CLINICAL STUDY PROTOCOL

A Double-blind, Placebo-Controlled, Randomized Withdrawal Following Open-label Therapy Study to Assess the Safety and Efficacy of Levoketoconazole (2S, 4R-ketoconazole) in the Treatment of Endogenous Cushing's Syndrome

Statistical Analysis Plan

Cortendo AB*

COR-2017-01 (LOGICS)

Covance Study ID: 39204

Document Version: Amendment 2

Document Date: August 13, 2020

Covance Inc. CDCS

Clinical Development Commercialization Services

*Cortendo AB is a subsidiary of Strongbridge Biopharma plc.

Table of Contents

TABLE OF TABLES

TABLE OF FIGURES

Approvals

The undersigned agree that all required reviews of this document are complete and approve this Statistical Analysis Plan as final. Programming of the tables, figures and listings based upon the specifications within this document can proceed.

Approved by:
Covance Lead Statistician

DocuSigned by:
Shahnaz Hawkins
080602C8D33C486...
Signature

13AUG2020
Date

Shahnaz Hawkins, Sr. Principal Biostatistician, Covance Inc.

Covance Management Approval:

DocuSigned by:
Dan Pericak
8BE6758C40E346F...
Signature

13AUG2020
Date

Dan Pericak, Director, Biostatistics, Covance Inc.

Cortendo AB Approval:

DocuSigned by:
Robbie Huang
B0AD258A6449455...
Signature

13AUG2020
Date

Robbie Huang, Associate Director, Biostatistics and Statistical Programming, Strongbridge Biopharma plc on behalf of Cortendo AB

DocuSigned by:
Fred Cohen
3BEC9ADD380C478...
Signature

13Aug2020
Date

Fredric Cohen, Chief Medical Officer, Strongbridge Biopharma plc on behalf of Cortendo AB

Reviewers

The following reviews of the Statistical Analysis Plan were conducted by Covance:

Name and Title	Role	Version Last Reviewed	Company/ Organization
Dan Pericak, Director, Biostatistics	Senior Statistical and QC Review	Draft 3, Version 0.1	Covance
Richard McNally, Statistical Fellow, Statistical Analysis Research Center	Senior Statistical and General Review	Draft 3, Version 0.1 (comments from sponsor)	Covance
Dan Pericak, Director, Biostatistics	Senior Statistical and QC Review, Updates as per the sponsor comments	Draft 3, Version 0.1 (comments from sponsor)	Covance
Dan Pericak, Director, Biostatistics	Senior Statistical and QC Review, Updates as per the sponsor comments	Draft 4, Version 0.1 (comments from sponsor)	Covance
Margaret, Connolly, Sr. Principal Biostatistician	Senior Statistical and QC Review, Updates as per the sponsor comments	Draft 5, Version 0.1 (comments from sponsor)	Covance
Margaret, Connolly, Sr. Principal Biostatistician	Senior Statistical and QC Review, Updates as per the sponsor comments	Draft 6, Version 0.1 (comments from sponsor)	Covance
Mark Evangelista, Biostatistician II	Peer Statistical and QC Review, Updates as per the sponsor comments	SAP Amendment version 1	Covance
Margaret, Connolly, Sr. Principal Biostatistician	Senior Statistical and QC Review, Updates as per the sponsor comments	SAP Amendment version 1	Covance
Dan Pericak, Director, Biostatistics	Senior Statistical and QC Review	SAP Amendment version 2	Covance

Version History

Version #	Description of Changes	Version Date
1	Initial version	12 November 2019
Amendment 1	1. Updates to Glossary of Abbreviations 2. Section 5.10 - Revision to the paragraph on the handling of missing data for the analysis of secondary efficacy endpoints by visit in the RW Phase and RES Phase to delete the reference to imputation to replace missing values. 3. Section 6.3 – Added text regarding subjects affected by COVID-19 related study disruption and protocol deviations arising from COVID-19 illness and/or public health control measures. 4. Section 6.4.1 - Updated the variables to be presented for demographic, baseline, and disease characteristics. 5. Sections 6.6.1 and 6.6.2 - Added details regarding subgroup analyses for the primary and secondary efficacy endpoints. 6. Section 6.6.2 - Clarified that End of RW Phase time point is the main time point of interest for secondary efficacy analyses assessed by visit during the RW Phase. 7. Section 6.6.2 - Elevation of the normalization of mUFC at the end of the RW Phase to the first among the 6 sets of secondary efficacy endpoints. 8. Section 6.6.2 - Deletion of statistical tests to compare the two treatment groups with respect to their mean percent changes from RW Baseline in the secondary efficacy endpoints by visit. 9. Section 6.6.2.1 - Revision to the definition of the normalization of mUFC at the end of RW Phase, using the last non-missing post-baseline mUFC result during the RW phase (before discontinuing or early rescue), instead of mUFC value at RW5 (or RW4 if RW5 value is missing). 10. Section 6.6.3 - Deletion of a sensitivity analysis on the secondary efficacy endpoints of change from RW Baseline in mUFC and biomarkers for CS comorbidities in the RW Phase that was supposed to exclude data collected after early rescue for subjects randomized to placebo, since there are no such data as the RW Phase concludes once subjects are early rescued. 11. Section 6.6.3 - Added sensitivity analyses to describe the handling of potential impact of Coronavirus 2019 on the SAP.	22 June 2020

Version #	Description of Changes	Version Date
	12. Section 6.6.5 - Expansion of the analysis of the exploratory endpoint normalization of mUFC at the end of TM Phase to also include normalization of mUFC at any time during the TM Phase 13. Sections 6.7.2, 6.7.4, and 6.7.5 - Added analysis of laboratory parameters, vital signs, and ECGs based on PCS criteria 14. Section 6.7.3 - Deletion of insulin from the OGTT analyses, because the blood samples for OGTT have not been analyzed for this parameter by the central laboratory; hence, insulin-related analysis cannot be performed. 15. Section 6.7.6 – This section has been expanded to provide more details on the data from the MRI central reads that will be summarized. 16. Section 6.8 – The reference to a top line analysis has been deleted, since all analyses, both top line and non-top line, will be included in the interim analysis study report. 17. Section 7 - New text added to describe changes from planned analyses in the protocol. Previously, this section stated that there were no changes from the planned analyses. 18. There were also minor edits and updates for accuracy and completeness.	
Amendment 2	1. Section 5.3 Table 3 – Added clarification for End of RW Phase analysis visit window in Table 3. 2. Section 5.4 – Various updates added to clarify how samples were taken for the S-C cohort, the differences in sample date for the first two samples versus the third confirmatory sample, and how different samples should be grouped together for mUFC calculation. 3. Section 5.6 – Various updates added to clarify how different samples were collected and how samples mapped to different visit should be analyzed. 4. Section 5.9 – Add clarification for missing grade for irregular menstruation. 5. Section 6.1 Table 4 – Added clarification of using Satterthwaite method in 2-sample t-test. 6. Section 6.6.2.1 – Added clarification on definition of mUFC normalization at the end of RW Phase; updated "Changes in mUFC and LNSC from RW Baseline" section with the addition of Wilcoxon rand sun test.	13 August 2020

Version #	Description of Changes	Version Date
	7. Section 6.6.2.2 – Added definition for HOMA-IR, as well as exclusions for related analysis. 8. Section 6.6.2.2 – Removed LDL, Total Cholesterol, and hsCRP from the repeated measures mixed effect model section as these parameters were only measured at Baseline (RW0) and RW5, therefore the model doesn't apply. 9. Section 6.6.5 – Updated analysis on medication changes. 10. Section 6.6.5 HOMA%B Analysis – Corrected an error and clarified exclusions for related analysis. 11. Section 6.7.2 Table 6 – Corrected errors in PCS criteria - SI unit. 12. Section 6.7.3 – Clarified that time to maximum glucose is calculated from the time of glucose administration 13. Section 6.7.5 – Clarified that QTcF results and time elapsed relative to study drug/PK draw will be summarized and listed, and rejected ECG readings will not be excluded from listings.	

Glossary of Abbreviations

Abbreviation	Term
ACTH	Adrenocorticotrophic Hormone
ASCVD	Atherosclerotic Cardiovascular Disease
AICC	Corrected Akaike Information Criterion
AE	Adverse Event
ALT	Alanine Aminotransferase
AP	Alkaline Phosphatase
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Concentration Time Curve
BID	Twice Daily
BDI	Beck Depression Inventory
BMI	Body Mass Index
CI	Confidence Interval
CD	Cushing's Disease
CL/F	Clearance
Cmax	Peak Concentration
COVID-19	Coronavirus Disease 2019
CS	Cushing's Syndrome
DBP	Diastolic Blood Pressure
DSMB	Data Safety and Monitoring Board
eCRF	Electronic Case Report Form
ECG	Electrocardiogram
eDISH	Evaluation of Drug-induced Serious Hepatotoxicity
GGT	Gamma-glutamyl Transferase
FPI	Fasting Plasma Insulin
FPG	Fasting Plasma Glucose
HbA1c	Hemoglobin A1c
HDL-C	High-Density Lipoprotein-Cholesterol
HOMA-IR	Homeostatic Model Assessment-Insulin Resistance
HOMA-%B	Homeostatic Model Assessment-Beta Cell Function
HR	Heart Rate
HRQoL	Health-related Quality of Life
hsCRP	High Sensitivity C-reactive Protein
IC50	Levoketoconazole Concentration Producing Half Maximal UFC Suppression
ICH	International Conference on Harmonization
Imax	Maximal Suppression of UFC
ITT	Intent-to-Treat
IFG	Impaired Fasting Glucose
Ka	Absorption Rate Constant
kg	Kilogram
KM	Kaplan-Meier
L	Liter
LDH	Lactate Dehydrogenase
LDL-C	Low-Density Lipoprotein-Cholesterol
LFT	Liver Function Tests
LLN	Lower Limit of Normal

L-N	Levoketoconazole-naïve
LNSC	Late Night Salivary Cortisol
Max	Maximum
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
Min	Minimum
mL	Millilitre
MRI	Magnetic Resonance Imaging
mUFC	Mean Urinary Free Cortisol
n	Number of subjects/observations
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OGTT	Oral Glucose Tolerance Test
PD	Pharmacodynamic
PK	Pharmacokinetics
PP	Per Protocol
PT	Preferred Term
QoL	Quality of Life
RBC	Red Blood Cell
RES	Restoration (Phase)
RW	Randomized Withdrawal (Phase)
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
S-C	SONICS-completer
SD	Standard Deviation
SE	Standard Error
SI	International System of Units
SOC	System Organ Class
TFLs	Tables, Figures, and Listings
TM	Dose Titration and Maintenance (Phase)
TSH	Thyroid-stimulating Hormone
t½	Half-life
UFC	Urinary Free Cortisol
ULN	Upper Limit of Normal
US	United States
V/F	Volume of Distribution
WBC	White Blood Cell
WHO	World Health Organization

1. Source Documents

The Statistical Analysis Plan (SAP) was written based on the following documentation:

Document	Date	Version
Protocol	17 April 2017	Original
Protocol Amendment 1	06 July 2017	Amendment # 1
Protocol Amendment 2	21 June 2018	Amendment # 2
Protocol Amendment 3	14 December 2018	Amendment # 3
Protocol Amendment 4	23 September 2019	Amendment # 4
Coronavirus Disease 2019 (COVID-19) Annex	14 April 2020	Version 1.0
Electronic Case Report Form (eCRF)	29 January 2020	Version 8.0

The planned statistical methodologies described in this SAP are in accordance with the principles outlined in the International Conference on Harmonization (ICH) E9.

2. Protocol Details

2.1 Study Objectives

The objectives of Study COR-2017-02 (also referred to as LOGICS) are:

- Primary:

  To determine the effect of withdrawing to placebo versus continuing treatment with levoketoconazole on the cortisol therapeutic response previously established during open-label levoketoconazole therapy.

- Secondary:

  1. To compare the effects of levoketoconazole with placebo on cortisol status (inferred from mean urinary free cortisol [mUFC] and/or late night salivary cortisol [LNSC]) during the Randomized Withdrawal (RW) Phase and the subsequent Restoration (RES) Phase;
  2. To compare the effects of levoketoconazole with placebo on changes in biomarkers of Cushing's Syndrome (CS) comorbidities (fasting glucose, fasting insulin, homeostatic model assessment-insulin resistance [HOMA-IR], hemoglobin A1c [HbA1c], total cholesterol, low-density lipoprotein-cholesterol [LDL-C], and high-sensitivity C-reactive protein [hsCRP]);
  3. To compare the effects of levoketoconazole with placebo on changes in health-related quality of life (QoL) and symptoms of depression;
  4. To compare the effects of levoketoconazole with placebo on changes in acne, hirsutism and peripheral edema;
  5. To assess the safety and tolerability of levoketoconazole;
  6. To evaluate the population pharmacokinetics (PK) of levoketoconazole in subjects with CS.

  NOTE: Secondary Objectives 5 and 6 are not subject to hypothesis tests.

- Exploratory:

  1. To assess changes in anti-diabetic, anti-cholesterol, anti-hypertensive, and chronic anti-inflammatory therapies;
  2. To describe the effects and durations of levoketoconazole action with respect to cortisol status and clinical signs and symptoms of CS other than acne, hirsutism and peripheral edema;
  3. To describe the dose-response relationship of levoketoconazole with respect to safety and tolerability;

4. To describe the effects of levoketoconazole on glucose tolerance among subjects with impaired fasting glucose (IFG).

2.2 Overall Study Design

The study is a double-blind, placebo-controlled, randomized withdrawal and restoration (or as needed early rescue) study in subjects with endogenous CS previously treated with single-arm, open-label levoketoconazole that will assess efficacy, safety, tolerability, and PK of levoketoconazole.

Around 45 sites in North America, Europe and the Middle East will enroll up to approximately 54 subjects into the RW Phase of the study.

Two populations (cohorts) will be eligible for participation in the study. One cohort is based on previous participation in the COR-2012-01 (aka SONICS) trial and usage of a stable Therapeutic Dose of levoketoconazole in the 12-week timeframe prior to screening. This cohort will represent a minority of randomized subjects. The second cohort includes subjects completely naïve to levoketoconazole (the largest proportion of randomized subjects) and those who previously participated in SONICS but were not on a stable Therapeutic Dose in the 12-week timeframe prior to screening (also a minority of randomized subjects). The cohort determination will be collected in the clinical database and will not be derived.

Table 1 Cohort Determination

Cohort 1: Levoketoconazole-naïve	(1) Subjects who did not participate in the prior clinical study of levoketoconazole (COR-2012-01 aka SONICS); or (2) Subjects who completed Visit Month 12 (M12) of SONICS (final scheduled visit prior to safety follow-up) but who have **not been treated** with a Therapeutic Dose of levoketoconazole throughout the 12-week period immediately prior to Screening in the current study.
Cohort 2: SONICS-completer	(1) Subjects who have completed all visits in the SONICS study through M12 (final scheduled visit prior to safety follow-up) at any time prior to the current study and **who have been treated** with a Therapeutic Dose of levoketoconazole throughout the 12-week period immediately prior to Screening.

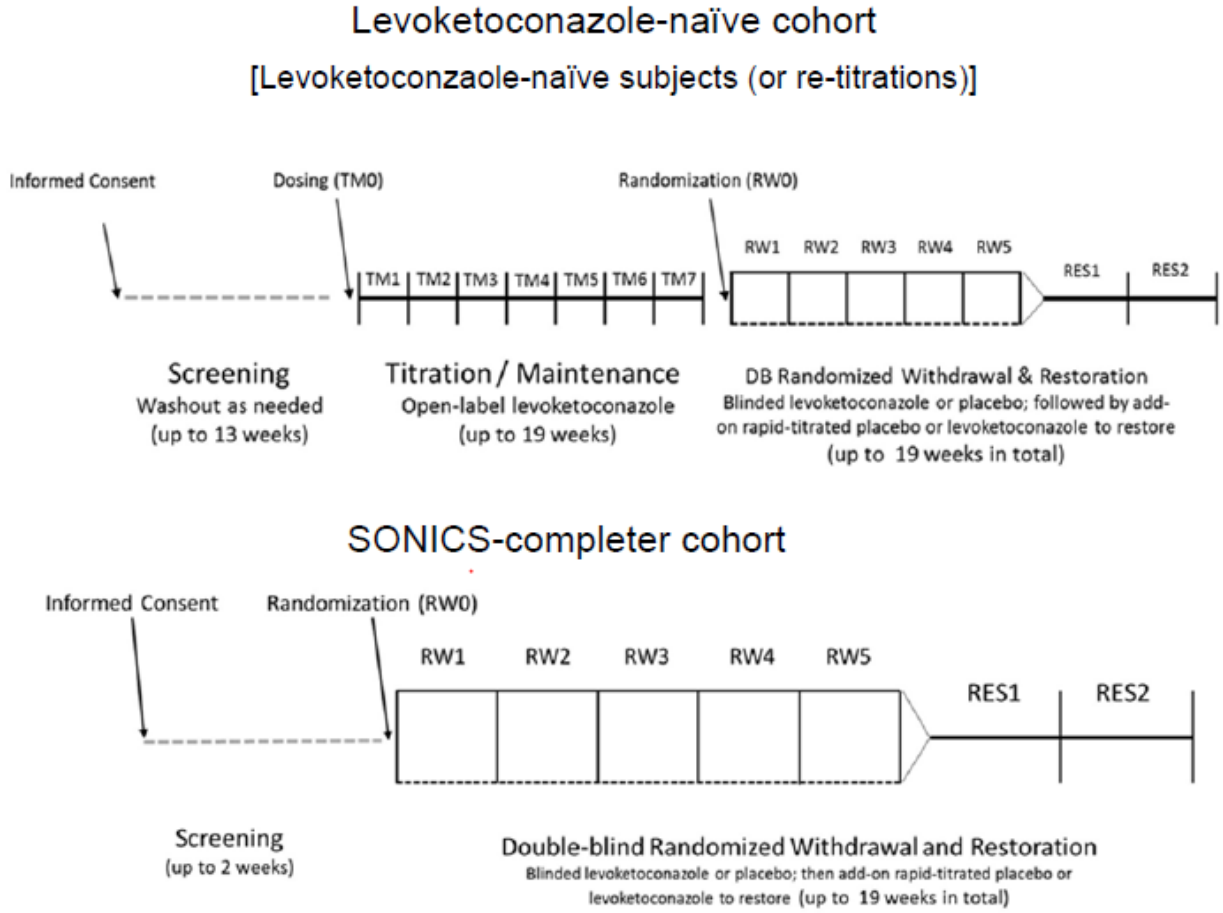
Total duration of the study will depend on the cohort.

Cohort 1: Subjects in the levoketoconazole-naïve (L-N) cohort will be enrolled for up to approximately 51 weeks due to the Screening and Dose Titration and Maintenance (TM) phases of the study.

