# STATISTICAL ANALYSIS PLAN

| | |
|---|---|
| **Study Title:** | A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of SPR001 (Tildacerfont) in Reducing Supraphysiologic Glucocorticoid Use in Adult Subjects with Classic Congenital Adrenal Hyperplasia |
| **Study Number:** | Study SPR001-204 |
| **Investigational Drug:** | SPR001 (Tildacerfont) |
| **Indication:** | Treatment of Congenital Adrenal Hyperplasia |
| **Investigators:** | Multicenter |
| **IND Number:** | 131761 |
| **Registries** | ClinicalTrials.gov: NCT04544410<br>EudraCT #: 2019-004765-40 |
| **Sponsor:** | Spruce Biosciences, Inc.<br>2001 Junipero Serra Blvd, Suite 640<br>Daly City, CA 94014 |
| **Plan Version:** | Version 2.0, 24 September 2024 |
| **Plan Prepared by:** | Kevin J Carroll, PhD<br>Statistical Consultant<br>KJC Statistics Ireland Limited |

# STATISTICAL ANALYSIS PLAN

## Study SPR001-204

A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of SPR001 (Tildacerfont) in Reducing Supraphysiologic Glucocorticoid Use in Adult Subjects with Classic Congenital Adrenal Hyperplasia

Plan Version: Version 2.0 (24 September 2024)

Signature Page

Author: DocuSigned by: Kevin Carroll E45953F2TAC... 25-Sep-2024
Kevin J Carroll, PhD
Statistical Consultant
KJC Statistics Ireland Limited
Date

Reviewer: Signed by: Elizabeth Ludington E37918A4... 29-Sep-2024
Elizabeth Ludington, PhD Statistician
PharmaStat, LLC
Date

Reviewer: DocuSigned by: Will Charlton 25-Sep-2024
Will Charlton, MD
Chief Medical Officer
Spruce Biosciences, Inc.
Date

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

| Abbreviation | Description |
|---|---|
| 17-OHP | 17-hydroxyprogesterone |
| A4 | androstenedione |
| ACTH | adrenocorticotropic hormone, corticotropin |
| ADaM | analysis data model |
| AE | adverse event |
| AESI | adverse event of special interest |
| ALP | alkaline phosphatase |
| ALT | alanine aminotransferase |
| ANCOVA | analysis of covariance |
| ASCVD | atherosclerotic cardiovascular disease |
| BID | twice daily |
| BMD | bone mineral density |
| BMI | body mass index |
| BSA | body surface area |
| BSAP | bone-specific alkaline phosphatase |
| C-SSRS | Columbia–Suicide Severity Rating Scale |
| CAH | congenital adrenal hyperplasia |
| CDISC | Clinical Data Interchange Standards Consortium |
| CGI-I | Clinical Global Impression – Improvement Scale |
| CI | confidence interval |
| CM | concomitant medication |
| CSR | clinical study report |
| C-SSRS | Columbia–Suicide Severity Rating Scale |
| CTCAE | Common Terminology Criteria for Adverse Events |
| CTX-1 | C-terminal telopeptide of type I collagen |
| CV | Cardiovascular (end point) |
| CV | coefficient of variation (statistics) |
| DBP | diastolic blood pressure |
| DXA | dual-energy X-ray absorptiometry |
| ECG | electrocardiogram |
| eCRF | electronic case report form |
| EOT | end of treatment |

| Abbreviation | Description |
|---|---|
| fCP | Female of childbearing potential |
| GC | glucocorticoid |
| GGT | gamma-glutamyl transferase |
| HADS | Hospital Anxiety and Depression Scale |
| HbA1c | hemoglobin A1c |
| HC | hydrocortisone |
| HCe | hydrocortisone equivalent |
| HDL | high-density lipoprotein |
| HEENT | head, eyes, ears, neck, and throat |
| HOMA-IR | homeostatic model assessment of insulin resistance |
| HR | heart rate |
| ICF | informed consent form |
| ICH | International Council for Harmonisation |
| INR | international normalized ratio |
| ITT | Intent to Treat (Population) |
| LDL | low-density lipoprotein |
| LLD | lower limit of detection |
| LS | least squares |
| MedDRA | Medical Dictionary for Regulatory Activities |
| mITT | modified Intent to Treat (Population) |
| MMRM | Mixed Model Repeated Measures |
| PGIC | Patient Global Impression of Change |
| PI | Primary investigator |
| PK | pharmacokinetic |
| PP | Per Protocol (Population) |
| PT | preferred term (adverse event) |
| PT | prothrombin time (liver function) |
| PTT | partial thromboplastin time |
| Q | quartile |
| QD | once daily |
| QTcF | Fridericia-corrected QT interval |
| RTSM | randomization and trial supply management |
| SAE | serious adverse event |
| SAP | statistical analysis plan |

| Abbreviation | Description |
|---|---|
| SBP | systolic blood pressure |
| SD | standard deviation |
| SDTM | standard data tabulation model |
| SE | standard error |
| SF-36 | Short Form 36 |
| SOC | system organ class |
| TART | testicular adrenal rest tumor |
| TEAE | treatment-emergent adverse event |
| TID | three times daily |
| TLF | table, listing, figure |
| ULN | upper limit of normal |
| WC | waist circumference |
| WHO DD | World Health Organization Drug dictionary |

# 1 RELATED DOCUMENTS: PROTOCOL AND CASE REPORT FORMS

| Version | Date |
|---|---|
| Protocol Version 1.0 | 19DEC2019 |
| Protocol Version 2.0 | 30JAN2020 |
| Protocol Version 3.0 | 14FEB2020 |
| Protocol Version 4.0 | 14APR2020 |
| Protocol Version 5.0 | 31AUG2020 |
| Protocol Version 6.0 | 31MAR2021 |
| Protocol Version 7.0 | 16MAR2022 |
| Protocol Version 8.0 | 14APR2024 |
| Case Report Forms Version 5.0 | 15Mar2024 |

# 2 COMMITMENT TO GOOD STATISTICAL PRACTICE

## 2.1 Definition of Good Statistical Practice

The International Council for Harmonisation Guideline on Statistical Principles for Clinical Trials (ICH E9) implicitly defines good statistical practice. Good statistical practice includes both appropriate statistical designs to minimize bias and maximize precision of analysis plus operational excellence to assure credibility of results. The scientific design associated with any clinical trial is found in the protocol. More detailed, pre-specified statistical analysis methods can be found in the statistical analysis plan.

We interpret the operational side of good statistical practice as a transparent, reproducible, and validated approach to acquiring and analyzing clinical trial data. Reproducible research depends upon process transparency and also provides auditability of the statistical analysis. Analysis transparency requires that a navigable electronic process chain exists from defining the objective of the analysis to creating the results.

## 2.2 Use of Standards

Data standards are foundational for creating an environment where tools can be leveraged at different points in the analysis process. Data standards for clinical development of drugs have been defined and are maturing under various initiatives through the Clinical Data Interchange Standards Consortium (CDISC). Spruce Biosciences uses standard data tabulation model (SDTM) data sets and Analysis Data Model (ADaM) statistical analysis files for producing

analysis results. Other applicable standards include regulatory guidance from the Food and Drug Administration (FDA) and ICH:

- ICH Guideline on the Structure and Content of Clinical Study Reports (ICH E3)
- ICH Guideline for Good Clinical Practice (ICH E6)

## 3 PURPOSE OF THE ANALYSIS PLAN

Tildacerfont is an investigational drug product designed to treat classic congenital adrenal hyperplasia (CAH) and is being evaluated in clinical study SPR001-204. This statistical analysis plan (SAP) pre-specifies the statistical analysis methods for the primary analysis (Study Week 24 = 24 weeks of randomized, double blind treatment = Placebo-Controlled Treatment Period) and final analysis (Study Week 76 = 52 weeks of open label tildacerfont treatment after 24 weeks of randomized, double blind treatment = Open Label Treatment Period) to support completion of the SPR001-204 clinical study report (CSR). This SAP will be used to analyze the safety and efficacy data collected during the main study period (Day 1 [baseline] to Week 76). The analysis of the optional open label extension period (Weeks 76 through 316) will be covered in a separate SAP. The analysis of pharmacokinetics (data derivation and summary of individual pharmacokinetic [PK] parameters) is outside the scope of this document and is not addressed here. The planned analyses identified in this SAP may be included in regulatory submissions and/or future manuscripts.

The analysis methods described in this plan are considered *a priori*, in that they have been defined prior to clinical database lock and treatment unblinding. Exploratory analyses that are not defined in this SAP may be performed to support the clinical development program. Any post-hoc or unplanned analyses that are performed for the CSR, but not defined in this SAP, will be documented in the CSR. Changes from the planned analyses stated in the study protocol are described in Section 14. Should the SAP and the protocol be inconsistent with respect to any further planned analyses, the language of the SAP is governing.

## 4 STUDY DESIGN

This is a multi-center, randomized, double-blind, placebo-controlled, clinical study to evaluate the efficacy and safety of tildacerfont in reducing supraphysiologic glucocorticoid (GC) use in

adult subjects with classic CAH who have androstenedione (A4) ≤ 2.5x upper limit of normal (ULN) and are on supraphysiologic doses of GC therapy (≥ 30 mg/day and ≤ 60 mg/day in hydrocortisone equivalents [HCe]) at baseline. This study consists of the following periods:

- ≤ 45-day Screening Period will confirm subject eligibility
- Week -12 to either Week -6 or Week -2 Day -1 [±3 days] GC Conversion Period, if applicable
- 76-week, 2-part Treatment Period
    - 24-week Placebo-Controlled Treatment Period
    - 52-week Open-Label Period
- 240-week, optional Open-Label Extension Period
- 30-day Follow-up Period

Per sample size calculations, a total of 90 adult subjects with classic CAH were planned to be randomized into the 76-week Treatment Period. Subjects on dexamethasone at the initial Screening Visit who agree to convert to a non-dexamethasone regimen will enter the GC Conversion Period for conversion to and stabilization on a non-dexamethasone regimen as determined by their physician. Subjects enrolled under protocol amendment ≤ version 6 could also complete the GC Conversion Period to convert and stabilize to Sponsor supplied GC regimens, hydrocortisone (3 times daily) TID or prednisolone (twice daily) BID, if the subject was not on one of these regimens at screening. Subjects enrolled under and after protocol amendment 7, who are not on dexamethasone at screening, will continue on their current GC regimen at screening without the need for conversion to a standardized GC regimen prior to entering the treatment period. Permitted GC regimens include either monotherapy or combination therapy of the following GC types, hydrocortisone, hydrocortisone acetate, prednisone, prednisolone, or methylprednisolone.

Upon completion of the Screening period and GC Conversion Period, if applicable, subjects who continue to meet eligibility criteria will enter the randomized, Double-Blind, Placebo-Controlled Treatment Period. Subjects will be randomized in a 1:1 ratio to receive either tildacerfont at a dose of 200 mg QD (once a day) or matching placebo for the 24-week, double-blind part of the study. During this part of the study, a subject's daily GC regimen may be adjusted

(increases/decreases in total daily dose or redistribution of doses across the day) dependent on a protocol algorithm based on A4 assessments and clinical judgement.

Subjects who complete the Placebo-Controlled Treatment Period may continue into the Open-Label Period to receive 200 mg Tildacerfont QD and be eligible for additional GC dose adjustments for a duration of 52 weeks in the main study. For subjects randomized to tildacerfont, this will include up to 76 weeks of treatment with tildacerfont in the main study. For subjects randomized to placebo, this will include up to 52 weeks of treatment with tildacerfont in the main study. Subjects who continue into the optional Open-Label Extension Period will receive 200 mg Tildacerfont QD and be eligible for GC dose adjustments for a duration of up to 240 weeks. At the end of treatment (EOT), subjects will maintain the GC dose regimen established during the course of the study until the 30-day follow-up visit.

Clinical visits during the GC Conversion Period, Treatment Periods, optional Open-Label Extension Period, and Follow-up will include efficacy, biomarkers, and safety assessments. Figure 1 depicts study visits from the Screening Period through the Follow-up Period in the main study. Figure 2 depicts the study visits during the optional Open-Label Extension Period. Assessments and procedures for evaluation of safety, efficacy, and biomarkers will be conducted per the protocol-specified schedule (see Table 1, Table 2, Table 3, and Table 4).

## 4.1 Randomization and Blinding

Subjects will be centrally randomized in a 1:1 ratio to receive tildacerfont at 200 mg QD or matching placebo using a randomization and trial supply management (RTSM) system. Randomization will be stratified by sex (male and female) and baseline GC dose level (< 40 mg/day and ≥ 40 mg/day in HCe). Investigators, subjects, and the Sponsor will be blinded to assigned study drug for the duration of the Placebo-Controlled Treatment Period.

## 4.2 Study Treatment

### 4.2.1 Study Drug Administration

In the Placebo-Controlled Treatment Period, subjects will receive randomized study treatment (200 mg QD tildacerfont or matching placebo). During the 52-week Open-Label Period and optional 240 week Open-Label Extension Period, all subjects will receive 200 mg QD

tildacerfont. Study drug will be taken orally each day between 6 PM and midnight with an evening meal.

### 4.2.2 Glucocorticoid Regimen

Subjects entering the study under protocol versions 7.0 and higher will use a Principal Investigator-prescribed supply of GCs. Subjects who entered the study under prior protocol versions will continue to use the GCs provided by the Sponsor through Week 24. After Week 24, subjects will use a Principal Investigator-prescribed supply of GCs.

Subjects will continue to use the same type of GC (hydrocortisone, hydrocortisone acetate, prednisone, prednisolone, or methylprednisolone) throughout the study. Changes to GC type during study will be considered a protocol deviation.

#### 4.2.2.1 Glucocorticoid Changes

During the Treatment Period, per the protocol A4 algorithm (Protocol Section 4.1.4.3, reproduced below), subjects with an A4 measurement ≤ ULN at a study visit will reduce their daily GC dose. Each reduction of the daily GC dose will be by 5 mg HCe, according to the protocol defined conversion ratios of GC regimens summarized in Table 11, down to a minimum of 15 mg HCe per day (approximately physiologic replacement level). Subjects may reduce to a dose below 15 mg HCe with medical monitor approval. GC reductions may stop at a higher GC level (>15 mg) based on Primary Investigator (PI) judgment as physiologic replacement levels are dependent on each subject. In all cases, clinical judgement supersedes the protocol algorithm. GC reductions may begin at Week 2 (based on the Day 1 A4 measurement). Increases to GC dose may begin after Week 6 (based on the Week 6 A4 measurement).

**A4-based algorithm for daily GC dose reduction**

**A4 Level ≤ULN**

- Reduce GC dose
- GC dose reductions may begin on Week 2 (based on Day 1 measurement) and thereafter at Week 6, 12, 18, 24, 32, 40, 52, and 64 study visits, or during the optional Open-Label Extension Period

**A4 Level >1.25x ULN**

- Increase GC dose
- GC dose increases can occur at Week 6, 12, 18, 24, 32, 40, 52, and 64 study visits, or during the optional Open-Label Extension Period

**A4 Level >ULN and ≤1.25x ULN**

- Maintain GC Dose

### 4.2.3 Stress Dosing of Glucocorticoid

The Sponsor will provide a supply of oral HC tablets for periods of stress dosing. During times of clinically significant stress, subjects are advised to take stress doses of HC, under advisement of their physician. Subjects who are currently stress dosing should reschedule efficacy and safety assessments until the event that initiated the stress dosing has resolved (see Protocol Section 4.1.4.4). Stress dosing will be recorded by the subject in a patient diary and by the PI on the concomitant medication eCRF using the verbiage "stress dose/dosing" in the indication variable.

## 4.3 Assessments

Table 1, Table 2, Table 3, and Table 4 show the schedule of events for the study.

### 4.3.1 Efficacy Assessments

#### 4.3.1.1 Hormone Assessments

All blood samples for hormones and androgens will be collected at approximately 8 AM (7- 9 am window) prior to administration of a subject's morning GC dose. A4, 17-hydroxyprogesterone (17-OHP), adrenocorticotropic hormone (ACTH), and testosterone will be measured. Hormone assessments will be collected at all study visits except for Week 3 (safety laboratory assessments only).

#### 4.3.1.2 Glucocorticoid Dose Process and Documentation

Documentation of GC doses will be completed using eCRFs. If an adjustment in GC dose or change in distribution of daily doses is warranted, a new record will be recorded on the Prior and Concomitant Daily Glucocorticoids/Mineralocorticoids eCRF for each change.

At Weeks 24, 52 and 76 the eCRFs will record the total daily prescribed GC dose and will serve as the source for the Primary Efficacy Endpoint and GC-related change endpoints. For example, the recorded GC dose at Week 12 and the A4 assessment taken at Week 12 will be used to meet the Primary Efficacy Endpoint definition.

Based on the A4 assessment from each visit and aligning with the primary investigator's clinical judgement, the primary investigator will confirm if the GC dose level will be adjusted or maintained until the next clinic visit. For every visit with a potential adjustment in GC, the Key Glucocorticoid Dosing eCRF will record the A4 algorithm outcome (decrease, increase, maintain dose), whether the PI followed the algorithm's outcome (yes, no), reason for not following the outcome (clinical judgement, subject at physiologic dose), and the PI final decision (decrease, increase, maintain dose).

For subjects enrolled under protocol amendment 6 or earlier and prior to Week 24, the RTSM system will dispense a new GC regimen. For subjects enrolled under or after protocol amendment 7 or for visits after Week 24 for subjects enrolled under protocol amendment 6 or earlier, a new prescription will be written by the PI for the adjusted daily GC dose regimen and relayed to the subject to fill at their local pharmacy. An eCRF Glucocorticoid Medication – Received will record the date the subjects received/filled the new Sponsor-provided GC regimen and the day of first dose from the new regimen.

Reconciliation across eCRFs will be performed to ensure the accuracy of the GC dose data.

#### 4.3.1.3 Metabolic Assessments

Metabolic assessments include assessments of hypertension (systolic blood pressure [SBP] and diastolic blood pressure [DBP]), waist circumference, fasting assessments of insulin resistance (hemoglobin A1c [HbA1c], fasting glucose, fasting insulin, and homeostatic model assessment of insulin resistance [HOMA-IR]), assessments of dyslipidemia (total cholesterol, low-density lipoprotein [LDL] cholesterol, high-density lipoprotein [HDL] cholesterol, and triglycerides).

Fasting HOMA-IR, fasting glucose, fasting insulin, total cholesterol, and LDL cholesterol will be assessed as part of clinical laboratory. SBP, DBP, BMI, body weight, and WC will be assessed as part of vital signs.

#### 4.3.1.4 Dual-Energy X-ray Absorptiometry Scan

DXA is a whole-body scan that measures both total body composition (including the mass of fat, bones, and lean tissue) and bone mineral density (BMD) (spine t- and z-scores, hip t- and z-scores,). DXA scans will be performed at Day 1 (baseline), Week 24, Week 52, and Week 76 of the Treatment Period; Week 124, Week 172, Week 220, Week 268, and Week 316 of the optional, Open-Label Extension Period.

#### 4.3.1.5 Testicular Adrenal Rest Tumors

Complete scrotal ultrasounds will be obtained for male subjects to detect and to evaluate the size (volume) and number of TARTs. The initial scrotal ultrasound may be scheduled for any time before the first dose of study drug and will be considered the baseline measurement for the TART assessment. Only subjects with detectable TART(s) at baseline will be followed for TART changes at Week 24, and Week 76 of the Treatment Period; Week 124, Week 172, Week 220, Week 268, and Week 316 of the optional, Open-Label Extension Period.

