will be based on the safety population.

8.5.3 Exposure to CR845

The exposure to CR845 (mg) will be calculated also using the patient daily dosing diary, however, the actual dose levels taken by the patient will be taken into account rather than the number of tablets (i.e. for a correctly dosed placebo patient, exposure to CR845 = 0mg).

8.6 Efficacy analyses

This section describes the analysis to be conducted on the primary, secondary and additional efficacy variables. All the efficacy analyses will be performed using the full analysis set and data that is collected on or prior to the date of treatment discontinuation/completion. An analysis of the primary and secondary efficacy variables for the per-protocol population will be performed if more than 20% of the FAS patients are excluded in any of the treatment groups. Each CR845 dose group, will be compared against placebo for all the endpoints. In addition, a separate analysis comparing all CR845 doses pooled against placebo will be performed with respect to the primary and key secondary variables. Only the results for the all CR845 doses pooled group and the comparison with placebo will be presented from these analyses. An analysis of all CR845 doses pooled versus placebo with respect to the rest of the secondary variables may also be considered. A dose response analysis will be performed with respect to the primary and key secondary variables. Results for the primary and secondary efficacy variables will be presented by renal disease status (moderate CKD non-dialysis; severe CKD non-dialysis; severe CKD on dialysis).

8.6.1 Analysis methods

8.6.1.1 Multiplicity

There will be no adjustment for multiple comparisons.

8.6.1.2 Treatment by center interaction analysis (multi-center study)

Given the large (~60) number of sites with few patients per site, no analysis will be conducted to assess the treatment-by-center interaction.

8.6.2 Analysis of primary efficacy endpoint

The primary efficacy endpoint is defined as the change from baseline to week 12 of the Treatment period with respect to the weekly mean of the daily 24 hour Worst Itching Intensity NRS score. Testing of the primary efficacy endpoint will be 2-sided and conducted at the 5% error level.

The study will be considered positive if at least one safe and efficacious dose is identified. Effectiveness will be evaluated based on results of the primary efficacy analysis either using the estimate of treatment effect or the p-value of the hypothesis tests.

The null hypotheses H_{01}, H_{02}, H_{03}, and H_{04} given below will be tested against the alternative hypotheses H_{A1}, H_{A2} and H_{A3} respectively:

H_{01}: $\mu_{D1} - \mu_P = 0$, H_{02}: $\mu_{D2} - \mu_P = 0$, H_{03}: $\mu_{D3} - \mu_P = 0$, H_{04}: $\mu_{D123} - \mu_P = 0$

H_{A1}: $\mu_{D1} - \mu_P \neq 0$, H_{A2}: $\mu_{D2} - \mu_P \neq 0$, H_{A3}: $\mu_{D3} - \mu_P \neq 0$, H_{A4}: $\mu_{D123} - \mu_P \neq 0$

Where μ_{D1}, μ_{D2} and μ_{D3} denote the mean change in Worst Itching Intensity NRS score from baseline to week 12 in the three active treatment dose groups and μ_P denotes the mean change from baseline to week 12 in the placebo group (placebo, P). D_1 denotes the CR845 0.25mg dose group, D_2 denotes the CR845 0.5mg dose group and D_3 denotes the CR845 1mg dose group and D_{123} denotes the doses (0.25, 0.5 and 1 mg) combined.

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 of the Treatment period (eg, Days 1 to 7, Days 8 to 14, Days 15 to 21, etc. till 78 to 84) divided by the number of non-missing scores for that week. If there are <4 non-missing daily worst itching scores during a specific week, the corresponding weekly mean worst itching score will be set to missing. The baseline score will be defined as the average of the daily 24-hour Worst Itching Intensity NRS scores over the last 7 days prior to randomization; at least 4 completed Worst Itching Intensity NRS worksheets will be required prior to randomization.

The primary efficacy variable 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; baseline score and renal disease status as covariates. For each

dose group, the treatment group difference vs placebo will be estimated as the simple contrast in the treatment effect at Week 12 of the Treatment Period. The treatment difference between all doses combined and placebo will be estimated using a similar approach.

For week 12, the adjusted least square means (LS means) and standard error (SE) from the model will be presented for each treatment group. The LS means estimate for the difference, standard error, 95% CI, and p-value will also be presented. The primary analysis will be conducted using the Full Analysis Set. A summary table including descriptive statistics (n, mean, SD, median, minimum, maximum) for the observed value and the change from baseline by treatment group will also be provided for the NRS score.

An unstructured covariance matrix structure will be used to model the within patient errors. If there are convergence issues, other structures that require fewer parameters will applied in the following order: 1) heterogeneous Toeplitz, 2) autoregressive, 3) compound symmetry until convergence is attained. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom.

For the MMRM analyses, the missing Weekly NRS score will be imputed using a Multiple Imputation (MI) approach, assuming that patients who discontinue study drug early would have similar Worst Itching Intensity NRS scores as other patients in their respective treatment arm that have complete data. The details of missing data imputations described as below.

- Intermittent missing weekly 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. Additional terms for baseline NRS score and randomization stratification category (as dummy 0/1 variables) will be added. This will be performed within each treatment group separately.