Cohort 2: Subjects in the SONICS-completer (S-C) cohort who have established a Therapeutic Dose level of levoketoconazole and will be enrolled for less time (up to 21 weeks).

The overall study design is presented in Figure 1 below.

Figure 1 Study Design Schematics



Note: RW5, the last visit of the Randomized Withdrawal marks the start of the Restoration Phase

TM Phase: After confirmation of eligibility, subjects in the L-N cohort (including subjects who completed SONICS but require re-titration) will enter the TM Phase. The first dose of open-label therapy for this phase will be at the TM0 visit. Subjects who did not complete Visit M12 of SONICS will always begin titration at Dose Level 1 (150 mg BID). Subjects who have completed Visit M12 in SONICS and have **not** been treated with a Therapeutic Dose of levoketoconazole throughout the 12-week period immediately prior to Screening must also re-establish a Therapeutic Dose and remain in the TM Phase for a full 14-week period, regardless of the duration needed to establish their Therapeutic Dose for this study (i.e. even if it takes less than 14 weeks).

The RW Phase and RES Phase of the study will be the same for both cohorts. All subjects have established their own Therapeutic Dose level prior to the start of the double-blind RW Phase.

- RW Phase: The subjects will be randomized on a double-blind basis to either continue levoketoconazole at the Therapeutic Dose level or receive matched placebo.
- RES Phase: Subjects who complete the RW Phase and do not meet any of the early rescue criteria will enter the double-blind RES Phase. The levoketoconazole Therapeutic Dose level is maintained for those subjects randomized to levoketoconazole and is restored for those subjects randomized to placebo. Those who have been randomized to levoketoconazole will receive matched placebo, while those who have been randomized to placebo will receive levoketoconazole additionally.

The treatment schema for the study is presented in the table below.

Table 2 Study Treatment Schema

Open-label levoketoconazole	Therapeutic Dose established from TM Phase for subjects in the L-N cohort, or from SONICS study for S-C cohort	
Randomization (RW0)	⇩	⇩
Blinded Randomized Withdrawal Phase	**Levoketoconazole** at Therapeutic Dose	**Placebo** at equivalent tablet count to Therapeutic Dose
Blinded Restoration Phase	**Placebo** added as 1 tablet twice daily (BID) at RW5 then add 1 tablet every 2 days, alternating AM and PM, until the total tablet count is double the count at RW0	**Levoketoconazole** added as 1 tablet BID at RW5 then add 1 tablet every 2 days, alternating AM and PM, until the total tablet count is double the count at RW0

Subjects who meet any of the early rescue criteria during the RW Phase and are rescued with open-label levoketoconazole may continue in the RES Phase, except that therapy will be administered as open-label therapy. Subjects who meet any of the early rescue criteria during the RW Phase, but are not rescued with open-label levoketoconazole, will be discontinued from the study. For the purposes of defining study completion, however, subjects who complete RW0 and met any of the early rescue criteria during the RW Phase are considered as study completers (see Section 5.10).

2.3 Sample Size and Power

For the purpose of sample size estimation, it is assumed that among subjects in the intent-to-treat (ITT) population who complete at least the RW4 visit in the RW Phase, 10% will lose therapeutic response in the levoketoconazole arm versus 75% in the placebo arm. The 10% loss of therapeutic response rate in the levoketoconazole arm is based on data from the SONICS study.

For the analysis of the primary endpoint, subjects in the ITT population who withdraw prior to RW4 are considered as losing therapeutic response. As such, the assumed loss of therapeutic response rate in the overall ITT population in each arm depends on the withdrawal rate prior to RW4 during the RW Phase. It is expected that the withdrawal rate will be no more than 10%, but up to 20% is considered in the sample size estimation. If the withdrawal rate is 10% per arm, a sample size of 46 subjects (23 per arm) randomized and treated during the RW Phase corresponds to 42 subjects (21 per arm) completing RW4 and provides approximately 99% power, using two-sided Fisher's Exact test, to detect the alternative hypothesis of a loss of therapeutic response rate of 17% in the levoketoconazole arm and 78% in the placebo arm versus the null hypothesis of no difference between the two arms. If the withdrawal rate is 20% per arm, a sample size of 54 subjects (27 per arm) randomized and treated during the RW Phase corresponds to 44 subjects (22 per arm) completing RW4 and provides approximately 98% power to detect the alternative hypothesis of a loss of therapeutic response rate of 26% in the levoketoconazole arm and 81% in the placebo arm versus the null hypothesis of no difference between the two arms.

Because enrollment into the TM Phase will be closed prior to the last subject being randomized and treated in the RW Phase, the actual sample size for the RW Phase is not prespecified. The timing of enrollment closure will be based on the rolling 4-week completion rate trend for the TM Phase. A minimum of 46 subjects (23 per arm) and a maximum of 54 subjects (27 per arm) will be targeted to be enrolled into the RW Phase based on the observed withdrawal rate trend for the phase.

A minority of randomized subjects are anticipated to be in the S-C cohort. Any subject who withdraws after randomization will not be replaced. Withdrawn subjects will not be re-entered into the study.

3. Study Endpoints

3.1 Primary Efficacy Endpoint

Loss of Therapeutic Response

The primary efficacy endpoint is proportion of subjects with loss of therapeutic response to levoketoconazole upon withdrawing to placebo compared with the proportion of subjects with loss of therapeutic response upon continuing treatment with levoketoconazole. Loss of therapeutic response (i.e., relapse) is inferred based on the mean of three 24-hour UFC measurements (mUFC) obtained at any visit from second through final randomized-withdrawal visits (RW1 through RW5 inclusive) when:

- mUFC is above 1.5X the upper limit of normal (ULN) of the central laboratory's reference range OR
- mUFC is more than 40% above the Baseline (RW0) value, if the Baseline value is above the ULN (i.e. > 1.0X ULN)[1] OR
- an early rescue criterion is met.

Early Rescue Criteria

Early rescue during the RW Phase must be considered when a subject demonstrates one or more of the following criteria at any time after randomization:

1. Relapse of hypercortisolemia (i.e. loss of therapeutic response). Defined as:
   - mUFC from **three** urine collections that is above 1.5X ULN
     OR
   - mUFC from **three** urine collections that increases by 40% or more above the Baseline (RW0) value and when the Baseline value is above 1X ULN, for the S-C cohort only.
2. Partial loss of cortisol response. Defined as:
   - mUFC from three urine collections that increases from Baseline (RW0) to greater than 1X ULN but no more than 1.5X ULN
     OR
   - LNSC that was normal at Baseline (RW0; both values ≤ULN) and becomes abnormal (>ULN) based on each of two LNSC collections obtained on different nights

[1] This category applies only to subjects who have completed SONICS, who may have a Therapeutic Dose established with mUFC above the ULN of the reference range. Levoketoconazole-naïve cohort must have mUFC at or below the ULN at RW0 to qualify for randomization.

AND

In addition to one of the above partial-loss criteria, the subject must exhibit clinically significant worsening of signs and symptoms of CS, as determined by:

- Clinically significant deterioration in disease-related signs or symptoms (using the signs and symptoms validated instrument and/or physical examination).

  OR

- Relevant deterioration in one or more disease-related biomarkers (protocol Section 4.3.1, Table 3).

3. In the absence of loss or partial loss of cortisol response as defined above, clinically significant deterioration in disease-related signs or symptoms (using the signs and symptoms validated instrument and/or physical examination) **AND** clinically relevant deteriorations in **two or more** disease-related biomarkers (see protocol Section 4.3.1, Table 3).

Early rescue that occurs without meeting one of the three pre-defined categories above will not be considered a loss of response for the purposes of the primary endpoint.

3.2 Secondary Efficacy and Safety Endpoints

- Changes from Baseline (RW0) in mUFC and LNSC at all post-baseline visits with these assessments through the final study visit (RES2)—applies to Secondary Objective 1;
- Proportion of subjects with normalization of mUFC at RES2—applies to Secondary Objective 1;
- Proportion of subjects with normalization of mUFC at the end of RW Phase (RW5) -applies to Secondary Objective 1;
- Changes from Baseline (RW0) in biomarkers of CS comorbidities at all post-baseline visits with these measurements through the final study visit (RES2)—applies to Secondary Objective 2;
- Changes from Baseline (RW0) in health-related QoL and symptoms of depression at all post-baseline visits with these assessments through the final study visit (RES2)—applies to Secondary Objective 3;
- Changes from Baseline (RW0) in acne, hirsutism and peripheral edema at all post-baseline visits with these assessments through the final study visit (RES2)—applies to Secondary Objective 4;

- Incidence and severity of adverse events (AEs), particularly adverse events of special interest (AESIs) during levoketoconazole open-label therapy in the TM Phase (L-N cohort) and during blinded therapy in the RW Phase and RES Phase (both cohorts)—applies to Secondary Objective 5.

3.3 Exploratory Efficacy and Safety Endpoints

- Frequency of usage and changes from Baseline (RW0) in frequency of usage of anti-diabetic, anti-cholesterol, anti-hypertensive, and chronic anti-inflammatory therapies at all post-baseline visits; changes in corresponding biomarkers accounting for changes in medication usage will also be explored—applies to Exploratory Objective 1;
- Time from RW0 to first time of loss of response, when:
  (1) mUFC is above 1.5X the ULN of the central laboratory's reference range, OR
  (2) mUFC is more than 40% above the RW0 value, if the RW0 value is above the ULN (i.e. >1.0X ULN)[2], OR
  (3) an early rescue criterion is met—applies to Exploratory Objective 2;
- Time to first normalization of mUFC beginning from RW5 (subset with mUFC above 1.5X ULN at RW5)—applies to Exploratory Objective 2;
- Time to first normalization of LNSC beginning from RW5 (subset with LNSC above ULN at RW5) —applies to Exploratory Objective 2;
- Proportion of subjects with normalization of mUFC at the end of TM Phase (TM7) - applies to Exploratory Objective 2;
- Proportion of patients with either normalization of mUFC or partial response (at least 50% decrease in mUFC) at the end of TM Phase (TM7) - applies to Exploratory Objective 2;
- Proportion of subjects with normalization of LNSC at RES2—applies to Exploratory Objective 2;
- Changes from Baseline (RW0) in serum cortisol and adrenocorticotrophic hormone (ACTH) at all post-baseline visits with these assessments through the final study visit (RES2)—applies to Exploratory Objective 2;
- Changes from Baseline (RW0) in clinical signs and symptoms of CS excluding acne, hirsutism and peripheral edema at all post-baseline visits with these

[2] This category applies only to subjects who have completed SONICS, who may have a Therapeutic Dose established with UFC above the ULN of the reference range. Levoketoconazole-naïve cohort must have UFC at or below the ULN at RW0 to qualify for randomization.

assessments through the final study visit (RES2)—applies to Exploratory Objective 2;

- Frequency and severity of common AEs and laboratory abnormalities in relation to dose of study drug administered at the time of the reported AE or laboratory abnormality—applies to Exploratory Objective 3;
- Shifts from normality and concentration changes from Baseline (RW0) in serum transaminases, alkaline phosphatase (AP), and total bilirubin at all post-baseline visits in relation to dose of study drug administered at the time of the shift or change—applies to Exploratory Objective 3;
- Durations and changes in durations from Baseline (RW0) of the QT interval corrected for heart rate (QTc interval) in relation to dose of study drug administered proximal to the measurement—applies to Exploratory Objective 3.
- Change from Baseline (RW0) in observed and derived glucose and insulin parameters during oral glucose tolerance test (OGTT) in the subset of subjects with IFG—applies to Exploratory Objective 4.

3.4 Pharmacokinetic/Pharmacodynamic Endpoints

- Estimates of the following PK parameters: clearance (CL/F), volume of distribution (V/F), absorption rate constant (Ka), with associated between-subject variability where feasible. These parameters will be used to calculate half-life (t½), area under the concentration time curve (AUC) and peak concentration (Cmax), if feasible—applies to Secondary Objective 6;
- Estimates of the following pharmacodynamic (PD) parameters: levoketoconazole concentration producing half maximal UFC suppression (IC50), maximal suppression of UFC (Imax) and associated estimates of between-subject variability, if feasible. UFC concentrations in relation to dose and plasma exposure will be explored—applies to Secondary Objective 6.

4. Analysis Populations

4.1 Intent-to-Treat (ITT) Population

The ITT population will include all subjects who are randomized and receive at least one dose of blinded study drug during the RW Phase. The ITT population will be used for the primary analysis of efficacy and secondary analyses of efficacy and safety. For efficacy, subjects will be analyzed by their randomized treatment assignment. For safety, subjects will be analyzed by the actual study drug they receive.

The ITT population for levoketoconazole in the RES Phase will include all subjects who receive at least one dose of levoketoconazole during the RES Phase. It will be referred to as the RES ITT population and will be used for the analysis of efficacy and safety during the RES Phase.

4.2 Per-Protocol (PP) Population

The PP population will consist of all subjects in the ITT population who have no major protocol deviations during the RW Phase that could affect the primary endpoint. This population will be used in supportive analyses of the primary endpoint. Selective secondary efficacy endpoints will also be analyzed using PP population if PP population < 90% of the ITT population.

A protocol deviation is defined as any change, divergence, or departure from the study design or procedures defined in the study protocol. Major protocol deviations are a subset of protocol deviations and may significantly impact the correctness, accuracy, and/or reliability of the study data or may significantly affect a subject's rights, safety, or well-being.

4.2.1 Major Protocol Deviations Leading to Exclusion from PP Population Analysis

All major (also known as important) protocol deviations occurring during the RW Phase of the study will be reviewed and approved by the Sponsor prior to the lock of the database containing data through the end of the RW Phase and prior to treatment unblinding. Of these major protocol deviations, those that will lead to exclusion of subjects from the PP population will be identified by the Sponsor prior to the first database lock that will occur after the last subject completes the RW Phase.

The list of categories of protocol deviations during the RW Phase that may be considered major by the Sponsor, subject to review of the actual deviations that occur, are provided in a separate document called the Study Specific Protocol Deviation List dated August 06, 2019.

4.3 Levoketoconazole-naïve(L-N) Population

The L-N population will consist of all subjects in the L-N cohort who receive at least one dose of study drug during the TM Phase. This population will be used for supportive evaluations of safety and for exploratory efficacy, safety, and PK analyses.

4.4 SONICS-completer Population

The S-C population will consist of all subjects in the S-C cohort who receive at least one dose of blinded study drug during the RW Phase. This population will be used in summaries of disposition, demographics and baseline characteristics, and prior medications by subject cohort. Because the size of this cohort is n=5, which is < 30% of the minimum expected ITT population size, there will be no additional efficacy, safety, and PK analyses for this population except for those explicitly specified in Sections 6.6.1 and 6.6.2.1 (for the primary efficacy endpoint and Set 1 of the secondary efficacy endpoints).

4.5 Safety Population

The Safety population will be determined based on study phase and the study overall. Subjects will be analyzed by the actual study drug they receive.

The Safety population for the study overall (i.e. all phases combined), referred to simply as the Safety population, will include all subjects who receive at least one dose of study drug during any of the 3 study phases.

The Safety population for levoketoconazole for the TM Phase will include all subjects who receive at least one dose of levoketoconazole during the TM Phase. This is the same as the L-N population defined in Section 4.3 and will be referred to as such.

The Safety population for the RW Phase will include all subjects who receive at least one dose of blinded study drug during the RW Phase. This is the same as the ITT population defined in Section 4.1 and will be referred to as such.

The Safety population for levoketoconazole for the RES Phase will include all subjects who receive at least one dose of levoketoconazole during the RES Phase. This is the same as the RES ITT population defined in Section 4.1 and will be referred to as such.

The Safety population for levoketoconazole for the RW Phase and RES Phase combined, will include all subjects who receive at least one dose of levoketoconazole during the RW Phase or the RES Phase.

The ITT population is the main Safety population of interest. All safety endpoints will be analyzed on this Safety population. The other definitions of Safety populations will be used for supportive evaluations of safety.

5. Data Handling

5.1 Baseline, First Dose Date, and Last Dose Date

There are three phases of interest in the study, as listed below. The phase of main interest is the RW Phase, with the RES Phase and the TM Phase of secondary interest. For some endpoints, the combined RW Phase and RES Phase are of interest. Baseline visit, Baseline value, first dose date, and last dose date are defined for each study phase, or combination thereof, as stated below.

- TM Phase (applicable to the L-N cohort only):
    - Baseline visit is the TM0 visit.
    - Baseline value is the pre-dose value at the TM0 visit; if this value is missing or post-dose, then the last non-missing value prior to the TM0 visit, if there is any, will be used as the Baseline value.
    - First dose date (and the start of the study phase) is the TM0 visit date, as recorded in the TM0 Visit Date eCRF.
    - Last dose date is the day before the RW0 visit date (i.e. RW0 visit date – 1 day) for subjects who continue on to the RW Phase, or the date of last dose recorded in the eCRF for those subjects who discontinue otherwise.
    - End of the study phase is the last dose date for subjects who continue on to the RW Phase. It is either the last dose date or the date of study discontinuation, whichever is later, for those subjects who discontinue prior to the RW Phase.
- RW Phase:
    - Baseline visit is the RW0 visit.
    - Baseline value is the pre-dose value at the RW0 visit; if this value is missing or post-dose, then the last non-missing value prior to the RW0 visit, if there is any, will be used as the Baseline value.
    - First dose date (and the start of the study phase) is the RW0 visit date, as recorded in the RW0 Visit Date eCRF.
    - Last dose date is the RW5 visit date, as recorded in the RW5 Visit Date eCRF, minus 1 day for subjects who continue on to the RES Phase, or the date of last dose recorded in the eCRF for those subjects who discontinue otherwise. This last dose date is also considered as the end of the RW Phase. The subjects who continue on to the RES Phase include not only

the subjects who complete all visits in the RW Phase but also the early rescue subjects who receive open-label levoketoconazole as early rescue treatment.

- End of the study phase is the RW5 visit date, prior to the first dose in the RES Phase for subjects who continue on to the RES Phase. It is either the last dose date or the date of study discontinuation, whichever is later, for those subjects who discontinue prior to the RES Phase.

- RES Phase:
  - Baseline visit is the RW5 visit.
  - Baseline value is the pre-dose value at the RW5 visit; if this value is missing or post-dose, then the last non-missing value prior to the RW5 visit will be used as the Baseline value.
  - First dose date (and the start of the study phase) is the RW5 visit date.
  - Last dose date is the date of last dose recorded in the eCRF. If this date is missing for a subject who completes the RES Phase, then the RES2 date will be used instead.
  - End of the study phase is either the last dose date or the date of study completion (for subjects who complete the RES Phase) or study discontinuation (for subjects who discontinue during the RES Phase), whichever is later.
- RW Phase and RES Phase Combined:
  - Baseline visit, Baseline value, and First dose date are the same as the ones for the RW Phase above.
  - Last dose date and end of the study phase are the same as the ones for the RW Phase for subjects who discontinue prior to the RES Phase and are the same as the ones for the RES Phase for subjects who continue on to the RES Phase.
- All Phases Combined (applicable to the L-N cohort only):
  - Baseline visit, Baseline value, and First dose date are the same as the ones for the TM Phase above.
  - Last dose date and end of the study phase are the same as the ones for the TM Phase for subjects who discontinue prior to the RW Phase, the same as the ones for the RW Phase for subjects who discontinue prior to the RES Phase, and the same as the ones for the RES Phase for subjects who continue on to the RES Phase.