#### 4.3.1.6 Quality of Life Assessments: Short Form 36

Quality of life will be measured using the SF-36, which will be administered on Day 1 (baseline), Week 24, and Week 76 of the Treatment Period; Week 124, Week 172, Week 220, Week 268, and Week 316 of the optional, Open-Label Extension Period.

### 4.3.2 Safety Assessments

Safety will be assessed by repeated clinical evaluations including adverse events (AEs), serious AEs (SAEs), AEs leading to discontinuation/withdrawal, AEs of special interest (AESIs), vital signs, electrocardiograms (ECGs), and clinical laboratory tests (serum chemistry, hematology, and urinalysis).

#### 4.3.2.1 Adverse Events

Investigators will collect information related to AEs throughout this clinical trial beginning at the signing of the first informed consent form (ICF) through the follow-up visit. Events that have not resolved by the end of the follow-up will be followed through resolution, if possible.

#### 4.3.2.2 Vital Signs

The following vital signs will be assessed after the subject has been sitting for approximately 5 minutes: blood pressure (systolic and diastolic; mmHg); pulse rate (beats per minute); respiration rate (breaths per minute); and body temperature (°C); weight (kg); waist circumference (cm). Vital signs will be obtained at Screening, and at every clinic visit during the Treatment Period and Open-Label Extension Period. Waist circumference is measured at Screening, Day 1 (baseline), Week 24, and Week 52, Week 76 of the Treatment Period, and every clinical visit of the Open-Label Extension Period.

#### 4.3.2.3 Physical Examinations

A full physical examination will be performed during the Screening Visit and at Day 1 (baseline), Week 24, and Week 76. Abbreviated physical examinations occur at Week 6, Week 12, Week 18, Week 32, Week 40, Week 52, Week 64 of the Treatment Period; Week 76, Week 124, Week 172, Week 220, Week 268, and Week 316 of the optional Open-Label Extension Period; and Follow-up visit.

Clinically significant adverse findings from the PE will be recorded on the AE eCRF.

#### 4.3.2.4 Electrocardiograms

All 12-lead ECG assessments will include: heart rate (HR), QRS, QT, and QTc intervals, using Fridericia's formula (QTcF). ECGs will be performed at the following visits: Day 1, Week 12, Week 24, Week 52, Week 76 of the Treatment Period; and Follow-up visit as shown in Table 3.

Clinically significant adverse findings from the ECG will be recorded on the AE eCRF.

#### 4.3.2.5 Clinical Laboratory and Urinalysis

Clinical laboratory assessments include hematology, clinical chemistry, coagulation, lipid panel, thyroid panel, and other hormones. Clinical lab assessments will be performed at all clinic visits.

Urinalysis and fasting assessments (hemoglobin A1c [HbA1c], HOMA-IR, fasting glucose, and fasting insulin) are collected at Day 1 (baseline), Week 24, Week 52, and Week 76 of the Treatment Period.

Clinically significant adverse findings from laboratory or urinalysis assessments will be recorded on the AE eCRF.

#### 4.3.2.6 Psychiatric Evaluations

##### 4.3.2.6.1 Columbia–Suicide Severity Rating Scale (C-SSRS)

The C-SSRS will be used during the study to monitor suicidal ideation and behavior. The Baseline/Screening Version of the C-SSRS, which assesses both lifetime history and history from the last 12 months, will be used at screening to determine subject eligibility. The Since Last Visit Version of the C-SSRS will be used at all subsequent visits specified: Day 1, Week 12, Week 24, Week 52, and Week 76 of the Treatment Period; Week 124, Week 172, Week 220, Week 268, and Week 316 of the optional, Open-Label Extension Period.

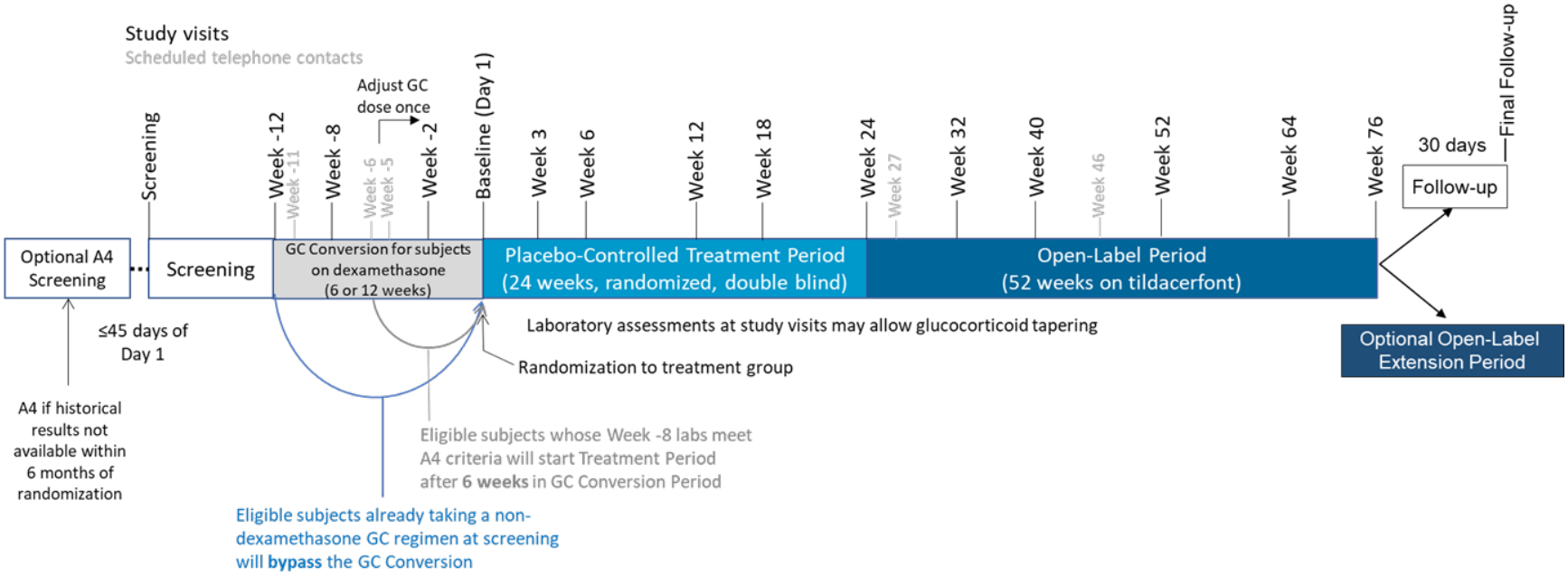
##### 4.3.2.6.2 Hospital Anxiety and Depression Scale (HADS)

Subject anxiety and depression will be monitored during the study using the HADS, assessed at Day 1, Week 12, Week 24, Week 52, and Week 76 of the Treatment Period.

### 4.3.3 Pharmacokinetic Measurements

Sparse plasma samples will be collected for evaluation of the pharmacokinetics of tildacerfont. A single plasma sample will be collected at approximately 8 am for measurement of tildacerfont concentration at each clinic visit (Week 3, Week 6, Week 12, Week 24, Week 52, Week 64, Week 76, and Follow-up) as indicated in Table 3.

**Figure 1 Schema of Study SPR001-204 Study Visits**




The 6- or 12-week GC conversion only applies to subjects on dexamethasone at the initial Screening Visit who agree to convert to a non-dexamethasone regimen as determined by their physician.

**Figure 2 Schema of Study SPR001-204 Optional Open-Label Extension Period**

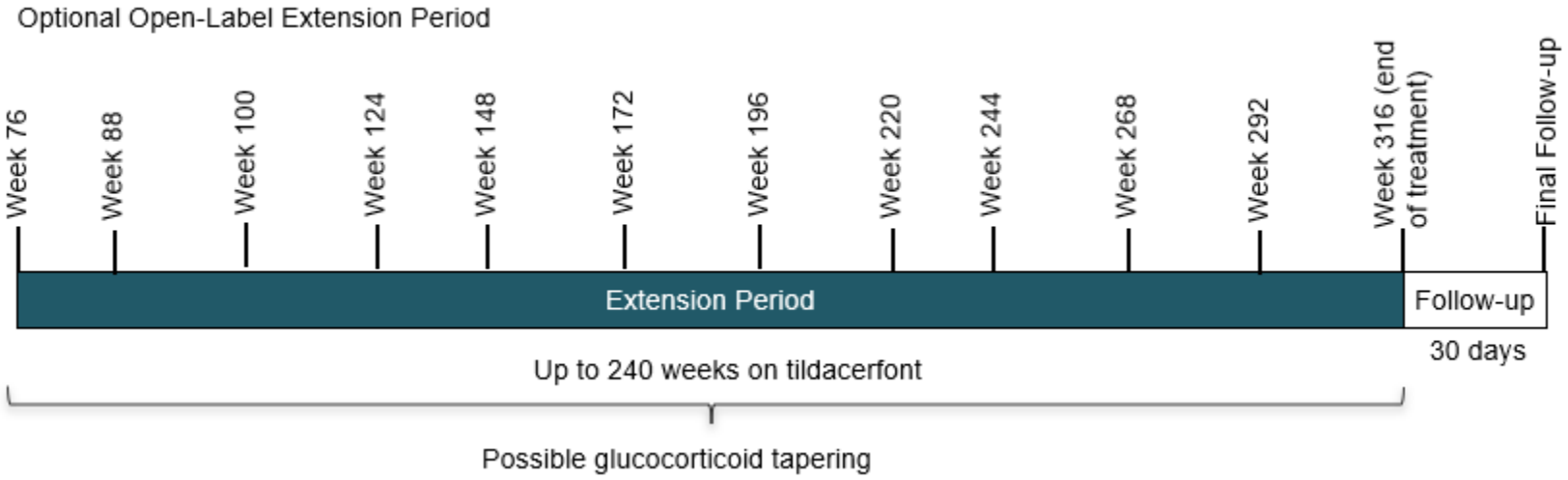


**Table 1 Study SPR001-204 Protocol Schedule of Events at Screening (Protocol Version 7.0 March, 2022)**

| | Screening Period | |
|---|---|---|
| **STUDY VISIT NUMBER** | **1a Optional A4 Screening** | **1b Screening** |
| | **In-Clinic Screening Visit** | **In-Clinic Screening Visit** |
| **STUDY DAY** | **≤45 days before Day 1** | **≤45 days before Day 1[1]** |
| Informed consent | X | X |
| Inclusion/exclusion criteria | | X |
| Demography | | X |
| Medical history | | X |
| As Prior medications from past year | | X |
| Concomitant medications | | X |
| Prior and current GC regimens[2] | | X |
| Vital signs[3], body weight | | X |
| Height | | X |
| Waist circumference | | X |
| Full physical examination | | X |
| C-SSRS | | X |
| HADS | | X |
| Hepatitis B & C and HIV tests | | X[4] |
| Urine drug screen | | X[4] |

[1] Screening information captured within 45 days of the start of Day 1 in this study (particularly screening information transferred from Spruce Biosciences Study SPR001-203) will be used to determine eligibility and fulfill screening requirements for this study.
[2] Subjects must be on a stable, supraphysiologic dose of GC replacement (defined as ≥30 mg/day and ≤60 mg/day in HCe) for ≥1 month before the Screening Visit. Information to be collected at screening about a subject's current and historical GC therapy during the past year include the type(s) of GC, the regimen(s), reason(s) that the subject is/was on a particular GC regimen, and any GC stress dosing during the past year.
[3] Vital signs consist of systolic and diastolic blood pressure, pulse rate, body temperature, and respiration rate and will be measured after a 5-min rest period.
[4] These clinic activities of the Screening Visit may be conducted remotely (at the subject's home) if approved by the site's IRB/EC.

| | Screening Period | |
|---|---|---|
| **STUDY VISIT NUMBER** | **1a**<br>**Optional A4 Screening** | **1b**<br>**Screening** |
| | **In-Clinic Screening Visit** | **In-Clinic Screening Visit** |
| **STUDY DAY** | **≤45 days before Day 1** | **≤45 days before Day 1**[1] |
| Serum pregnancy test for FCP | | X[4] |
| Hormones from blood | X[5] | X[4, 6] |
| Clinical laboratory[7] | | X[4] |
| Urinalysis | | X[4] |
| HbA1c | | X[4] |
| 12-lead ECG | | X[4] |
| Report subject's status to site | | X[4] |

17-OHP = 17 hydroxyprogesterone; A4 = androstenedione; ACTH = adrenocorticotropic hormone; ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; C SSRS = Columbia–Suicide Severity Rating Scale; EC = Ethics Committee; ECG = electrocardiogram; eGFR = estimated glomerular filtration rate; FCP = female of childbearing potential; FSH = follicle-stimulating hormone; GC = glucocorticoid; GGT = gamma-glutamyl transferase; HADS = Hospital Anxiety and Depression Scale; HbA1c = hemoglobin A1c; HCe = hydrocortisone equivalents; HIV = human immunodeficiency virus; INR = international normalized ratio; IRB = Institutional Review Board; LH = luteinizing hormone; PT = prothrombin time; PTT = partial thromboplastin time; SHBG = sex hormone–binding globulin

[5] A4 and 17-OHP will be measured. The blood draw for sample will be obtained at 8 AM ± 1 hour prior to a morning GC dose.
[6] A4, 17-OHP, ACTH, and testosterone will be measured prior to a morning GC dose.
[7] Clinical laboratory assessments include hematology, clinical chemistry (including liver function tests such as ALT, AST, ALP, GGT, total and direct bilirubin, and total bile acids), coagulation (PT/INR, and PTT), lipid panel, thyroid panel, LH, FSH, SHBG, renin, aldosterone, inhibin B for males only, and estradiol, prolactin, and progesterone for females only. eGFR for screening will be calculated from blood creatinine measured as part of screening clinical chemistry.

**Table 2 Study SPR001-204 Protocol Schedule of Glucocorticoid Conversion Period Activities**

| | **Glucocorticoid Conversion Period [1]** | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1st 6 weeks** | | | | **2nd 6 weeks, if GC dose adjustment necessary** | | |
| **STUDY VISIT NUMBER** | **2** | | **3** | | | | **4** |
| **STUDY WEEK** | **-12** | **-11** | **-8** | **-6** | | **-5** | **-2** |
| **STUDY DAY [2]** | **-84** | **-77** | **-56** | **-42** | **GC dose adjustment** | **-35** | **-14** |
| **Visit Window (days)** | | **±3** | **±3** | **±3** | | **±3** | **±3** |
| **Study Visit (V)/ Telephone Contact (T)** | **V** | **T** | **V** | **T** | | **T** | **V** |
| Urine pregnancy test for FCP | X | | | | | | |
| Hormones from blood [3,4] | X | | X | | | | X |
| Hematology [3] | | | X | | | | X |
| Clinical chemistry [3] | X [5] | | X | | | | X |
| Renin [3] | X | | X | | | | X |
| Vital signs [6], body weight | X | | X | | | | X |
| GC accountability and return of used GC to site | X | | X | | | | X |
| Report subject's status to site | X | | X | | | | X |
| Concomitant medications | X | X | X | X | | X | X |
| Monitor background GC dosing | X [7] | X | X | X | | X | X |
| Determine whether subject requires GC dose adjustment/2nd 6 weeks OR can proceed to Treatment Period | | | | X | | | |
| Background GC accountability | | | X | | | | X |

[1] If a subject fails to meet eligibility criteria or cannot tolerate the non-dexamethasone GC regimen during or at the end of the Glucocorticoid Conversion Period, the investigational site will schedule an abbreviated ET Visit to review AEs and concomitant medications and perform an abbreviated physical examination (preferably in the clinic but also permissible via home visit).

[2] Study days are numbered relative to Day 1 (Week 0), the first day of the Treatment Period, when the subject will take the first dose of study drug. There is no Day 0 (i.e., study days go directly from Day -1 to Day 1).

[3] Samples for these lab assessments will be obtained at 8 AM ± 1 hour, after an overnight fast (nothing to eat since the previous midnight) and before any morning dose of GC medication. On the evenings before scheduled laboratory assessments, subjects should take their evening GC dose before 10 PM.

[4] A4, 17-OHP, ACTH, testosterone, and background GC levels will be measured.

[5] Total bile acids will not be included in clinical chemistry at the Week -12 visit.

[6] Vital signs consist of systolic and diastolic blood pressure, pulse rate, body temperature, and respiration rate and will be measured after a 5-min rest period.

[7] On the first day of the Glucocorticoid Conversion Period, subjects will discontinue their dexamethasone GC regimen and begin taking a non-dexamethasone GC regimen.

| | **Glucocorticoid Conversion Period [1]** | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1st 6 weeks** | | | | **2nd 6 weeks, if GC dose adjustment necessary** | | |
| **STUDY VISIT NUMBER** | **2** | | **3** | | | | **4** |
| **STUDY WEEK** | **-12** | **-11** | **-8** | **-6** | | **-5** | **-2** |
| **STUDY DAY [2]** | **-84** | **-77** | **-56** | **-42** | **GC dose adjustment** | **-35** | **-14** |
| **Visit Window (days)** | | **±3** | **±3** | **±3** | | **±3** | **±3** |
| **Study Visit (V)/ Telephone Contact (T)** | **V** | **T** | **V** | **T** | | **T** | **V** |
| Dispense HC for stress dosing and perform accountability [8] | X | | X | | | | X |
| Review study diary/GC adherence [9] | X | X | X | X | | X | X |
| Review adverse events | X | X | X | X | | X | X |
| Physical examination (abbreviated) | X | | X | | | | X |

17-OHP = 17 hydroxyprogesterone; A4 = androstenedione; ACTH = adrenocorticotropic hormone; EDC = electronic data capture; ET = early termination; FCP = female of childbearing potential; GC = glucocorticoid; T = telephone contact; V = study visit

[8] Bottles of HC for stress dosing will be dispensed starting at the beginning of the Glucocorticoid Conversion Period and thereafter as needed to replace opened bottles. Accountability for bottles of HC will be performed at every study visit.

[9] Subjects will use an electronic study diary to document background GC adherence. Female subjects will also use the study diary to record menstrual information (including start and stop dates of menses).