- The monotone missing weekly NRS values will then be multiply imputed with the SAS MI procedure using the monotone regression method. The regression model will include terms for treatment group, baseline NRS score, and randomization stratification category; all previous non-missing post-baseline NRS scores will be included in the model. Should convergence issues occur due to small cell size for the categorical covariates at either stage, they will be removed from the model.

- The change from baseline at Week 12 with respect to the weekly mean of the daily 24-hour Worst Itching Intensity will be analyzed for each imputed dataset using an MMRM model.

- One hundred imputations will be performed.

- Results of the MMRM analysis on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure.

The details of MI SAS procedure described in Appendix F.

In addition, sensitivity analyses of the primary efficacy endpoint will be conducted to evaluate the robustness of study results under different assumptions and imputation algorithms, as described below:

Sensitivity analysis 1 (No imputation; Missing at Random [MAR]):

In this sensitivity analysis, missing weekly worst itching scores will not be imputed. Assuming the data are MAR, the estimates of the treatment differences calculated from the MMRM model described above are unbiased. Details of analysis using SAS described under Appendix G.

Sensitivity analysis 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. This will be performed within each treatment group separately.

- For patients who discontinue study drug due to adverse events, data missing after discontinuation will be imputed using the distribution of the baseline value of all patients' weekly worst itching score with the same renal status assuming a trimmed normal (from 5 to 10).

- For patients who discontinue study drug due to reasons other than adverse event, monotone missing weekly NRS values will be multiply imputed with the SAS MI procedure using the monotone regression method applied to all patients with data available at the particular visit (including those withdrawing due to an adverse event). The regression model will include terms for treatment group, baseline NRS score, and randomization stratification category; all previous non-missing post-baseline NRS scores. Should convergence issues occur due to small cell size for the categorical covariates at either stage, they will be removed from the model.

- Results of the MMRM on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure.

- The details of MNAR based imputation using SAS described under Appendix H.

Descriptive statistics for the primary efficacy endpoint will also be provided by treatment group for each renal disease category (moderate CKD non-dialysis; severe CKD non-dialysis; severe CKD on dialysis). In addition, an analysis of the primary efficacy endpoint within each renal disease status will be performed using an MMRM model similar to the one implemented for the primary efficacy analysis but without renal disease status as a fixed effect. As the renal disease status was included a factor in the multiple missing data imputation for the primary analysis, the same multiple imputed datasets as created for the primary analysis and methods to combine results will be used for these analyses.

A linear trend test across the doses based on a MMRM model will also be presented to explore the dose response curve. The model will contain treatment as continuous variable (Placebo = 0 mg and other doses 0.25, 0.5, and 1.0 mg), week, and treatment-by-week interaction as fixed effects; baseline score and renal disease status as covariates. The estimate and p-value for overall effect of dose will be provided. An unstructured covariance matrix structure will be used to model the within patient errors.

Listings will be provided for Worst Itching Intensity NRS score for all the available visits.

8.6.3 Analysis of secondary efficacy endpoint(s)

The key secondary efficacy endpoints are:

- Change from baseline in itch-related quality of life at the end of Week 12, as assessed by the total Skindex-10 scale score;
- Change from baseline in itch-related quality of life at the end of Week 12, as assessed by the total 5-D Itch scale score;
- Proportion of patients achieving an improvement from baseline ≥3 points with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score at Week 12 of the Treatment period.

The Skindex-10 scale total score and the 5-D Itch total scale score will be analyzed using MMRM that will contain treatment, week, and treatment-by-week interaction as fixed effects; baseline score and renal disease status as covariates. Repeated measures will include values that reflect the Skindex-10 and 5-D total score at the end of weeks 4, 8, 10, and 12. The total score derivation for Skindex-10 and 5-D Itch scale detailed in Appendix B and Appendix C respectively.

The baseline value will be defined as the value of the Skindex-10 or 5-D Itch total score collected on Day 1 prior to randomization. Missing Skindex-10 Itch scale and 5-D Itch scale

total scores will be imputed in a similar fashion as the primary endpoint for the MMRM analyses using the MI approach assuming data is missing at random. The mean treatment difference between each CR845 dose and placebo, and between all CR845 doses combined and placebo will be estimated as the simple contrast in the treatment effect at Treatment period week 12.

Descriptive statistics will be provided for the observed value and change from baseline at week 12 by treatment group for Skindex-10 scale and the 5-D Itch total scores. The adjusted LS means and SE from the model will be presented for each treatment group. The LS means estimate for the difference, standard error, 95% CI, and p-value will be presented. The analysis will be conducted using the Full Analysis Set.

An unstructured covariance matrix structure will be used to model the within patient errors. If there are convergence issues, other structures that require fewer parameters will applied in the following order: 1) heterogeneous Toeplitz, 2) autoregressive, 3) compound symmetry until convergence is attained. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom.

Descriptive statistics for the change from baseline with respect to the Skindex-10 and the 5-D Itch total score will also be provided by treatment group for each renal disease category (moderate CKD non-dialysis; severe CKD non-dialysis; severe CKD on dialysis). In addition, an analysis of these key secondary endpoints by renal disease status will be performed using an MMRM model and a treatment of missing data similar to the one described above.