As there are 3 Baselines defined above, they will be designated as TM Baseline, RW Baseline, and RES Baseline. RW Baseline is the Baseline of main interest. Changes from Baseline will be calculated relative to the TM Baseline for the TM Phase, relative to the RW Baseline for the RW Phase, relative to the RES Baseline for the RES Phase, and relative to the RW Baseline for the RW Phase and RES Phase, combined.

5.2 Study Day and Relative Day

Study day is calculated within each study phase based on the first dose date within the study phase.

- If the date of assessment is before the date of first dose, then Study Day = Date of assessment – First dose date;
- If the date of assessment is on or after the first dose date, then Study Day = Date of assessment – First dose date + 1 day;
- If the first dose was never given in a study phase, then the Study day will be missing.

Relative day is calculated in an analogous manner to Study day, based on the date of interest (e.g. AE onset date) and the date of reference (e.g. date of first dose).

Day 1 within a study phase is defined as the day of first dose of study drug for that study phase. The day prior to Day 1 is Day -1. As the study days are relative to a study phase, they will be designated as TM Day x, RW Day x, and RES Day x.

5.3 Time Points and Visit Windows

Efficacy and safety endpoints for the TM Phase will not be analyzed by visit, because the L-N cohort who completed SONICS but require re-titration may begin dose titration at their current or more recently received dose of levoketoconazole, which may be higher than Dose Level 1 (300 mg/day). As such, each TM visit will have subjects at differing dose levels, which would make inferences of efficacy and safety results by visit during this phase untenable. With the exception of AEs which will be analyzed for incidence rates during the entire phase, safety endpoints will be analyzed by: 1) the worst results (for shift or threshold analyses only), 2) the last non-missing results, and 3) the non-missing results at the highest dose received during the TM Phase. Efficacy endpoints will be analyzed by: 1) the last non-missing results and 2) the non-missing results at the highest dose received during the TM Phase.

For the RW Phase and RES Phase, Table 3 below presents the visit windows that will be used for the by-visit analysis of efficacy and safety. It also presents the visit window that will be used for the End of RW Phase time point analysis of efficacy and safety.

Table 3 RW Phase and RES Phase – Visit Windows for by-visit Analysis of Efficacy and Safety

Nominal Visit/Time Point	Target Study Day (range)	Study Day Analysis Visit Window Range
RW1/D10	10 (8 to 10)[1]	RW Days 2 to 15[1]
RW2/D20	20 (18 to 20)[1]	RW Days 16 to 25[1]
RW3/D30	30 (26 to 30)[1]	RW Days 26 to 35[1]
RW4/D40	40 (38 to 40)[1]	RW Days 36 to 45[1]
RW5/D58	58 (54 to 58)[1]	RW Days 46[1] to RES Day 1[2] (or the end of the RW Phase [see Section 5.2]) for subjects who do not enter the RES Phase)
End of RW Phase	No specific target day (range) specified, so long as it is during the RW Phase.	No specific study day analysis visit window range specified, so long as it is during the RW Phase. The non-missing result from the last analysis visit window in the RW Phase for a subject will be used for the End of RW Phase time point. Note: This time point is the main analysis time point of interest for secondary efficacy endpoints that are assessed during the RW Phase.
RES1	28+/-5 days after RW5	RES Days 2 to 42[2]
RES2	28+/-5 days after RES1	RES Days 43 to 70[2,3]

[1]Study Days are relative to RW Day 1 (= RW0 visit date)
[2]Study Days are relative to RES Day 1 (=RW5 visit date)
[3]Data from assessments performed after RES Day 70 will not be considered in any by-visit analysis of efficacy and safety but will be included in the data listings.

For all populations, multiple visits within the same window will be dealt with as follows:

- If both scheduled and unscheduled visits fall within the same visit window, then the scheduled visit will be used for analysis, except for by-visit safety endpoints (e.g. laboratory evaluations) when they are analyzed with respect to normal ranges (e.g. shift tables) or prespecified thresholds; the worst results among all scheduled and unscheduled visits within the same visit window will then be used instead. For vital signs (specifically blood pressure and heart rate) and electrocardiogram (ECG) intervals (specifically QT, QTcF, QTcB, PR, and QRS), the worst result for a subject in a study phase or visit window is defined as the maximum value for that subject in that study phase or visit window. For the safety laboratory tests, the worst result for a subject in a study phase or visit window can be the maximum or minimum value, depending on the laboratory test. Refer to Appendix 1 for details on the identification of the worst results for safety laboratory tests. If multiple scheduled visits occur within a single visit window, then the visit closest to the target day of the visit window will be used in the analysis. If there is a tie, then the later scheduled visit will be used in the analysis.

- If multiple unscheduled visits occur within a single visit window (with no scheduled visit within the window), then the unscheduled visit closest to the target day of the visit window will be used in the analysis. If there is a tie, then the later unscheduled visit will be used in the analysis.

A blinded review of the analysis visit assignments for the RW Phase based on the analysis visit window ranges above will be performed by the Contract Research Organization (CRO) Lead Statistician (and confirmed by the Sponsor Lead Statistician) prior to the first database lock of all data through the end of the RW Phase, which will occur after the last subject completes the RW Phase. A blinded review of analysis visit assignments for the RES Phase based on the analysis visit windows above will be performed by the CRO Lead Statistician (and confirmed by the Sponsor Lead Statistician) prior to the final database lock after the end of the study. The reviews will check for illogical analysis visit assignments. The blinded listing that will be produced for the review will list the following data: subject id, parameter or endpoint, visit name and number in the database, whether the visit is scheduled or unscheduled, visit date (and time, if applicable), study day, and assigned analysis visit. The following parameters/endpoints will be included in the blinded listing for review: Primary and secondary efficacy endpoints, liver function tests (LFTs), ECGs, and vital signs. Only visits with non-missing data for the parameter or endpoint will be included in the listing. Any change to the analysis visit window ranges specified in this SAP as a result of the blinded review(s) will be documented in a SAP amendment prior to database lock and (in the case of the blinded reviews during the dry runs) prior to treatment unblinding.

5.4 Urinary Free Cortisol – Determination of Adequacy and Calculation of Mean

24-hour urine samples will be collected for determination of UFC during the 3 study phases. Depending on the visit, 2 or 3 samples will be collected over sequential days per the study days indicated in Appendix A (Time and Events Schedules) of the protocol, prior to the actual scheduled visit. At the Screening, TM0, TM7, RW0, RW5, and RES2 visits, 3 samples will be collected over sequential days prior to the visit. For subjects in the S-C cohort for which their first 2 samples for the Screening visit are from their Month 12 visit in SONICS, their third sample will be collected just prior to their first visit in LOGICS. At the TM1 to TM6 visits, 2 samples will be collected over sequential days prior to the visit. If the mUFC from the first 2 samples is ≤ ULN, a third sample will be collected as a confirmatory sample as soon as possible following the visit to confirm UFC normalization. At the RW1 to RW4 visits, 2 samples will be collected over sequential days prior to the visit. A third sample will be collected as a confirmatory sample as soon as possible following the visit if UFC is being used to establish the need for early rescue therapy during the RW Phase. At the RES1 visit,

2 samples will be collected over sequential days prior to the visit. In general, the first 2 samples will be on consecutive days, but the third sample may be several days after the second sample. This includes the case of subjects in the S-C cohort and their samples corresponding to the Screening visit. This also includes the case of any subject where the third sample for a visit is a confirmatory sample (i.e. either to confirm UFC normalization or the need for early rescue therapy).

For the analyses of the primary efficacy endpoint and other UFC-related endpoints, only UFC values from adequate urine collections will be included. The urine creatinine excretion rate will be measured from the 24-hour urine samples as a marker of the adequacy of the urine collections.

A urine sample for UFC is considered to be adequate for analysis purposes if its corresponding urine creatinine excretion is ≥ 4.5 mg/kg/day for females and ≥ 6.2 mg/kg/day for males. The thresholds of 4.5 mg/kg/day for females and 6.2 mg/kg/day for males were based on the mean minus (2 x SD) of all urine creatinine excretion values in the SONICS study by sex, which meant that approximately 2.5% of the urine creatinine excretion values from SONICS were < 4.5 mg/kg/day for females and < 6.2 mg/kg/day for males. It is anticipated that LOGICS will have about the same overall means by sex and, therefore, using these thresholds will result in excluding about 2.5% of all samples from analyses.

At each visit, the mUFC corresponding to that visit will be calculated using all UFC values: 1) from adequate 24-hour urine collections and 2) with urine sample start dates within the analysis window for that visit. The mUFC for a subject at a visit will be calculated only if there are <u>at least two</u> UFCs that meet these two requirements; otherwise, mUFC will be considered as missing for that subject at that visit. For the purpose of determining loss of therapeutic response for the primary endpoint when a subject was not early rescued, mUFC must be based on <u>three</u> adequate samples. In the case of subjects in the S-C cohort and their samples corresponding to the Screening visit, the first 2 samples may be recorded at one visit (e.g. Screening) and the third sample at another visit (e.g. RW0). In the case of any subject where a third sample is collected as a confirmatory sample, the third sample may be recorded at a different CRF visit than the first 2 samples, because it will be collected following the visit instead of prior to the visit. In these cases, the third sample will be included in the calculation of the mUFC for that visit, so long as it was an adequate sample.

Additional 24-hour urine samples for UFC other than those specified in Appendix A of the protocol will not be included in the calculation of the mUFC for a visit.

To ensure that samples for UFC that are included in the calculation of a mUFC for a subject are always samples corresponding to the same visit, blinded reviews of the urine collections included in the calculation of the mUFCs will be conducted in conjunction with the blinded review of the analysis visit assignments, described in

Section 5.3. The blinded listing will list only the following data: subject id, visit name and number in the database, whether the visit is scheduled or unscheduled, urine collection start and stop dates and times, study day start and stop, whether the sample is adequate or not based on the Sponsor's determination, assigned analysis visit, and a flag variable that indicates if the UFC (not provided in the listing) corresponding to the sample is used in the calculation of mUFC. Only urine collections with non-missing UFC will be included in the listing.

Sensitivity analyses that allow for the use of a single adequate urine sample that is within the analysis visit window to calculate mUFC for that visit or the imputation of a missing mUFC due to all missing or inadequate samples for that visit in the RW Phase will be performed Section 6.6.3. Sensitivity analyses where UFC values from all urine samples, regardless of adequacy, will be used to calculate mUFC will also be performed (see Section 6.6.3).

5.5 Early Rescue Criteria – Programmed Verification of Quantitative Criteria

The study Medical Monitor is to be informed by the site when early rescue for a subject in the RW Phase is deemed necessary by the Investigator. Prior to the site's initiation of early rescue therapy, the Medical Monitor is supposed to review the relevant data to confirm that the subject meets the early rescue criteria as defined in Section 3.1 of this SAP and Table 3 in Protocol Section 4.3.1.

The early rescue criteria met by a subject are recorded in the Early Rescue Study Completion eCRF. For the quantitative element(s) that is (are) marked in this eCRF as having met by the subject, a blinded programmed verification using the electronic laboratory data will be performed prior to the first database lock that will occur after the last subject completes the RW Phase. If the programmed results do not confirm that the quantitative element(s) is (are) met for one or more early rescue subjects, a sensitivity analysis of the primary efficacy endpoint will be performed with the subject(s) excluded from the analysis (see Section 6.6.3).

5.6 Late Night Salivary Cortisol – Calculation of Means

Salivary samples will be collected between 11 PM and midnight for the determination of LNSC during the 3 study phases. Depending on the visit, 1 or 2 samples will be collected prior to the actual scheduled visit. At the Screening, TM0, TM7, RW0, RW5, and RES2 visits, 2 samples will be collected over sequential days. For subjects in the S-C cohort for which the first sample for the Screening visit is from their Month 12 visit in SONICS, their second sample will be collected just prior to their first visit in LOGICS; therefore, there may be a gap of several days between the 2 samples. At

the TM1 to TM6 and RES1 visits, 1 sample will be collected. At the RW1 to RW4 visits, 1 sample will also be collected, and a second sample may be collected as soon as possible following the visit if LNSC is being used to establish the need for early rescue therapy during the RW Phase. As such, there may be a gap of several days between the 2 samples.

For the analysis of LNSC at a visit, only adequate sample(s), defined as those collected between 10 PM and midnight (i.e., a 1-hour allowance earlier), and with collection date(s) within the analysis window for that visit will be included. If there are 2 such samples, the mean LNSC will be calculated and used in the analysis. In the case of subjects in the S-C cohort and their samples corresponding to the Screening visit, the first samples may be recorded at one visit (e.g. Screening) and the second sample at another visit (e.g. RW0). In the case of any subject where a second sample is collected to establish the need for early rescue therapy, the second sample may be recorded at a different CRF visit than the first sample, because it will be collected following the visit instead of prior to the visit. In these cases, the second sample will be included in the calculation of the mean LNSC for that visit, so long as it is adequate and within the same analysis visit window.

5.7 Blood Pressure, Heart Rate, and Abdominal Girth – Calculation of Means

Blood pressure, heart rate, and abdominal girth are measured in triplicate at each visit. The mean of the measurements will be used as the result for each visit. The mean of the measurements will be calculated only if there are at least 2 measurements.

5.8 Cushing's Syndrome Quality of Life and Beck Depression Inventory (BDI) II – Calculation of Total Scores

The CS QoL questionnaire is a general questionnaire of 12 items that assesses subject health-related quality of life. Answers to QoL are based on Likert scales with five response categories, rated on a scale of 1-5, where '1' corresponds to 'Always' or 'Very much' and '5' to 'Never' or 'Not at all'. The lower the score, the lower the health-related quality of life (HRQoL). The total score for QoL is the sum of all the item responses and can range from 12 (worst score) to 60 points (best score). A standardized total score on a scale from 0 (worst HRQoL) to 100 (best HRQoL) will be calculated with the following formula:

$$\text{Standardized total score} = \frac{TotalScore - \min}{\max - \min} \times 100$$

where 'min' is the minimum (min=12 if all 12 item responses are non-missing), and 'max' is the maximum possible score (max=60 if all 12 item responses are non-missing).

If more than 3 items have a missing response, then the total score will be considered missing for that visit. The standardized total score can be calculated (as above) if the number of unanswered items does not exceed 3 (25% of the questions). In this case, the standardized total score will have min = the number of item responses that are non-missing and max = 5 x the number of items responses that are non-missing.

The BDI-II inventory comprises 21 questions, with a score from 0 to 3. Scores for each question are summed to create a total score. A total score between 0 and 13 indicates minimal depression, 14 to 19 mild depression, 20 to 28 moderate depression, and 29 to 63 severe depression. Scoring information for the 21 questions is in Appendix N of the protocol. If one question has a missing score, then the missing score will be imputed as the average of the non-missing scores for that question in the ITT population. If there are more than 25% of the questions (i.e. 6 or more questions out of 21) with a missing score at a given visit, then the total score will be set to missing.

5.9 Acne, Hirsutism, Peripheral Edema, and Other Clinical Signs and Symptoms of Cushing's Syndrome – Calculation of Total Scores

Acne Global Score

Each type of lesion is graded with a value depending on the severity: 0 - No Lesions, 1 - Greater than or equal to one comedone, 2 - Greater than or equal to one papule, 3 - Greater than or equal to one pustule, 4 - Greater than or equal to one nodule. Factor is dependent on the location of the lesion: 2 – Forehead, 2- Right Cheek, 2- Left Cheek, 1 – Nose, 1 – Chin and 3 - Chest and Upper Back.

The score for each area (Local score) is calculated using the formula:

Local score = Factor × Grade (0-4)

The global score is the sum of local scores as follows:

Global score = Sum (all local scores)

The global score can range from 0 to 44. If one area has a missing local score, then the missing local score will be imputed as the average of the non-missing local scores for that area in the ITT population. If there are more than 25% of the areas (i.e. 2 or more areas out of 6) with a missing local score at a given visit, then the global score will be set to missing.

Hirsutism Total Score (for females only)

Hirsutism will be assessed for females only. The total score is the sum of all the scores obtained in each of the 9 locations. Total score can vary between 0 and 36. A high score means that the hirsutism is more visible. If one location has a missing score, then the missing score will be imputed as the average of the non-missing scores for that location in the ITT population. If there are more than 25% of the locations (i.e. 3 or more locations out of 9) with a missing location score at a given visit, then the total score will be set to missing.

Peripheral Edema Total Score

The total score is the sum of all the scores obtained in each of the 3 locations. Total score can range between 0 and 12. A high score means that the edema is more pronounced. If one or more locations has a missing score, then the total score will be set to missing.

Total Score for the Other Clinical Signs and Symptoms of Cushing's Syndrome

The other clinical signs and symptoms, defined in Appendix M of the protocol, are moon facies, facial plethora, striae, bruising, supraclavicular fat, irregular menstruation (females only), and dysmenorrhea (females only). Each symptom, if present, is graded as Mild (=1), Moderate (=2), or Severe (=3). If not present, the symptom is graded as None (=0). If the grade is missing for irregular menstruation or dysmenorrhea for a female subject who is post-menopausal or surgically sterile, the missing grade will be set to 0. Total severity score (or simply, total score) is the sum of the severities for each sign or symptom assessed. Since only 5 of the 7 signs and symptoms apply to males, their total score is divided by 5 and then multiplied by 7 to standardize it to the same scale as the total score for females. The total score can range from 0 (none) to 21 (worst). If more than 25% of the individual signs and symptoms (i.e. 2 or more signs and symptoms out of 5 for males and out of 7 for females) have missing scores at a given visit for a subject, then the total score will be set to missing. If 4 of 5 signs and symptoms for males and 6 of 7 signs and symptoms for females are actually assessed, the total score will be standardized by dividing it by the number of non-missing signs and symptoms and then multiplying by 7.

5.10 Handling of Dropouts, Missing Data, and Outliers

Definition of Study Completion

Study completion for purposes of assessing the primary endpoint is defined as either of the following:

1. having completed Visits RW0 (RW Baseline) and RW5 (i.e. final scheduled visit in RW Phase) regardless of UFC status, OR,
2. having completed RW0 and having met at least one of the three criteria for use of early rescue treatment with open-label levoketoconazole (or another rescue medication) after RW0.