**Table 3 Study SPR001-204 Protocol Schedule of Treatment Period and Follow-up Activities**

| | Treatment Period | | | | | | | | | | | | | Follow-up[1] | ET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Placebo-Controlled Treatment Period | | | | | | Open-Label Period | | | | | | | | |
| **STUDY VISIT NUMBER** | **5 Baseline** | **6** | **7** | **8** | **9** | **10** | | **11** | **12** | | **13** | **14** | **15** | **16** | |
| **STUDY WEEK** | **0** | **3** | **6** | **12** | **18** | **24** | **27** | **32** | **40** | **46** | **52** | **64** | **76** | **80** | |
| **STUDY DAY[2]** | **1** | **22** | **43** | **85** | **127** | **169** | **190** | **225** | **281** | **323** | **365** | **449** | **533** | **Last dose +30 days** | |
| **Study Visit (V) / Telephone Contact (T)[3]** | **V** | **V[4]** | **V** | **V** | **V** | **V** | **T** | **V** | **V** | **T** | **V** | **V** | **V** | **V** | **V** |
| Urine pregnancy test for FCP | X | | X | X | X | X | | X | X | | X | X | X | X | X |
| Hormones from blood[5,6] | X | | X | X | X | X | | X | X | | X | X | X | X | X |
| Clinical laboratory[5,7] | X | X | X | X | X | X | | X | X | | X | X | X | X | X |
| Urinalysis[5] | X | | | | | X | | | | | X | | X | | X |
| HbA1c, fasting glucose and insulin, HOMA-IR[5] | X | | | | | X | | | | | X | | X | | X |

[1] Subjects who do not continue to the optional Open-Label Extension Period upon completion of the Treatment Period will enter a 30-day Follow-up Period.
[2] All study visits and telephone contacts should be performed on the indicated study days. In cases where adherence to the foregoing schedule is not possible, all activities for study visits and telephone contacts must be completed within a +6-day window after the indicated study days. The day of the baseline visit (and first dose of study drug) will be considered Day 1 (Week 0), and all other study days are counted relative to Day 1. There is no Day 0 (ie, study days go directly from Day -1 to Day 1).
[3] During scheduled telephone contacts, sites will record any AEs and concomitant medications. After receiving A4 results during the Treatment Period, sites will make telephone contacts to applicable subjects regarding GC dose adjustment (detailed later in this Schedule of Activities). Subjects should be instructed to telephone sites if they have any concerns about their health. Telephone contacts initiated by sites and telephone contacts initiated by subjects should be captured in the EDC system as "unscheduled" telephone contacts.
[4] This visit for safety assessments may be scheduled for any day of Week 3, for any time of day, at the subject's convenience. Clinical laboratory can be measured at any time of day for this visit and need not occur before the morning dose of GC.
[5] Samples for these lab assessments will be obtained at 8 AM ± 1 hour, after an overnight fast (nothing to eat after midnight) and before any morning dose of GC medication. On the evenings before scheduled laboratory assessments, subjects should take their evening GC dose before 10 PM.
[6] A4, 17-OHP, ACTH, and testosterone will be measured.
[7] Clinical laboratory assessments include hematology, clinical chemistry (including liver function tests such as ALT, AST, ALP, GGT, total and direct bilirubin, and total bile acids), coagulation (PT/INR, and PTT), lipid panel, thyroid panel, LH, FSH, SHBG, renin, aldosterone, inhibin B for males only, and estradiol, prolactin, and progesterone for females only.

| | Treatment Period | | | | | | | | | | | | | Follow-up[1] | ET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Placebo-Controlled Treatment Period | | | | | | Open-Label Period | | | | | | | | |
| **STUDY VISIT NUMBER** | **5 Baseline** | **6** | **7** | **8** | **9** | **10** | | **11** | **12** | | **13** | **14** | **15** | **16** | |
| **STUDY WEEK** | **0** | **3** | **6** | **12** | **18** | **24** | **27** | **32** | **40** | **46** | **52** | **64** | **76** | **80** | |
| **STUDY DAY[2]** | **1** | **22** | **43** | **85** | **127** | **169** | **190** | **225** | **281** | **323** | **365** | **449** | **533** | **Last dose +30 days** | |
| **Study Visit (V) / Telephone Contact (T)[3]** | **V** | **V[4]** | **V** | **V** | **V** | **V** | **T** | **V** | **V** | **T** | **V** | **V** | **V** | **V** | **V** |
| BSAP, P1NP, CTX-1, U-NTx, Osteocalcin | X | | | | | X | | | | | X | | X | | X |
| PK[5,8] | | X | X | X | | X | | | | | X | | X | X | X |
| Genetic sample[9] | X | | | | | | | | | | | | | | |
| Vital signs[10], body weight | X | X | X | X | X | X | | X | X | | X | X | X | X | X |
| Waist circumference | X | | | | | X | | | | | X | | X | | X |
| 12-lead ECG | X | | | X | | X | | | | | X | | X | X | X |
| HADS | X | | | X | | X | | | | | X | | X | | X |
| SF-36 | X | | | | | X | | | | | | | X | | X |
| PGIC | | | | | | X | | | | | | | X | | X |
| Study treatment accountability and return of used treatment to site | X | X | X | X | X | X | | X | X | | X | X | X | | |
| Scrotal ultrasound for males | X[11] | | | | | X[12] | | | | | | | X[12] | | X[12] |
| DXA scan for body composition and bone mineral density[13] | X | | | | | X | | | | | X | | X | | X |

[8] A single blood sample will be drawn for PK measurement at each specified visit.
[9] Where local regulations permit and subject to discretionary approval from each site's IRB/EC and to subject consent, a voluntary blood sample may be collected for DNA analysis.
[10] Vital signs consist of systolic and diastolic blood pressure, pulse rate, body temperature, and respiration rate and will be measured after a 5-min rest period.
[11] The initial scrotal ultrasound may be scheduled for any time before the first dose of study drug and will be considered the baseline measurement for TART.
[12] Scrotal ultrasounds will only be conducted in subjects who had TART(s) at baseline.
[13] DXA scans should be performed for all subjects as indicated. The initial DXA scan may be scheduled for any time before the first day of study drug and will be

| | Treatment Period | | | | | | | | | | | | | Follow-up[1] | ET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Placebo-Controlled Treatment Period | | | | | | Open-Label Period | | | | | | | | |
| **STUDY VISIT NUMBER** | **5 Baseline** | **6** | **7** | **8** | **9** | **10** | | **11** | **12** | | **13** | **14** | **15** | **16** | |
| **STUDY WEEK** | **0** | **3** | **6** | **12** | **18** | **24** | **27** | **32** | **40** | **46** | **52** | **64** | **76** | **80** | |
| **STUDY DAY[2]** | **1** | **22** | **43** | **85** | **127** | **169** | **190** | **225** | **281** | **323** | **365** | **449** | **533** | **Last dose +30 days** | |
| **Study Visit (V) / Telephone Contact (T)[3]** | **V** | **V[4]** | **V** | **V** | **V** | **V** | **T** | **V** | **V** | **T** | **V** | **V** | **V** | **V** | **V** |
| Inclusion/exclusion criteria | X | | | | | | | | | | | | | | |
| Concomitant medications | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| C-SSRS | X | | | X | | X | | | | | X | | X | | X |
| CGI-I | | | | | | X | | | | | | | X | | X |
| Randomization to study drug | X | | | | | | | | | | | | | | |
| Dispense study drug[14] | X | | X | X | X | X | | X | X | | X | X | | | |
| Study drug accountability | | X | X | X | X | X | | X | X | | X | X | X | | X |
| Monitor background GC dosing | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| GC dose adjustment telephone contacts[15] | X | | X | X | X | X | | X | X | | X | X | | | |

considered the baseline measurement for the bone mineral density endpoint and body composition assessments. Subjects should not take calcium supplements within the 24 hours before a DXA scan.

[14] Either the subject will pick up study drug in the clinic or study drug will be shipped directly to the subject. Study drug will be taken daily between 6 PM and midnight, with an evening meal. The evening meal should contain <50% fat content. Study drug may be consumed up to 30 minutes after completing the evening meal, if necessary.

[15] During the Treatment Period, sites will make telephone contacts to subjects who are eligible for a GC dose adjustment within 2 weeks after each applicable study visit and based on their A4 level at that visit. Subjects who are not eligible for a GC change will not be telephoned. The site will direct the subject to adjust his/her GC dose by no more than 5 mg/day HCe increments if the subject has not experienced any change in clinical status since the previous study visit. If a subject's GC dose is adjusted, the site will telephone the subject again 7 days after the GC dose adjustment to assess for the occurrence of AEs and the initiation of new concomitant medications. GC reductions may begin at Week 2 (based on the Day 1 A4 measurement) and increases in GC dose may begin based on the A4 measurement at Week 6. GC dose may be reduced to a minimum of 15 mg HCe per day (approximately physiologic replacement level).

| | Treatment Period | | | | | | | | | | | | | Follow-up[1] | ET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Placebo-Controlled Treatment Period | | | | | | Open-Label Period | | | | | | | | |
| **STUDY VISIT NUMBER** | **5 Baseline** | **6** | **7** | **8** | **9** | **10** | | **11** | **12** | | **13** | **14** | **15** | **16** | |
| **STUDY WEEK** | **0** | **3** | **6** | **12** | **18** | **24** | **27** | **32** | **40** | **46** | **52** | **64** | **76** | **80** | |
| **STUDY DAY[2]** | **1** | **22** | **43** | **85** | **127** | **169** | **190** | **225** | **281** | **323** | **365** | **449** | **533** | **Last dose +30 days** | |
| **Study Visit (V) / Telephone Contact (T)[3]** | **V** | **V[4]** | **V** | **V** | **V** | **V** | **T** | **V** | **V** | **T** | **V** | **V** | **V** | **V** | **V** |
| Dispense HC for stress dosing and perform accountability[16] | X | X | X | X | X | X | | X | X | | X | X | X | X | X |
| Review study diary/drug adherence[17] | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Review adverse events | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Physical examination[18] | X | | X | X | X | X | | X | X | | X | X | X | X | X |

[16] Bottles of HC for stress dosing will be dispensed starting at Day 1 and thereafter as needed to replace opened bottles. Accountability for bottles of HC will be performed at every study visit.
[17] Subjects will use an electronic study diary to document study drug and background GC adherence. Female subjects will also use the study diary to record menstrual information (including start and stop dates of menses).
[18] A full physical examination will be conducted at Baseline (Day 1), Week 24, and Week 76. The full physical examination may exclude rectal, genitourinary, and breast exams. An abbreviated physical examination will be conducted at all other visits indicated. If the physical exam is performed at the subject's home, the qualified medical professional will report any changes to the subject's health to the site to determine whether further evaluation is needed via an unscheduled visit.

**Table 4 Study SPR001-204 Protocol Schedule of Optional Open-Label Extension Period Activities**

| | Extension Treatment Period | | | | | | | | | | | | Follow-up | ET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **STUDY VISIT NUMBER** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | |
| **STUDY WEEK** | **76** | **88** | **100** | **124** | **148** | **172** | **196** | **220** | **244** | **268** | **292** | **316** | **320** | |
| **STUDY DAY**[1] | **533** | **617** | **701** | **869** | **1037** | **1205** | **1373** | **1541** | **1709** | **1877** | **2046** | **2213** | **Last dose +30 days** | |
| **Study Visit (V)** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** |
| Urine pregnancy test for FCP | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Hormones from blood[2,3] | X | X | X | X | X | X | X | X | X | X | X | X | | X |
| Clinical laboratory[2] | X | X | X | X | X | X | X | X | X | X | X | X | | X |
| Vital signs[4], body weight, waist circumference | X | X | X | X | X | X | X | X | X | X | X | X | | X |
| SF-36 | | | | X | | X | | X | | X | | X | | X |
| C-SSRS | | | | X | | X | | X | | X | | X | | X |
| Scrotal ultrasound for males [5] | | | | X | | X | | X | | X | | X | | X[5] |
| DXA scan for body composition and bone mineral density[6] | | | | X | | X | | X | | X | | X | | X |
| Concomitant medications | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Review adverse events | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Abbreviated physical examination | X | | | X | | X | | X | | X | | X | | X |
| Dispense study drug[7] | X | X | X | X | X | X | X | X | X | X | X | X | | |
| Study drug accountability | X | X | X | X | X | X | X | X | X | X | X | X | | X |

[1] All study visits should be performed within a +30-day window after the indicated study days.
[2] Samples for these lab assessments will be obtained at 8 AM ± 1 hour, after an overnight fast (nothing to eat after midnight) and before any morning dose of GC medication. On the evenings before scheduled laboratory assessments, subjects should take their evening GC dose before 10 PM.
[3] A4 and 17-OHP and testosterone will be measured.
[4] Vital signs consist of systolic and diastolic blood pressure, pulse rate, body temperature, and respiration rate and will be measured after a 5-min rest period.
[5] Scrotal ultrasounds will only be conducted in subjects who had TART(s) at baseline.
[6] DXA scans should be performed for all subjects as indicated. Subjects should not take calcium supplements within the 24 hours before a DXA scan.
[7] Either the subject will pick up study drug in the clinic or the study drug will be shipped directly to the subject. Study drug will be taken daily between 6 PM and midnight, with an evening meal. The evening meal should contain <50% fat content. Study drug may be consumed up to 30 minutes after completing the evening meal, if necessary.

| | Extension Treatment Period | | | | | | | | | | | | Follow-up | ET |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **STUDY VISIT NUMBER** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | |
| **STUDY WEEK** | **76** | **88** | **100** | **124** | **148** | **172** | **196** | **220** | **244** | **268** | **292** | **316** | **320** | |
| **STUDY DAY**[1] | **533** | **617** | **701** | **869** | **1037** | **1205** | **1373** | **1541** | **1709** | **1877** | **2046** | **2213** | **Last dose +30 days** | |
| **Study Visit (V)** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** | **V** |
| GC dose adjustment telephone contacts[8] | X | X | X | X | X | X | X | X | X | X | X | | | |
| Dispense HC for stress dosing and perform accountability[9] | X | X | X | X | X | X | X | X | X | X | X | | | X |

17-OHP = 17 hydroxyprogesterone; A4 = androstenedione; ACTH = adrenocorticotropic hormone; ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BSAP = bone-specific alkaline phosphatase; CGI-I; Clinical Global Impression – Improvement Scale; C-SSRS = Columbia–Suicide Severity Rating Scale; CTX-1 = C-terminal telopeptide of type 1 collagen; DXA = dual-energy x-ray absorptiometry (scan); EC = Ethics Committee; ECG = electrocardiogram; EDC = electronic data capture; ET = early termination; FCP = female of childbearing potential; FSH = follicle-stimulating hormone; GC = glucocorticoid; GGT = gamma-glutamyltransferase; HADS = Hospital Anxiety and Depression Scale; HbA1c = hemoglobin A1c; HC(e) = hydrocortisone (equivalents); HOMA-IR = homeostatic model assessment of insulin resistance; INR = international normalized ratio; IRB = Institutional Review Board; LH = luteinizing hormone; LLD = lower limit of detection; P1NP = procollagen type 1 N terminal propeptide; PGIC = Patient Global Impression of Change; PK = pharmacokinetics; PRN = as needed; PROs = patient-reported outcome measures; PT = prothrombin time; PTT = partial thromboplastin time; SF-36 = Short Form 36; SHBG = sex hormone–binding globulin; T = telephone contact; TART = testicular adrenal rest tumor; ULN = upper limit of normal; U-NTx = urinary N-linked telopeptide of type 1 collagen; V = study visit

[8] During the Treatment Period, sites will make telephone contacts to subjects who are eligible for a GC dose adjustment within 2 weeks after each applicable study visit and based on their A4 level at that visit. Subjects who are not eligible for a GC change will not be telephoned. The site will direct the subject to adjust his/her GC dose by no more than 5 mg/day HCe increments if the subject has not experienced any change in clinical status since the previous study visit. If a subject's GC dose is adjusted, the site will telephone the subject again 7 days after the GC dose adjustment to assess for the occurrence of AEs and the initiation of new concomitant medications. GC dose may be reduced to a minimum of 15 mg HCe per day (approximately physiologic replacement level).
[9] Bottles of HC for stress dosing will be dispensed as needed. Accountability for bottles of HC will be performed at every study visit.

# 5 STUDY OBJECTIVES AND ENDPOINTS

The objectives of this study are to evaluate the efficacy and safety of tildacerfont in adult subjects with classic CAH who have A4 ≤ 2.5 x ULN and are on supraphysiologic doses of GC therapy (≥ 30 mg/day and ≤ 60 mg/day in HCe) at baseline.

This study will evaluate the potential of tildacerfont to reduce GC burden (total daily GC dose, total GC exposure over 24 weeks) in adult subjects with classic CAH. Study SPR001-204 will characterize clinical outcomes after 1) 24 weeks of double-blind treatment, 2) 52 weeks of open label treatment with tildacerfont and 3) 76 weeks of treatment with tildacerfont to support the clinical relevance of observed changes in GC dose.

The optional Open-Label Extension Period will provide additional characterization of up to 240 weeks of treatment with tildacerfont on GC dose changes and changes in clinical outcomes. See Table 5 for a summary of the study objectives and endpoints.

**Table 5 SPR001-204 Study Objectives and Endpoints**

| Objectives | Endpoints | | Analysis |
|---|---|---|---|
| ***1. Primary Efficacy*** | | | |
| To evaluate the mean absolute GC change in subjects with classic CAH over the 24-week, Double Blind, Placebo-Controlled Treatment Period | 1.1 | Absolute change from baseline in GC dose in HCe at Week 24 | Primary Analysis |
| ***2. Secondary Efficacy*** | | | |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 2.1 | Proportion of subjects with GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 2.2 | Proportion of subjects with baseline GC dose ≤ 35mg HCe who achieve GC dose ≤11 mg/m2/day in HCe and A4 ≤ 1.2x baseline or ≤ ULN at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing cardiovascular risk in subjects with CAH | 2.3 | Proportion of subjects with improvement in at least one cardiovascular risk factor at Week 24 | Primary Analysis |
| ***3. Exploratory Efficacy*** | | | |
| To evaluate the percentage change in GC use in subjects with CAH | 3.1 | Percent change from baseline in GC dose at Week 24 | Primary Analysis |

| Objectives | Endpoints | | Analysis |
|---|---|---|---|
| To evaluate the effect of tildacerfont in reducing the cumulative HCe dose in subjects with CAH | 3.2 | Change in total cumulative GC dose in HCe at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in improving HOMA-IR in subjects with CAH | 3.3 | Change from baseline in the HOMA-IR at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in improving HOMA-IR after 52 weeks of tildacerfont treatment in subjects with CAH | 3.4 | Change from baseline in HOMA-IR after 52 weeks of tildacerfont treatment | Final Analysis |
| To evaluate the effect of tildacerfont in reducing cardiovascular risk in subjects with CAH | 3.5 | Proportion of subjects with improvement in at least one cardiovascular risk factor at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 3.6 | Proportion of subjects with GC dose $\leq$ 11 mg/m$^2$/day and A4 $\leq$ 1.2x baseline or $\leqslant$ ULN at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 3.7 | Proportion of subjects with baseline GC dose $\leq$ 35mg HCe who achieve GC dose $\leq$11 mg/m2/day in HCe and A4 $\leq$ 1.2x baseline or $\leq$ ULN at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in improving quality of life in subjects with CAH | 3.8 | Change from baseline at Week 24 in the SF-36 domains and composite scores | Primary Analysis |
| To evaluate the effect of tildacerfont in improving quality of life in subjects with CAH | 3.9 | Change from baseline at Week 52 in the SF-36 domains and composite scores | Final Analysis |
| To evaluate the effect of tildacerfont on BMI after 24 weeks in subjects with CAH | 3.10 | Percent change from baseline in BMI at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont on BMI after 52 weeks of tildacerfont treatment in subjects with CAH | 3.11 | Percent change from baseline in BMI after 52 weeks of tildacerfont treatment | Final Analysis |
| To evaluate the effect of tildacerfont on waist circumference after 24 weeks of tildacerfont treatment in subjects with CAH | 3.12 | Change from baseline in waist circumference after 24 weeks of tildacerfont treatment | Primary Analysis |
| To evaluate the effect of tildacerfont on waist circumference after 52 weeks of tildacerfont treatment in subjects with CAH | 3.13 | Change from baseline in waist circumference after 52 weeks of tildacerfont treatment | Final Analysis |
| To evaluate the effect of tildacerfont in improving body composition after 24 weeks of tildacerfont treatment in subjects with CAH | 3.14 | Change from baseline in percent total fat mass and total percent lean mass after 24 weeks of tildacerfont treatment | Primary Analysis |
| To evaluate the effect of tildacerfont in improving body composition after 52 weeks of tildacerfont treatment in subjects with CAH | 3.15 | Change from baseline in percent total fat mass and total percent lean mass after 52 weeks of tildacerfont treatment | Final Analysis |
| To evaluate the effect of tildacerfont in improving BMD after 52 weeks of tildacerfont treatment in subjects with CAH | 3.16 | Change from baseline in BMD after 24 and 52 weeks of tildacerfont treatment | Primary (Week 24) and Final Analysis |

| Objectives | Endpoints | | Analysis |
|---|---|---|---|
| To evaluate the effect of tildacerfont in reducing TART(s) in male CAH subjects with TART(s) at baseline after 24 weeks of tildacerfont treatment in subjects with CAH | 3.17 | Reduction in TART volume at Week 24 in male subjects who had TART(s) at baseline | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing TART(s) in male CAH subjects with TART(s) at baseline after 52 weeks of tildacerfont treatment in subjects with CAH | 3.18 | Reduction in TART volume at Week 52 in male subjects who had TART(s) at baseline | Final Analysis |
| To evaluate the effect of tildacerfont in reducing GC use while maintaining normal A4 levels at Week 24 in subjects with CAH | 3.19 | Proportion of subjects with at least a 5 mg/day HCe reduction from baseline (Day 1) in GC dose and A4 ≤ 1.2x baseline or ⩽ ULN at week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing GC use while maintaining normal A4 levels at Week 52 in subjects with CAH | 3.20 | Proportion of subjects with at least a 5 mg/day HCe reduction from baseline (Day 1) in GC dose and A4 ≤ 1.2x baseline or ≤ ULN at week 52 | Final Analysis |
| ***4. Exploratory Efficacy (Optional Open-Label Extension Period) (Separate SAP)*** | | | |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels in subjects with CAH | 4.1 | Proportion of subjects with GC dose ≤ 11 mg/m$^2$/day in HCe and A4 ≤ ULN at EOT | Separate SAP |
| To evaluate the percentage change in GC use in subjects with CAH | 4.2 | Percent change from baseline in GC dose at EOT | Separate SAP |
| To evaluate the effect of tildacerfont in reducing cardiovascular risk in subjects with CAH and at least one cardiovascular risk factor at baseline | 4.3 | Proportion of subjects with improvement in at least one cardiovascular risk factor at EOT | Separate SAP |
| To evaluate the effect of tildacerfont in improving BMD in subjects with CAH | 4.4 | Change from baseline in BMD at EOT | Separate SAP |
| To evaluate the effect of tildacerfont in reducing TART(s) in male CAH subjects with TART(s) at baseline | 4.5 | Proportion of male subjects with reduction TART(s) at EOT who had TART(s) at baseline | Separate SAP |
| ***5. Safety*** | | | |
| To evaluate the safety of tildacerfont in subjects with CAH | 5.1 | AEs, SAEs | Primary Analysis |