A linear trend test across the doses based on a MMRM model will also be presented to explore the dose response curve. The model will contain treatment as continuous variable (Placebo = 0 mg and other doses 0.25, 0.5, and 1.0 mg), week, and treatment-by-week interaction as fixed effects; baseline score and renal disease status as covariates. The estimate and p-value for overall effect of dose will be provided. An unstructured covariance matrix structure will be used to model the within patient errors.

Patients achieving improvement from baseline ≥3 points will be summarized using count and percentage. The proportion of patients achieving an improvement from baseline ≥3 points with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score at week 12 of the treatment period will be analyzed using logistic regression with terms for treatment group, baseline NRS score, and renal disease status. Odds ratio, and P-value based on logistic regression comparing difference in patients achieving NRS response between CR845 treatment group and placebo and between all CR845 doses combined and placebo at week 12 will be presented. In case of small counts resulting in separation issues of the log-likelihood, Firth's correction will be applied.

Missing NRS data will be imputed using the same methodology described for the primary efficacy endpoint. Specifically,

- The proportion of patients 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 derived for the primary efficacy analysis. Differences between each CR845 treatment group and placebo will be compared using a logistic regression model containing terms for treatment group, baseline NRS score and renal disease status. P-value will be presented.
- Results of the logistic regression on the multiply imputed data sets will be summarized by the SAS MIANALYZE procedure (see Appendix I).
- For comparisons between treatments at end of treatment (not using MI), the same logistic regression will be performed on observed data with additional non-parametric analyses evaluating pairwise treatment differences will be tested using the Cochran-Mantel-Haenszel test, adjusting for strata (i.e renal disease status). An exact test of whether the common odds ratio across strata equals 1 will be used if the observed count in a particular cell is small (ie, <5) (see Appendix J).

The proportion of patients achieving an improvement from baseline ≥3 points by each renal disease category (moderate CKD non-dialysis; severe CKD non-dialysis; severe CKD on dialysis) will be presented. In addition, treatment differences within each subgroup will be investigated using logistic regression, including a Fisher's exact test for the end of treatment (not using MI) comparison.

The proportion of patients achieving an improvement from baseline ≥3 points with respect to the weekly mean of the daily 24-hour Worst Itching Intensity NRS score at week 12 of the Treatment period will be also be analyzed using logistic regression with terms for treatment group as continuous variable (Placebo = 0 mg and other doses 0.25, 0.5, and 1.0 mg), baseline NRS score, and renal disease status to evaluate the dose response curve. Overall effect of treatment group will be presented.

Listings will be provided for all the collected Skindex-10 scale score and the 5-D Itch Scale score separately.

8.6.4 Analysis of additional efficacy endpoint(s)

Itch-intensity Measures:

- Proportion of patients who have an improvement from baseline at week 12 of the treatment period with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS scores ≥1, ≥2, and ≥4 points will be calculated. Treatment differences

will be compared using a logistic regression model with terms for treatment group, baseline NRS score and renal disease status. This analysis will be conducted separately for each NRS score categories (≥1 vs <1 (no or less improvement compared to baseline, or, ≥2 vs <2, and ≥4 vs <4 points). The calculation of the proportions will be based on the NRS data using an MI approach for the missing data as in the primary efficacy analysis. Results of the logistic regression on the multiple imputed data sets will be summarized by the SAS MIANALYZE procedure. In addition, a graph showing the proportion of patients who have an improvement from baseline in NRS scores at week 12 that are ≥1, ≥2, ≥3, and ≥4 will be prepared. For comparisons between treatments at end of treatment (not using MI), the same logistic regression will be performed on observed data with additional non-parametric analyses evaluating pairwise treatment differences pairwise treatment differences will be tested using the Cochran-Mantel-Haenszel test, adjusting for strata (i.e renal disease status). An exact test of whether the common odds ratio across strata equals 1 will be used if the observed count in a particular cell is small (ie, <5).

- Proportion of "complete responders" with respect to Worst Itching Intensity score NRS at week 12 and end of treatment. A patient that has ≥ 80% of the non-missing daily NRS scores in a week equal to 0 or 1 is considered a complete responder at that week. Patients with <4 non-missing NRS scores within a week will be derived as missing at that week. Treatment differences will be compared using a logistic regression model with terms for treatment group, baseline NRS score and renal disease status. In addition, pairwise treatment differences will be tested using the Cochran-Mantel-Haenszel test, adjusting for strata (i.e renal disease status). An exact test of whether the common odds ratio across strata equals 1 will be used if the observed count in a particular cell is small (ie, <5).

- Change from baseline in the weekly mean of the 24-hour Worst Itching Intensity NRS score at each post baseline week (except week 12) of the treatment period will be summarized. Treatment differences between CR845 and placebo at each post baseline time point will be evaluated and presented using the same MMRM detailed in primary efficacy analysis (see section 8.6.2). Analysis will be based on the NRS data using an MI approach for the missing data as in the primary efficacy analysis.