Definition of Non-completion

Subjects withdrawn from the study without meeting the definition of study completer above will be considered as prematurely withdrawn (i.e. non-completers).

Missing or Inadequate UFC Collections

Refer to Section 5.4 for the handling of missing or inadequate UFC collections.

Handling of Missing Data for the Analysis of the Primary Endpoint

The following rules will be applied for determining loss of therapeutic response when some or all mUFC data are missing:

1. A subject withdraws prematurely during the RW Phase (i.e. Visit RW5 is not completed nor is the subject in need of early rescue treatment) and there is no post-randomization data available for determination of adequate mUFC. In this case, the subject will be assumed to have lost therapeutic response; OR
2. A subject withdraws prematurely during the RW Phase after RW0 and before RW5, with no early rescue, and at least one post-randomization mUFC determination is available. In this case, if the latest available adequate mUFC determination is the one for RW4, all available mUFC determination(s) will be used to determine if loss of therapeutic response occurred for that subject, but if the latest available mUFC determination is for a visit before RW4, the subject will be assumed to have lost therapeutic response; OR
3. A subject completes the RW Phase with one or more missing adequate mUFC determinations, with at least one post-randomization adequate mUFC determination available. In this case, if the latest available adequate mUFC determination is the one for RW4 or RW5, all available mUFC determination(s) will be used to determine if loss of therapeutic response occurred for that subject, but if the latest available mUFC determination is for a visit before RW4, the subject will be assumed to have lost therapeutic response.

Sensitivity analysis of the primary endpoint that requires at least two RW visits, rather than one, with adequate mUFC determination during the RW Phase to be available in #2 and #3 above will be performed (see Section 6.6.3).

Handling of Missing Data for the Analysis of Secondary Efficacy Endpoints by Visit in the RW Phase and RES Phase

For the secondary and exploratory efficacy endpoints assessed at multiple visits, at-visit analyses will be performed using by-visit univariate methods (for the RW and RES Phases) and longitudinal models (for the RW Phase) without imputation. For the RW Phase, in addition to the by-visit analysis, analysis of the secondary endpoints will be performed at the End of RW Phase time point (see Table 3), where the last non-missing value during the RW Phase will be used as the End of RW Phase value, if there is no non-missing value at RW5.

Partially Missing Data for Triplicate Measurements of Blood Pressure and Abdominal Girth

Refer to Section 5.7 for the handling of partially missing data for triplicate measurements of blood pressure and abdominal girth.

Partial and Missing Dates

Imputations of missing and partial dates are given below and to be used only for the following: assessment of treatment emergence status of an AE and determination of the study phase or study day of onset and duration of the AE (if applicable); for definitions of prior and concomitant medications by study phase and for the study overall; and for calculation of time since diagnosis of historical conditions of interest (e.g. CS).

Missing or Partial Adverse Event and Prior/Concomitant Medication Start Dates

- If the day is missing but both the month and year are non-missing, then the missing day will be imputed as 01. For an AE, however, if the month and year are the same as the month and year of the date of the first dose of study drug, then the missing day will be imputed to be the day part of the first dose date.
- If both the day and month are missing but the year is non-missing, then the missing day and month will be imputed as 01JAN. For an AE, however, if the year is the same as the year of the date of the first dose of study drug, then the missing day and month will be imputed to be the day and month parts of the first dose date.
- If the date is completely missing, then the start date will be imputed to be the earlier of the date of the first dose of study drug and the AE or medication end date. If the end date is partial, then impute the end date first before imputing the start date.

Missing or Partial Adverse Event and Prior/Concomitant Medication End Dates

If the AE or medication is reported as ongoing (or for an AE, not resolved/not recovered or recovering/resolving), then the end dates should be blank; otherwise, the following rules for imputation should be followed:

- If the day is missing but both the month and year are non-missing, then the missing day will be imputed as the last day of the month.
- If both the day and month are missing but the year is non-missing, then the missing day and month will be imputed as 31DEC. For an AE, however, if the year is the same as the year of the date of the last dose of study drug, then the missing day and month will be imputed to be the day and month parts of the last dose date
- If the date is completely missing, then the end date will be imputed to be the date of the first scheduled visit that is after the AE or medication start date or the subject's last dose date, if there is no scheduled visit that is after the AE or medication start date. If the start date is partial, then impute the start date first before imputing the end date.

If the imputation rules above result in an end date being earlier than the start date, where one or both dates are imputed, the original dates and the imputed dates will be reviewed by the CRO statistician on a blinded basis to adjust the imputed dates just enough to make the end date no earlier than the start date without changing the non-missing parts of a partial date.

Missing or Partial Diagnosis Dates of Historical Conditions of Interest

- If the day is missing but both the month and year are non-missing, then the missing day will be imputed as 01.
- If both the day and month are missing but the year is non-missing, then the missing day and month will be imputed as 01JAN.
- If the date is completely missing, then it will not be imputed.

Outliers

No rules for outlier detection are planned.

6. Statistical Methods

6.1 General Principles

All data processing, summarization and analyses will be performed using SAS Version 9.4 or later of the SAS® statistical software package.

The following principles will be applied to all TFLs unless otherwise stated.

Table 4 General Principles for the Presentation of Results in Tables, Figures, and Listings

	Principles or Values
Study phase labels and order presented	Dose Titration and Maintenance Phase Randomized Withdrawal Phase Restoration Phase Randomized Withdrawal Phase and Restoration Phase Combined (for study drug exposure and shifts from Baseline in safety endpoints, where specified, and while on levoketoconazole only) All Phases Combined (for study drug exposure and AEs only and while on levoketoconazole only) For the Restoration Phase, in addition to the summaries by treatment group (see below), separate summaries will be presented for subjects who required early rescue during the Randomized Withdrawal Phase and continued with open-label levoketoconazole during the Restoration Phase versus those who did not require early rescue. For the Randomized Withdrawal Phase and Restoration Phase Combined, data from subjects who required early rescue will not be included following the date of the early rescue determination in the summaries, because subjects requiring early rescue are expected to experience adverse events resulting from recurrence of Cushing's syndrome. Excluding data from these subjects in RW will allow for comparisons of subject AE incidence between those subjects with AEs that are more likely related to disease recurrence (i.e. early-rescue subjects in RES Phase) versus those with AEs that are less likely related to disease recurrence (i.e. those taking study drug during RW and RES who did not require early rescue).
Visit labels and order presented	Screening (in listings only) For the Dose Titration and Maintenance Phase: In tables and figures - TM Baseline, Worst Result (for shift or threshold analyses of safety endpoints only), Last Non-

	Principles or Values
	missing Result, and Non-missing Result at the Highest Dose In listings – TM0, TM1, TM2, TM3, TM4, TM5, TM6, TM7 For the Randomized Withdrawal Phase: RW Baseline (in tables and figures), RW0 (in listings) RW1, RW2, RW3, RW4, RW5 For the Restoration Phase: RES Baseline (in tables and figures), RW5 (in listings) RES1, RES2
Subject cohort labels (in order of presentation)	For study and study phase disposition, demographics and other background characteristics, and prior medications: Levoketoconazole-naïve SONICS-completer All Subjects For the Dose Titration and Maintenance Phase and for All Phases Combined (where specified): Levoketoconazole-naïve All Subjects Wherever the SONICS-completer cohort is presented in a TFL, there will be an asterisk (*) next to the term and the following footnote will be included: * Subjects who completed SONICS and directly randomized without titration in LOGICS.
Treatment group labels (in order of presentation)	For the Randomized Withdrawal Phase and Restoration Phase: Levoketoconazole Placebo All Subjects (where applicable) For all other study phases: All Subjects
Tables	For the Randomized Withdrawal Phase, Restoration Phase, and Randomized Withdrawal Phase and Restoration Phase Combined: Data in summary tables will be presented by treatment group and visit (wherever applicable). For subjects in the Randomized Withdrawal Phase only, summaries will be presented by Therapeutic Dose (i.e. the dose at RW0) and for all dose groups combined where applicable. For the Dose Titration and Maintenance Phase, except for AEs:

	Principles or Values
	Data in summary tables will be presented by the levoketoconazole dose received corresponding to the last non-missing results or corresponding to the worst results (where specified). For the Dose Titration and Maintenance Phase for AEs only: Data in summary tables will be presented by the levoketoconazole dose received at the day of onset of the AE. For All Phases Combined (where specified): Data in summary tables will be presented for All Subjects on levoketoconazole only.
Figures	In general, data will be presented as described above for tables, unless otherwise specified
Listings	All data collected, as well as derived data used in the analyses, will be presented in listings by population, subject cohort, treatment group (where applicable), Site, subject, study phase, and visit (where applicable), unless otherwise specified.
Units for quantitative measures	All laboratory test results will be received from the central laboratories, and the results will be provided in both International System of Units (SI) and conventional units. For the TFLs, the results will be summarized or presented in SI units, with one exception. For HbA1c, the results will be summarized in the tables and figures in the conventional unit (i.e. in %) but will be presented in the listings in both SI and conventional units. Refer to Appendix 1 for the SI unit corresponding to each laboratory test.
Precision levels for laboratory values	Refer to Appendix 2 for the precision level in which each laboratory test is reported by the central laboratories.
Descriptive summary statistics for continuous variables	Number of subjects/observations (n), mean, standard deviation (SD), standard error (SE; where applicable), median, minimum and maximum For a numeric result presented in the database with a "<", "≤", ">", or "≥" preceding the number, the value at (if "≤" or "≥") or just below (if "<') or just above the numeric part will be used in the summary. For example, if the result is "<5", and the data for the variable are measured to one decimal place, then the result will be summarized as 4.9.
Descriptive summary statistics for categorical variables	Frequency counts and percentages [n (%)], where percentages will be presented to one decimal place

	Principles or Values
	Percentages greater than 0 and less than 0.1 will be presented as "<0.1". All categories specified in the table and figure shells will be presented in the actual tables and figures, even if frequency is zero; in this case, no percentage will be presented. Note that the terms "proportion" and "percentage" of subjects are used interchangeably in this SAP. In the descriptive summaries, percentages (e.g. 85.1%) will be presented rather than proportions (e.g. 0.851).
Denominator for percentages	Number of subjects in the analysis population, unless stated otherwise in table shell(s)
Include "Missing" as category	Yes, when the number missing is greater than zero for at least one treatment group, unless otherwise specified.
Display for 0 percentages	Blank
Display to the same number of decimal place as the collected value	Minimum Maximum
Display to one more decimal place than collected value	Mean Median Confidence Interval (CI)
Display to two more decimal places than collected value	SD
Limit of precision for displays	3 decimal places, unless otherwise specified
Date Format	DDMMMYYYY
Significance tests and CIs	Two-sided and use a 5% significance level P-values will be presented to four decimal places, except values that are less than 0.0001 will be presented as "<0.0001" and values that are greater than 0.9999 will be presented as ">0.9999". In the two-sample t-test, p-value and 95% CI from the Satterthwaite method (instead of pooled method) will be used, since the Satterthwaite method doesn't require assumption of equal variance.
Source footnotes	Each table will have a footnote that lists the source data listing(s). Each figure will have a footnote that list the source table(s), if available, or the source data listing(s), otherwise.
Dictionary names and versions	The dictionary names and versions will be included in a footnote in all AE and prior or concomitant medication TFLs that present coded terms from the dictionaries.

6.2 Subject Disposition

The disposition summary for all subjects will include:

- number of subjects in each analysis population
- number of subjects randomized
- number and percentage of subjects who entered, completed, and discontinued the study and their discontinuation reason; overall and by study phase (TM Phase, RW Phase, and RES Phase) will also be displayed.

The subject disposition summary will be presented for the Safety population by subject cohort and ITT population by treatment group.

A subject is considered to have been treated during the study when the subject has received at least one dose of study medication (Levoketoconazole or Placebo) during any phases. A subject is considered to have been treated during a study phase when the subject has received at least one dose of study medication during the corresponding phase.

Overall study completion is recorded in the clinical database. Overall study completion is defined as follows:

- Subjects who complete the end of the RES Phase (Visit RES2) of the study, regardless of their response status, will be considered to have completed the study.
- Subjects who receive early rescue medication during the RW Phase will also be considered to have completed the study.

Subjects who entered the study but did not complete the study as defined above are considered as discontinuing from the study.

Study completion by phase is defined as follows:

- TM Phase: Subjects who complete the TM7 visit and do not discontinue in the TM Phase.
- RW Phase: Subjects who complete the RW5 visit and do not discontinue in the RW Phase. These include subjects who receive early rescue medication during the RW Phase.
- RES Phase: Subjects who complete the RES2 visit and do not discontinue in the RES Phase.

For the RW Phase and the RES Phase, and the study overall, study completion status will be presented separately for both Early Rescue used, and Early Rescue not used. For the RW Phase and the study overall, early rescue subjects are considered as

completers. For the RES Phase, the reason for discontinuation will also be presented separately for Early Rescue used versus Early Rescue not used.

Subjects who entered a study phase but did not complete the phase as defined above are considered as discontinuing from that phase. The reasons for discontinuation will be summarized for each phase.

Subject disposition with respect to the presence of non-missing and adequate 24-hour urine samples for UFC determination will be summarized by visit. For each visit, the number of subjects still in the study as of that visit will be presented. The number and percentage of subjects for each sample corresponding to the visit will determined as follows:

- Sample 1 – adequate, not adequate, or no sample taken
- Sample 2 – adequate, not adequate, or no sample taken
- Sample 3 – not required, required and adequate, required but not adequate, or required but no sample taken

In addition, a separate summary will also be provided for each of the following categories for all screened subjects:

- Subject disposition for the screening period by subject cohort and overall
- Subject counts by study phase and by site, country, and region (United States [US], non-US)

6.3 Protocol Deviations

A summary of major protocol deviations will be presented for the Safety population by subject cohort (by study phase and overall) and ITT population by treatment group (by RW Phase, RES Phase, and combined). Both major and minor deviations will be provided in a listing.

Subjects affected by COVID-19 related study disruption will be listed by investigational site, including descriptions of how the individual's participation was altered. Protocol deviations due to COVID-19 illness and/or COVID-19 public health control measures will be clearly identified in the protocol deviation listing and the summary of major protocol deviations (if any). Any major protocol deviations related to COVID-19 will be included in the review of major protocol deviations for the purpose of identifying subjects to be excluded from the PP population (see Section 4.2.1).

6.4 Demographics and Other Background Characteristics

6.4.1 Demographics and Baseline Characteristics

Demographics, baseline, and disease characteristics will be summarized for the Safety population by subject cohort, ITT population by treatment group, and PP population by treatment group.

Standard descriptive statistics will be presented for the continuous variables of:

- age (years)
- baseline weight (kg)
- baseline height (cm)
- baseline body mass index (BMI) (kg/m^2)
- baseline mUFC
- baseline LNSC
- baseline HbA1C
- time since diagnosis of CS (mos)
- time since last surgery for the management of CS
- baseline systolic and diastolic blood pressure
- baseline QTcF.

The total counts and percentages will be presented for the categorical variables of:

- sex
- childbearing potential (females only)
- ethnicity
- race
- classification of CS
- prior treatment for CS
- receiving anti-diabetic medication at baseline (determined from concomitant medications)
- receiving anti-hypertension medication at baseline (determined from concomitant medications)

No formal tests of statistical significance will be performed on the demographics, baseline, and disease characteristics.

Baseline value for the L-N cohort is the TM Baseline. Baseline for the other summaries is the RW Baseline.

6.4.2 Medical History, Surgical History, and Prior Radiotherapy

Medical history and surgical history (including both CS related surgeries and not CS related ones) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 21.1. Medical history and surgical history will be summarized by system organ class (SOC) and preferred term (PT) for the Safety population by subject cohort and ITT population by treatment group.

Prior radiotherapy for management of CS will be summarized by the number of prior radiotherapies, location (pituitary, adrenal gland, and other), time since the end of the last radiotherapy (months), and outcome of the last radiotherapy (complete response, partial response, no response, and unknown) for the Safety population by subject cohort and ITT population by treatment group. For location and radiotherapy type, a subject who had more than one radiotherapy will be counted once under each different location and different radiotherapy type.

Medical history, surgical history and prior radiotherapy for management of CS will be presented in listings.

6.4.3 Prior and Concomitant Medications

Incidence of prior and/or concomitant medications will be presented by drug class (i.e. Anatomical Therapeutic Chemical (ATC) Level 3 classification). Coding will be done using the World Health Organization (WHO) Drug September 2018 dictionary. Prior medications will be summarized by subject cohort. Concomitant medications will be presented for TM Phase by subject cohort, and for RW and RES Phase (and RW/RES Phases combined) by treatment group. For the RES Phase, summaries will also be presented for Early Rescue used versus Early Rescue not used.

In addition, separate summaries for prior and concomitant medications will be presented by the five medication categories prespecified in the CRF, namely, anti-diabetic, anti-hypertensive, cholesterol medication, anti-inflammatory, and Cushing's medication.

If the medication record satisfies the concomitant criteria, then the record will be summarized as concomitant. A medication record will not be classified as both prior and concomitant within the same study phase.

TM Phase:

- Prior medications are medications taken (either started or ongoing) within 3 months (i.e. 91 days) of the date of informed consent and ended prior to the TM0 visit date.

- Concomitant medications are medications taken during the study phase, including those started before but ongoing at first dose (TM0 visit), up through end of phase (as defined in Section 5.1).
- To be summarized for the L-N population.

RW Phase:

- For the S-C population only, prior medications are medications taken (either started or ongoing) within 3 months (i.e. 91 days) of the date of informed consent and ended prior to the RW0 visit date.
- For all subjects in the RW Phase, concomitant medications are medications taken during the study phase, including those started before but ongoing at first dose (RW0 visit), up through end of phase (as defined in Section 5.1).

RES Phase:

- Concomitant medications are medications taken during the study phase, including those started before but ongoing at first dose (RW5 visit), up through end of phase (as defined in Section 5.1).

RW Phase and RES Phase Combined:

- Concomitant medications of interest for the two phases combined are those taken during levoketoconazole dosing. As such, for subjects on placebo during the RW Phase, only their concomitant medications during the RES Phase, when they are on levoketoconazole, will be included.
- Concomitant medications are medications taken during the two study phases combined, including those started before but ongoing at first dose (RW0 visit), up through end of the subject's last phase.

6.5 Study Drug Exposure and Compliance

Dose Levels and Titration Status during the TM Phase

Summary of levoketoconazole titration status will be presented by visit and dose level established at each visit during and at the end of the TM Phase for the L-N population. Titration status at the current visit will be the change in dose level since the last visit (up-titration, no change, or down-titration). Percentages are calculated based on the number of total subjects still in the study as of the current visit.