## 5.1 Efficacy Objectives and Endpoints

### 5.1.1 Primary Efficacy Endpoint

The Primary Objective of the study is to evaluate the change in total daily GC dose in subjects with classic CAH over the 24-week, double blind, Placebo-Controlled Treatment Period. The Primary Efficacy Endpoint is the absolute change from baseline (Day 1) in GC dose in HCe at Week 24. The analysis of absolute change in total daily GC dose in HCe will include data from Weeks 3, 6, 12, 18 and 24 in a mixed model.

### 5.1.2 Secondary Efficacy Endpoints

The Secondary Efficacy Objectives and associated Secondary Efficacy Endpoints are provided in Table 5. The Secondary Efficacy Endpoints are further described in the following sections.

#### 5.1.2.1 GC use while maintaining androgen control at Week 24 with subjects with CAH

Secondary Efficacy Endpoint 2.1 evaluates the response to randomized treatment in terms of reducing GC to ≤ 11mg mg/m2/day in HCe while maintaining androgen control, defined as A4 levels at approximately baseline (≤ 1.2x baseline) or within normal limits (≤ ULN) at Week 24 in subjects with CAH.

Secondary Efficacy Endpoint 2.2 is similar to 2.1 but evaluated in those subjects with baseline GC ⩽35 mg/m2/day in HCe.

#### 5.1.2.2 Improving Cardiovascular Risk Factors at Week 24

Secondary Efficacy Endpoint 2.3 evaluates the effect of randomized treatment in reducing CV risk factors in classic CAH subjects at Week 24. The Secondary Efficacy Endpoint is the proportion of subjects who had improvement in at least one CV risk factor at Week 24. Only subjects CV risk factors at baseline will be included in the endpoint analysis.

### 5.1.3 Exploratory Efficacy Endpoints

#### 5.1.3.1 Percent Change in HCe Dose at Week 24

Exploratory Efficacy Endpoint 3.1 evaluates the effect of randomized treatment in reducing GC HCe dose in classic CAH subjects at Week 24 in mean percentage terms. This analysis will be achieved by examination of the change in daily GC HCe dose from baseline (Day 1) to post baseline visits on the log scale. Results will be back transformed to present the outcome as a percentage change.

#### 5.1.3.2 Cumulative HCe Dose at Week 24

Exploratory Efficacy Endpoint 3.2 evaluates the effect of randomized treatment in reducing the cumulative HCe dose (total exposure in HCe mg) in classic CAH subjects over 24 Weeks. The endpoint is the cumulative GC dose in HCe from baseline (Day 1) at Week 24. Similar to endpoint 3.1, the endpoint will be analyzed on the log scale.

#### 5.1.3.3 Change in HOMA-IR at Weeks 24 and 52

Exploratory Efficacy Endpoints 3.3 and 3.4 evaluate the effect of randomized treatment in improving HOMA-IR in classic CAH subjects at Weeks 24 and 52. The associated endpoints are the change in fasting HOMA-IR from baseline (Day 1) to Week 24 and Week 52.

#### 5.1.3.4 Improving Cardiovascular Risk Factors at Week 52

Exploratory Efficacy Endpoint 3.5 evaluates the effect of randomized treatment in reducing CV risk factors in classic CAH subjects at Week 52. The endpoint is similar to Secondary Efficacy Endpoint 2.3, being the proportion of subjects who had improvement in at least one CV risk factor at Week 52. Only subjects CV risk factors at baseline will be included in the endpoint analysis.

#### 5.1.3.5 GC use while maintaining androgen control at Week 52

Exploratory Efficacy Endpoints 3.6 and 3.7 are similar to the Secondary Efficacy Endpoints 2.1 and 2.2, but now assessed at 52 weeks. Secondary Efficacy Endpoints 3.6 is the response to randomized treatment in terms of reducing GC to ≤ 11mg mg/m2/day in HCe while maintaining

androgen control, defined as A4 levels at approximately baseline (≤ 1.2x baseline) or within normal limits (≤ ULN) at Week 52 in subjects with CAH.

Exploratory Efficacy Endpoint 3.7 is similar to 3.6 but evaluated in those subjects with baseline GC ⩽35 mg/m2/day in HCe.

#### 5.1.3.6 Quality of Life at Week 24 and Week 52

Exploratory Efficacy Endpoints 3.8 and 3.9 evaluate the effect of randomized treatment in improving the quality of life as quantified by the Short Form-36 total score in classic CAH subjects at Weeks 24 and 52. The endpoints are the change in the SF-36(8 subscales using the population-based normalized values, as well as PCS and MCS population-based normalized values) from baseline (Day 1) at Week 24 and Week 52.

#### 5.1.3.7 Change in BMI at Week 24 and Week 52

Exploratory Efficacy Endpoints 3.10 and 3.11 evaluate the effect of randomized treatment on BMI in classic CAH subjects. The Secondary Efficacy Endpoint is the percent change in BMI from baseline (Day 1) to Week 24 and Week 52.

#### 5.1.3.8 Waist Circumference at Week 24 and Week 52

Exploratory Efficacy Endpoints 3.12 and 3.13 evaluate the effect of randomized treatment in decreasing waist circumference in classic CAH subjects at Weeks 24 and 52. The endpoints are the change in waist circumference from baseline (Day 1) to Week 24 and Week 52.

#### 5.1.3.9 Change in Body Composition at Week 24 and Week 52

Exploratory Efficacy Endpoints 3.14 and 3.15 evaluate the effect of randomized treatment in improving percent total fat mass and total percent lean mass in classic CAH subjects at Week 24 and Week 52. The associated endpoints are the change in percent total fat mass and total percent lean mass from baseline (Day 1) to Week 24 and Week 52.

#### 5.1.3.10 Change in Bone Mineral Density at Week 24 and Week 52

Exploratory Efficacy Endpoint 3.16 evaluates the effect of randomized treatment in improving BMD in classic CAH subjects at Week 52. The endpoint is the change in total BMD from baseline (Day 1) to Week 24 and to Week 52.

#### 5.1.3.11 Change in TART volume at Week 24 and Week 52

Change in TART volume is assessed by Exploratory Efficacy Endpoints 3.17 and 3.18. These endpoints evaluate the effect of randomized treatment in reducing TART volume in classic CAH subjects at Week 24 and Week 52. Both endpoints will only include those subjects whom had TART(s) at baseline. The associated endpoints are thus (i) the change in total TART volume from baseline (Day 1) to Week 24 in subjects whom had TART(s) at baseline and (ii) the change in total TART volume from baseline (Day 1) to Week 52 in subjects whom had TART(s) at baseline.

#### 5.1.3.12 GC use while maintaining androgen control at Week 24 and Week 52 in subjects with CAH

To further evaluate the effect of randomized treatment in reducing GC use while maintaining androgen control, defined as normal A4 levels at Week 24 in subjects with CAH, Exploratory Efficacy Endpoints 3.19 and 3.20 evaluate response to randomized treatment in terms of the proportion of subjects with at least a 5 mg/day HCe reduction from baseline (Day 1) in GC dose and A4 $\leqslant$1.2x baseline or A4 $\leqslant$ ULN at (i) Week 24 and (ii) Week 52.

### 5.2 Safety Objective

The safety objective of the study is to evaluate the safety of tildacerfont in classic CAH patients using safety assessments described in Section 4.3.1.6.

## 6 SAMPLE SIZE AND POWER

The original sample size calculation was based the binary endpoint of percent of subjects with a 5 mg/day HCe reduction from baseline in GC dose at Week 24 with A4 $\leq$ ULN. Assuming a

placebo group response of 30%, N = 45 subjects per group would provide at least 90% power to detect a between group relative difference of 33% at the 0.05 two-sided type I error rate.

Additionally, it was computed that a sample size of N=45 subjects per group would provide at least 90% power to detect a mean change in daily GC dose from baseline and placebo of 7.25 mg/day in HCe, assuming a standard deviation (SD) of 10.5 mg/day HCe and the two-sided type I error is 0.05. Examination of group blinded data, provides a conservative SD estimate of 6.9 mg/day. On this basis, a sample size of N=40 subjects per group provides at least 90% power to detect a mean change in daily GC dose from baseline and placebo of 5 mg/day in HCe with a two-sided type I error of 0.05.

# 7 ANALYSIS SETS

## 7.1 Intent-To-Treat Analysis Set

The Intent-To-Treat (ITT) Analysis Set will include all randomized subjects (based on the RTSM system as described in Section 4.1) regardless of Treatment Period eligibility or completion. The ITT Analysis Set will be the basis for demographics, baseline characteristics, and efficacy.

## 7.2 Modified Intent-To-Treat Analysis Set

The modified ITT (mITT) Analysis Set will include all randomized subjects who receive at least 1 dose of study drug (tildacerfont or placebo) and have a baseline A4 assessment. The mITT population will be used for supportive analyses of the Primary and Secondary Efficacy Endpoints if it differs from the ITT population.

## 7.3 Per-Protocol Analysis Set

The Per-Protocol (PP) Analysis Set for each part of the Treatment Period will include all eligible subjects who have no major protocol deviations that would affect the analysis of efficacy data, defined prior to the database lock as the following:

- Subject did not meet efficacy-based inclusion criteria (hormone levels, GC regimen)

- Subject met efficacy-based exclusion criteria (inadequate adherence to GC regimen, night-shift worker)
- Subject's adherence to study drug or the subject-specific GC regimen, determined by pill count and diary data compliance, was < 80% over the Treatment Period
- Subjects having laboratory data specifically hormone levels during stress dosing periods instead of waiting 5 days per protocol.

The PP population will be used for supportive analyses of the Primary and Secondary Efficacy Endpoints only.

### 7.4 Safety Analysis Set

The Safety (SAF) Analysis Set is defined as all subjects who received at least one dose of study drug (placebo or tildacerfont). The safety analysis set, which will be based on the actual treatment received, will be used for evaluation of general medical history, study drug exposure, concomitant medication, and safety. Placebo-treated subjects who received any amount of tildacerfont during the Open-Label Period will be assigned to the tildacerfont treatment group from the time when tildacerfont was received.

### 7.5 DXA Analysis Set

The dual-energy X-ray absorptiometry (DXA) analysis set is defined as all ITT subjects who have a baseline and post-baseline evaluable scan. The DXA analysis set will be used to analyze bone mineral density efficacy assessments.

## 8 GENERAL CONSIDERATIONS

Data summarization and presentation conventions are documented in the Appendix A for mock-shells. All analyses and summaries will be produced using SAS® version 9.4 or higher.

### 8.1 Presentation of Summary Statistics

For most summary statistics, data will be displayed in tabular format. Unless otherwise specified, continuous adrenal biomarkers (A4, 17-OHP, ACTH) will be summarized using a 11-point descriptive statistics (i.e., n, mean, standard deviation [SD], median, 25% quartile [Q1], 75% quartile [Q3], minimum, maximum, geometric mean, geometric coefficient of variance [CV%],

95% confidence interval [CI] for geometric mean [including geometric mean ratio and its 95% CI]). All other continuous variables will be summarized using an 8-point descriptive summary (n, mean, SD, median, Q1, Q3, minimum, and maximum). The same number of decimal places as in the observed value will be presented when reporting minimum and maximum; 1 additional decimal place than in the observed value will be presented when reporting mean, median, Q1, Q3, geometric mean, 95% CI; 2 additional decimal places than in the observed value will be presented when reporting SD. Geometric CV% will be reported to one decimal place.

All categorical/qualitative data will be presented using the frequency of events and percentages. All percentages will be presented to 1 decimal place, unless otherwise specified. Percentages equal to 100 will be presented as 100% and percentages will not be presented for zero frequencies. Where individual variable values are missing, categorical data will be summarized based on reduced denominators (i.e., only subjects with available data will be included in the denominators).

For summaries of AEs and concomitant medications (CMs), the percentages will be based on the number of subjects who received study drug.

## 8.2 Presentation of p-values

Results of statistical analyses will be reported using summary tables, listings, and figures (TLFs). The ICH of Technical Requirements for Pharmaceuticals for Human Use numbering convention will be used for all TLFs. The following conventions will be followed:

Unless otherwise noted, all statistical testing will be two-sided and will be performed at the 0.05 significance level. Tests will be declared statistically significant if the calculated p-value is $< 0.05$, unless otherwise specified. P-values will be reported with 4 significant digits except when reporting p-values less than 0.001, reported as $< 0.001$.

## 8.3 Definitions and Derived Variables

### 8.3.1 Screened and Enrolled Subjects

#### 8.3.1.1 Screened Subjects

Subjects who signed an informed consent form are considered Screened Subjects

#### 8.3.1.2 Enrolled Subjects

Subjects who received at least one dose of Sponsor-provided standardized GC regimen or Principal investigator-prescribed supply of GCs or were randomized.

### 8.3.2 Study Day

As treatment will not begin until randomization into the Treatment Period, Treatment Study Day 1 is defined as the first double-blinded dose date in the Placebo-Controlled Treatment Period. Study Day, which follows the CDISC SDTM standard, is defined as (Assessment date – date of first study drug dosing + 1 day), where the assessment date is on or after the date of first study drug dosing; (Assessment date – date of first study drug dosing), where the assessment date is before the date of first study drug dosing.

### 8.3.3 Enrollment Day

In this study, subjects may receive a Sponsor-provided, standardized GC regimen or Principal Investigator-prescribed supply of GCs during the 6- to 12-week Glucocorticoid Conversion Period. Enrollment Day 1 will be the first dose date of the GC regimen in the Glucocorticoid Conversion Period, following the same format as study day. For subjects who do not require the GC conversion period, enrollment day will be the same as Study Day 1

### 8.3.4 Tildacerfont Treatment Day

Tildacerfont Treatment Day 1 is defined as the first dose date of tildacerfont in the study, following the same format as study day. For subjects randomized into the tildacerfont treatment group in the Placebo-Controlled Treatment Period, the first dose date of tildacerfont is defined as the first double-blinded dose date in the Placebo-Controlled Treatment period. For subjects

randomized into the placebo treatment group in the Placebo-Controlled Treatment Period, the first dose date of tildacerfont is defined as the first dose date in the Open-Label Period. Tildacerfont Treatment Day will be included in select efficacy listings.

### 8.3.5 End of Study Treatment Definition

A subject is considered to have completed study treatment if the subject has completed all 76 weeks of the Treatment Period.

### 8.3.6 End of Study Definition

A subject is considered to have completed the study if the subject has completed study treatment and the follow-up visit.

### 8.3.7 Body Mass Index

Body mass index ($kg/m^2$) is derived as weight (kg) / [height (m) × height (m)]. Height measured at Screening will be used to calculate BSA. Body mass index may be summarized using the classifications defined in Table 6.

**Table 6 Classification of Body Mass Index**

| Parameter | BMI ($kg/m^2$) |
|---|---|
| Underweight | < 18.8 |
| Normal Range | 18.5 -< 25 |
| Overweight | 25 -< 30 |
| Obese Class I | 30 -< 35 |
| Obese Class II | 35 - < 40 |
| Obese Class III | ≥ 40 |

### 8.3.8 Body Surface area

Body surface area (BSA) is calculated per the Mosteller formula as sqrt((height (cm) × weight (kg))/3600). Height measured at Screening will be used in all BSA calculations.

### 8.3.9 Baseline Values

#### 8.3.9.1 Baseline at Randomization

Baseline at randomization values are defined as the last non-missing assessment prior to the first dose of randomized, double-blind study drug in Placebo-Controlled Treatment Period of the study. Baseline at randomization will be used in the Primary Efficacy Endpoint, Secondary Efficacy Endpoint, and exploratory endpoint analyses that explore treatment at Week 24. Demographics and summaries of change from baseline will use baseline at randomization unless otherwise specified.

#### 8.3.9.2 Baseline Prior to Treatment with Tildacerfont

Baseline values prior to treatment with tildacerfont are defined as the last non-missing assessment prior to the first dose of tildacerfont. For subjects randomized into the tildacerfont treatment group of the Placebo-Controlled Treatment Period, the first dose of tildacerfont is defined as the first double-blinded dose date in the Placebo-Controlled Treatment period. For subjects randomized into the placebo treatment group of the Placebo-Controlled Treatment Period, the first dose of tildacerfont is defined as the first dose in the Open-Label Period. The baseline prior to treatment with tildacerfont will be used in Secondary Efficacy and Exploratory Efficacy Endpoint analyses that explore effects after 52 weeks of tildacerfont treatment.

### 8.3.10 Change from Baseline

Change from baseline is calculated as the post-baseline assessment subtracted by the baseline assessment.

### 8.3.11 Percent Change from Baseline

Percent change from baseline is calculated as the change from baseline divided by the baseline assessment and multiplied by 100%. This value will be rounded to two decimal places unless otherwise stated.