- Proportion of patients who rate their itch condition as "Very much improved" or "Much improved" at the end of week 12 of the treatment period and end of treatment, as measured by the Patient Global Impression of Change (PGIC). Treatment differences will be compared using a logistic regression model with terms for treatment group, baseline Patient Global Impression Worst Severity score and renal disease status. In addition, pairwise treatment differences will be tested using the

Cochran-Mantel-Haenszel test, adjusting for strata (i.e renal disease status). An exact test of whether the common odds ratio across strata equals 1 will be used if the observed count in a particular cell is small (ie, <5). The PGIC responses at week 12 and end of treatment will also be presented with count and percentage.

- The number and percentage of patients in each of the 5 categories of the Patient Global Impression of Worst Itch Severity at baseline and at each scheduled post baseline week of the treatment period will be tabulated. In addition, the proportion of patients who have at least 1-point improvement or more from baseline at each week of the Treatment Period, as measured by the Patient Global Impression of Worst Itch Severity will be calculated. Similarly, the proportion of patients who indicate a Patient Global Impression of Worst Itch Severity of 'None' will be derived at Week 12 and end of treatment. Treatment differences will be compared using a logistic regression model with terms for treatment group, baseline Worst Itch Severity and renal disease status. In addition, pairwise treatment differences will be tested using the Cochran-Mantel-Haenszel test, adjusting for strata (i.e renal disease status). An exact test of whether the common odds ratio across strata equals 1 will be used if the observed count in a particular cell is small (ie, <5).

Itch-related Quality-of-Life Measures:

- Change from baseline in itch-related quality of life at each scheduled post baseline week of the Treatment period, as assessed by the total score of the 5-D itch scale. Treatment differences between each CR845 dose group and placebo at each post baseline time points will be evaluated and presented (except week 12) using the MMRM fitted for the secondary efficacy analysis section 8.6.3.
- The proportion of patients who either rate the degree of itching intensity as 'Not Present' or the direction of itching as 'Completely Resolved' over the past 2 weeks from the 5-D Itch Scale will be calculated at Week 12 and end of treatment. Treatment differences will be compared using a logistic regression model with terms for treatment group, baseline 5-D Itch Total Score and renal disease status. In addition, pairwise treatment differences will be tested using the Cochran-Mantel-Haenszel test, adjusting for strata (i.e renal disease status). An exact test of whether the common odds ratio across strata equals 1 will be used if the observed count in a particular cell is small (ie, <5).
- The proportion of patients with itching present at each body part will be presented at Week 12 and end of treatment along with the shift from baseline (Present or Not Present combinations). The presence of itching at each body part at end of treatment will be compared between treatments adjusting for strata (i.e renal disease status) and presence of itch at baseline (Present/Not Present). An exact test of whether the

common odds ratio across strata equals 1 will be performed.

- Change from baseline in itch-related quality-of-life at each week of the Treatment period, as assessed by the total score of the Skindex-10 Scale. Treatment differences between each CR845 dose group and placebo at each scheduled post baseline time points will be evaluated and presented (except week 12) using the same MMRM fitted for the secondary efficacy analysis section 8.6.3.
- The total Skindex-10 score is subdivided into 3 domain scores, which are sums of the scores of the following consideration of questions (Appendix B): disease domain (questions 1 to 3), mood/emotional distress domain (questions 4 to 6), and social functioning domain (questions 7 to 10). Change from baseline in itch-related quality of life at each of the post baseline weeks of the treatment period with respect to each of the 3 domains of the Skindex-10 Scale will be calculated. Treatment differences with respect to each domain will be evaluated using a model similar to the MMRM fitted for the analysis of the change from baseline in the total score of the Skindex-10 Scale.
- Change from baseline in itch-related quality of life at week 12 of the treatment period and at each of the remaining weeks of the treatment period with respect to each of the domains of 5-D Itch scale. Treatment differences with respect to each domain will be evaluated using a model similar to the MMRM fitted for the analysis of the change from baseline in the total score of the 5-D Itch Scale. The domain score derivation will be considered as detailed in the Appendix C.

Listings will be provided for the variables used in the additional efficacy endpoints (which are not mentioned in the primary or secondary analysis).

8.7 Safety analyses

The following assessments will be used to evaluate the overall safety of CR845:

- Adverse events
- Vital signs
- 12-lead ECGs
- Clinical laboratory evaluations.

All safety analysis will be performed using the safety population. All safety endpoints will be summarized by treatment group and overall. Baseline value defined as the last lab value obtained prior to treatment.

8.7.1 Adverse events

All adverse events (AE) will be coded using MedDRA dictionary to the corresponding MedDRA system organ class (SOC) and Preferred Term (PT) for standardization and summary purposes. Only treatment emergent adverse events (TEAEs) will be included in summary tables, except for a table presenting an overall summary of all AES and of REAEs. Summaries will provide results overall and by treatment.

TEAEs are defined as adverse events (AEs) with an onset date on or after the first dose of the study drug up to the End of Treatment/Early Termination visit or 7 days after the last dose, whichever is later.

The incidence of TEAEs will be presented using counts and percentages of patients with adverse events and tabulated by SOC and preferred term. System organ class will be sorted alphabetically and preferred term within SOC will be sorted by descending frequency based on the incidence across all patients in the safety population. A patient will be counted only once in the incidence count for a MedDRA SOC or preferred term, although a patient may have multiple occurrences (start and stop) of an event associated with a specific MedDRA preferred term or SOC.