Therapeutic Dose Levels at Start of RW Phase

Summary of Therapeutic Dose Level (mg/day) at the end of the TM Phase and at start of the RW Phase will be presented. Percentages are calculated based on the number

of total subjects still in study at the start of each phase or combined phases, respectively.

Exposure to Study Drug

Study drug exposure will be summarized for Safety population and L-N population by each of the three phases, the RW Phase and RES Phase combined, and overall (all three phases combined); and for ITT population by randomized treatment. The total study drug duration (days), total days on study drug, the cumulative dose of study drug (mg), and the average daily dose of study drug (mg/day) will be presented. For the RW Phase and RES Phase Combined and the All Phases Combined, only the exposure while on levoketoconazole will be summarized.

Total study drug duration (days) will be calculated for each subject as the last study drug stop date minus first study drug start date plus one day.

Total days on study drug will be calculated for each subject as the actual total number of days the subject received study drug (i.e. gaps in study drug intake are not included in the calculation). If only one out of two daily dosing (i.e. A.M. or P.M. but not both) of the study drug is taken, it counts as 0.5 day.

Actual cumulative dose (mg) of study drug will be calculated as the total number of tablets dispensed minus the total number of tablets returned, then multiplying the result by 150 mg.

Average daily dose of study drug (mg/day) will be calculated as the actual cumulative dose of study drug divided by total study drug duration.

Compliance to Study Drug

Study drug compliance will be summarized separately for Safety population and L-N population by each of the three phases, the RW Phase and RES Phase combined, and overall (all three phases combined); and for ITT population by randomized treatment. Study drug compliance will be calculated by dividing the actual cumulative dose of study drug (as described above) by the expected cumulative dose of study drug and multiplying the result by 100. For the RW Phase, the expected cumulative dose of study drug is the total study drug duration for the RW Phase x 150 mg x the expected number of tablets for the subject's dose level in the RW Phase. For all other phases, including combined phases, the expected cumulative dose is first calculated at each dose level received and then summed up across all dose levels received.

For summaries of the RW Phase and RES Phase combined and for the study overall, placebo exposure for the subjects who receive placebo during the RW Phase will be excluded.

6.6 Efficacy

Descriptive summaries of all efficacy endpoints will be presented for the L-N population (for the TM Phase), and by treatment group (for the RW Phase and RES Phase), and visit (where applicable for RW Phase and RES Phase). The by-visit summaries in the RW Phase will include summaries at the individual visits and summaries at the End of RW Phase time point. In addition, separate summaries will be presented for RES Phase for Early Rescue used versus Early Rescue not used. For continuous variables, summaries will be presented for the observed values and for the absolute and percent changes from the relevant Baseline. Statistical testing, whether within-treatment or between-treatment, will only be for the all levoketoconazole dose groups combined.

For the TM Phase, efficacy endpoints measured at multiple visits will be summarized by the levoketoconazole dose received corresponding to the last non-missing results and by the highest levoketoconazole dose received with non-missing results.

The primary, secondary, and many of the exploratory efficacy analyses will be based on the RW Phase and RES Phase. These efficacy analyses will be based on the ITT population and for the primary and secondary efficacy endpoints only, supportive analyses will use the PP population.

There are also exploratory analyses of efficacy during the TM Phase, and these will be based on the L-N population.

For secondary efficacy endpoints that are continuous variables, the within-treatment mean changes from the relevant Baseline (see Section 5.1) to each visit in the RW Phase and RES Phase and at the End of RW Phase time point will be assessed for nominal significance using paired t-test.

For all efficacy analyses that involve the RES Phase only, the analysis population is the subset of ITT that receive at least one dose of study drug in the RES Phase.

6.6.1 Primary Efficacy Analysis

Loss of Therapeutic Response

The primary efficacy analysis will determine the significance of the difference in proportions of subjects losing therapeutic response between the two treatment groups. In addition to the primary analysis based on the ITT population, a supportive analysis will be conducted on the PP population.

Statistical significance testing will be conducted using a logistic regression model containing fixed effect terms for treatment group (levoketoconazole, placebo) and subject cohort. The results including the estimated proportion and standard error (SE) for each treatment group, the estimated difference and SE between the two

treatment groups, and associated 95% CI and p-value will be derived from the model (see SAS STAT Procedure details in Appendix 3).

A supportive analysis that does not require statistical modeling will be conducted to confirm the results of the primary analysis. The proportions between the treatment groups will be compared using a two-sided Fisher's Exact test and an exact unconditional two-sided 95% CI of the difference will be calculated (see SAS STAT Procedure details in Appendix 3).

The primary efficacy analysis will be repeated for the following subgroups:

1. S-C cohort versus L-N cohort
2. Subjects who completed SONICS (regardless of their need to titrate in LOGICS) versus those who did not participate in SONICS

The same Fisher's Exact test as mentioned above will be conducted for subgroups with a minimum size of at least 30% of the ITT population only.

In addition, the number and percentage of subjects who met each of the criteria for loss of therapeutic response will be summarized by treatment group and by therapeutic dose level, respectively. No statistical testing will be performed.

Sensitivity analysis of the primary endpoint that allows subjects who came from the SONICS study but are in the L-N cohort, to be included in S-C cohort in the logistic regression model described above will be performed (see Section 6.6.3).

6.6.2 Secondary Efficacy Analyses

Inferences derived from secondary efficacy analyses will be gated on results from the primary efficacy analysis. Secondary efficacy analyses will be hierarchically structured to ensure control of the family-wise type I error rate at the 0.05 level. Hypothesis tests for secondary efficacy endpoints will be based on null hypotheses that assume no *a priori* differences between placebo and levoketoconazole treatments.

A closed or sequential testing procedure will be used for testing the key secondary efficacy variables. With this hierarchical procedure, the key secondary efficacy variables will only be tested if the primary efficacy analysis is statistically significant.

Analyses will be carried out for the key secondary variables, 'Endpoint set 1', 'Endpoint set 2', 'Endpoint set 3', 'Endpoint set 4', 'Endpoint set 5', and 'Endpoint set 6'.

The analyses will be carried out on the change from RW Baseline to the End of RW Phase time point for the following efficacy parameters in hierarchical order:

Set 1 - Normalization of mUFC at the end of RW Phase

Set 2 – mUFC (change from RW Baseline)

Set 3 – biomarkers of CS comorbidities (fasting glucose, fasting insulin, HOMA-IR, HbA1c, total cholesterol, LDL-C, and hsCRP)

Set 4 – CS QoL total score and BDI-II total score

Set 5 – Acne Global Score, Hirsutism Total Score (for females only), and Peripheral Edema Total Score

Set 6 - LNSC

The secondary efficacy endpoint of proportion of subjects with normalization of mUFC at RES is not part of the hierarchical testing since it cannot be analyzed until the end of the study, while the endpoints above are at RW5.

Within each set with multiple endpoints (i.e. Sets 3, 4, and 5), the statistical significance level will be adjusted using the Hochberg method, controlling the familywise error rate at 5%. All endpoints within a set with multiple endpoints have to be statistically significant in favor of levoketoconazole compared to placebo, after the Hochberg adjustment, in order for the next set in the hierarchy to be tested. The testing is only stopped when all of the alpha is 'exhausted'. The sequential testing will stop at the first endpoint set where levoketoconazole does not demonstrate statistical superiority over placebo.

In addition to the comparisons on the changes from RW Baseline to the End of RW Phase time point, comparisons between the two treatments will also be performed on the changes from RW Baseline to each of RW1 to RW5. The p-values from these comparisons will be used for descriptive purposes only, i.e. no statistical inferences will be made and as such, no multiplicity adjustments.

For the above efficacy parameters, the absolute and percent changes from RW Baseline and from RES Baseline to each of the 2 visits in RES will be summarized (no imputation). Although summaries will be presented separately by the randomized treatment group (in addition to all subjects together), there will be no treatment comparisons.

Line plots of the mean +/- SE of actual and change from RW Baseline values by treatment group and visit during the RW Phase will be produced for efficacy endpoints that are assessed by visit. Line plots of least squares mean changes from RW Baseline (where applicable) during the RW Phase will also be presented.

6.6.2.1 UFC and LNSC Analyses

Normalization of mUFC at the end of RW Phase

The number and proportion of subjects in the RW Phase with mUFC normalization at the end of RW Phase will be compared between treatment groups using Fisher's Exact test (with corresponding 95% CIs). mUFC normalization at the end of RW Phase will be based on the subject's mUFC corresponding to the End of RW Phase time point

defined in Section 5.3. In addition, a subject must not have been early rescued during the RW Phase in order to be considered as achieving mUFC normalization at the end of RW Phase.

In addition, this analysis will be repeated for the following subgroups:

1. S-C cohort versus L-N cohort
2. Subjects who completed SONICS (regardless of their need to titrate in LOGICS) versus those who did not participate in SONICS

Fisher's Exact test results will only be provided for subgroups with a minimum size of at least 30% of the ITT population.

Side-by-side needle plots for the two treatment groups of the individual subject mUFC from RW Baseline to the end of the RW Phase will be produced, with the ULN as a horizontal reference line.

Changes in mUFC and LNSC from RW Baseline

Changes from RW Baseline to the End of RW Phase time point in mUFC and LNSC will be compared between treatment groups using two-sample t-test (with corresponding 95% CIs). To allow for possible departures from the assumption of normal distribution of the data, the treatment groups will also be compared using Wilcoxon rank sum test as a supportive analysis.

Supportive analyses using a repeated measures mixed effects model with fixed effect terms for treatment group, subject cohort, time (i.e. visit), and treatment-by-time interaction and with Baseline value as a covariate, with subject as random effect (depending on covariance matrix structure) will also be performed on the changes from RW Baseline to RW1 to RW5.

The following covariance matrix structures will be considered for the model: spatial power, toeplitz with heterogeneity, and toeplitz. The final covariance matrix structure will be selected as the one resulting in a statistical model that converges and has the minimum Corrected Akaike information criterion (AICC) (i.e., minimum loss of information). The treatment groups will be compared using least squares mean differences between treatment groups and associated 95% CIs for each visit will be derived from the model. These CIs will be regarded as descriptive statistics (see SAS STAT Procedure details in Appendix 3).

Two-sample t-tests (with corresponding 95% CIs) will also be performed, as supportive analyses, to compare the two treatment groups for the mean change from RW Baseline to each nominal visit during the RW Phase.

In addition to the line plots by visit (in Section 6.6.2), spaghetti plots of the individual subjects' mUFC by visit in the RW Phase will be produced with a horizontal line at ULN.

Normalization of mUFC during RES Phase

The number and proportion of subjects with normalization of mUFC at RES2 is defined as the proportion of subjects with abnormal mUFC at RW5 whose mUFC is at or below the ULN of the reference range at RES2. Normalization of mUFC at RES2 will be calculated for each treatment group and presented as shift tables from RW5 to RES2 using three strata of mUFC status: at or below ULN, more than ULN and no more than 1.5X ULN, above 1.5X ULN.

6.6.2.2 CS Comorbidity Biomarkers

The mean changes from RW Baseline in individual biochemical markers of CS comorbidities (fasting glucose, fasting insulin, HOMA-IR, HbA1c, total cholesterol, LDL-C, and hsCRP) during the RW Phase will be assessed for treatment group differences using two-sample t-test at the End of RW Phase time point and at each visit. HOMA-IR is calculated as multiplying fasting plasma insulin (FPI) by fasting plasma glucose (FPG), then dividing by the constant 22.5, i.e. HOMA-IR = (FPI×FPG)/22.5. Subjects who are using insulin as a concomitant medication during the RW phase will be excluded from HOMA-IR summaries.

In addition, shifts from the relevant Baseline (with regards to laboratory normal range) in individual biochemical markers of CS comorbidities will be summarized at each visit during the RW Phase and RES Phase and at the End of RW Phase time point.

In addition for biomarkers fasting glucose, fasting insulin, and HOMA-IR, shifts from the relevant Baseline will be presented by diabetes status at Baseline (diabetic, non-diabetic) by treatment group, where the biomarker values are categorized by tertiles, defined within each diabetes status based on the Baseline values (also stated in Section 6.6.4.4).

For CS comorbidities measured at more than one visit after RW0 during the RW Phase, changes from RW Baseline during the RW Phase will also be analyzed using a repeated measures mixed effects model with fixed effect terms for treatment group, time (i.e. visit), and treatment-by-time interaction, the Baseline value and concurrent CS medical condition (where applicable) as baseline covariates and with subject as a random effect (depending on Covariance structure). The concurrent CS medical condition is the the presence of diabetes at screening for the biomarkers fasting glucose, fasting insulin, HOMA-IR, and HbA1c. The choice of covariance structure for the model and the use of least squares means to compare treatment groups are as described above for mUFC and LNSC (see SAS STAT Procedure details in Appendix 3). For CS comorbidities measured at more than one visit after RW0 during the RW Phase, the analysis at the End of the RW Phase time point is the main analysis, and the analyses by visit and using the statistical model above are supportive.

6.6.2.3 Cushing's Syndrome Quality of Life Questionnaire

The mean changes from RW Baseline in the Cushing QoL total score during the RW Phase will be assessed for treatment group differences using two-sample t-test at the End of RW Phase time point and at RW5.

A summary table for Cushing QoL total score will be presented to display the change (in categories: <10.1, ≥10.1) from the relevant Baseline at each visit during the RW Phase and RES Phase.

6.6.2.4 BDI-II

The mean changes from RW Baseline during the RW Phase will be assessed for treatment group differences using two-sample t-test at the End of the RW Phase time point and at RW5.

A shift table for BDI-II total score will be presented to display the shift in the range categories (minimal, mild, moderate and severe depression) from the relevant Baseline at each visit during the RW Phase and RES Phase.

6.6.2.5 Acne, Hirsutism and Peripheral Edema

The mean changes from RW Baseline in acne, hirsutism (for females only) and peripheral edema total scores during the RW will be assessed for treatment groups differences using two-sample t-test at the End of RW Phase time point and at each visit.

Shifts from the relevant Baseline for acne, hirsutism and edema total scores will also be summarized in the RW Phase and RES Phase, using the following categories:

- Acne global score: 0, 1-18, >18 (GAGS, Doshi 1997) = none, mild, moderate to severe
- Hirsutism total score: 0-7, 8-15, >15 (Ferriman-Gallwey scale) = none to mild, mild to moderate, moderate to severe
- Peripheral edema total score: 0-3, 4-6, >6 = none to mild, mild to moderate, moderate to severe

6.6.3 Sensitivity Analyses

The following sensitivity analyses of the primary efficacy endpoint (loss of therapeutic response in the RW Phase) and the secondary efficacy endpoints of normalization of mUFC at the end of RW Phase and change from RW Baseline in mUFC at RW5 (where specified) will be conducted on the ITT population (only if the original analysis of the efficacy endpoint is statistically significant):

1) Sensitivity analyses of the primary endpoint and the secondary mUFC (normalization at the end of RW Phase and change from RW Baseline) endpoints

that allow for the use of a single adequate urine sample that is within the analysis visit window (RW1-RW5) to calculate mUFC for that visit or the imputation of a missing mUFC due to all missing or inadequate samples for that visit in the RW Phase (see Section 5.4). In the latter case, the imputed value will be last non-missing mUFC for that subject prior to that visit.

2) Sensitivity analyses of the primary endpoint and the secondary mUFC (normalization at the end of RW Phase and change from RW Baseline) endpoints where UFC values from all samples, regardless of adequacy, will be used to calculate mUFC (see Section 5.4).

3) Sensitivity analysis of the primary endpoint that excludes early rescue subjects identified through programmed verification as not having met the quantitative element(s) of the early rescue criteria marked as met in the Early Rescue Study Completion eCRF (see Section 5.5).

4) Sensitivity analysis of the primary endpoint that requires more than one RW visit with adequate mUFC determination during the RW Phase (RW1-RW5) to be available (see Section 5.10).

5) Sensitivity analysis of the primary endpoint allowing subjects who came from the SONICS study but are in the L-N cohort, to be included in the S-C cohort instead in the logistic regression model (see Section 6.6.1).

6) A tipping point analysis will be applied to the primary endpoint. This analysis will explore the effect of missing data on the reliability of the efficacy results by determining the extent the missing data have to change for the results of the study to tip from statistically significant to not. The following approach for the tipping point analysis will be undertaken:

   The number of placebo subjects losing therapeutic response will be incrementally decreased by 1 until all placebo subjects originally categorized as losing response due to missing mUFC values during the RW Phase (see Section 5.10) are re-categorized as not losing response, and the number of levoketoconazole subjects losing therapeutic response will be incrementally increased by 1 until all levoketoconazole subjects originally categorized as not losing response based on an incomplete set of mUFC values during the RW Phase (see Section 5.10) are re-categorized as losing response, or until the p-value for the logistic regression model (see Section 6.6.1) exceeds 0.05.

7) For the secondary efficacy endpoint of change from RW Baseline to the RW5 visit in mUFC, if the value at RW5 is missing and the subject had loss of therapeutic response, the value from the visit where the subject first exhibits the loss of therapeutic response will be imputed for the RW5 value. This will be analyzed for the RW5 data using an analysis of covariance (ANCOVA) with

treatment group and subject cohort as fixed effects and Baseline mUFC as a covariate. Subjects who are missing data at RW5 for reasons other than loss of therapeutic response or early rescue will not be included in this analysis.

8) In order to assess the impact of COVID-19, a sensitivity analysis of the primary efficacy endpoint at the End of RW Phase will be performed where subjects who dropped out before RW4 specifically due to COVID-19 are excluded from the analysis (rather than being considered as loss of response in the main analysis). If there are subjects rescued early due to, in part, clinically significant worsening of signs and symptoms of CS, as described in Section 3.1, and source data validation by clinical monitors to confirm the worsening cannot be performed due to COVID-19 public health control measures, a sensitivity analysis of the primary efficacy endpoint with these subjects excluded will be performed.

6.6.4 Additional Subgroup Analyses

In addition to the subgroup analyses described in Section 6.6.1 and Section 6.6.2.1, the following subgroup analyses will be performed (unless specified otherwise) for the primary efficacy endpoint and the secondary endpoints of change from RW baseline to End of RW Phase for biomarkers of CS comorbidities, and normalization of mUFC at the end of RW Phase, for a minimum subgroup size of at least 30% of the ITT population. Univariate statistical tests (i.e. Fisher's Exact tests and two-sample and paired t-tests) will be performed for descriptive purposes only.

6.6.4.1 Prior Surgery for Cushing's Syndrome

Subgroup displays will be generated for subjects who enter the study as surgery-naïve versus surgery-experienced, for prior surgeries that were performed for the management of CS.

6.6.4.2 Prior Radiotherapy

Subset displays will be generated for subjects who have not received radiotherapy versus those who have previously received radiotherapy at any time prior to the study.