### 8.3.12 Upper Limit of Normal and Targets by Efficacy Biomarker

The ULN and target values that are displayed in Table 7 will be used in data analysis. The age at screening will be used to determine those age specific ULN values.

**Table 7 Upper Limit of Normal Values by Subgroup and Key Biomarker**

| Baseline Characteristic | ACTH ULN | 17-OHP ULN | 17-OHP Target | A4 ULN | A4 Target (1.25xULN) |
|---|---|---|---|---|---|
| Males: 18-30 | 63.3 ng/dL | 200 ng/dL | 1200 ng/dL | 220 ng/dL | 275 ng/dL |
| Males: 31-50 | 63.3 ng/dL | 200 ng/dL | 1200 ng/dL | 190 ng/dL | 238 ng/dL |
| Males: 51+ | 63.3 ng/dL | 200 ng/dL | 1200 ng/dL | 220 ng/dL | 275 ng/dL |
| Females: pre-menopausal | 63.3 ng/dL | 200 ng/dL | 1200 ng/dL | 230 ng/dL | 288 ng/dL |
| Females: post-menopausal | 63.3 ng/dL | 200 ng/dL | 1200 ng/dL | 75 ng/dL | 94 ng/dL |

### 8.3.13 Cardiovascular Risk Factors

CV risk factors are defined by gender in the table below. Subjects with values exceeding the threshold are considered for being at elevated risk for a CV event.

**Table 8 CV Risk Factors by Gender**

| Risk Factor | Female | Male |
|---|---|---|
| HOMA-IR[1,2,3] | ≥ 2.5 | ≥ 2.5 |
| Waist circumference[1,4,5] | ≥ 88 cm | ≥ 102 cm |
| Total cholesterol[6] | ≥ 170 mg/dL | ≥ 200 mg/dL |
| LDL cholesterol[6] | ≥ 110 mg/dL | ≥ 130 mg/dL |
| Systolic blood pressure[7] | ⩾ 130 mmHg | ⩾ 130 mmHg |
| Diastolic blood pressure[7] | ≥ 70 mmHg | ≥ 70 mmHg |
| BMI[8] | ≥ 30 kg/m$^2$<br>OR<br>≥ 27 kg/m$^2$ and at least one other risk factor | ≥ 30 kg/m$^2$<br>OR<br>≥ 27 kg/m$^2$ and at least one other risk factor |

1 (Torkey, et al., 2021); 2 (Matthews, et al., 1985); 3 (Muniyappa, Lee, Chen, & MJ, 2008); 4 (Flint, et al., 2010); 5 (WHO, 2008); 6 (Grundy, et al., 2019); 7 (Whelton, et al., 2018); 8 (FDA, 2007)

### 8.3.14 Responder Definitions

Where Efficacy Endpoints are proportions of responders based on serum hormone assessments, GC dose, and/or CV risk factors, indicator variables will be created as Table 9.

**Table 9 Responder Definition by Endpoint**

| Endpoint | Definition |
|---|---|
| Proportion of subjects with GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN at Week 24 | $1 = Week\ 24\ GC\ dose \leq 11mg/m2/day\ in\ HCe$ AND ($A4 \leq$ 1.2x baseline or A4 $\leq ULN$)<br>$0 = otherwise$ |
| Proportion of subjects with baseline GC dose ≤35mg HCe who achieve GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN at Week 24 | $1 = B$aseline GC dose $\leq$ 35mg HCe AND Week 24 $GC\ dose \leq 11mg/m2/day\ in\ HCe$ AND ($A4 \leq$ 1.2x baseline or A4 $\leq ULN$)<br>$0 = otherwise$ |
| Proportion of subjects with GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN at Week 52 | $1 = Week\ 52\ GC\ dose \leq 11mg/m2/day\ in\ HCe$ AND ($A4 \leq$ 1.2x baseline or A4 $\leq ULN$)<br>$0 = otherwise$ |
| Proportion of subjects with baseline GC dose ≤35mg HCe who achieve GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN at Week 52 | $1 = B$aseline GC dose $\leq$ 35mg HCe AND Week 52 $GC\ dose \leq 11mg/m2/day\ in\ HCe$ AND ($A4 \leq$ 1.2x baseline or A4 $\leq ULN$)<br>$0 = otherwise$ |
| Proportion of subjects with at least a 5 mg/day HCe reduction from baseline (Day 1) in GC dose and A4 ≤ 1.2x baseline or ≤ ULN at week 24 | $1 = Week\ 24\ GC\ dose\ reduction\ from\ baseline \leq -5/day\ in\ HCe$ AND ($A4 \leq$ 1.2x baseline or A4 $\leq ULN$)<br>$0 = otherwise$ |
| Proportion of subjects with at least a 5 mg/day HCe reduction from baseline (Day 1) in GC dose and A4 ≤ 1.2x baseline or ≤ ULN at week 52 | $1 = Week\ 52\ GC\ dose\ reduction\ from\ baseline \leq -5/day\ in\ HCe$ AND ($A4 \leq$ 1.2x baseline or A4 $\leq ULN$)<br>$0 = otherwise$ |
| Proportion of subjects with ≥ 1 CV risk factor at baseline achieving an improvement in at least one CV risk factor at Week 24 | $\begin{cases} 1 & improvement\ in \geq 1\ baseline\ CV\ risk\ factor\ at\ Week \\ 0 & no\ improvement \end{cases}$ |
| Proportion of subjects with ≥ 1 CV risk factor at baseline achieving an improvement in at least one CV risk factor at Week 52 | $\begin{cases} 1 & improvement\ in \geq 1\ baseline\ CV\ risk\ factor\ at\ Week \\ 0 & no\ improvement \end{cases}$ |

### 8.3.15 Quality of Life Assessments and Scoring

#### 8.3.15.1 HADS Anxiety and Depression Scores

The scale consists of 14 items, 7 items each for anxiety and depression. The anxiety score is the sum of the odd-numbered items (questions 1, 3, 5, 7, 9, 11, 13), and the depression score is the sum of the even-numbered items (questions 2, 4, 6, 8, 10, 12, 14). Each item is rated on a 4-point scale based on the frequency of symptoms over the preceding week and ranging from 0 (not at all) to 3 (very often). The sub-scores are the sum of individual item scores within each subcategory (anxiety or depression) and ranges from 0 to 21, with higher scores corresponding to higher levels of anxiety or depression. The total score of the individual item scores and ranges from 0 to 42, with higher scores corresponding to higher levels of anxiety or depression.

#### 8.3.15.2 SF-36 Domain Scores and Component Scores

The SF-36 consists of 36 items in the following 8 domains: physical functioning, physical role functioning (limitations in usual role activities because of physical health problems), bodily pain, general health perceptions, vitality, social functioning, emotional role functioning (limitations in usual role activities because of emotional problems), and mental health; these domain scores are abbreviated PF, RP, BP, GH, VT, SF, RE, MH respectively. Domain scores range from 0 to 100, with higher scores corresponding to better subjective health status; these are then normalized to scores obtained in the 2009-general population, such that scores above 50 are higher than the general population and scores less than 50 are lower than those for the general population. In addition, and overall component score representing physical functioning (PCS) and mental functioning (MCS) are calculated, which are also normalized to the general population. The SF-36 scores will be calculated and provided by Medpace Reference Laboratories.

#### 8.3.15.3 Patient Global Impression of Change and Clinical Global Impression – Improvement (PGIC)

The PGIC is a 1-question survey that asks subjects to evaluate whether there has been an improvement in overall subjective health status. Subjects select a response on a 7-point Likert scale. The CGI-I is a 1-question survey that requires the clinician to assess how much a patient's

illness has improved or worsened relative to a baseline state. Clinicians select a response on a 7-point Likert scale.

### 8.3.16 Homeostatic Model Assessment of Insulin Resistance

HOMA-IR is calculated using the following equations:

$$HOMA\text{-}IR = \frac{glucose\ (mg/L) * insulin(IU/L)}{405}\ or\ \frac{glucose\ (mmol/L) * insulin(IU/L)}{22.5}$$

### 8.3.17 Testicular Adrenal Rest Tumor Volume

For each TART assessment, the total number of TARTs will be captured as well as the size (lesion volume) of each TART individually. The lesion volume for each detectable TART will be calculated as follows using the formula for an ellipsoid (length × width × thickness × 0.52). The total lesion volume for all detectible TARTs will be the endpoint for analysis.

### 8.3.18 Study Treatment Assignment to Data Records

This study consists of a GC Conversion Period, a two-part Treatment Period (placebo-controlled period and open-label period), and an optional Open-Label Extension Period. For a data record, the planned treatment will be based on a specific study period as follows:

**Table 10 Treatment Assignment by Period**

| Period | Treatment Assignment |
|---|---|
| GC Conversion Period: | None |
| Placebo Controlled Treatment Period: | Randomized treatment assignment (placebo or 200 mg tildacerfont QD) |
| Open-Label Treatment Period: | 200 mg tildacerfont QD |
| Open-Label Extension Period | 200 mg tildacerfont QD |

The actual study drug treatment will be the study drug a subject was on when a data record was measured/assessed, unless otherwise specified in this section.

#### 8.3.18.1 Actual Treatment for an Adverse Event Record

AEs are those AEs with onset date from the signing of the ICF until the end of the follow up period. No treatment will be assigned to AEs from the initiation of the ICF through the GC Conversion Period AE records.

Treatment emergent adverse events (TEAEs) are those AEs with onset date from the date of first dose of randomized study drug until 30 days after the last dose of study drug. The actual study treatment will be assigned to these records.

If a subject discontinues study drug but remains in the study, AEs with onset date after 30 days of last dose of study drug until the end of the follow up period will be considered post-treatment adverse events. The last treatment that the subject was on before study drug discontinuation will be assigned to these records.

#### 8.3.18.2 Actual Treatment for a Medication Record

Since a medication could be a prior medication and be taken during the GC Conversion Period and/or Treatment Periods, flag variables will be used to identify a medication record that was taken while a subject received placebo or 200 mg tildacerfont. A medication record will be assigned to the first study treatment when the subject was on and was taking the medication during the Treatment Periods.

### 8.3.19 Study Drug Exposure Variables

#### 8.3.19.1 Study Drug Compliance

Study drug compliance is defined as the percentage of study drug actually taken compared to what was expected based on the randomized dose for time period *i* (where subject is receiving a constant dose). Using the drug accountability data, study drug compliance is calculated as follows:

$$\%Compliance_i = 100 \times \left(\frac{\#\ Dispensed - \#\ Returned - \#Lost\ or\ Destroyed}{\#\ Expected\ tablets\ to\ be\ taken}\right)_i$$

and total % compliance over a treatment period is calculated as:

$$\%Compliance = 100 \times \left( \frac{\sum_{i=1}^{k} (\#Dispensed - \#Returned - \#Lost\ or\ Destroyed)_i}{\sum_{i=1}^{k} (\#Expected\ tablets\ to\ be\ taken)_i} \right)$$

where
Σ represents the summation operator and the value in parentheses is summed over the sequence *i* = 1 to *k*,

where

*k* = number of time intervals subject is receiving a constant dose, and possible constant dose levels are placebo or 200 mg QD tildacerfont.

The expected number of tablets to be taken is the number of expected tablets per day, based on the randomized dose, multiplied by the number of days in time period *i*. Within each time period *i,* if there is no documented number of returned, lost, or destroyed tablets, a conservative approach to the calculation of compliance will assume all tablets were lost and therefore not taken, resulting in a compliance calculation of zero for the given time period. Compliance over 24 weeks of randomized treatment will be summarized in terms of the proportion of subjects with 100% compliance, ≥80 to <100% compliance, ≥50 to <80% compliance and <50% compliance.

#### 8.3.19.2 Total Tildacerfont Dose (mg) Taken

Total tildacerfont dosage (mg) is defined as the dosage of tildacerfont taken over a specific study period *i* (where subject is receiving a constant dose), calculated as follows:

$$Total\ Tildacerfont\ Dose_i = \left( (\#Dispensed - \#Returned - \#Lost\ or\ Destroyed) \times Tablet\ Strength \right)_i$$

and total tildacerfont dose over the entire study period is calculated as:

$$Total\ SPR001\ Dose = \sum_{i=1}^{k} (Total\ SPR001\ Dose)_i$$

where
Σ represents the summation operator and the value in parentheses is summed over the sequence *i* = 1 to *k*,

where
*k* = number of time intervals subject is intended to receive study drug, and Tablet strength is 0 mg if placebo or 200 mg tildacerfont if active drug.

#### 8.3.19.3 Mean Daily Tildacerfont Dose (mg/day)

Mean Daily tildacerfont Dose (mg/day) is calculated as Total tildacerfont Dose (mg) divided by Treatment Duration (days) for each study period and over the entire study.

#### 8.3.19.4 Study Drug Exposure Duration (days)

The duration (days) is calculated as (LASTDAY - FIRSTDAY+ 1 day), where LASTDAY is the date of last dose during a treatment period, and FIRSTDAY is the date of first dose during the treatment period.

#### 8.3.19.5 Diary Compliance

For daily diary entries that have a double entry on a given day: if the prior day is missing an entry, then the first of the double day entry applies to the previous day. If the prior day is not missing and two entries exist, use last entry for the given date.

##### 8.3.19.5.1 Study Drug Diary Compliance

Total study drug diary compliance is calculated as the total count of 'Yes' responses to taking study drug as directed, divided by (LASTDAY – FIRSTDAY+ 1 day), where LASTDAY is the date of last dose, and FIRSTDAY is the date of first dose.

Weekly diary compliance is calculated as the total count of number of 'Yes' responses to taking study drug as directed within the duration {Study Day (7n+1) to Study Day (7[n+1])}, divided by 7. Where n has values {0, 1, 2, ..., 76}. Missing entries are counted as non-compliance.

#### 8.3.19.5.2 Food Diary Compliance

Food diary compliance is calculated as the total count of ‘Yes, with evening meal’ to taking a meal with the study drug, using the same methods of study drug diary compliance.

#### 8.3.19.5.3 Glucocorticoid Diary Compliance

Total GC diary compliance is calculated as the total count of number of 'Yes' or ‘Took Stress Dosing’ responses to taking GC as directed and number of stress dosing days, divided by (LASTDAY – FIRSTDAY+ 1 day), where LASTDAY is the date of last dose, and FIRSTDAY is the date of first dose.

#### 8.3.19.5.4 Composite Diary Compliance

Aggregate diary compliance is calculated as (study drug compliance + food compliance + GC diary compliance) divided by 3. Compliance versus each component (study drug, GC compliance and food compliance) will also be presented and summarized.

Weekly compliance for GCs and aggregate compliance is calculated using the same methodology as weekly study drug compliance.

### 8.3.20 Glucocorticoid Therapy

GC therapies are concomitant medications that fall under the Anatomical Therapeutical Chemical (ATC) classification is “Systemic Hormonal Preparations, Excl. Sex Hormones and Insulins” or “Glucocorticoids.”

#### 8.3.20.1 Duration (Days) on Current Background Glucocorticoid Therapy

Duration (days) on the Sponsor-provided or Principal Investigator-prescribed supply of standardized GC regimen is defined as the date of first dose of study drug (GC Conversion Period) - the date of start of the current background GC therapy + 1 day.

#### 8.3.20.2 Glucocorticoid Dose in Hydrocortisone Equivalents

Table 11 defines the potencies of various GCs relative to HC in treating CAH. These conversion factors will be used to determine whether a subject's daily GC dose satisfies study eligibility criteria and provide general guidelines for GC tapering.

**Table 11 Relative Potencies of Glucocorticoids in Hydrocortisone Equivalents**

| Glucocorticoid | Potency in CAH (HCe) |
|---|---|
| Hydrocortisone | 1 |
| Prednisone | 4 |
| Prednisolone, Prednisolone Sodium Phosphate, Methylprednisolone | 5 |
| Fludrocortisone | 0 |
| Dexamethasone | 70 |
| Cortisone Acetate | 0.8 |

#### 8.3.20.3 Glucocorticoid Total Daily Dose in Hydrocortisone Equivalents

The total daily GC dose of a given subject is calculated as the sum of doses per day of concomitant GC therapies using the conversion in Table 11, excluding stress dosing.

#### 8.3.20.4 Stress Event

A stress event is defined as an entry in the diary indicating a stress dose was taken or a record in the CM eCRF if the stress dosing was not recorded in the diary in a timely manner. Consecutive days of stress dosing are counted as a single stress event.

#### 8.3.20.5 Cumulative GC Dose at Week 24

The cumulative GC dose from Study Day 1 to Week 24 is defined as:

- Sum of daily GC exposures (total daily dose multiplied by days on a regimen) from CM eCRF
- Sum of stress doses (HCe mg) (from CM eCRF where the indication for the stress dosing is "STRESS DOSE/DOSING")

For subjects who discontinue study drug treatment prior to the Week 24 visit, the cumulative dose will be summed over the duration of Study Day 1 to the date of last dose of study drug.

### 8.3.21 Concomitant and Prohibited Medications

All medication verbatim terms reported on the eCRFs will be mapped according to the World Health Organization (WHO) Drug Dictionary (WHO DD March 2020 B3). The medications will be mapped to Anatomical/Therapeutic/Chemical (ATC) class and preferred names.

A prior medication is considered to be any medication that is taken within a year prior to screening.

#### 8.3.21.1 Concomitant Medications

A concomitant medication (CM) is considered to be any medication that is continued during Screening through last dose +24 hours.

CMs that stop prior to the first dose of study drug will be reported as screening concomitant medications and will be summarized with the prior medications.

CMs that are continued after the first study drug dosing, or with start dates or stop dates post first dose of study drug through last dose date + 24 hours, missing CM end date, or ongoing are considered on-treatment CMs.

Post medications is considered to be any medication taken after the last study drug dose date + 24 hours, missing CM end date, or ongoing.

#### 8.3.21.2 Prohibited Concomitant Medications

Prohibited and cautionary concomitant medications are those medications with their potential for metabolic interactions with tildacerfont. Specific definitions can be found in Protocol Sections 6.5.1 and 13.1.

### 8.3.22 Treatment-Emergent Adverse Events

All AE verbatim terms reported on the eCRFs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA® version 23.0). Treatment-emergent adverse events (TEAEs)

are defined as any AEs that have an onset or worsening in severity on or after the first dose of double-blinded study drug until 30 days after the last dose of study drug. Non-TEAEs include those events occurred during the screening and/or GC conversion Period and after 30 days of the last dose of study drug. Non-TEAEs will be listed in AE listings but excluded from TEAE summaries.

Related AEs are those reported by investigators as possibly related, probably related, or definitely related to the study drug or GC regimen.

AESIs are events that do not meet SAE criteria but must be monitored on an ongoing basis as defined in Protocol Section 8.3.8 and identified in the AE eCRF page. Serious AEs (SAEs) are defined according to the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE 2017) version 5.0, as described in Protocol Section 8.3.3.1. Adverse events

### 8.3.23 Geographic Regions

Approximately 75 investigative sites will participate in this study. Three geographic regions will be classified by the countries where these study centers are located: United States, Europe, and rest of world.

## 8.4 Analysis Windows

Clinical visits may occur outside protocol-specified windows. Therefore, instead of relying solely on visit labels in the clinical database, analysis visits and their windows are defined using Enrollment Day or Study Day. For the purposes of data analysis and summary, assessments and/or measurements will be flagged based on the collection date/time that is closest to the protocol-scheduled time point (or Target Study Day). Analysis visit windows are presented in by type of assessments and/or measurements.