The following specific summary tables will be generated:

- An overall summary showing the number and percentage of patients with
    - Any AE (both TEAE and non-TEAE)
    - TEAE
    - Serious TEAE
    - Serious non-TEAE
    - Related TEAE
    - Severe TEAE
    - TEAE leading to drug interruption
    - TEAE leading to dose reduction
    - TEAE leading to study drug discontinuation
    - Number of any events and number of TEAEs (Count only)
- TEAEs by SOC and preferred term
- TEAEs related to study drug by SOC and preferred term
- TEAEs by SOC, preferred term and maximum severity
- TEAEs occurring in ≥5% of patients in at least 1 treatment group
- Treatment-related TEAEs occurring in ≥5% of patients in at least 1 treatment group

- Serious TEAEs by SOC and preferred term
- TEAEs leading to study drug discontinuation by SOC and preferred term

If the relationship to the study drug of an adverse event (AE) is missing, the AE will be categorized as "related" to the study drug. If a patient reports 2 or more AEs that code to the same preferred term, the event with the maximum relationship will be included in the table.

If the severity to the study drug of an AE is missing, the AE will be categorized as "severe". If a patient reports 2 or more adverse events that code to the same preferred term, the event with the maximum severity will be included in the table.

All adverse events will be listed in chronological order, including patient identifier, age, race, gender, a flag indicating whether the event was treatment-emergent, 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 drug, and outcome).

In addition, separate listings will be generated for serious adverse events (SAEs), deaths, and adverse events leading to study drug discontinuation.

A coding list of SOC, high level, preferred and lowest level terms and the verbatim text associated with them will also be produced additionally. TEAEs and non-TEAEs will be listed separately.

8.7.2 Clinical laboratory evaluations

Blood samples for clinical laboratory tests, including hematology, serum chemistry and serum pregnancy will be collected during screening and day 1 before study drug administration, on week 6 and week 12 (end of treatment).

All clinical laboratory data will be reported in Système International units.

Summary statistics for each scheduled time point measured and mean changes from baseline to each time point will be presented for hematology and serum chemistry laboratory results. The baseline value is defined as the last observation prior to or on the date of the first dose of study drug whether scheduled or unscheduled (see section 7.1).

Hematology and serum chemistry laboratory test results will be assigned as Low (L), Normal (N) or High (H) classification according to whether the value was below (L), within (N), or above (H) the laboratory parameter's reference range.

Comparisons will be based on 3×3 tables (shift tables) that, for a particular laboratory test, compare the baseline L/N/H classification to the highest and/or lowest L/N/H classification during the treatment period.

Clinically important laboratory values based on the below specified criteria will also be summarized.

- Alanine aminotransferase (ALT), aspartate aminotransferase (AST),
  - >3 × ULN
  - >5 × ULN
  - ALT>3xULN or AST>3xULN
- Alkaline phosphatase (ALP)
  - >1.5 × ULN
- Total Bilirubin (TBILI)
  - >2 × ULN
- AST and Total Bilirubin
  - AST >3xULN and TBILI>2xULN
- ALT and Total Bilirubin
  - ALT>3xULN and TBILI>2xULN
- AST or ALT and Total Bilirubin
  - AST or ALT>3xULN and TBILI>2xULN

A summary of treatment emergent sodium findings will also be provided, categorizing results into those <146 mmol/L, 146 - 150 mmol/L, and >150 mmol/L at the 6 and 12 week assessments. An additional summary will present these categories for subjects over the full study period presenting the highest category attained (at any on-treatment assessment) and sustained (at any on-treatment assessment and all subsequent on-treatment assessments); this analysis will include both scheduled and unscheduled laboratory test results.

All clinical laboratory values will be presented in a listing. Additionally, ALT, AST, bilirubin, ALP and sodium will be presented in a separate listing, with values >3 × or >5 × ULN flagged for alanine aminotransferase and aspartate aminotransferase; 2 × ULN flagged for total bilirubin, 1.5 × ULN flagged for alkaline phosphatase and ≥ 146 mmol/L flagged for sodium. In each listing, all the data for the patient across the study will be presented for any parameter flagged post-baseline.

8.7.3 Vital signs

Vital sign parameters include systolic blood pressure (SBP), diastolic blood pressure (DBP), body temperature and heart rate. For each scheduled timepoint, descriptive summary statistics will be calculated for both the actual and the change from baseline values for all

vital sign parameters. The baseline value is defined as the last observation prior to or on the date of the first dose of study drug whether scheduled or unscheduled (see section 7.1).

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 Treatment period. This analysis will include results from both scheduled and unscheduled assessments.

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

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

8.7.4 Electrocardiograms

ECG assessment will be done at screening, day 1 and end of treatment (EOT).

ECG results include an overall interpretation of 'normal,' 'abnormal but not clinically significant,' or 'abnormal and clinically significant.' These results will be tabulated at each time point using a shift table.

Electrocardiogram results will be listed for all available visits, including visit, whether ECG was performed (yes/no), explanation (if not performed), assessment date/time, overall interpretation.