6.6.4.3 Hypertensive Subjects

Subgroup displays will be generated for subjects who enter the study while being prescribed anti-hypertensive medications versus not and subjects who enter the study with a baseline systolic blood pressure (SBP) > 130 mmHg or diastolic blood pressure (DBP) ≥ 90 mmHg regardless of anti-hypertensive medication status versus those with lower blood pressure measurements.

6.6.4.4 Pre-diabetic/Diabetic Subjects

Subgroup displays will be generated for subjects who enter the study as pre-diabetic (Baseline fasting glucose greater or equal to 100 mg/dL (5.6 mmol/L) and less than 126 mg/dL (7.0 mmol/L) without concomitant use of anti-diabetic medication) versus those with normal fasting glucose versus those who were diabetic (Baseline fasting glucose 126 mg/dL or above or lower fasting glucose while being prescribed anti-diabetic medications).

In addition, for biomarkers fasting glucose, fasting insulin, HOMA-IR and Homeostatic Model Assessment-Beta Cell Function (HOMA-%B), shifts from the RW and RES Baselines will be presented by baseline glycemia status (diabetic, prediabetic, non-diabetic) by treatment group, where the biomarker values are categorized by tertiles, defined within each diabetes status based on the Baseline values.

If there are < 30% of the ITT population of the subjects in the pre-diabetic subgroup, the pre-diabetic subjects and subjects with normal fasting glucose will be combined into one subgroup versus those in the diabetic subgroup.

6.6.4.5 Region

Subgroup displays will be generated for subjects enrolled in US sites versus those in non-US sites.

6.6.4.6 Sex

Subgroup displays will be generated for female versus male subjects.

6.6.4.7 Age

Subgroup displays will be generated for subjects with age <= the median age (of the Safety population for the study overall) versus those with age > the median age.

6.6.5 Exploratory Analyses

Medication Changes

Frequency of usage and changes from Baseline in frequency of usage of each of the concomitant medication categories (anti-diabetic, anti-cholesterol, anti-hypertensive, and chronic anti-inflammatory therapies) will be assessed over each study phase separately. TM Phase analysis will be presented for all subjects in the L-N cohort, while RW and RES Phase analysis will be presented for ITT population by treatment group.

The number and proportion of subjects with medication changes from the TM/RW Baseline will be presented. Subjects whose usage is ongoing or started at the TM/RW Baseline will be categorized as follows for each study phase separately:

1) New and clinically significant medication – New medication is started during the study phase and its use was potentially clinically significant

2) Medication dose increased – Dose is increased compared to the TM/RW Baseline AND the increase is considered potentially clinically significant
3) Medication dose restarted after a gap – The medication is stopped after TM/RW Baseline but is restarted during the study phase AND the restart is considered potentially clinically significant
4) Clinically insignificant change in medication – The change represents either the same dose in a different formulation or a pharmaceutically equivalent dose of a different drug (i.e. a substitution)
5) No medication change from the TM/RW Baseline
6) Medication dose decreased – Dose is decreased compared to the TM/RW Baseline AND the decrease is considered potentially clinically significant

New and clinically significant medication changes, medication dose increased or restarted after a gap, clinically insignificant change in medication, and medication dose decreased (Items #1,2,3,4,6 above) will be evaluated prior to the first database lock of all data through the end of the RW Phase by a medical monitor blinded to treatment assignment and again for the RES Phase prior to the final database lock after the end of the study.

Subjects with no usage ongoing or started at the TM/RW Baseline will be categorized as follows:

1) Medication added after TM/RW Baseline
2) No medication used.

If a subject has more than one medication in the same medication class that falls under different usage categories above, the subject will be counted under the "worst" of these categories, i.e. the first applicable category that is listed above.

Biomarkers Change Associated with Medication Usage Changes

Changes in the corresponding biomarkers of CS comorbidities at End of RW Phase relative to changes in medication usage over the entire RW Phase, as described above, will be explored through HIGHLOW plots (produced by SAS® software), where the x-axis corresponds to the medication categories and the y-axis corresponds to the highest and lowest changes in the corresponding biomarkers joined by a line with a tick mark for the mean changes.

The following associations will be represented: fasting glucose, fasting insulin, HOMA-IR, and Hb1Ac with anti-diabetic medications; total cholesterol and LDL-C with anti-cholesterol medications; blood pressure with anti-hypertensive medications; hsCRP with chronic anti-inflammatory therapies.

Time-to-Event Analysis

The time will be calculated from RW0 to the first time when:

- The mUFC is above 1.5X the ULN of the central laboratory's reference range; OR
- The mUFC is over 40% above the Baseline (RW0) value if the RW0 value is above the ULN (i.e. >1.0X ULN) for the S-C cohort subjects; OR
- An early rescue criterion is met.

Time from RW0 to documented loss of therapeutic response will be defined as event date (i.e. earliest date a subject meets one of the criteria above) minus the date of the first dose of blinded study drug (levoketoconazole or placebo) + 1 day.

For each of the first two criteria, the date for meeting the criterion is considered to be the date of the last adequate sample included in the calculation of the mUFC. For the third criterion for early rescue subjects, the date for meeting the criterion is considered to be the date in which the Investigator withdraws blinded study drug in favor of other treatments, including open-label levoketoconazole. This date is the day after the date of last dose of blinded study drug recorded in the Early Rescue Study Completion eCRF.

In the absence of loss of therapeutic response where subjects do not meet any of the criteria above, subjects will be censored at the last sample date of the last mUFC during the RW Phase (i.e. either from RW4 or RW5, see Section 5.10). Subjects with no post-RW Baseline mUFC will be considered as losing therapeutic response. Time (day) from RW0 to documented loss of therapeutic response for these subjects is set to 1 (day).

Time to first normalization of mUFC beginning from RW5 (for the subset of subjects with mUFC above 1.5X ULN at RW5 and received at least one dose of study drug during the RES Phase) is defined as the time from the first dose of add-on study drug (levoketoconazole or placebo) at RW5 to the time when a subject has confirmed mUFC no more than ULN (i.e. <=ULN) for the assay reference range (inclusive of both dates). In the absence of normalization of mUFC, a subject will be censored at the date of the last adequate sample included in the calculation of the last non-missing mUFC during the RES Phase.

Time to first normalization of LNSC beginning from RW5 (for the subset of subjects with LNSC above ULN at RW5 and received at least one dose of study drug during the RES Phase) will be estimated as the time from the first dose of add-on study drug (levoketoconazole or placebo) at RW5 to the time when a subject has an LNSC result no more than the ULN for the assay reference range (inclusive of both dates). In the absence of normalization of LNSC, a subject will be censored at the

date of last sample for the assessment of LNSC during the RES Phase. If there are two such samples (in which case the mean LNSC is calculated), the subject will be censored at the date of the second sample.

Time from TM0 to first normalization of mUFC (during TM Phase) is defined as the time from the first dose in TM Phase to the time when a subject has confirmed mUFC no more than ULN (i.e. <=ULN) for the assay reference. In the absence of normalization of mUFC, a subject will be censored at the date of the last adequate sample included in the calculation of the last non-missing mUFC during the TM Phase.

Each time-to-event endpoint will be calculated in days as end date minus start date + 1 day, where the start date is the RW first dose date for the first endpoint, the RES first dose date for the second and third endpoints, and the TM first dose date for the fourth endpoint, and the end date is as defined for each endpoint above. Each time-to-event endpoint will be analyzed using Kaplan-Meier (KM) methods. Median time to each event along with the associated two-sided 95% CI will be presented by treatment group. The time-to-event distributions will be compared statistically with a two-sided log-rank test. The CI will be regarded as a descriptive statistic. Corresponding KM curves for each time-to-event endpoint will also be presented by treatment group (see SAS STAT Procedure details in Appendix 3).

Normalization of mUFC during the TM Phase

The number and proportion of subjects meeting the following criteria will be summarized by Therapeutic Dose, and for All Subjects (along with 95% CI for the proportion):

- Achieved normalization of mUFC (<=ULN for the assay reference range) at the end of TM Phase
- Achieved normalization of mUFC (<=ULN for the assay reference range) at any time during the TM Phase
- Achieved either normalization of mUFC or partial response (at least 50% decrease in mUFC from TM Baseline) at the end of TM Phase
- Achieved either normalization of mUFC or partial response (at least 50% decrease in mUFC from TM Baseline) at any time during the TM Phase

Subject's last non-missing mUFC result during the TM Phase will be used to determine normalization at the end of TM Phase.

A needle plot of the individual subject mUFC from TM Baseline to the end of the TM Phase (i.e. non-missing mUFC at the last dose level; otherwise, non-missing mUFC at the penultimate dose level) will be produced, with the ULN as a horizontal reference line.

Normalization of LNSC during RES Phase

The number and proportion of subjects in the RES Phase with normalization of LNSC at RES2 will be analyzed in the same way as the corresponding normalization of UFC at RES2 analyses described in Section 6.6.2.1.

Serum Cortisol and ACTH

Serum cortisol and ACTH are primarily considered to be safety endpoints and will be descriptively summarized. Refer to Section 6.7.2 for details.

Clinical Signs and Symptoms Excluding Acne, Hirsutism and Peripheral Edema

Individual clinical signs and symptoms, which are assessed at each visit, will be summarized by number and percentage of subjects with each sign/symptom. Individual clinical sign and symptoms severity scores will be summarized using shift tables from the relevant Baseline. Total severity score and changes from the relevant Baseline of this total score will be summarized.

Exploratory Efficacy Endpoint Analysis for the TM Phase, RW Phase, and RES Phase

HOMA-%B

HOMA-%B is calculated as (20 x FPI)/(FPG - 3.5), where FPI=fasting plasma insulin (mU/L) and FPG=fasting plasma glucose (mmol/L). HOMA-%B is only calculated whenever FPI and FPG are collected on the same day. Collections from different days cannot be combined to calculate HOMA-%B. Subjects who are using insulin during the study phase will be excluded from all HOMA-%B summaries.

For TM Phase, HOMA-%B will be summarized for the L-N population, by the levoketoconazole dose received corresponding to the last non-missing results and by the highest levoketoconazole dose received with non-missing results.

For the RW Phase and RES Phase, observed values as well as absolute and percent changes from Baseline will be summarized by treatment group. Shift table by screening diabetes status will be presented as described in Section 6.6.4.4. No statistical testing will be performed.

Other Exploratory Efficacy Endpoint Analyses for the RW Phase and RES Phase

Partial loss of response during the RW Phase

The proportion of subjects with partial loss of response (mUFC that increases from Baseline (RW0) to above normal but no more than 1.5X ULN, or LNSC that was normal at Baseline (RW0; both values ≤ULN) and becomes abnormal (>ULN) based on each of two LNSC collections obtained on different nights) AND clinically significant

worsening of signs and symptoms of CS during the RW Phase will be compared between treatment groups using a two-sided Fisher's Exact test and a 95% CI of the difference will be calculated.

Shifts in mUFC and LNSC

Shifts from RW Baseline in mUFC to RW5 (or last post-RW Baseline visit) and RES2 (or RES1 if RES2 value is missing) will be summarized based on the following categories:

- Less than lower limit of normal (LLN) (< LLN)
- Within the normal range ([LLN; ULN])
- Greater than ULN and less than or equal to 2x ULN ((ULN; 2 x ULN])
- Greater than 2x ULN and less than or equal to 5x ULN ((2 x ULN; 5 x ULN])
- Greater than 5x ULN (>5 x ULN)

Shifts from RW Baseline in LNSC to RW5 (or last post-RW Baseline visit) and RES2 (or RES1 if RES2 value is missing) will be summarized based on the following categories: >ULN or ≤ULN.

Other Exploratory Efficacy Endpoint Analysis for the TM Phase for the L-N Cohort

Selective Efficacy Endpoints at the Last and the Highest Dose Received with Non-missing Result

Observed values as well as change from TM Baseline will be presented for mUFC, LNSC, biomarkers of CS comorbidities, CS QoL total score, BDI-II total score, acne global score, hirsutism (for females only) total score, and peripheral edema total scores, at the last observed result by corresponding dose level, and at the highest dose level received with non-missing result by corresponding dose level during the TM Phase.

6.6.6 Pharmacokinetics Analysis

The Pharmacokinetics analysis will be detailed in a separate analysis plan.

6.6.7 Pharmacodynamics Analysis

The Pharmacodynamics analysis will be detailed in a separate analysis plan.

6.7 Safety

Descriptive summaries of all safety endpoints will be presented by treatment group and visit (RW1-RW5 and RES1-RES2; where applicable) for the RW Phase and RES Phase and at the End of RW Phase time point. For continuous variables, summaries

will be presented for the observed values and for the absolute and percent changes from the relevant Baseline.

For the TM Phase, safety endpoints measured at multiple visits will be summarized by the levoketoconazole dose received corresponding to the last non-missing results, by the highest levoketoconazole dose received with non-missing results, and by the levoketoconazole dose received corresponding to the worst results (where specified). For subjects in the RW Phase, summaries will also be presented by Therapeutic Dose and for all dose groups combined. Statistical testing, whether within-treatment or between-treatment, and only where specified, will only be for the all levoketoconazole dose groups combined.

For the RW Phase and RES Phase Combined and the All Phases Combined, where applicable, only the safety results while on levoketoconazole will be summarized.

The Safety population will be determined based on study phase and the study overall (i.e. all phases combined). Refer to Section 4.5 for details on the analysis populations to use.

6.7.1 Adverse Events

All adverse events (AEs) will be collected from the date that informed consent was signed. Verbatim terms for each AE will be coded using MedDRA version 21.1 and graded per the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.03 [NCI CTCAE, 2010].

Table 5 The Severity of AEs and SAEs per NCI CTCAE, Version 4.03

Grade 1	Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated
Grade 2	Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living
Grade 3	Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care activities of daily living
Grade 4	Life-threatening consequences; urgent intervention indicated
Grade 5	Death related to AE

An AE of scientific and medical concern specific to the study treatment will be identified as AESI. AESIs will be reported regardless of seriousness or causality. There are 3 types of AESIs: Persistent QTc prolongation, Potential hepatic events, and Adrenal insufficiency.

All AE summaries will be restricted to Treatment-Emergent AEs (TEAEs), which are defined as those AEs with an onset on or after the date of the first study drug dose and up to 30 days after the last dose of study drug. TEAE summaries will be presented

by study phase including TM Phase (L-N population only), RW Phase, RES Phase, the RW Phase and RES Phase Combined, and All Phases Combined. For each phase, only TEAEs that occur on or after the start of that phase up to the end of that phase will be summarized. For subjects who withdraw during or at the end of the phase, any reported TEAE that starts after withdrawal but within 30 days of the last dose will be included in the summaries for that phase. For the TEAE summaries for the RW Phase and RES Phase Combined and All Phases Combined, where specified below, only the TEAEs while on levoketoconazole will be included (i.e. the TEAEs during the RW Phase for subjects randomized to placebo will be excluded).

For the RW Phase and RES Phase, TEAEs will be summarized by treatment group. For the TM Phase summary, TEAEs will be summarized by dose level. To determine the dose level that is associated with each AE record, the start date of the AE will be compared to the dose level start and end dates. If dose level start date <= AE start date <= dose level end date, then the dose level will be assigned to the AE record.

Details of start and end date for each study phase are provided in Section 5.1. Details of summary presentations for each phase are provided in Section 6.1.

The number and percentage of subjects will be summarized by SOC and PT and will be ordered by descending order of incidence of SOC and PT within each SOC. A subject with more than one occurrence of the same TEAE in a particular SOC or PT will be counted only once in in that particular SOC or PT. If a subject experience the same TEAE at more than one severity, or with more than one relationship to study drug, the most severe rating or the stronger causal relationship to study drug will be given precedence.

The following TEAE summaries will be produced:

- For the overview summary of all TEAEs - Total number of TEAEs, number of subjects with at least one TEAE, number of subjects with at least one treatment-emergent SAE, number of subjects with at least one TEAE with special interest, number of subjects with at least one serious TEAE of special interest, number of subjects with at least one severe TEAE (CTCAE grade =3), number of subjects with at least one life-threatening TEAE (CTCAE grade =4), number of subjects with at least one TEAE leading to discontinuation of study drug, and number of subjects with at least one TEAE leading to death. For the number of TEAEs, a unique event for a subject is defined as a unique combination of SOC, PT, onset date, CTCAE grade, and relationship to study drug.
- For the overview summary of all study drug-related TEAEs – Same as in the previous bullet point.
- Number of subjects with TEAEs by MedDRA SOC and PT.

- Number of subjects with TEAEs by MedDRA SOC, PT, and closest relationship to study drug (Related/Not Related). Related events are defined as events that are definitely or probably related to study drug. AEs with a missing relationship are considered definitely related. At each level of subject summarization, a subject is classified according to the closest relationship if the subject reported one or more events.
- Number of subjects with TEAEs by MedDRA SOC, PT, and worst reported severity. Worse severity is defined as first Death, then life-threatening followed by Severe, Moderate and by Mild as the lowest level. AEs with missing severity are categorized as severe in the TEAE summary tables. At each level of subject summarization, a subject is classified according to the highest severity if the subject reported one or more events.
- Number of subjects with treatment-emergent SAEs by MedDRA SOC and PT.
- Number of subjects of the most common TEAEs (frequency above 5% of all subjects in the TM Phase and 5% for at least one treatment group in the RW Phase and RES Phase) by MedDRA PT.
- Number of subjects with TEAEs of special interest (AESI) by category, MedDRA SOC and PT. The AESI categories are QTc prolongation, potential hepatic events, and adrenal insufficiency.
- Number of subjects with TEAEs of AESI with grades >=3 by category, MedDRA SOC and PT.
- Number of subjects with serious TEAEs of special interest by AESI category, MedDRA SOC and PT.
- Number of subjects with TEAEs leading to study drug discontinuation by MedDRA SOC and PT.
- Descriptive summary of time from start of treatment to first onset of most common TEAEs (frequency above 5% of all subjects in the TM Phase and 5% for at least one treatment group in the RW Phase and RES Phase).

For summaries of overall TEAEs, TEAES by SOC and PT, SAEs, AESIs, AEs leading to study drug discontinuations, combined summaries of the RW Phase and RES Phase and all three phases will also be presented for subjects while they are on levoketoconazole. For the RW Phase and RES Phase Combined, data after early rescue will not be included in the summaries.

Listings of all AEs, SAEs, AESIs, AEs leading to discontinuation of study drug, AEs leading to withdrawal from the study, and AEs leading to death will be presented by study phase, cohort, treatment group, subject, detailing verbatim term given by the Investigator or designee, SOC, PT, onset date, resolution date, severity, seriousness,

AESI category, action taken, outcome, and drug relatedness. The event onset will also be shown relative (in number of days) to date of first dose within each phase and across all phases.

TEAE summaries by subgroups of interest will be performed as described in Section 6.7.7.