**Table 12 Visit Analysis Windows (Clinic visits only)**

| Protocol Specified Visit Number | Analysis Visit | Reference Study Day | Target Study Day | Start (days) | Stop (days) |
|---|---|---|---|---|---|
| Visit 1 (Screening) | Screening | Enrollment Day | -114 | low | -7 |
| Visit 2 | Week -12 | Enrollment Day | -84 | -6 | 15 |
| Visit 3 | Week -8 | Enrollment Day | -56 | 16 | 50 |
| Visit 4 | Week -2 | Enrollment Day | -14 | 51 | high |
| Visit 5 | Day 1 | Study Day | 1 | -6 | 7 |
| Visit 6 | Week 3 | Study Day | 22 | 8 | 32 |
| Visit 7 | Week 6 | Study Day | 43 | 33 | 64 |
| Visit 8 | Week 12 | Study Day | 85 | 65 | 106 |
| Visit 9 | Week 18 | Study Day | 127 | 107 | 148 |
| Visit 10 | Week 24 | Study Day | 169 | 149 | 197 |
| Visit 11 | Week 32 | Study Day | 225 | 198 | 253 |
| Visit 12 | Week 40 | Study Day | 281 | 254 | 323 |
| Visit 13 | Week 52 | Study Day | 366 | 324 | 408 |
| Visit 14 | Week 64 | Study Day | 450 | 409 | 492 |
| Visit 15 or Extension Visit 1 | Week 76 | Study Day | 534 | 493 | high or 575 |
| Extension Visit 2 | Week 88 | Study Day | 617 | 576 | 659 |
| Extension Visit 3 | Week 100 | Study Day | 701 | 660 | 785 |
| Extension Visit 4 | Week 124 | Study Day | 869 | 786 | 953 |
| Extension Visit 5 | Week 148 | Study Day | 1037 | 954 | 1121 |
| Extension Visit 6 | Week 172 | Study Day | 1205 | 1122 | 1289 |
| Extension Visit 7 | Week 196 | Study Day | 1373 | 1290 | 1457 |
| Extension Visit 8 | Week 220 | Study Day | 1541 | 1458 | 1625 |
| Extension Visit 9 | Week 244 | Study Day | 1709 | 1626 | 1793 |
| Extension Visit 10 | Week 268 | Study Day | 1877 | 1794 | 1961 |
| Extension Visit 11 | Week 292 | Study Day | 2046 | 1962 | 2129 |
| Extension Visit 12 | Week 316 | Study Day | 2213 | 2130 | 2228 |
| Visit 16 or Extension Visit 13 (follow up) | Follow-up | Study Day | | Last dose +1 | Last dose +30 |

# 9 STATISTICAL AND ANALYSIS ISSUES

## 9.1 Adjustments for Covariates

The categorical, randomization stratification variables sex (0 = male, 1 = female) and baseline HCe group (<40 mg/day and ≥40 mg/day) will be used in all efficacy analyses. Additional categorical covariates may include baseline daily GC type (0 = short-acting, 1 = intermediate acting or combination) and CAH Type (0 =simple virilizing, 1=salt-wasting). Baseline daily GC dose in HCe or BSA adjusted dose may also be included as a continuous covariate. The baseline value of continuous efficacy endpoints will be included as a continuous covariate in statistical models, such mixed model repeated measures (MMRM), analysis of covariance (ANCOVA), or logistic regression.

## 9.2 Assessing Normality

Modest departures in the distribution of subject level data from Normality are of no consequence when sample size is high, for example, above a total N ≥10, as Normality of the test statistics, e.g the difference in treatment group means, is assured by the Central Limit Theorem. When based on a total N ≥10, the moment generating function for any test statistic derived from any underlying (non-Normal) probability distribution falls within less than 5% of the ideal moment generating function for the Standard Normal distribution, thus guaranteeing valid inference based upon the Normal approximation. This figure rises to within less than 2.5% of the ideal moment generating function with a total N ≥ 20.

If, however, departures from Normality are encountered, especially when the effective sample size may be low, a supportive analysis may be performed based on rank values in place of original values.

## 9.3 Handling Dropouts or Missing Data

Every effort will be made by the Sponsor to ensure completeness of data collection. In the subject listings, all collected and all imputed values, if any, will be presented. However, and despite best efforts, some missing data are inevitable.

### 9.3.1 Handling of Efficacy Data

A sensitivity analyses will be conducted for the Primary Efficacy Endpoint and Secondary Efficacy Endpoints under the assumption of missing at random and missing not at random.

### 9.3.2 Handling of Laboratory Data

For laboratory data, a missing baseline value will be replaced with the last available assessment, generally the screening assessment. A retest value will be used if the first test is invalidated, e.g., specimen hemolyzed. For non-PK laboratory values that are continuous in nature but are presented as either above or below the respective quantitation limits, the following imputations will be made for the purposes of summarization:

- If value is listed as < X, then the imputed value will be X/2
- If a value is listed as ≤ X, then the imputed value will be X
- If value is listed as > X, then the imputed value will be X+1
- If a value is listed as ≥ X, then the imputed value will be X

### 9.3.3 Handling of Safety Data

#### 9.3.3.1 Adverse Events

If the time of onset (before or after intake of study drug) cannot be determined whether an AE is treatment-emergent because of a partial onset date, the event will be counted as a TEAE.
Adverse events with incomplete start dates will be considered TEAEs, if:

- Day and month are missing and the year is equal to or after the year of the first date of study drug dosing;
- Month is missing and the year is equal to or after the year of the first date of study drug dosing;
- Day is missing and the year is equal to the year of the first date of study drug dosing and the month is equal to or after the month of the first date of study drug dosing; or
- Year is missing.

If severity or relationship of an AE to study drug is not recorded, the severity or relationship will be imputed as “severe” or relationship as “possibly related,” for analysis purposes. All efforts

will be made to ensue no missing severity or relationship of an AE to study drug prior to database lock finalization.

#### 9.3.3.2 Concomitant Medications

If start date of a medication is missing, the medication will be considered to have started prior to the study. Such a medication may also be considered concomitant, depending on the stop date or lack thereof. If the stop date of a concomitant medication is missing, then the medication will be treated as ongoing. If the start date of a medication is missing, the stop date will be used to determine whether or not it is concomitant. Medications with other incomplete start dates or end dates will be imputed as follows:

Incomplete medication start date/time:
- If only have a YEAR, impute as Jan. 1.
- Else if only have YEAR and MONTH, impute as Day 1 of month.
- Otherwise missing, no imputation.

Incomplete medication end date/time:
- If only have a YEAR, impute as December 31.
- Else if only have YEAR and MONTH, then impute to last day of the month.
- Otherwise missing, no imputation.

### 9.4 Primary Analyses and Data Monitoring

The primary efficacy analysis will be conducted when all subjects complete the Week 24 visit of the Placebo-Controlled Treatment Period and the data has been hard-locked for analysis. The Final Analysis will occur after the last subject has completed the Week 76 visit or if the study is discontinued by the Sponsor.

An external, independent data safety monitoring board (DSMB) will review the available safety data (both blinded and unblinded) from this study periodically. The DSMB will advise the Sponsor of any trends or safety issues that may impact the study or study subjects and provide recommendations to the Sponsor if needed.

## 9.5 Multicenter Considerations

This trial will be conducted at approximately 75 study centers in the United States, Europe, and other regions. Data from all study centers will be pooled for efficacy and safety analyses.

Because the number of subjects at each center is likely to be small, no analyses will be performed by center.

## 9.6 Type I Error Control

Type I Error control for the Secondary Efficacy Endpoints will be achieved by hierarchical testing. If the Primary Efficacy Endpoint is met, then Secondary Efficacy Endpoints displayed in Table 5 will be evaluated using the Hochberg (1988) closed test procedure with an alpha of 0.05.

## 9.7 Examination of Subgroups

Subjects will be categorized into the following common subgroups for the purposes of evaluating the Primary and Secondary Efficacy Endpoints.

- Sex (Male, Female)
- Baseline GC dose in HCe
  - < 40 mg/day, ≥ 40 mg/day
  - < 35 mg/day (For endpoint 2.2 and 3.7 only)
- Baseline A4 levels (A4 ≤ ULN, A4 > ULN)

In addition, all potential subgroup x subgroup combinations are also pre-specified. If the primary analysis is positive, exploratory subgroup analysis for age, race, ethnicity, and region will be used to further evaluate the Primary and Secondary Efficacy endpoints.

Comparisons will only be provided for subgroups where each subgroup has at least 10 subjects. Additional subgroups specific to only select endpoints are specified in the appropriate section in the SAP.

# 10 STUDY SUBJECTS

## 10.1 Subject Enrollment and Disposition

### 10.1.1 Screened and Enrolled Subjects

Screened and enrolled subjects will be summarized enrolled by country and by investigator for treatment each treatment group and overall.

Screening, enrollment, and disposition will be summarized for all screened subjects. Percentages will be based on total subjects who enrolled the study. Disposition will also be summarized for all subjects in the ITT analysis set. The percentages will be based on total subjects who were randomized in the study. A subject listing will be provided with the above information, as well as geographic region, country, and site ID for subjects who were screen failures and for subjects enrolled in the study.

### 10.1.2 Randomized Subjects

Dispositions of randomized subjects will be summarized by treatment group and overall for the Placebo-Controlled Treatment Period and the Open-Label Treatment Period.

A listing of disposition will be provided for all randomized subjects. A separate listing will present subjects who are excluded from SAF, mITT, PP, PK, and/or DXA analysis sets, along with reasons for the exclusions.

## 10.2 Protocol Deviations

Protocol deviations will be listed by deviation type (e.g., major) and category (e.g, noncompliance with study procedures or restrictions). All deviations will be identified prior to database lock and will be summarized and presented in listing(s); listings will include flags for deviation type (major or minor). Major protocol deviations are defined per *Guideline for Industry: Structure and Content of Clinical Study* from International Council for Harmonisation (ICH E3) guidance. Major analysis protocol deviations include, but not limited to:

- Randomized subjects who did not satisfy efficacy inclusion and exclusion criteria,
- Randomization error,

- Subjects who received the wrong tildacerfont dose,
- Subjects who received an excluded concomitant treatment or medication,
- Subjects who are less than 80% compliant with aggregate study medication measured by pill count or by e-diary

All protocol deviations and major protocol deviations will be summarized by deviation category and type. In addition, a listing of all deviations and major protocol deviations will be provided.

## 10.3 Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized and listed for all randomized subjects (ITT).

### 10.3.1 Demographics

Demographic characteristics will include age, age category, sex, females with child-bearing potential, race, and ethnicity, type of CAH diagnosis, and time since CAH diagnosis. A subject listing of demographics will be provided.

### 10.3.2 Baseline Characteristics

The following baseline characteristics will be summarized. They will be listed in subject listings in the following categories: CV risk assessments; glucocorticoid regimen; selected serum hormones; and TART(s)

## 10.4 Medical History

Medical history and CAH history will be summarized by randomized treatment group for all enrolled subjects. Individual subject listings will include subject identification number and randomized treatment/dose level (if a subject is randomized), along with general history and CAH history data collected from eCRF, for all enrolled.

### 10.4.1 General Medical History

General medical history, besides CAH, will be mapped to preferred terms and system organ classes using the Medical Dictionary for Regulatory Activities (MedDRA®) Dictionary (version 23.0). General medical history will be summarized by system organ class (SOC) and preferred

term (PT). A subject listing of general medical history will include start date/end date, verbatim medical history term, SOC, PT, and ongoing status.

### 10.4.2 Congenital Adrenal Hyperplasia History

#### 10.4.2.1 Signs and Symptoms

The most bothersome signs and/or symptoms of CAH will also be summarized. Signs and/or symptoms of CAH will be included in a subject listing.

#### 10.4.2.2 Glucocorticoid History

The Baseline GC dose level in HCe, the duration (years) on the current background GC therapy, GC therapy medication(s), reasons for change to the current background GC therapy (if changed therapy) will be summarized for all enrolled subjects. A subject listing will include all summarized variables, as well as any higher/lower dose of GC during adulthood, reason to reduce/increase to the current GC dose.

#### 10.4.2.3 Hospital and Stress Dosing History

Hospitalization and stress dosing history will be summarized separately and will include the frequency and percentage of subjects with at least one hospitalization for adrenal insufficiency/crisis in the past 5 years, at least one hospitalization for adrenal insufficiency/crisis in the past 12 months, and/or at least one stress dosing event in the past 12 months. A subject listing will be included.

## 11 STUDY DRUG AND OTHER MEDICATIONS

### 11.1 Exposure to Study Drug and Study Treatment Compliance

Summaries of study drug exposure and compliance will be summarized overall and separately for those taken during Placebo-Controlled Treatment Period and the Open-Label Treatment Periods of the study and overall for the Safety Analysis Set and the PP Analysis Set. The source for study drug dosing and compliance is drug accountability.

The summaries will include the total number of doses taken, mean daily dose (mg/day), duration (days) of exposure, and study drug compliance by treatment group.

The drug accountability information, including number of tablets dispensed, number of tablets returned, number of tablets lost or destroyed, and reasons/comments for incomplete or missed doses, along with all summarized variables, will be listed in subject listings. Subject dosing diary information, along with diary compliance, will also be listed in a separate listing.

### 11.2 Glucocorticoid Therapy

Changes in GC therapy will be summarized overall and separately for those taken during each Treatment Period (Placebo-Controlled, Open-Label) of the study and overall. The source of GC regimen change is the concomitant medication eCRF.

The following items will be summarized, proportion of subjects with a stress dosing event, number of unique stress dosing events, duration of stress dosing events. The source of GC stress dosing is the GC daily diary. All the above information will be included in a subject listing.

### 11.3 Prior and Concomitant Medications

Prior and concomitant medications, other than glucocorticoids, will be summarized using WHO DD ATC class and preferred name. The summary results will be presented by Treatment Period (Placebo-Controlled, Open-Label). Prior medications will be summarized by randomized treatment group. Concomitant medications will be summarized by actual dose taken concomitantly during each treatment period.

These summaries will present the number and percentage of subjects using each medication. Subjects may have more than one medication within an ATC class and preferred name. At each summary level subjects are counted once if they reported one or more medications at that level. Each summary will be ordered by descending frequency of incidence of ATC class and preferred name within each ATC class.

All medications will be provided in a subject listing including, but not limited to: Start date/end date/ongoing, medication name, ATC class and preferred name, indication, dose, unit, form, frequency, and route.

### 11.3.1 Prohibited Concomitant Medications

Any use of prohibited and cautionary concomitant medications during the study will be identified. A summary table of prohibited concomitant medications will be presented similar to that of concomitant medications. A separate subject listing of prohibited and cautionary concomitant medications will be provided.

## 12 EFFICACY ANALYSES

### 12.1 Primary Efficacy Analyses

The Primary Analysis is planned after all subjects have completed the Week 24 visit at the end of the Placebo-Controlled Period of the study. The data from Baseline through Week 24 will be hard-locked prior to the Primary Analysis.

During this first 24-week of double-blind, placebo-controlled treatment period, the scheduled visits included in this treatment period include Day 1, Week 3, Week 6, Week 12, Week 18, and Week 24.

Efficacy analyses will compare the tildacerfont dose group to the placebo group with respect to

- The Primary Efficacy Endpoint (see Table 5)
- Secondary Efficacy Endpoints (see Table 5)
- and Exploratory Efficacy Endpoints (see Table 5)

These efficacy analyses will be performed using the ITT Analysis Set. The PP and mITT Analysis Sets will be used as supplemental for the primary and Secondary Efficacy Endpoint analyses. The Primary Efficacy Endpoint, select Secondary Efficacy Endpoints and select Exploratory Efficacy Endpoints (as outlined in Table 5) will be analyzed at the Primary Analysis.

## 12.2 Final Efficacy Analyses

The Final Analysis will occur after the last subject has completed the Week 76 visit or if the study is discontinued by the Sponsor. The data from Week 24 through Week 76 + follow-up will be hard-locked prior to the Final Analysis.

The scheduled visits included in this treatment period include Day 1, Week 3, Week 6, Week 12, Week 18, Week 24, Week 32, Week 40, Week 52, Week 64, Week 76 and Follow-up visits. Exploratory Efficacy Endpoints along with select Secondary Efficacy Endpoints (as outlined in Table 5) will be analyzed at the Final Analysis.

## 12.3 Primary Efficacy Endpoint Analysis

The primary objective of the study is to evaluate GC dosage change in HCe in subjects with CAH over the 24-week, double blind, Placebo-Controlled Treatment Period.

### 12.3.1 Definition of Primary Efficacy Endpoint Estimand

The primary "treatment policy" estimand is the absolute change from baseline in GC dose in HCe at Week 24 post randomization. In the original study protocol, the primary endpoint was defined in a binary fashion as the proportion of subjects with at least a 5 mg/day HCe reduction from baseline in GC dose and A4 ≤ULN at Week 24. Consistent with the emerging analysis of competitor trial data, and also realizing the loss of information and power associated the reduction of continuous data to a simple binary form, the primary endpoint definition was changed from binary to a continuous change from baseline and, hence, the associated use of MMRM analysis below. Nevertheless, binary evaluation of the primary endpoint is still retained in the SAP in the form of supportive secondary analyses.

**Population:** The analysis population is the ITT analysis set, including all classic CAH subjects who were randomized into the study

**Variable:** The analysis variable is the absolute change from baseline to Week 24 in GC dose in HCe.

**Intercurrent events:** For the Primary Efficacy Endpoint analysis, intercurrent events such as changes in GC type or lack of treatment compliance or GC compliance, treatment interruption or discontinuation, use of prohibited medications will be ignored.

### 12.3.2 Primary Endpoint Statistical Hypotheses

For the Primary Efficacy Endpoint, the null hypothesis is that there is no difference in the mean change from baseline to Week 24 in GC dose in HCe. The alternative hypothesis will be that there is a difference between the treatment groups.

$$H_0: \mu_{200\ mg\ QD\ tildacerfont} - \mu_{placebo} = 0$$
$$H_A: \mu_{200\ mg\ QD\ tildacerfont} - \mu_{placebo} \neq 0$$

### 12.3.3 Statistical Modeling and Testing

An MMRM model will be used to analyze the Primary Efficacy Endpoint. The model will comprise the change from baseline as the dependent variable; randomized treatment group (0: placebo, 1: tildacerfont), visit (Weeks 3, 6, 12, 18 and 24), randomized treatment by visit interaction as fixed effects; baseline value as a continuous covariate; within subject error will be modelled using an unstructured covariance matrix; sex (0: male, 1: female), baseline GC type (0: short-acting, 1: intermediate-acting or combination) and CAH Type (0 =simple virilizing, 1=salt-wasting) will also be included as fixed effects. The estimation method will be restricted maximum likelihood (REML). The principal treatment effect of interest is at 24 weeks for which the associated LSmeans, difference in LSmeans, CI and 2-sided p-value will be extracted from the model; in addition, the same will be extracted for each earlier visit.