8.8 Other analysis

8.8.1 Subgroup analysis

The following endpoints will be summarized by strata- patient's renal disease status: moderate CKD non-dialysis; severe CKD non-dialysis; and severe CKD on dialysis.

- All primary and secondary efficacy endpoints
  - 24-hour Worst Itching Intensity numerical rating scale (NRS) score
  - Total Skindex-10 Scale score
  - 5-D Itch Scale score

- Safety endpoints
  - Overall Adverse Events
  - AEs by SOC and PT

8.9 Interim analysis

An unblinded Interim Analysis (IA) will be conducted when approximately 50-60% of the planned 240 patients have been randomized and have either completed the 12-week Treatment Period or have discontinued study drug early. The primary goal of the IA is to identify dose(s) of oral CR845 that are both safe and efficacious across all strata. Doses that are found to be unsafe or poorly tolerated will be dropped from the study. No dose will be dropped for futility reasons. The sample size will not be increased past 480 total patients (120 patients per treatment group).

8.9.1 General Strategy

Since there are 3 CR845 doses, the decision rule for sample size re-estimation to be followed for the IA, and described below is designed to address multiple comparisons and protect the type 1 error

The conditional power of each CR845 dose to separate from placebo at the end of study (based on the planned sample size of 60 patients per group) will be calculated, assuming that estimated treatment effect at interim analysis is the true effect.

- If the conditional power is >= 80% for at least one dose then the sample size will not be increased in any of the doses
- If the conditional power is < 80% across all doses and ≥ 20% in at least one dose,

then the following process will be followed

- For each CR845 dose group the sample size ($size_{new}$) needed to achieve an 80% power given the observed treatment effect will be calculated
- The calculated sample sizes will be ordered from lowest to highest as $min(size_{new)}$, $int(size_{new})$ and $max\ (size_{new})$
- If the smallest re-estimated sample size $min(size_{new})$ is > 120 then the recommendation will be to increase the sample size to 120 for all treatment groups
- If $min(size_{new}) \leq 120$ but $\{int(size_{new})$ and $max(size_{new})\} > 120$ (i.e. only the minimum sample size is ≤ 120) then the recommendation will be to increase the sample size to $min(size_{new)}$ for all treatment groups
- If $min(size_{new})$ and $int(size_{new})$ are both ≤ 120 (i.e. 2 out 3 re-estimated sample sizes are ≤ 120) then the recommendation will be to increase the sample size for all treatment groups to
    - $int(size_{new})$ if $[int(size_{new}) - min(size_{new})] \leq 10$

    or

    - $min(size_{new})$ if $[int(size_{new}) - min(size_{new})] > 10$
- If $max(size_{new}) \leq 120$ (i.e. 3 out 3 re-estimated sample sizes are ≤ 120), then the recommendation will be to increase the sample size for all treatment groups to
    - $max(size_{new})$ if $[max(size_{new}) - min(size_{new)}] \leq 10$

    or

    - $min(size_{new})$ if $[max(size_{new}) - min(size_{new)}] > 10$

o If the conditional power is < 20% for all doses the sample size will not be increase for any dose

Conditional power will be calculated for each CR845 dose assuming that the estimated treatment effect at interim analysis is the true effect as proposed by formula in the article by Mehta and Pocock[3].

$$CP_{\hat{\delta}_1}(z_1, \tilde{n}_2) = 1 - \Phi\left(\frac{z_\alpha\sqrt{n_2} - z_1\sqrt{n_1}}{\sqrt{\tilde{n}_2}} - \frac{z_1\sqrt{\tilde{n}_2}}{\sqrt{n_1}}\right).$$

Where Z_1 is observed test-statistic calculated using the sample included in the IA, n_2 is the total planned size (i.e. 120 = 60 per group), n1 is the interim size (approximately 50% of the target n_2), ň2 is the difference (n2 – n1), and delta hat is the observed treatment difference.

8.9.2 Interim Analysis Statistical Analysis

The following variables/summaries will be included in the Interim Analysis

8.9.2.1 Subject disposition including study population

Subject disposition information will be summarized by treatment group and overall (across CR845 doses and across all treatments). The number of patients who are enrolled, the number and percent of screen failures, of randomized and treated patients will be summarized. In addition, the reason for screen failure will be provided. In this summary table, percentages will be calculated using the number of enrolled patients as the denominator.

A separate display will present the number and percent of patients who complete the 12-week study treatment, who discontinue from double-blind treatment, (and the reason for study treatment discontinuation), and who complete the follow-up. For this table, percentages will be calculated using the number in patients in the safety population.

Additionally, the above table will summarize the number and percent of randomized patients in the following populations:

- Full Analysis set;
- Safety population;

8.9.2.2 Demographic and Baseline Characteristic

Summary of demographic and baseline variable described in section 8.4.1 and 8.4.2.

8.9.2.3 Summary and analysis of NRS score

The change from baseline in the weekly mean of the 24-hour Worst Itching Intensity NRS score at each post baseline week of the treatment period will be summarized. Treatment differences between CR845 and placebo at each post baseline time point will be evaluated and presented using mixed effects model with repeated measures (MMRM). Refer to section 8.6.2 for the details of NRS data analysis.

Summary statistics will be provided for the observed and the change from baseline with respect to the weekly mean of the 24-hour Worst Itching Intensity NRS score by strata- renal disease status (moderate CKD non-dialysis; severe CKD non-dialysis; and severe CKD on dialysis).

8.9.2.4 Summary and analysis of Skindex-10 and 5-D Itch total scale score and Responder Analysis

The secondary efficacy endpoints mentioned below will be summarized for IA.

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

Refer to section 8.6.3 for details of the analysis for the above endpoints.

8.9.2.5 Safety Analysis

The following summaries will be provided to evaluate the safety profile of CR845. Analysis details are provided in section 8.7.1. Except for the first display listed below, only TEAE during treatment period will be summarized.

- Overall summary of Adverse Event
- AEs summary by SOC and PT
- AEs related to study drug by SOC and PT
- AEs leads to study drug discontinuation by SOC and PT
- AEs by Maximum severity
- Serious adverse event summary
- Summary of Clinically important laboratory values
- Summary of Sodium abnormal findings (<146 mmol/L, >= 146 mmol/L, <=150 mmol/L, >150 mmol/L)

Line plot for the below parameters will also be provided.

- Change from baseline in selected Laboratory Tests over Time (hemoglobin, creatinine, GFR, albumin, phosphate, calcium, sodium)
- Blood Pressure Measurements (Systolic and diastolic blood pressure)
- Heart Rate Measurements

In addition, the following listings will also be provided to support safety data.

- Listing of Adverse Events
- Listing of death
- Listing of Serious Adverse Events
- AEs leading to study drug discontinuation
- Listings for clinically important laboratory values

9 CHANGES TO PLANNED ANALYSIS FROM STUDY PROTOCOL

- Antipruritic medications are not identified in the CRF or the database using a Yes/No flag. The use antipruritic medications will need to be derived based on the indication entered on the CRF. Therefore, a specific analysis of antipruritic medications by ingredient will not be completed since it will not be different than the analysis summarizing medications by preferred term.

- Section 8.6.1 of the protocol defines the primary endpoint as follows: "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 during the Treatment Period (eg, Days 2 to 8, Days 9 to 15, Days 16 to 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.".
  However, to facilitate patient diary dates anomalies which may result in duplicate dates (e.g. patient completes diary shortly after midnight and then again on same date) and the fact that only the diary date is used for mapping to a specific week as per section 7.1, of this SAP, the underlined text above is replaced with "…divided by the number of non-missing scores for that week. If there are <4 non-missing daily worst itching scores during a specific week…". In addition, because in this study the Day 1 NRS score is planned to be taken after dosing, weeks are re-defined in this SAP one day earlier than stated in the protocol (eg. Days 1 to 7, Days 8 to 14, Days 15 to 21, etc.).

- Section 8.6.2 of the protocol states that missing Skindex-10 Itch Scale and 5-D Itch Scale total scores will not be imputed. To ensure a more unbiased assessment of these secondary endpoints, which does not ignore missing data, analyses in this SAP state that the same MI method with MAR assumption will be applied to these endpoints in similar fashion to the primary endpoint.

- Section 8.6.3 of the protocol stated that the study would be considered positive if the null hypothesis of no treatment difference is rejected in favor of the alternative that patients randomized to CR845 experience significantly less itching as measured by the change from baseline in the weekly mean of the Worst Itching Intensity score at Week 12. In fact, since this study compares several CR845 doses versus placebo, the study will be considered positive if at least one dose is found to be safe and effective. Effectiveness will be evaluated based on results of the primary efficacy analysis either using the estimate of treatment effect or the p-value of the hypothesis tests.

10 REFERENCES

1. ICH Topic E9: Statistical Principles for Clinical Trials (CPMP/ICH/363/96 – adopted March 1998).
2. Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. Br J Dermatol. 2010;162(3):587-93.
3. Mehta C. and Pocock J Adaptive increase in sample size when interim results are promising: A practical guide with example.

11 APPENDICES

Appendix A: Worst Itching Intensity Numerical Rating Scale

This is a representation of the content of the instrument to be used. Please refer to the Study Reference Manual for the instrument to be administered to patients and instructions.

INSTRUCTIONS

Please indicate the intensity of the **WORST ITCHING** you experienced over the past 24 hours by marking the box with the number that best describes it. After completing the scale below, please provide your initials in the **SUBJECT INITIALS** box indicating that you completed the scale by yourself and the **DATE** and **TIME** you completed the scale.

Worst Itching Over the Past 24 Hours

Please indicate the intensity of the **WORST ITCHING** you experienced over the past 24 hours.

0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐

NO ITCHING — WORST ITCHING IMAGINABLE

Date Completed: ☐☐ ☐☐☐ 2 0 ☐☐ (D D M M M Y Y Y Y)

Time: ☐☐ : ☐☐ ☐ AM ☐ PM

SUBJECT INITIALS (First Middle Last) ☐☐☐

Appendix B: Skindex-10 Scale

This is a representation of the content of the instrument to be used. Please refer to the Study Reference Manual for the instrument to be administered to patients and instructions.

INSTRUCTIONS: During the past **WEEK**, how often have you been bothered by:

	0 (Never bothered)	1	2	3	4	5	6 (Always bothered)
1. Your itching	☐	☐	☐	☐	☐	☐	☐
2. The persistence/reoccurrence of your itching	☐	☐	☐	☐	☐	☐	☐
3. The appearance of your skin from scratching	☐	☐	☐	☐	☐	☐	☐
4. Frustration about your itching	☐	☐	☐	☐	☐	☐	☐
5. Being annoyed about your itching	☐	☐	☐	☐	☐	☐	☐
6. Feeling depressed about your itching	☐	☐	☐	☐	☐	☐	☐
7. Feeling embarrassed about your itching	☐	☐	☐	☐	☐	☐	☐
8. The effects of your itching on your interactions with others (for example: interactions with family, friends, close relationships, etc.)	☐	☐	☐	☐	☐	☐	☐
9. The effects of your itching on your desire to be with people	☐	☐	☐	☐	☐	☐	☐
10. The effect of your itching making it hard to work or do what you enjoy	☐	☐	☐	☐	☐	☐	☐

Date Completed: ☐☐ ☐☐☐ 2 0 ☐☐ (D D M M M Y Y Y Y)

Time: ☐☐ : ☐☐ ☐ AM ☐ PM

SUBJECT INITIALS (First Middle Last) ☐☐☐

Patients will be asked to 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 total score is the sum of the numeric value of each answered question. 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).

Appendix C: 5-D Itch Scale

This is a representation of the content of the instrument to be used. Please refer to the Study Reference Manual for the instrument to be administered to patients and instructions.

1.	DURATION:	During the last 2 weeks, how many hours a day have you been itching?				
		Less than 6 hrs/day ☐	6-12 hrs/day ☐	12-18 hrs/day ☐	18-23 hrs/day ☐	All day ☐
2.	DEGREE:	Please rate the intensity of your itching over the past 2 weeks				
		Not present ☐	Mild ☐	Moderate ☐	Severe ☐	Unbearable ☐
3.	DIRECTION:	Over the past 2 weeks has your itching gotten better or worse compared to the previous month?				
		Completely resolved ☐	Much better, but still present ☐	Little bit better, but still present ☐	Unchanged ☐	Getting Worse ☐
4.	DISABILITY:	Rate the impact of your itching on the following activities over the last 2 weeks				
		Never affects sleep	Occasionally delays falling asleep	Frequently delays falling asleep	Delays falling asleep and occasionally wakes me up at night	Delays falling asleep and frequently wakes me up at night
	Sleep	☐	☐	☐	☐	☐
		N/A	Never affects this activity	Rarely affects this activity	Occasionally affects this activity	Frequently affects this activity	Always affects this activity
	Leisure/Social	☐	☐	☐	☐	☐	☐
	Housework/ Errands	☐	☐	☐	☐	☐	☐
	Work/School	☐	☐	☐	☐	☐	☐

5.	DISTRIBUTION:	Mark whether itching has been present in the following parts of your body over the last 2 weeks. If a body part is not listed, choose the one that is closest anatomically.			
		Head/Scalp	☐	Soles	☐
		Face	☐	Palms	☐
		Chest	☐	Tops of Hands/Fingers	☐
		Abdomen	☐	Forearms	☐
		Back	☐	Upper Arms	☐
		Buttocks	☐	Points of Contact w/ Clothing (e.g waistband, undergarment)	☐
		Thighs	☐		
		Lower legs	☐	Groin	☐
		Tops of Feet/Toes	☐		

Date Completed: ☐☐ ☐☐☐ 2 0 ☐☐ (D D M M M Y Y Y Y)

Time: ☐☐ : ☐☐ ☐ AM ☐ PM

SUBJECT INITIALS (First Middle Last) ☐☐☐

The scores of each of the five domains are achieved separately and then summed together to obtain a total 5-D Itch score. 5-D Itch scores can potentially range between 5 (no pruritus) and 25 (most severe pruritus)[2].

Single-item domain scores (duration, degree and direction) are equal to the value indicated below the response choice (range 1–5)[2]. 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.

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.

Appendix D: Patient Global Impression of Change

This is a representation of the content of the instrument to be used. Please refer to the Study Reference Manual for the instrument to be administered to patients and instructions.

Since the start of the study, my itch is:

1. ☐ Very much improved
2. ☐ Much improved
3. ☐ Minimally improved
4. ☐ Not changed
5. ☐ Minimally worse
6. ☐ Much worse
7. ☐ Very much worse

Date Completed:	Time:	SUBJECT INITIALS
D D M M M 2 0 Y Y Y Y	: ☐ AM ☐ PM	First Middle Last

Appendix E: Patient Global Impression of Change

This is a representation of the content of the instrument to be used. Please refer to the Study Reference Manual for the instrument to be administered to patients and instructions.

Please indicate the severity of the **WORST ITCHING** you experienced over the <u>past 24 hours</u> by marking the box with the category that best describes it.

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very severe

Date Completed:	Time:	SUBJECT INITIALS
D D M M M 2 0 Y Y Y Y	: ☐ AM ☐ PM	First Middle Last

Appendix [redacted]

[redacted]