6.7.2 Laboratory Evaluations

The following safety laboratory analytes will be analyzed:

Liver Function Tests: Aspartate Aminotransferase (AST) (SGOT), Alanine Aminotransferase (ALT) (SGPT), Gamma-glutamyl transferase (GGT), Alkaline phosphatase (ALP), Total bilirubin.

Hematology: Platelet Count, Red Blood Cell (RBC) Count, White Blood Cell (WBC) Count (absolute), Hemoglobin, Hematocrit, RBC indices (Mean corpuscular volume (MCV), Mean corpuscular hemoglobin (MCH), Mean corpuscular hemoglobin concentration (MCHC)), WBC Differentials (Neutrophils, Lymphocytes, Monocytes, Eosinophils, Basophils).

Clinical Chemistry, excluding Liver Function Tests: Blood Urea Nitrogen, Creatinine, Sodium, Potassium, Chloride, Total CO_2, Calcium, Magnesium, Phosphate, Uric Acid, Albumin, Total Protein, and Lactate dehydrogenase (LDH).

Coagulation Tests: INR, PT, PTT

Urinalysis: Specific gravity, pH, glucose, protein, blood and ketones. In addition, microscopic examination if blood or protein is > trace positive by dipstick.

Safety Hormones and Biomarkers: Free and Total Testosterone, Thyroid-Stimulating Hormone (TSH)/Free T4, Prolactin, High-Density Lipoprotein-Cholesterol (HDL-C), LDL-C:HDL-C ratio, Triglycerides

Other Laboratory Tests: Serum cortisol, ACTH, urine albumin: creatinine (ratio from spot urine)

Laboratory results will be presented as received from the central laboratory, no data conversions will be performed. Results will be summarized by the SI units (see Section 6.1).

Descriptive summaries by phase and visit (where applicable), as described in Section 6.7, will be produced for the laboratory tests. In addition, the number and percentage of subjects with abnormal laboratory values will be summarized.

Shifts (low/normal/high) from the relevant Baseline tables based on the normal ranges will be constructed. For the TM Phase, the shifts will be from the TM Baseline to the last non-missing results, to the non-missing results while on the highest levoketoconazole dose, and to the worst results during the phase. For each of the RW

Phase and RES Phase separately, the shifts will be from the relevant Baseline to each post-Baseline visit and to the End of RW Phase time point and to the worst results during the phase. For the RW Phase and RES Phase Combined, the shifts will be from the RW Baseline to the worst results during the two phases, including only the results while the subjects were on levoketoconazole.

For LFTs, tables will also be created to summarize the incidence at each visit and at the End of RW Phase time point and the shifts from the relevant Baseline for ALT, AST and total bilirubin (similar to the shifts described above) using the following categories:

ALT and AST: <LLN, within the normal range, > ULN and ≤ 3X ULN, > 3X ULN to ≤ 5X ULN, > 5X ULN to ≤ 10X ULN and > 10X ULN.

Total bilirubin: <LLN, within the normal range, > ULN and ≤ 1.5X ULN, >1.5X ULN and ≤2X ULN, and >2X ULN.

In addition, summaries for LFTs using the categories described above at the worst post-baseline result and the last observed result, as well as the shift from baseline to the worst post-baseline result and the last observed result, will be presented for all phases combined while on levoketoconazole.

An evaluation of drug-induced serious hepatotoxicity (eDISH) plot will be produced to assess the liver enzyme abnormalities that occur. ALT, AST, and Total bilirubin during each study phase will be plotted as multiples of the ULN, on a logarithmic scale. Those results that are greater than 2x ULN for Total bilirubin and greater than 3x ULN for ALT (or AST) will be easily identified through the use of reference lines. The following 6 eDISH plots will be produced with the first variable on the x-axis and the second variable on the y-axis, and where the Baseline is the relevant Baseline for the study phase being plotted:

- Baseline ALT (expressed as x ULN) x Maximum Post-RW Baseline ALT (expressed as x ULN)
- Maximum Post-baseline ALT (expressed as x ULN) x Maximum Post-baseline AST (expressed as x ULN)
- Maximum Post-baseline ALT (expressed as x ULN) x Maximum Post-baseline Total Bilirubin (expressed as x ULN)
- Baseline AST (expressed as x ULN) x Maximum Post-baseline AST (expressed as x ULN)
- Maximum Post-baseline AST (expressed as x ULN) x Maximum Post-baseline Total Bilirubin (expressed as x ULN)

- Baseline Total Bilirubin (expressed as x ULN) x Maximum Post-baseline Total Bilirubin (expressed as x ULN)

A spaghetti plot will be produced for each study phase to show the results for ALT, AST, and Total bilirubin relative to the ULN for each subject versus time since start of study drug. The results will be plotted as multiples of the ULN, on a logarithmic scale. Box and whisker plots presenting actual and change from Baseline values in the three LFTs for each visit in all three phases will be produced.

Potential clinically significant (PCS) results for selected laboratory tests will be listed, according to the criteria listed in Table 6 below:

Table 6 Laboratory Test PCS Criteria

Category	Analyte	PCS Criteria – Conventional Units	PCS Criteria – SI Units
Hematology	Eosinophils	≥ 700/µL	≥ 0.7×10^9/L
	Hematocrit (Hct)	Female: ≤ 32%; Male: ≤ 37%	Female: ≤ 0.32; Male: ≤ 0.37
	Hemoglobin (Hgb)	Female: ≤ 9.5 g/dL; Male: ≤ 11.5 g/dL or Decrease of ≥ 20%	Female: ≤ 95 g/L; Male: ≤ 115 g/L or Decrease of ≥ 20%
	Leukocytes	Low: ≤ 2800/ µL High: ≥ 16000/µL	Low: ≤ 2.8×10^9/L High: ≥ 16×10^9/L
	Neutrophils	≤ 1000/µL	≤ 10^9/L
	Platelets	Low: ≤ 75×10^3/µL High: ≥ 700×10^3/µL	Low: ≤ 75×10^9/L or High: ≥ 700×10^9/L
Blood Chemistry	Albumin	< 2.5 g/dL	< 25 g/L
	Blood Urea Nitrogen (BUN)	> 30 mg/dL	> 10.71 mmol/L
	Calcium	Low: < 7 mg/dL; High: > 12 mg/dL	Low: < 1.75 mmol/L; High: > 3 mmol/L
	Cholesterol	> 300 mg/dL	> 7.76 mmol/L
	Creatinine	> 2 mg/dL	> 176.8 umol/L
	Glucose	Low: < 50 mg/dL; High: >250 mg/dL	Low: < 2.8 mmol/L; High: >13.9 mmol/L

Category	Analyte	PCS Criteria – Conventional Units	PCS Criteria – SI Units
	LDH	> 3x ULN	> 3x ULN
	Phosphate	Low: < 2.0 mg/dL; High: > 5.0 mg/dL	Low: < 0.65 mmol/L; High: > 1.62 mmol/L
	Potassium	Low: < 3.0 mEq/L; High: > 5.5 mEq/L	Low: < 3.0 mmol/L; High: > 5.5 mmol/L
	Sodium	Low: < 130 mEq/L; High: > 150 mEq/L	Low: < 130 mmol/L; High: > 150 mmol/L
	Triglycerides	High: > 500 mg/dL	> 5.65 mmol/L
	Uric Acid	Female: > 8.0 mg/dL; Male: > 10.0 mg/dL	Female: > 475.8 umol/L; Male: > 594.8 umol/L
Liver Safety	ALT	High: > 3x ULN	High: > 3x ULN
	ALP	High: > 3x ULN	High: > 3x ULN
	AST	High: > 3x ULN	High: > 3x ULN
	Total Bilirubin	High: > 2x ULN	High: > 2x ULN
	GGT	High: > 3x ULN	High: > 3x ULN
Urinalysis	Blood	Increase of ≥ 2 units	
	Glucose	Increase of ≥ 2 units	
	Protein	Increase of ≥ 2 units	

Testosterone

In addition to the summaries described above for laboratory tests in general, all summaries for testosterone levels will be done separately by sex (also stated in Section 6.7.7). The mean changes from RW Baseline during the RW Phase in testosterone levels by sex will be assessed for treatment group differences using two-sample t-test at each visit and at the End of RW Phase time point. Within-treatment mean changes from the relevant Baseline to each visit in the RW Phase and RES Phase and at the End of RW Phase time point will be assessed for nominal significance using paired t-test.

Serum Cortisol

In addition to the summaries described above for laboratory tests in general, the mean changes from RW Baseline during the RW Phase in serum cortisol will be assessed for

treatment group differences using two-sample t-test at each visit and at the End of RW Phase time point. Within-treatment mean changes from the relevant Baseline to each visit in the RW Phase and RES Phase and at the End of RW Phase time point will be assessed for nominal significance using paired t-test.

ACTH

All summaries for ACTH will be presented for subjects with Cushing's Disease (CD) only. In addition to the summaries described above for laboratory tests in general, the following will be done for ACTH:

- ACTH will be transformed to fold-ULN (i.e. multiples of the ULN) and summaries of ACTH as fold-ULN will be produced.
- The mean changes from RW Baseline, using both the actual values and the fold-ULN values, during the RW Phase will be assessed for treatment group differences using two-sample t-test at each visit and at the End of RW Phase time point. Within-treatment mean changes from the relevant Baseline to each visit in the RW Phase and RES Phase and at the End of RW Phase time point will be assessed for nominal significance using paired t-test.
- Subgroup summary: ACTH summary of actual value and change from RW Baseline to all post-RW Baseline visits during the RW Phase and at the End of RW Phase time point will be presented by sex (also stated in Section 6.7.7).
- Needle plots of individual subject ACTH values plotted as fold-ULN from RW Baseline to the end of the RW Phase (i.e. last non-missing post-RW Baseline ACTH during the phase, before early rescue in cases where early rescue occurred) will also be produced.
- Additional exploratory analysis to examine correlations between ACTH levels and the normalization of mUFC during the RES Phase may be also performed.

6.7.3 Oral Glucose Tolerance Test (for Pre-diabetic Subjects)

Pre-diabetic subjects are defined as those who had a Baseline fasting glucose greater or equal to 100 mg/dL (5.6 mmol/L) and less than 126 mg/dL (7.0 mmol/L) without concomitant use of anti-diabetic medication. The diabetic and pre-diabetic status is recorded on the Oral Glucose Tolerance Test (OGTT) eCRF.

All pre-diabetic subjects with impaired fasting glucose (subjects who have not been previously diagnosed with diabetes and are currently pre-diabetic and not receiving medication to lower blood glucose) have an OGTT at Screening, TM7, RW0, RW5, and RES2. Plasma glucose values before (i.e. 0 min) and 30 min, 60 min, 90 min, and 120 min after glucose administration are to be determined. Glucose tolerance before and after the OGTT is to be determined based on 120-min glucose level as the following:

Pre-Test:

- Normal: glucose level< 100 mg/dL (5.6 mmol/L)
- Impaired fasting glucose: >= 100 mg/dL (5.6 mmol/L) and < 126 mg/dL (7.0 mmol/L)

Post-Test:

- Normal: 2-hour glucose level below 140 mg/dL
- Impaired glucose tolerance (pre-diabetic): glucose level at least 140 mg/dL and less than 200 mg/dL
- Provisional diagnosis of diabetes (diabetic): glucose level at least 200 mg/dL

  Note: If 120-min glucose level is missing, then the non-missing 90-min glucose level will be used; otherwise, the glucose level will be considered as missing for the purposes of determining glucose tolerance.

The following summaries of OGTT glucose values will be produced:

- Number and percentage of subjects for each pre-test category at Screening and TM7 by last dose received at TM7, and at RW0, RW5, End of RW Phase, and RES2 by treatment group.
- Number and percentage of subjects for each post-test category of 120-min glucose level at Screening, TM7 by last dose received at TM7, and RW0, RW5, End of RW Phase, and RES2 by treatment group.
- Shifts from Screening to TM7 by last dose received at TM7, from RW0 to RW5, from RW0 to End of RW Phase, and from RW5 to RES2 in glucose post-test category by treatment group.
- Descriptive statistics will be determined for the maximum glucose values during OGTT at Screening, TM7 (actual and change), RW0, RW5 (actual and change), End of RW Phase (actual and change), and RES2 (actual and change). Paired t-test will be used to assess the changes in maximum glucose values during OGTT from Screening to TM7, from RW0 to RW5, from RW0 to End of RW Phase, and from RW5 to RES2.
- Descriptive statistics will be determined, and paired t-test will be performed, as described above, for time (from glucose administration, in minutes) to the maximum glucose values during OGTT.
- Descriptive statistics will be determined, and paired t-test will be performed, as described above, for total area under curve (AUC) and incremental AUC of glucose values in OGTT.

- For the determination of Maximum, Time to Maximum, AUC and incremental AUC for individual subjects – there is no imputation of missing values. These variables will be determined only if there is no more than one time point with a missing value and the time point with a missing value cannot be the 0-min and 120-min time points.

The calculated indices and their changes from Screening to TM7, from RW0 to RW5, from RW0 to End of RW Phase, and from RW5 to RES2 will be summarized by each treatment group.

6.7.4 Vital Signs

Descriptive summaries by phase and visit (where applicable) and at the End of RW Phase time point, as described in Section 6.7 will be produced for vital signs (blood pressure, heart rate, and temperature). The mean changes from RW Baseline in blood pressure during the RW Phase will be assessed for treatment group differences using two-sample t-test at each visit and at the End of RW Phase time point. Within-treatment mean changes in blood pressure from the relevant Baseline to each visit in the RW Phase and RES Phase and at the End of RW Phase time point will be assessed for nominal significance using paired t-test.

The number and percentage of subjects by category of vital sign clinical importance, presented in Table 7, will be summarized by phase and visit (where applicable) and at the End of RW Phase time point, as described in Section 6.7.4.

Table 7 Vital Sign Clinical Importance Criteria

Vital Sign	Criteria	Flag
Heart Rate (HR)	below 44 bpm	Low (L)
	44-100 bpm	Normal
	101-120 bpm	High (H)
	above 120 bpm	Very High (VH)
SBP	below 90 mm Hg	Low (L)
	90-139 mm Hg	Normal
	140-169 mm Hg	High (H)
	170 mm Hg and above	Very High (VH)
DBP	below 50 mm Hg	Low (L)
	50-89 mm Hg	Normal
	90-109 mm Hg	High (H)
	110 mm Hg and above	Very High (VH)

Shifts from the relevant Baseline tables based on the categories above will be constructed. For the TM Phase, the shifts will be from the TM Baseline to the last non-missing results, to the non-missing results while on the highest levoketoconazole dose, and to the worst results during the phase. For each of the RW Phase and RES Phase separately, the shifts will be from the relevant Baseline to each post-Baseline

visit and at the End of RW Phase time point and to the worst results during the phase. For the RW Phase and RES Phase Combined, the shifts will be from the RW Baseline to the worst results during the two phases, including only the results while the subjects were on levoketoconazole.

Shifts from the relevant Baseline in blood pressure classification as defined in Table 8 below, will be summarized in the same manner as described above for the shifts based on vital sign values of clinical importance.

Table 8 Blood Pressure Classification (Adults Over 18 Years of Age)

Blood Pressure Classification	
SBP (mmHg)	**DBP (mmHg)**
< 120	and < 80
120-139	or 80-89
140-159	or 90-99
≥ 160	or ≥ 100

Subjects with PCS results in vital signs will be listed, according to the criteria listed in Table 9 below:

Table 9 Vital Sign PCS Criteria

Vital Sign Parameters	PCS Criteria
Pulse	Increase of ≥ 15 bpm from baseline and ≥ 120 bpm
	Decrease of ≥ 15 bpm from baseline and ≤ 50 bpm
Systolic Blood Pressure	Increase of ≥ 20 mmHg from baseline and ≥ 180 mmHg
	Decrease of ≥ 20 mmHg from baseline and ≤ 90 mmHg
Diastolic Blood Pressure	Increase of ≥ 15 mmHg from baseline and ≥ 105 mmHg
	≤ 50 mmHg
Temperature	Increase of ≥ 2.0 ºC from baseline and ≥ 38.0 ºC
Weight	Decrease of ≥ 10% from baseline
	Increase of ≥ 10% from baseline

6.7.5 Electrocardiograms

The following quantitative ECG measurements will be taken during the study:

- heart rate (bpm);
- RR interval (msec);
- PR interval (msec);
- QRS interval (msec);
- Bazett corrected QT (QTcB) interval (msec)
- Fridericia corrected QT (QTcF) interval (msec)

Descriptive summaries by study phase and visit (where applicable) and at the End of RW Phase time point, as described in Section 6.7 will be produced for the quantitative ECG measurements listed above and the qualitative overall ECG interpretation (normal versus abnormal). Within-treatment mean changes in QTcF from the relevant Baseline to each visit in the RW Phase and RES Phase and at the End of RW Phase time point will be assessed for nominal significance using paired t-test.

The number and percentage of subjects by category of QTc interval clinical importance, presented in Table 8, will be summarized by phase and visit (where applicable) and at the End of RW Phase time point, as described in Section 6.7.5. The worst categorical actual and worst categorical change from the relevant Baseline in post-Baseline values within each study phase will be summarized. Using the categories for actual values, shifts from the relevant Baseline will be constructed. For the TM Phase, the shifts will be from the TM Baseline to the last non-missing results, to the non-missing results while on the highest levoketoconazole dose, and to the worst results during the phase. For each of the RW Phase and RES Phase separately, the shifts will be from the relevant Baseline to each post-Baseline visit and at the End of RW Phase time point and to the worst results during the phase, and the within-treatment shifts from Baseline will be assessed for nominal significance using McNemar's test. For the RW Phase and RES Phase Combined, the shifts will be from the RW Baseline to the worst results during the two phases, including only the results while the subjects were on levoketoconazole.

In addition, summary of categorical ECG results at the worst post-baseline result and the last observed result, as well as shift from baseline to the worst post-baseline result and the last observed result, will be presented for all phases combined while on levoketoconazole. QTcF assessments and the time elapsed relative to study drug administration and PK blood draws will also be presented in listing. The elapsed time from dosing to each QTcF interval measurement with a valid time stamp was summarized by visit and overall within each study phase through descriptive statistics, as well as through frequency distributions where the elapsed times were categorized into: <1 hour post-dose, 1 to <2 hours post-dose, 2 to <3 hours post-dose, 3 to <4 hours post-dose, and ≥4 hours post-dose.

ECG readings that have been rejected after Spaulding central reader review will not be included in any analysis tables.