In the event the MMRM model with an unstructured covariance structure does not converge, the following covariance structures will be used as substitution in the order below. Each subsequent covariance structure will be used only if each previous covariance structure was used and the model did not converge:

a) Toeplitz covariance structure (assuming measurements from samples taken closer together in time are more highly correlated than those from samples taken farther apart)

b) First order of auto-regressive [AR(1)] covariance structure (assuming measurements from samples taken closer together in time are more highly correlated than those from samples taken farther apart)
c) Compound symmetry covariance structure (assuming equal correlation for measurements from a subject, regardless of how far apart in time when they were taken)

#### 12.3.3.1 Multiple Imputation for the Primary Efficacy Endpoint Analysis

Missing data can have an adverse effect on the estimation of endpoints. As such, multiple imputation for missing data will be used for the Primary Efficacy Endpoint analysis. Imputation will be performed under the assumption of missing at random (MAR) and supported by contemporary imputation under the assumption of missing not at random (MNAR). In addition, retrieved drop-out based multiple imputation will be explored whereby data from subjects who prematurely discontinue study drug treatment but remain in the study and still have assessments collected at the remaining scheduled visits in the study as used as the basis for imputation.

Thus, in the event of missing GC HCe values at Weeks 3, 6, 12 18 and 24, data will be imputed as follows:

(a) MAR Imputation

Step 1: Non-monotone missing data will be imputed first based on the MAR assumption via a multivariate joint Gaussian imputation model using Markov chain Monte Carlo (MCMC) method using the MCMC statement in the SAS® PROC MI procedure. Twenty imputed datasets will be produced. As a result, each imputed dataset will only have missing data at the end of subjects' records, (i.e. a monotone missing data pattern). The MCMC method in the MI procedure will be used with multiple chains (option CHAIN=MULTIPLE), 100 burn-in iterations, and a non-informative prior to imputed the remaining missing data. A separate imputation model will be used for each treatment arm. The imputation models will include GC HCe values at baseline and Weeks 3, 6, 12 and 18, sex, baseline GC type and baseline and CAH Type. In case

of non-convergence or non-estimability issues, a ridge prior and a single model will be considered with treatment arm added as explanatory variable to the model.

Step 2: The remaining, monotone, missing data for all subjects who discontinue study prematurely will be imputed using sequential regression multiple imputation model by randomised treatment arm. Each sequential regression model (i.e., for imputation of values at a given time point) will include the same explanatory variables as in Step 1 and all prior GC HCe values.

Step 3: The change from baseline in GC HCe to each scheduled post-baseline visit will be calculated, based on observed and imputed data. Each of the imputed complete datasets from Step 2 will be analysed with the same MMRM model used for the Primary Efficacy Endpoint (Section 12.3.3).

Step 4: The results of the analysis of each imputed dataset, i.e., treatment differences and their standard errors, will be combined using Rubin's imputation rules (Rubin, 1987) to produce pooled estimates of treatment differences, its 95% confidence interval and a pooled p-value. This will be done using SAS® MIANALYZE procedure.

(b) MNAR Imputation

Step 1: as (a)

Step 2: The remaining, monotone, missing data for all subjects who discontinue study prematurely will be imputed using sequential regression multiple imputation model estimated based on data from the *placebo arm only*. This will be achieved via the MNAR option in PROC MIANALYZE/ Each sequential regression model (i.e., for imputation of values at a given time point) will include the same explanatory variables as in Step 1 and all prior GC HCe values.

Step 3: as (a)

Step 4: as (a)

(c) Tipping Point MNAR Imputation

In the event the Primary Efficacy Endpoint is met, a Tipping Point analysis will be conducted. This will assess the robustness of the Primary Efficacy Endpoint analysis by determining the degree of penalty this is required to be applied to the active treatment arm to lose statistical significance. An increasingly punitive penalty will be determined as a small fraction of the 24 Week treatment effect estimate from the Primary Efficacy Endpoint analysis; i.e., if the treatment effect estimate from the Primary Efficacy Endpoint analysis of GC HCe is $\hat{\theta}$, then the penalty parameter, $\delta_{min}$, will be set as $\delta_{min} = 0.05 \cdot \hat{\theta}$ ; and the maximum penalty parameter will be set as $\delta_{max} = 40 \cdot \delta_{min} = 2 \cdot \hat{\theta}$. Missing data will then be imputed as in (a) above, firstly applying no penalty (=Primary Efficacy Endpoint analysis assuming MAR) and then applying and increasingly punitive penalty ($\delta_{min}, \delta_{max}$) at intervals of $0.05$ creating twenty imputed datasets. Each of these datasets will be analyzed as per Steps 3 and 4 in (a) above.

The p-values outputted from the application of each penalty value will be displayed along with the treatment effect estimates and associated CIs so that the tipping point resulting in loss of significance for the Primary Efficacy Endpoint can be identified. If the initial penalty parameters tested do not identify the tipping point, the process will be repeated with different penalty parameter values (e.g., $\delta_{min} = 0.04 \cdot \hat{\theta}, \ldots., \delta_{min} = 0.01 \cdot \hat{\theta}$) until the tipping point is identified.

(d) Retrieved Dropout Imputation

Step 1: as (a)

Step 2: If n ≥ 5 subjects in a treatment group have discontinued in each treatment arm, the remaining monotone missing data will be multiply imputed using a retrieved dropout approach via fully conditional specification and predictive mean matching assuming missing not at random (MNAR). The Markov chain Monte Carlo method with a non-informative prior (Jeffreys) will be used to perform a monotone-data imputation (IMPUTE=MONOTONE) by randomized treatment group.

Step 3: Week 24 GC HCe dose will be imputed using a set of predictors resulting in a set of linear coefficients *b,*

$$\text{Week 24 GC HCe dose} = Treatment + Sex + Visit + Treatment * Visit + \log(Baseline\ A4) + Baseline\ GC\ Dose + Baseline\ GC\ Type + Baselibe\ CAH\ Type\ + Last\ available\ \text{GC HCe dose}$$

Step 4: Do a random draw from the *posterior predictive distribution* of *b* producing a new set of coefficients $b^*$. Use $b^*$ to generate predicted values for the Week 24 for all subjects.

Step 5: For each missing Week 24 value, identify a set of candidates with observed Week 24 values whose predicted values are close to the predicted value for the case with missing data. From the candidates, randomly choose one Week 24 value and assign its observed value to the missing value. Repeat process for every subject with missing data.

The analysis methods described in Section 12.3.3 will be applied to the imputed datasets.

### 12.3.4 Reporting Results

#### 12.3.4.1 Summary of Descriptive Statistics

The change from baseline in GC HCe dose will be summarized by treatment group at Weeks 3, 6, 12, 18 and 24 using descriptive statistics; and the same will also be provided for the percent change from baseline in GC HCe dose.

##### 12.3.4.1.1 Summary of Analysis Results

For Section 12.3.3.1, (a) through (d), the resulting LS mean change from baseline, SE of the LS mean, the difference in LSmean values, tildacerfont – placebo, together with its associated SE and 2-sided p-value will be presented Weeks 3, 6, 12, 18 and 24.

## 12.4 Primary Efficacy Endpoint Additional Analyses

### 12.4.1 Per-Protocol and Modified ITT Analysis Sets

The PP and mITT Analysis Sets will provide for additional analyses of the Primary Efficacy Endpoint; and, further, the mITT Analysis Set will provide for additional analyses of the Secondary Efficacy Endpoints. In both instances, only inferences at the Week 24 time point will be presented in these analyses.

### 12.4.2 Subgroup Analyses

The Primary Efficacy Endpoint Analysis will be repeated using the planned subgroups, by Sex, by Baseline GC dose, by Baseline GC type, by Baseline A4 ULN group.

## 12.5 Secondary Efficacy Endpoint Analyses

The analysis population for the Secondary Efficacy Endpoint analyses is the ITT analysis set, including all classic CAH subjects who were randomized into the study.

**Table 13 Secondary Efficacy Endpoints**

| Objectives | Endpoints | | Analysis |
|---|---|---|---|
| ***Secondary Efficacy*** | | | |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 2.1 | Proportion of subjects with GC dose ≤11 mg/m2/day in HCe and A4 ≤ 1.2x baseline or A4 ≤ ULN at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 2.2 | Proportion of subjects with baseline GC dose ≤ 35mg HCe who achieve GC dose ≤11 mg/m2/day in HCe and A4 ≤ 1.2x baseline or ≤ ULN at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing cardiovascular risk in subjects with CAH | 2.3 | Proportion of subjects with improvement in at least one cardiovascular risk factor at Week 24 | Primary Analysis |

### 12.5.1 Statistical Hypotheses for Secondary Efficacy Endpoints

For Secondary Efficacy Endpoint 2.1 the null hypothesis is that the difference between randomized treatment groups in the proportion of subjects achieving GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN at Week 24 is equal to zero. The alternative hypothesis is that the difference is not equal to zero.

For Secondary Efficacy Endpoint 2.2 the null hypothesis is that in the subset of subjects with a baseline GC dose ⩽ 35mg the difference between randomized treatment groups in the proportion of subjects achieving GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN at Week 24 is equal to zero. The alternative hypothesis is that the difference is not equal to zero.

Secondary Efficacy Endpoint 2.3 is the proportion of classic CAH subjects who were randomized into the study who improve at least one baseline CV risk factor at 24 weeks from randomization. This proportion is calculated for each treatment group (e.g., placebo, tildacerfont) as the number of subjects at Week 24 who had at least one CV risk factor at baseline (i.e subjects without any CV risk factors will not be included in the analysis) and no longer meets one or more of the baselines CV risk factors at Week 24, divided by the total number of subjects in the ITT analysis who were randomized to that specific treatment and had at least one CV risk factor at baseline.

#### 12.5.1.1 Statistical Modeling and Testing

The Secondary Efficacy Endpoints are all binary ‘response’ endpoints will be analyzed using logistic regression modelling.

The model for each endpoint will include the responder indicator as the dependent variable; treatment group (0: placebo, 1: tildacerfont); the randomization strata: sex (0: male, 1: female) and baseline GC dose (in HCe, 0: < 40 mg/day, 1: ≥ 40 mg/day); and baseline GC type (0: short-acting, 1: intermediate-acting or combination) as categorical explanatory variables; and, for

Secondary Efficacy Endpoint 2.3, the baseline number of CV risk factors will also be included as a categorical covariate.

An exact score test derived from the logistic regression model, equivalent to an exact CMH test on small sample sizes, will be used to provide the 2-sided p-value to determine whether proportion of responders is different between the tildacerfont treatment group and the placebo treatment group at Week 24. If the logistic model does not converge, a Fisher's exact tests corresponding to the tildacerfont treatment group compared to placebo will be utilized.

#### 12.5.1.2 Multiple Imputation for Missing Response Data

Subjects with missing Secondary Efficacy Endpoint data, data imputed as follows:

Step 1: The missing observations are imputed twenty times to generate complete datasets using SAS® PROC MI with a monotone logistic regression model based on the multiple imputation method. The model shall be:

$$Responder = Treatment + Sex + Baseline\ GC\ Dose\ group + Baseline\ GC\ Type + CAH\ Type + Response\ at\ Week\ 3, 6, 12\ and\ 18$$

Step 2: The analysis methods described in Section 12.3.3 will be applied to impute the remaining monotone missing data.

Step 3: The results from the twenty completed datasets are then combined using SAS® PROC MIANALYZE via Rubin's combination rules.

#### 12.5.1.3 Reporting Results

**Summary of Descriptive Statistics**

For each treatment group, the number and proportions (expressed as percentages) of responders at each scheduled time point will be calculated.

**Summary of Analysis Results**

The parameter estimates (i.e., least squares [LS] proportions, odds ratios, the 95% CI of odds ratios, nominal p-values from the logistic model will be reported at Week 24.

#### 12.5.1.4 Supportive Analysis CV Risk Secondary Efficacy Endpoint

The responder criterion is systematically varied in order to identify the criteria under which the treatment effect is different. If Secondary Efficacy Endpoint 2.3 is significant, then each successive alternative responder definition will be tested using the methods described in Section 12.5.1.1. For Endpoints 2.3A and 2.3C, subjects who are missing data required to assess CV risk factors at Week 24 will be considered to be non-responders.

**Table 14 Alternate Secondary Efficacy Endpoint CV Risk Factor Responder Definitions**

| Def | Endpoint | Alternative Definition |
|---|---|---|
| 2.3A | Proportion of subjects with improvement in at least 1 CV risk factor **AND no worsening** in a CV factor | $\begin{cases} 1 & improvement\ in\ \geq 1\ baseline\ CV\ risk\ factor\ at\ Week\ 24 \\ 0 & no\ improvement\ \ OR\ worsening \end{cases}$ |
| 2.3B | Proportion of subjects with improvement in at least 2 CV risk factor | $\begin{cases} 1 & no\ longer\ meets\ \geq 2\ baseline\ CV\ risk\ factor\ criteria\ at\ Week\ 24 \\ 0 & no\ improvement \end{cases}$ |
| 2.3C | Proportion of subjects with improvement in at least 2 CV risk factor **AND no worsening** in a CV factor | $\begin{cases} 1 & improvement\ in\ \geq 2\ baseline\ CV\ risk\ factor\ at\ Week\ 24 \\ 0 & no\ improvement\ \ OR\ worsending \end{cases}$ |

### 12.5.2 Subgroup Analyses

Subgroup analyses for Secondary Efficacy Endpoints 2.1, 2.2 and 2.3 will be performed using the planned subgroups, by Sex, by Baseline GC dose, by Baseline GC type, by Baseline A4 ULN group.

## 12.6 Exploratory Efficacy Endpoint Analyses

The analysis population for Exploratory Efficacy Endpoint analyses is the ITT analysis set, including all classic CAH subjects who were randomized into the study.

The following Exploratory Efficacy Endpoints will be described in this section:

**Table 15 Exploratory Efficacy Endpoints**

| Objectives | Endpoints | | Analysis |
|---|---|---|---|
| ***Exploratory Efficacy*** | | | |
| To evaluate the percentage change in GC use in subjects with CAH | 3.1 | Percent change from baseline in GC dose at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing the cumulative HCe dose in subjects with CAH | 3.2 | Change in total cumulative GC dose in HCe at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in improving HOMA-IR in subjects with CAH | 3.3 | Change from baseline in the HOMA-IR at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in improving HOMA-IR after 52 weeks of tildacerfont treatment in subjects with CAH | 3.4 | Change from baseline in HOMA-IR at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in reducing cardiovascular risk in subjects with CAH | 3.5 | Proportion of subjects with improvement in at least one cardiovascular risk factor at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 3.6 | Proportion of subjects with GC dose ≤ 11 mg/m$^2$/day and A4 ≤ 1.2x baseline or ⩽ ULN at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in reducing GC use to near-physiologic levels while maintaining androgen control in subjects with CAH | 3.7 | Proportion of subjects with baseline GC dose ≤ 35mg HCe who achieve GC dose ≤11 mg/m2/day in HCe and A4 ≤ 1.2x baseline or ≤ ULN at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in improving quality of life in subjects with CAH | 3.8 | Change from baseline at Week 24 in the SF-36 domains and composite scores | Primary Analysis |
| To evaluate the effect of tildacerfont in improving quality of life in subjects with CAH | 3.9 | Change from baseline at Week 52 in the SF-36 domains and composite scores | Final Analysis |
| To evaluate the effect of tildacerfont on BMI after 24 weeks in subjects with CAH | 3.10 | Percent change from baseline in BMI at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont on BMI after 52 weeks of tildacerfont treatment in subjects with CAH | 3.11 | Percent change from baseline in BMI at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont on waist circumference after 24 weeks of tildacerfont treatment in subjects with CAH | 3.12 | Change from baseline in waist circumference at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont on waist circumference after 52 weeks of tildacerfont treatment in subjects with CAH | 3.13 | Change from baseline in waist circumference at Week 52 | Final Analysis |

| Objectives | Endpoints | | Analysis |
|---|---|---|---|
| To evaluate the effect of tildacerfont in improving body composition after 24 weeks of tildacerfont treatment in subjects with CAH | 3.14 | Change from baseline in percent total fat mass and total percent lean mass at Week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in improving body composition after 52 weeks of tildacerfont treatment in subjects with CAH | 3.15 | Change from baseline in percent total fat mass and total percent lean mass at Week 52 | Final Analysis |
| To evaluate the effect of tildacerfont in improving BMD after 52 weeks of tildacerfont treatment in subjects with CAH | 3.16 | Change from baseline in BMD at Week 24 and Week 52 | Primary (Week 24) and Final Analysis |
| To evaluate the effect of tildacerfont in reducing TART(s) in male CAH subjects with TART(s) at baseline after 24 weeks of tildacerfont treatment in subjects with CAH | 3.17 | Reduction in TART volume at Week 24 in male subjects who had TART(s) at baseline | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing TART(s) in male CAH subjects with TART(s) at baseline after 52 weeks of tildacerfont treatment in subjects with CAH | 3.18 | Reduction in TART volume at Week 52 in male subjects who had TART(s) at baseline | Final Analysis |
| To evaluate the effect of tildacerfont in reducing GC use while maintaining normal A4 levels at Week 24 in subjects with CAH | 3.19 | Proportion of subjects with at least a 5 mg/day HCe reduction from baseline (Day 1) in GC dose and A4 $\leq$ 1.2x baseline or $\leqslant$ ULN at week 24 | Primary Analysis |
| To evaluate the effect of tildacerfont in reducing GC use while maintaining normal A4 levels at Week 52 in subjects with CAH | 3.20 | Proportion of subjects with at least a 5 mg/day HCe reduction from baseline (Day 1) in GC dose and A4 $\leq$ 1.2x baseline or $\leq$ ULN at week 52 | Final Analysis |

### 12.6.1 Analysis for Change and Percent Change from Baseline at Week 24 Exploratory Efficacy Endpoints 3.1, 3.3, 3.4 and 3.8 to 3.15

A MMRM or ANCOVA model will be used to analyze Exploratory Efficacy Endpoints 3.1, 3.3, 3.4 and 3.8 to 3.15 in the same manner as described in Section 12.3.3; MMRM will be for endpoints with more than one post-baseline assessment, and ANCOVA will be used for endpoints with only one post-baseline assessment at Week 24. The dependent variable will be the absolute or percent change from baseline and the principal treatment effect of interest is at 24 weeks. The LS mean of change from baseline, SE of the LS mean, two-sided 95% CI of the LS mean, LS mean difference of change from baseline between the tildacerfont group and the

placebo (i.e., tildacerfont – placebo), SE of the LS mean difference, and 95% CI of the LS mean difference and its 2-sided p-value will be extracted from the model.

In terms of descriptive statistics, the change from baseline will be reported by treatment group for the reported value and change from baseline value, and percent change from baseline value (if applicable) will be presented by scheduled time point (i.e., Baseline [reported value only] and Week 24).

#### 12.6.1.1 Sensitivity Analyses

Missing data sensitivity analyses to Exploratory Efficacy Endpoints 3.1, 3.3, 3.4 and 3.8 to 3.15 may be performed as per imputation strategies (a) and (b) as described Section 12.3.3.1.

#### 12.6.1.2 Subgroup Analyses

Subgroup analyses to Exploratory Efficacy Endpoints 3.1, 3.3, 3.4 and 3.8 to 3.15 will be performed as per Section 12.5.2.

### 12.6.2 Analysis for Change and Percent Change from Baseline at Week 52 Exploratory Efficacy Endpoints 3.9, 3.11, 3.13, 3.15 and 3.16

An ANCOVA model will be used to analyses Exploratory Endpoints 3.9, 3.11, 3.13, 3.15 and 3.16. The absolute or change or percent change from baseline (placebo-controlled baseline; see section 8.3.9) to 52 weeks will be the dependent variable. The ANCOVA model will include terms for treatment group (0: placebo, 1: tildacerfont); the randomization strata: sex (0: male, 1: female) and baseline GC dose (in HCe, 0: < 40 mg/day, 1: ≥ 40 mg/day); and baseline GC type (0: short-acting, 1: intermediate-acting or combination) as categorical explanatory variables. The difference in LSmeans between randomised treatment groups at Week 52 along with the associated CI and 2 sided p-value will be extracted.