Table 10 QTcF Interval Clinical Importance Criteria

QTcF Interval	Criteria (msec)
Change from Baseline	below 30
	30-60
	above 60
Actual Value	<450
	451-480
	481-500
	above 500
Actual Value by Sex	>450 for Males >470 for Females

Subjects with PCS results in ECGs will be listed, according to the criteria listed in Table 11 below:

Table 11 Electrocardiogram (ECG) PCS Criteria

Parameter	Criteria
QTcF	Female: > 470 msec
	Male: > 450 msec
PR Interval	> 200 msec
Heart Rate	< 48 bpm
	> 96 bpm

6.7.6 Pituitary Magnetic Resonance Imaging

The results from the pituitary magnetic resonance imaging (MRI) will be summarized for CD subjects only. The number and percentage of subjects with macroadenoma and those with microadenoma will be summarized at Screening, RW Baseline, RW5, and End of RW Phase time point. The observed values in pituitary tumor size (in millimeters of maximum diameter), tumor height (sagittal and coronal views), tumor length (sagittal view), and tumor width (coronal view) and the changes in these values from Screening to RW Baseline, RW5, and End of RW Phase time point (for the L-N cohort) and from RW Baseline to RW5 and End of RW Phase time point (for the S-C cohort) will be summarized by treatment group. In addition to all CD subjects, the observed and changes values for tumor maximum diameter at each visit will be

summarized for the subset of CD subjects with a measurable tumor (ie, those with macroadenoma or microadenoma) at that visit.

The status of overall pituitary size and the status of pituitary tumor size/activity from Screening to RW Baseline, RW5, and End of RW Phase time point (for the L-N cohort) and from RW Baseline to RW5 and End of RW Phase time point (for the S-C cohort) will be summarized by treatment group, with the categories for the status as follows: Increase, Stable, Decrease, Unknown or Not Measurable, and Not Applicable. The basis of tumor status assessment will also be summarized with the categories as follows: a) Overall tumor appearance change; b) Maximum tumor diameter change; c) Inference from pituitary gland change; both a) and c); both b) and c); All of a), b), and c); No change; and Not applicable. In addition to all CD subjects, the status of pituitary tumor size/activity at each visit will be summarized separately for the subset of CD subjects with macroadenoma and those with microadenoma at Screening (for the L-N cohort) and at RW Baseline (for the S-C cohort).

6.7.7 Subgroup Analyses for Safety

The following safety summaries for the TM Phase and the RW Phase will be produced separately for the subgroup of subjects who completed SONICS (regardless of their therapeutic response at the end of SONICS) versus those who did not participate in SONICS:

- Overview summary of all TEAEs and summary of TEAEs by MedDRA SOC and PT (see Section 6.7.1)
- Shift from the relevant Baseline in ALT, AST, and Total bilirubin using the categories defined in Section 6.7.2.
- Summary of QTc intervals by the clinically important categories defined in Section 6.7.5.

The same set of summaries for TEAEs and QTc intervals will be produced in the RW Phase for subgroups of subjects (with a minimum subgroup size of at least 30% of the ITT population) based on the following:

- baseline glycemia status (pre-diabetic, diabetic, normal, as defined in Section 6.6.4.4)
- receiving anti-hypertension medication at Baseline/presence of high blood pressure at screening (Yes/No, as defined in Section 6.6.4.3)
- CD status (Yes/No)
- Region (US sites versus non-US sites)
- Sex (female versus male)

- Age (≤median age of the Safety population in the study overall versus > median age)

Descriptive summaries by subgroup of interest will be produced in each of the three study phases for the following laboratory tests:

- Testosterone levels by sex
- ACTH by sex (for the subset of subjects with CD only).

6.8 Interim Analysis

An interim analysis, also referred to as the End of RW Phase analysis, will be conducted after all enrolled subjects have either completed the RW Phase or discontinued prior to the end of the RW Phase. The interim analysis will include all data up to the end of the RW Phase. Data from the RES Phase will not be analyzed. Prior to the database lock and treatment unblinding for the interim analysis, the data up to the end of the RW Phase will be cleaned and reviewed in a blinded manner to resolve data queries, the major protocol deviations will be identified, including those that will lead to exclusion of subjects from the PP population, and the compositions of the analysis populations (excluding those populations that are based on dosing in the RES Phase) will be determined.

Treatment unblinding will occur after major protocol deviations are identified and used to finalize the analysis populations and following database lock for the interim analysis. After the database lock, a request for Unblinding Form, signed and approved by the Sponsor and CRO unblinded Biostatistics and unblinded Statistical Programming, will be sent to Bracket. Upon receipt of the approved Unblinding Form, Bracket will deliver the unblinded treatment codes (as a SAS dataset) to the designated CRO unblinded lead biostatistician via secure FTP server. Next, the unblinded treatment information received from Bracket will be integrated into the SDTM and ADaM datasets for generation of top line results first, and then the rest of the interim analysis results. As a result, CRO biostatistics and programming team will be unblinded.

The subjects, Investigators, study site personnel, representatives of the Sponsor and CRO involved in the monitoring of the study, and representatives of the CRO involved in the data management of the study will continue to be blinded at the individual subject level until after the database lock at the end of the study. The names of individuals who will be unblinded at the individual subject level for the interim analysis will be documented by the CRO Lead Statistician (and confirmed by the Sponsor Lead Statistician). Results of the interim analysis at the unblinded treatment group level (i.e. tabulated results only) may be shared with some of those who remain blinded at the individual subject level, if approved by the Sponsor.

No changes to the design or conduct of the remainder of the study nor the pre-specified analyses in the SAP will be made based on the interim analysis results. The results from the interim analysis for the primary efficacy endpoint and the secondary efficacy endpoints during the RW Phase will be considered as final. There will be no statistical adjustments of the results at the final analysis that will be conducted after the study ends.

Data up to the end of the RW Phase will be locked for the interim analysis and will remain locked for the remainder of the study. In the unlikely event that new or additional data become available after the interim analysis that impacts locked data, any data changes will be reviewed on a blinded basis and approved by the Sponsor. Such data changes will be discussed in the final study report.

An independent DSMB reviews all SAEs and AESIs on a rolling basis and assesses benefits and risk of therapy (unblinded) systematically at approximately 6-month intervals. There are no plans to stop the study prior to completion based on these reviews, and therefore no stopping rules are created. The DSMB will not report its findings other than to indicate whether the protocol must be changed for the study to continue as planned or otherwise make recommendations designed to ensure patient safety and study integrity. Details of the DSMB procedures are described in a separate charter and a DSMB statistical analysis plan.

7. Changes from Planned Analyses in the Protocol

In Section 12.5.2.6 of the protocol, the post-test glucose category of "Impaired glucose tolerance (pre-diabetic)" is defined as: "glucose level at least 140 mg/dL and no more than 200 mg/dL". This definition includes 200 mg/dL, which is an error as the next category of "diabetic" is defined as "glucose level at least 200 mg/mL" (ie, it also includes 200 mg/dL). In Section 6.7.3 of this SAP, this error has been corrected, and the pre-diabetic category is defined as "glucose level at least 140 mg/dL and *less* than 200 mg/dL".

In Section 12.5.6.1 of the protocol, it is stated that insulin values during the OGTT will be summarized. However, the blood samples for OGTT have not been analyzed for insulin by the central laboratory; therefore, the analysis for insulin cannot be performed.

In Section 12.5.6.1 of the protocol, it is stated that "Subgroup analyses by number of prior surgeries, time since last surgery, and category of prior medical treatment will also be explored depending on subgroup size." These subgroup analyses are no longer planned. Only those subgroup analyses that are defined in Sections 6.6.1, 6.6.2.1, 6.6.4, and 6.7.7 will be performed.

8. References

1 Doshi A, Zaheer A, Stiller MJ. A comparison of current acne grading systems and proposal of a novel system. *Int J Derm*. 1997;36:416-418.

2 ICH. *Statistical Principles for Clinical Trials*, Guideline E9, 1998. Available at http://www.emea.eu.int/pdfs/human/ich/036396en.pd

3 Arnett DK, Blumenthal RS, Albert MA1, Buroker AB2, Goldberger ZD3, Hahn EJ1, Himmelfarb CD1, Khera A1, Lloyd-Jones D1, McEvoy JW1, Michos ED1, Miedema MD1, Muñoz D1, Smith SC Jr1, Virani SS1, Williams KA Sr1, Yeboah J1, Ziaeian B4. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2019 Sep 10;140(11):e596–e646

9. APPENDIX 1 Safety Laboratory and Vital Signs Tests – Identifying Worst Values

The table below presents the direction of interest in determining the worst values for safety laboratory tests.

Laboratory Test	Direction of interest for worst case values	Direction to choose in case of both Low and High Values
HEMATOLOGY		
Hemoglobin	Low and High	Low
Hematocrit	Low and High	Low
RBC Count	Low and High	Low
MCV	Low and High	Low
MCH	Low	
MCHC	Low and High	Low
WBC Count	Low and High	Low
Basophils	High	
Eosinophils	High	
Neutrophils	Low and High	Low
Monocytes	High	
Lymphocytes	Low and High	High
Platelet Count	Low and High	Low
CHEMISTRY		
Calcium	Low and High	High
Magnesium	Low	
Direct Bilirubin	High	
Phosphate	Low	
Sodium	Low and High	High
Potassium	Low and High	Low
BUN (Blood Urea Nitrogen)	Low and High	High
Creatinine	High	
Albumin	Low	
Chloride	Low	
Total CO2	High	
Uric Acid	High	
Total Protein	Low and High	Low
LDH	High	
LIVER SAFETY TESTS		
Total Bilirubin	High	
AST (SGOT)	High	
ALT (SGPT)	High	
GGT	High	
ALP	High	
COAGULATION TESTS (INR/PT/PTT)	High/High/High	

Laboratory Test	Direction of interest for worst case values	Direction to choose in case of both Low and High Values
SAFETY HORMONES AND BIOMARKERS		
Free Testosterone	Low	
Total Testosterone	Low	
TSH	Low and High	Low
Free T4	High	
HDL-C	High	
LDL-C:HDL-C (ratio)	High	
Total Cholesterol	High	
Triglycerides	High	
hsCRP	High	
URINALYSIS		
Specific Gravity	Low and High	Low
pH	Low and High	Low
Protein		
Ketones		
Creatinine	High	
OTHER MEASURES		
Serum Cortisol	High	
ACTH	High	
Albumin: Creatinine (ratio)	Low and High	High

The table above will be used in determining the worst values for safety laboratory tests within a study phase or a visit window for a subject (see Section 5.3). If there are multiple such values for a laboratory test, then the worst value will be identified as follows:

- If all values are within the normal range (NR), then the lowest one will be used if 2nd column = Low. The highest one will be used if 2nd column = High. The one farthest from the midpoint of the NR will be used if 2nd column = Low and High. If two such values are equidistant from the midpoint, the lower one will be used if 3rd column = Low, and the higher one if 3rd column = High.
- If one or more values are within the NR and one value is outside the NR, then the value outside the NR will be used.
- If two or more values are outside the NR, then the lowest one among these will be used if 2nd column = Low. The highest one will be used if 2nd column = High. If the 2nd column = Low and High, then the lowest one will be used if 3rd column = Low, and the highest one if 3rd column = High.

The table below presents the direction of interest in determining the worst values for Vital Signs assessments tests.

Vital Signs Test	Direction of interest for worst case values	Direction to choose in case of both Low and High Values
HR	Low and High	Low
SBP	Low, High and Very High	High
DBP	Low, High and Very High	High
Temperature	High	

10. APPENDIX 2 Laboratory Tests – Precision Levels

Category	Analyte	SI Unit	Precision (Decimal Point)
Hematology	Hemoglobin	g/L	0
Hematology	Hematocrit	1	3
Hematology	RBC Count	10^{12}/L	1
Hematology	MCV	fL	1
Hematology	MCH	Pg	1
Hematology	MCHC	g/L	0
Hematology	WBC Count	10^9/L	1
Hematology	Basophils	%	1
Hematology	Abs. Basophils	10^9/L	2
Hematology	Eosinophils	%	1
Hematology	Abs.Eosinophils	10^9/L	2
Hematology	Neutrophils	%	1
Hematology	Abs.Neutrophils	10^9/L	2
Hematology	Monocytes	%	1
Hematology	Abs.Monocytes	10^9/L	2
Hematology	Lymphocytes	%	1
Hematology	Abs.Lymphocytes	10^9/L	2
Hematology	Platelet count	10^9/L	0
Chemistry	Calcium	mmol/L	2
Chemistry	Magnesium	mmol/L	2
Chemistry	Phosphate	mmol/L	2
Chemistry	Sodium	mmol/L	0
Chemistry	Potassium	mmol/L	1
Chemistry	Blood Urea Nitrogen (BUN)	mmol/L	1
Chemistry	Creatinine	umol/L	1
Chemistry	Albumin	g/L	0
Chemistry	Chloride	mmo/L	0
Chemistry	Total CO2	mmol/L	0
Chemistry	Uric Acid	umol/L	1
Chemistry	Total Protein	g/L	0
Chemistry	Glucose	mmol/L	1
Liver Safety	LDH	U/L	0
Liver Safety	Total Bilirubin	umol/L	1
Liver Safety	Direct Bilirubin	umol/L	1
Liver Safety	AST(SGOT)	U/L	0
Liver Safety	ALT(SGPT)	U/L	0
Liver Safety	GGT	U/L	0
Liver Safety	ALP	U/L	0
Liver Safety	LDH	U/L	0
Coagulation	INR	Ratio	1
Coagulation	PT	Seconds	1
Coagulation	APTT	Seconds	1
Safety Hormones and Biomarkers	Testosterone, Free	nmol/L	2

Category	Analyte	SI Unit	Precision (Decimal Point)
Safety Hormones and Biomarkers	Testosterone. Total	nmol/L	2
Safety Hormones and Biomarkers	TSH	mU/L	2
Safety Hormones and Biomarkers	T-4 (Thyroxine) Free	pmol/L	1
Safety Hormones and Biomarkers	Prolactin	pmol/L	1
Safety Hormones and Biomarkers	HDL-C	mmol/L	2
Safety Hormones and Biomarkers	LDL-C	nmol/L	2
Safety Hormones and Biomarkers (CS Comorbidity)	LDL-C:HDL-C (ratio)	Ratio	2
Safety Hormones and Biomarkers (CS Comorbidity)	Total Cholesterol	mmol/L	2
Safety Hormones and Biomarkers (CS Comorbidity)	Hemoglobin A1C	%	1
Safety Hormones and Biomarkers (CS Comorbidity)	Glucose, Plasma (Fasting) during OGTT*	mmol/L	1
Safety Hormones and Biomarkers (CS Comorbidity)	Insulin, Plasma (Fasting) during OGTT*	pmol/L	2
Safety Hormones and Biomarkers	Triglycerides	mmol/L	2
Urinalysis	Specific Gravity		3
Urinalysis	pH		1
Biochemical Marker	Serum Cortisol	nmol/L	1
Biochemical Marker	LNSC	nmol/L	1
Other-Biochemical Marker	HOMA-IR		1
Other-Biochemical Marker	ACTH	pmol/L	1
Other-Biochemical Marker	hsCRP	mg/L	1
Other-Biochemical Marker	Albumin: Creatinine (ratio)	g/mol	0
Efficacy - Urinalysis	Cortisol, 24-hour	nmol/D	1
Efficacy - Urinalysis	Mean Cortisol, 24-hour	nmol/D	1
Efficacy – Urinalysis	Creatinine Excretion (Urine Creatinine)	nmol/D	1

11. APPENDIX 3 SAS STAT Procedure Details

The SAS codes provided in this section best represent the planned analyses of endpoints at the time of the development of this SAP. Without deviating from the planned statistical models and tests, the SAS codes may be subject to change (e.g. additional options added to a statement) to provide all the results needed in the planned tables, as well as to interpret these results.

1. Primary Analysis of the Primary Endpoint

Statistical significance testing will be conducted using a logistic regression model containing fixed effect terms for treatment group (levoketoconazole, placebo).

Example of SAS STAT procedure:

```
proc genmod data = datain;
   class trt01pn cohort;
   model  avalc (event='Y') = trt01pn cohort/
   dist= binomial link=logit;
   lsmeans trt01pn/cl pdiff ilink obsmargins;
run;

where avalc = 'Y' if there is loss of therapeutic response for the
subject and 'N' if there is no loss of therapeutic response for the
subject
trt01pn= treatment group (levoketoconazole, placebo)
cohort = Cohorts 1 and 2 (levoketoconazole-naïve, SONICS-completer)
```

2. Supportive Analysis of the Primary Endpoint

The proportions between the treatment groups will be compared using a two-sided Fisher's Exact test and a 95% CI of the difference will be calculated.

Example of SAS STAT procedure:

```
proc freq data=datain order=data;
  tables trt01pn*avalc /Fisher riskdiff(cl=exact);
  exact riskdiff;
run;
where avalc and trt01pn are as defined above
```

3. Sensitivity Analysis of Key Secondary Efficacy Endpoint: mUFC

Change from Baseline in imputed mUFC at RW5 will be analyzed using an analysis of covariance with treatment group and subject cohort as fixed effects and Baseline mUFC as a covariate.

Example of SAS STAT procedure:

```
proc mixed data=param noclprint;
class trt01pn cohort;
```

```
model chg = trt01pn cohort base/ solution cl;
lsmeans trt01pn / pdiff cl;
run;
trt01pn and cohort are as defined above
base = RW Baseline value of the mUFC
chg = Change from RW Baseline in mUFC at RW5
```

4. Supportive Analysis of Key Secondary Efficacy Endpoints: mUFC, LNSC, and CS Comorbidity Biomarkers

Example of SAS STAT procedure for the change from Baseline using repeated measures mixed effects model*:*

```
proc mixed data=param noclprint;
 class trt01pn cohort avisitn conccond;
 model chg = trt01pn cohort avisitn trt01pn*avisitn base conccond /
solution cl ddfm=KR;
repeated avisitn / subject= usubjid type=<Model Structure>;
lsmeans trt01pn / pdiff cl obsmargins;
lsmeans trt01pn* avisitn/ pdiff cl obsmargins;
run;

where avisitn = analysis visit (RW1 - RW5)
trt01pn and cohort are as defined above
base = Baseline value of the efficacy parameter
conccond = concurrent CS medication condition (applicable to the CS
comorbidity biomarkers only, where specified)

chg = Change from RW Baseline (in the efficacy parameter)

where Model Structure = spatial power, toeplitz with heterogeneity, and
toeplitz

for spatial power option (i.e., REPEATED / type=SP(time), use days
variable (e.g., LBDY) for time.

Note: For the model with spatial power the following will be added to
the SAS code:

   Random intercept / subject=usubjid;
```

5. Hochberg Adjustment

```
proc multtest inpvalues=DATAIN hoc;
run;

Note: DATAIN contains endpoints and their associated raw p-values.
```

6. Analysis of Time-to-Event Exploratory Endpoints:

Example of SAS STAT procedure:

```
proc lifetest data=datain plots=survival(atrisk);
time aval*cnsr(1);
```

```
strata trt01pn;
run;
where aval = Time to event in days
cnsr = 0 for event, 1 for censored
trt01pn is as defined above
```