Further, a paired t-test will be used to compare the change from tildacerfont-baseline to 52 weeks on treatment with tildacerfont.

In terms of descriptive statistics, for each type of analysis, the change from baseline will be reported by treatment group for the reported value and change from baseline value, and percent change from baseline value (if applicable) will also be presented.

#### 12.6.2.1 Sensitivity Analyses

Missing data imputation strategies (a) and (b) may be applied to Exploratory Efficacy Endpoints 3.9, 3.11, 3.13, 3.15 and 3.16 as described Section 12.3.3.1.

#### 12.6.2.2 Subgroup Analyses

Subgroup analyses may be performed for Exploratory Efficacy Endpoints 3.9, 3.11, 3.13, 3.15 and 3.16 as per Section 12.5.2.

### 12.6.3 Exploratory Efficacy Endpoints 3.5 to 3.7 and 3.17 to 3.20, Responder Analyses

Exploratory Efficacy Endpoints 3.5 to 3.7 and 3.17 to 3.20 are responder endpoints and will be analyzed in the same manner as Secondary Efficacy Endpoint 2.1.

#### 12.6.3.1 Multiple Imputation for Missing Data

Subjects with missing response data Efficacy Endpoints 3.5 to 3.7 and 3.17 to 3.20 may have missing data imputed as per Section 12.5.1.2.

#### 12.6.3.2 Subgroup Analyses

Subgroup analyses may be performed for Exploratory Efficacy Endpoints 3.5 to 3.7 and 3.17 to 3.20 as per Section 12.5.2.

### 12.6.4 Exploratory Efficacy Endpoint 3.2, Cumulative GC Dose

Cumulative GC dose in HCe over the period baseline to 24 weeks from randomization will be analysed via ANCOVA on the log scale. The model will include log cumulative GC dose in HCe over baseline to 24 weeks as the dependent variable; treatment group (0: placebo, 1: tildacerfont); the randomization strata: sex (0: male, 1: female) and baseline GC dose (in HCe, 0: $< 40$ mg/day, 1: $\geq 40$ mg/day); and baseline GC type (0: short-acting, 1: intermediate-acting or combination) as categorical explanatory variables. The difference in LSmeans between

randomised treatment groups along with the associated CI and p-value will be extracted. The LSmeans and CI will be back transformed for the purposes of presentation.

The descriptive statistics in terms of the cumulative GC dose will be reported by treatment group and will include the geometric mean, log scale SE, Q1, Q3, min and max.

#### 12.6.4.1 Sensitivity Analyses: Completers

In order to further evaluate the impact of early discontinuation on GC exposure, the cumulative GC dose in HCe analysis at 24 weeks will be repeated using only subjects who completed the full 24 weeks of treatment.

### 12.6.5 Additional Analyses

#### 12.6.5.1 Additional BMD parameters

For the change from baseline in bone mineral density, Exploratory Efficacy Endpoint 3.16, the following supportive analyses will be performed by repeating the analysis with the following alternative dependent variables:

1. Spine t-score
2. Spine z-score
3. Hip t-score
4. Hip z-score

Further, bone biomarkers, bone-specific alkaline phosphatase (BSAP), P1NP, C-terminal telopeptide of type I collagen (CTX-1), U-NTx, and Osteocalcin, will be summarized using standard lab summaries with p-values for change from baseline (paired t-test) and difference from placebo (Wilcoxon rank sum).

#### 12.6.5.2 DXA Body Composition Parameters

Under Exploratory Efficacy Endpoints 3.14 and 3.15, changes in body composition as measured by DXA from baseline to Week 24 and Week 52 will be compared between tildacerfont and placebo; these parameters are percent total fat, and total percent lean mass .

### 12.6.5.3 Additional Adrenal Biomarker Analyses

The following Week 24 and Week 52 analyses will provide additional context to aid in the interpretation of adrenal biomarker data.

**Table 16 Adrenal Biomarker Responder Definitions**

| Endpoint | Definition |
|---|---|
| Proportion of subjects with A4 ≤ ULN | $1 \quad A4 \leq ULN$<br>$0 \quad A4 > ULN \text{ or missing}$ |
| Proportion of subjects with A4 ≤ 1.5 x ULN | $1 \quad A4 \leq 1.5 \times ULN$<br>$0 \quad A4 > 1.5 \times ULN \text{ or missing}$ |
| Proportion of subjects with A4 ≤ 2 x ULN | $1 \quad A4 \leq 2 \times ULN$<br>$0 \quad A4 > 2 \times ULN \text{ or missing}$ |
| Proportion of subjects with 17OHP ≤ Target | $1 \quad 17OHP \leq Target$<br>$0 \quad 17OHP > Target \text{ or missing}$ |

In keeping with the analysis of Secondary Efficacy Endpoint 2.1, a logistic regression model will be used to evaluate the endpoints described in Table 17. The model will include the responder indicator as the dependent variable; treatment group (0: placebo, 1: tildacerfont); the randomization strata: sex (0: male, 1: female) and baseline GC dose (in HCe, 0: < 40 mg/day, 1: ≥ 40 mg/day); and baseline GC type (0: short-acting, 1: intermediate-acting or combination) as categorical explanatory variables. The baseline value for the adrenal biomarker will be included as a variable.

### 12.6.5.4 Additional PGIC and CGI-I Analyses

Both patients' (PGIC) and clinician's (CGI-I) global impression of improvement scales will be summarized at Weeks 24 and 76. Summaries will include the count and proportion of subjects at each point on the Likert scale as well as an 8-point descriptive summary. At Week 24, a t-test will be used to derive a p-value on the mean score of the placebo and tildacerfont groups.

#### 12.6.5.5 Key Laboratory and Vital data used in Clinical Outcome Components

Key laboratory and vitals parameters that are used to build key clinical outcomes will be summarized by treatment and post-baseline scheduled time point, based on randomized treatment received at each scheduled time point and will include p-values reporting change from baseline using paired-tests and difference from placebo using Wilcoxon rank sum tests.

- Fasting glucose, HbA1c, and fasting insulin
- Total cholesterol, LDL, HDL, triglycerides
- Blood pressure (systolic and diastolic)

#### 12.6.5.6 Compliance to A4 algorithm

Principal investigators may deviate due to medical or patient reasons from the protocol specified A4 GC change algorithm. A summary will be provided to compare adherence to the GC dosing algorithm by summarizing the proportion of assessment outcomes that followed the protocol algorithm by nominal time point, by treatment group and overall.

### 12.7 Figures

Efficacy data will be plotted for primary and key secondary endpoints. The following plots are planned:

- LS Means and 95% CI of Change from Baseline in GC dose by Visit
- Proportion of subjects with GC dose ⩽11 mg/m2/day in HCe and A4 ⩽ 1.2x baseline or A4 ⩽ ULN through Week 24
- Proportion of subjects with baseline GC dose ≤ 35mg HCe who achieve GC dose ≤11 mg/m2/day in HCe and A4 ≤ 1.2x baseline or ≤ ULN at Week 24
- Proportion of subjects with improvement in at least one cardiovascular risk factor at Week 24

Additional figures may be produced for other endpoints as deemed appropriate. The following plots are planned, and additional types (forest plots; cumulative distribution plots) may be provided to facilitate illustration of the study results.

# 13 SAFETY ANALYSES

Safety will be evaluated by AEs, SAEs, AESIs, clinical laboratory, vital, ECG findings, monitoring for suicide risk and depression/anxiety, as well as physical examination. All analyses of safety data will be performed using the safety analysis set, based on the actual treatment a subject received. All descriptive statistics will be presented by treatment group and treatment period.

## 13.1 Adverse Events

All reported AEs (including non-TEAEs) will be listed. Separate listings will be provided for SAEs, AESIs, and TEAEs leading to study drug discontinuation. If a subject reported the same TEAE on multiple occasions, the highest severity (severe > moderate > mild) or study drug relationship (related > probable > possible > unlikely > unrelated) recorded for the event will be summarized. All TEAE summary tables will present the number and percentages of subjects reporting TEAEs, unless otherwise specified.

### 13.1.1 Overall Summary of TEAEs

Overall summary of TEAEs will be presented by study period (i.e., Placebo-Controlled Treatment, and Open-Label Treatment) and treatment group. Subjects will be counted only once at the highest severity when summarizing TEAE by severity (severe > moderate > mild).

### 13.1.2 Summary of TEAEs by System Organ Class and Preferred Term

In addition, TEAEs will be summarized by MedDRA SOC and PT. Each summary table will be presented by study period (i.e., Placebo-Controlled Treatment, and Open-Label Treatment) and treatment group as well as overall by treatment group. A subject experiencing multiple occurrences of an adverse event will be counted only once for each PT. Subjects with multiple PTs within a SOC will be counted only once for that SOC. The summaries will be presented for

the following; if a summary does not include more than 20 events, summaries by PT only (Section 13.1.3) may be deemed to be sufficient:

- TEAEs by SOC and PT
- Severe TEAEs by SOC and PT
- Study drug-related TEAEs by SOC and PT
- Severe Study drug-related TEAEs by SOC and PT
- AESIs by SOC and PT
- AESI of liver Events
- Serious TEAEs by SOC and PT
- Serious Study drug-related TEAEs by SOC and PT
- TEAEs leading to study drug discontinuation by SOC and PT
- TEAEs leading to death by SOC and PT

### 13.1.3 Summary of TEAEs by Preferred Term

In addition, TEAEs will be summarized by MedDRA PT. Each summary table will be presented by study period (i.e., Placebo-Controlled Treatment, and Open-Label Treatment) and treatment group as overall by treatment group. A subject experiencing multiple occurrences of an adverse event will be counted only once for each PT. The summaries will be presented for:

- TEAEs by PT
- Severe TEAEs by PT
- Study drug-related TEAEs by PT
- Severe Study drug-related TEAEs by PT
- AESIs by PT
- Serious TEAEs by PT
- Serious Study drug-related TEAEs by PT
- TEAEs leading to study drug discontinuation by PT
- TEAEs leading to death by PT

## 13.2 Clinical Laboratory Evaluation

Table 17 shows clinical laboratory assessments (hematology, chemistry, lipid panel, thyroid panel, urinalysis, and other tests) will be performed. Laboratory parameters will be summarized

by treatment and post-baseline scheduled time point, based on actual treatment received at each scheduled time point.

**Table 17 Clinical Laboratory Parameters**

| Chemistry | Hematology | Urinalysis | Thyroid |
|---|---|---|---|
| Glucose, fasting | Platelet count | Specific gravity | T3 |
| Potassium | RBC count | pH | T4 |
| Calcium | MCV | Glucose | TSH |
| Sodium | MCH | Protein | |
| BUN | % reticulocytes | Blood | **Other** |
| Creatinine | Hemoglobin | Ketones | LH |
| Total protein | Hematocrit | Bilirubin | FSH |
| ALP | WBC count | Urobilinogen | SHBG |
| ALT/SGPT | Neutrophils | Nitrite | Renin |
| AST/SGOT | Lymphocytes | Microscopic examination | Aldosterone |
| GGT | Monocytes | | *Male:* Inhibin B |
| Total bilirubin | Eosinophils | **Lipid** | *Female:* Estradiol |
| Direct bilirubin | Basophils | Total cholesterol | *Female:* Prolactin |
| Total bile acids | | LDL | *Female:* Progesterone |
| PT/INR | | HDL | Hy's Law |
| PTT | | Triglycerides | |

Numeric results will be summarized using (8-point) descriptive statistics at baseline and at each scheduled post-baseline time point, unless otherwise specified. Changes from baseline at randomization will also be summarized by scheduled time point.

In addition, the number and percentage of subjects with potentially clinically significant abnormalities (PCSA, see Table 18) in selected chemistry and hematology parameters (i.e., above the specified upper limit or below the specified limit) will be summarized by scheduled time point.

The number and percentage of subjects who meet criteria for potential drug-induced liver injury as defined by Hy's law will also be summarized.

**Table 18 Potentially Clinically Significant Abnormalities by Laboratory Parameter**

| Parameter | PCSA Low | PCSA High |
|---|---|---|
| Chemistry | | |
| ALT | | >3 x ULN |
| AST | | >3 x ULN |
| ALP | | >1.5 x ULN |
| Total Bilirubin | | >1.5 x ULN |
| Total bile acids | | > 10 umol/L |
| Creatinine | | > 1.8 mg/dL or > 0.4 mg/dL increase from baseline (*> 159.16 µmol/L or > 35.37 µmol/L increase from baseline)* |
| Sodium | < 130 mEq/L (*< 130 mmol/L)* | >150 mEq/L (*> 150 mmol/L)* |
| Glucose | < 3 mmol/L *(< 54 mg/dL*) | |
| Potassium | < 3 mEq/L (*< 3 mmol/L)* | ≥5.5 mEq/L (*> 5.5 mmol/L)* |
| Calcium | | >11.4 mg/dL (*> 2.84 mmol/L)* |
| BUN | | > 30 mg/dL and > 10 mg/dL increase form baseline (*> 10.71 mmol/L and > 3.57 mmol/L)* |
| INR | | >1.5 x ULN |
| Total Bile Acids | | >5 x ULN or >3 x ULN if ALT is >3 x ULN |
| FSH | | >15 IU/L if baseline was <5 IU/L |
| Hematology | | |
| Hemoglobin | ≤10 g/dL and >2 g/dL reduction from baseline (*≤ 100 g/L and > 20 g/L)* | |
| Leukocytes | <500/µL (*< 0.5 x $10^9$/L)* | |
| Neutrophils | <1800/µL and >500/µL reduction from baseline | |
| Platelets | <50,000/µL (*< 50 x $10^9$/L)* | |

All laboratory results will be listed with reference ranges and range indicator (Low, High, or Clinically Significantly Abnormal). In addition, laboratory results that meet or exceed the pre-specified PCSA levels (Table 19) will be flagged.

### 13.3 Vital Signs

Descriptive statistics for heart rate (HR), respiratory rate, temperature, including baseline values and change from baseline values, will be summarized by treatment group and scheduled time

point. In addition, the number and percentage of subjects with potentially clinically significant abnormalities (PCSA, see Table 19) in selected vital sign parameters (i.e., above the specified upper limit or below the specified limit) will be summarized. All vital signs parameters will be listed.

**Table 19 Potentially Clinically Significant Abnormalities by Vital Sign Parameter**

| Vital Sign Parameters | PCSA Low | PCSA High |
|---|---|---|
| Heart Rate (bpm) | < 40 | > 110 |
| Systolic Blood Pressure (mmHg) | < 85 | > 160 |
| Diastolic Blood Pressure (mmHg) | < 45 | > 100 |

## 13.4 12-Lead Electrocardiogram

Electrocardiogram (ECG) data, such as clinical interpretation of ECGs, heart rate (HR) values, and interval assessments of QRS duration, QT interval, and the Fridericia's corrected value of the interval between the Q and T waves on the ECG tracing (QTcF) will be listed. Descriptive statistics for observed values and change from baseline at each scheduled time point will be presented for these 12-lead ECG interval and HR assessments.

In addition, the number and percentage of subjects with any abnormal values (i.e. outside a pre-specified threshold) will be summarized by treatment and scheduled time point. The pre-specified levels of ECG QTc thresholds are provided by Spruce (See Table 20 below).

**Table 20 Pre-Specified Threshold Levels for ECG Parameters**

| ECG Parameter | Pre-Specified Level |
|---|---|
| Heart Rate (bpm) | < 40, > 120, > 130 |
| Heart Rate Change from Baseline (bpm) | > 20, > 30 |
| QRS Interval (msec) | > 120 |
| QTcF (msec) Normal<br>Borderline<br>Prolonged | Males: ≤ 430, Females: ≤ 450<br>Male: > 430 to 450, Female > 450 to 470<br>Males: > 450, Females: > 470<br>Male and Female: > 450, > 457, > 458, > 500 |
| QTcF change from Baseline (msec) | ≤ 30, > 30 to 60, > 60 |

All ECG parameters will be listed with flags for the above pre-specified level. A separate listing of subjects with values of QTcF > 500 msec or an increase > 60 msec will be provided, as necessary.

### 13.5 Physical Examination

Abnormal clinically significant findings from physical examinations are reported as AEs.

### 13.6 Psychiatric Evaluations

#### 13.6.1 The Hospital Anxiety and Depression Scale

The HADS is a self-assessment questionnaire for detecting states of anxiety and depression in subjects in clinical trials. The total score of HADS, change from baseline in total score will be summarized by time point, treatment group, and study period.

#### 13.6.2 Columbia–Suicide Severity Rating Scale

All responses Baseline/Screening C-SSRS and Since Last Visit C-SSRS will be listed by subject and by visit. Tables will include results from the Suicidal Ideation and Suicidal Behavior sections of the Since Last Visit C-SSRS. Number and percentage of subjects with a response of "Yes" at any point on the Suicidal Ideation and Suicidal Behavior items will be summarized by treatment group and scheduled time point, if sufficient number (i.e., > 5 in any dose level) of subjects responded "Yes".

## 14 CHANGES RELATIVE TO THE PROTOCOL-SPECIFIED ANALYSIS

TBD when complete

**REFERENCES**

FDA. (2007). *Guidance for Industry Developing Products for Weight Management.* FDA.

Flint, A., Rexrode, K., Hy, F., Glynn, R. C., Manson, J., Willett, W., & Rimm, E. (2010). Body mass index, waist circumference, and risk of coronary heart disease: a prospective study among men and women. *Obes Res Clin Pract*, e171-81.

Grundy, S., Stone, N., Bailey, A., Beam, C., Birtcher, K., Blumenthal, R., . . . Yeboah, J. (2019). 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*, e1082-1143.

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. (1996). *Guideline for Industry: Structure and Content of Clinical Study Reports (ICH E3).* Rockville, MD: Federal Register.

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. (1998). *Guideline for Industry: Statistical Principles for Clinical Trials (ICH E9).* Rockville, MD: Federal Register.

Little, R., & Rubin, D. (2002). *Statistical Analysis with Missing Data.* Hoboken, NJ: John Wiley & Sons.

Matthews, D., Hosker, J., Rudenski, A., Naylor, B., Treacher, D., & RC, T. (1985). Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*, 412-9.

Muniyappa, R., Lee, S., Chen, H., & MJ, Q. (2008). Current approaches for assessing insulin sensitivity and resistance in vivo: advantages, limitations, and appropriate usage. *Am J Physiol Endocrinol Metab*, E15-26.

Shetty, P., & Kumanyika, S. (2008). *Waist circumference and waist-hip ratio: report of a WHO expert consultation.* WHO.

Torkey, A., Sinaii, N., Jha, S., Desai, J., El-Mouche, D., Mallappa, A., & Merke, D. (2021). Cardiovascular Disease Risk Factors and Metabolic Morbidity in a Longitudinal Stduy of Congential Adrenal Hyperplasia. *JCEM*.

Whelton, P., Carey, R., Aronow, W., Casey, D., Collins, K., Dennison Himmelfarb, C., Wright Jr, J. (2018). 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American College of Cardiology/American Heart Association Task F. *Hypertension*, 1269-1324.

WHO. (2008). *Waist circumference and waist-hip ratio: report of a WHO expert consultation.* WHO.

Rubin, D.B. (1987), Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc.