Statistical Analysis Plan

Title: A Randomized, Double-Blind, Active-Controlled, Phase 3 Study of Chronocort Compared with Immediate-Release Hydrocortisone Replacement Therapy in Participants Aged 16 Years and Over with Congenital Adrenal Hyperplasia

Protocol Number: DIUR-014

Protocol Version: 6.0 / 09 February 2023

Sponsor: Diurnal

SAP Version 4.0

SAP Issue Date: 22 Feb 2024

SAP Authors: [redacted]

Revision History

Version	Issue Date	Revision / Addition	Rationale
1.0	30-Nov-2021	Initial Release Version	
2.0	12-Jul-2022	Version 2.0	Updated for Protocol Version 5.0. In addition, definition of Efficacy Evaluable Set updated to only exclude subjects with major protocol deviations that are deemed to have an effect on the assessment of efficacy.
3.0	26-Jun-2023	Version 3.0	Updated for Protocol Version 6.0.
4.0	22-Feb-2024	Version 4.0	Update for Protocol Version 7.0, updates prior to the Database Lock

I confirm that I have reviewed this document and agree with the content.

Approvals		
Syneos Health Approval		
[redacted]		
Name, Title	Signature	Date (DD-Mmm-YYYY)
[redacted]		
[redacted]		
Name, Title	Signature	Date (DD-Mmm-YYYY)
[redacted]		
Diurnal Ltd Approval		
[redacted]		
Name, Title	Signature	Date (DD-Mmm-YYYY)
[redacted]		

Table of Contents

1 INTRODUCTION

This document details the planned statistical analyses for the Diurnal, protocol "DIUR-014" study titled "A Randomized, Double-Blind, Active-Controlled Phase 3 Study of Chronocort Compared with Immediate-Release Hydrocortisone Replacement Therapy in Participants Aged 16 Years and Over with Congenital Adrenal Hyperplasia".

The proposed analyses are based on the contents of the final version 7.0 of the protocol (dated 30 October 2023).

Congenital adrenal hyperplasia (CAH) is most commonly caused by 21-hydroxylase deficiency that results in cortisol deficiency, with or without aldosterone deficiency, and androgen excess. Patients are currently treated with glucocorticoid replacement therapy, but immediate release formulations fail to replicate the natural cortisol circadian rhythm, and patients often struggle to take a midday dose, resulting in patients being poorly controlled. Chronocort® is a newly developed modified release oral formulation of hydrocortisone, and it is designed to closely replicate the normal serum levels of the endogenous cortisol circadian rhythm, offering the prospect of an improved treatment outcome and a novel treatment paradigm.

The proposed study will evaluate whether a twice daily dosing regimen of Chronocort®, with two-thirds of the total daily dose given at night and one-third in the morning (which can more closely replicate circadian cortisol levels) will improve control of adrenal androgen production (as measured by 17-hydroxyprogesterone [17-OHP] and androstenedione [A4]) compared to a similarly titrated regimen of twice daily immediate release hydrocortisone replacement therapy (IRHC) in conditions broadly resembling routine clinical practice. The study design is presented in the schema in Figure 1.

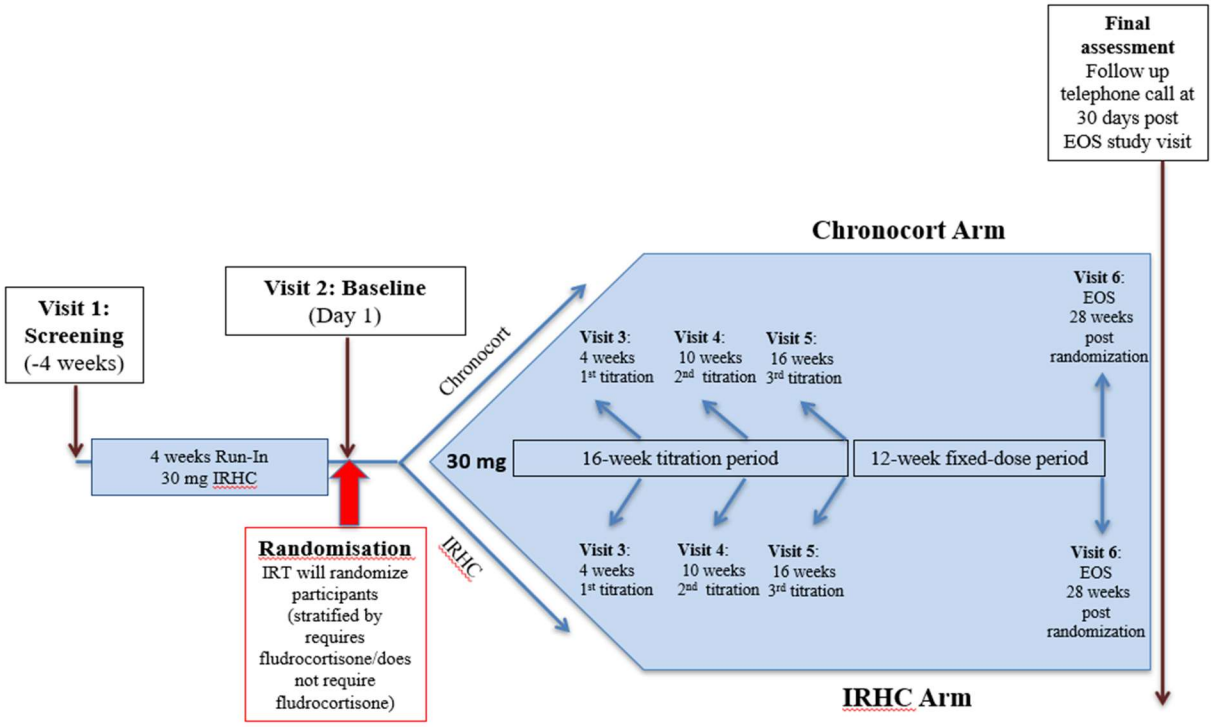
This is a randomized, double blind, active controlled, titrated, parallel arm, multicenter study. After a 4-week run-in period, participants will be randomized on a 1:1 basis to either Chronocort® or IRHC (Cortef®, Pfizer Inc.).

Version 1.0 of the statistical analysis plan (SAP) was developed and finalized before the first participant is screened and contained a full description of the statistical methods to be used. Versions 2.0 and 3.0 have updated the SAP in line with the subsequent protocol amendments.

Version 4.0 incorporated additional considerations for statistical analysis based on blinded review of the data prior to the database lock. Any changes from the analyses planned in Version 1.0 of the SAP are detailed in Section 9.

Figure 1

EOS = end of study; IRHC = immediate-release hydrocortisone; IR



T

= interactive response technology

Note: timelines not to scale

2 STUDY OBJECTIVES & ENDPOINTS/OUTCOME VARIABLES

Objectives	Outcome Variables and Analyses Supporting the Objectives
Primary Efficacy	
• To compare Chronocort to IRHC in terms of biochemical responder rate after 28 weeks of randomized treatment.	• The primary efficacy outcome variable is <u>whether or not the participant is a biochemical responder after 28 weeks of randomized treatment.</u> A biochemical responder is a participant who: i) is in biochemical control at the 08:00 hours assessment after 28 weeks of randomized treatment (wherein biochemical control is defined as both a 17-OHP concentration equal to or below the upper limit for optimal control and an A4 concentration equal to or below the upper limit of the reference range), and ii) is receiving after 28 weeks of randomized treatment a total daily dose of hydrocortisone of not more than 25 mg (if the participant was in biochemical control at baseline) or not more than 30 mg (if the participant was not in biochemical control at baseline). The difference (Chronocort minus IRHC) between the proportion of participants in each treatment arm who are biochemical responders 28 weeks after randomization will be estimated in the Full Analysis Set (FAS). Participants who die or withdraw from treatment before 28 weeks will be classified as not a biochemical responder. Participants receiving rescue medication under the 'stress dosing rules' within 5 days before the scheduled visit at 28 weeks after randomization will have their visit delayed appropriately[1]. Biochemical non-inferiority of Chronocort to IRHC will be declared if the 95% confidence interval (CI) for the difference in proportions lies wholly above minus 15 percentage points.
Key Secondary Efficacy	

• 1) To compare Chronocort to IRHC in terms of dose responder rate after 28 weeks of randomized treatment.	• The first key secondary efficacy outcome variable is <u>whether or not the participant is a dose responder after 28 weeks of randomized treatment.</u> A dose responder is a participant who: i) is receiving after 28 weeks of randomized treatment a total daily dose of hydrocortisone of not more than 25 mg, and ii) is in biochemical control at the 08:00 hours assessment after 28 weeks of randomized treatment (where in biochemical control is defined as both a 17-OHP concentration equal to or below the upper limit for optimal control and an A4 concentration equal to or below the upper limit of the reference range). The difference (Chronocort minus IRHC) between the proportion of participants in each treatment arm who are dose responders 28 weeks after randomization will be estimated in the FAS. Participants who die or withdraw from treatment before 28 weeks will be classified as not a dose responder. Participants receiving rescue medication under the 'stress dosing rules' within 5 days before the scheduled visit at 28 weeks after randomization will have their visit delayed appropriately[1]. Dose superiority of Chronocort to IRHC will be declared if the 95% CI for the difference in proportions lies wholly above zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared under the primary efficacy objective.
• 2) To compare Chronocort to IRHC in terms of total daily dose after 28 weeks of randomized treatment.	• The second key secondary efficacy outcome variable is <u>the total daily dose (mg) after 28 weeks of randomized treatment.</u> The difference (Chronocort minus IRHC) between the mean total daily dose after 28 weeks of randomized treatment in each treatment arm will be estimated in the FAS. Superiority of Chronocort to IRHC with respect to total daily dose after 28 weeks of randomized treatment will be declared if the 95% CI for the difference in means lies wholly below zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared under the primary

	efficacy objective and dose superiority of Chronocort to IRHC has been declared under the first key secondary efficacy objective.
Other Secondary Efficacy	
• To compare Chronocort to IRHC in terms of biochemical responders at 4, 10 and 16 weeks after randomization.	• The outcome variables whether or not the participant is a biochemical responder at 08:00 hours at 4, 10 and 16 weeks after randomization are compared between treatment arms by calculating the difference in proportion of participants responding. These outcome variables are to be analyzed in the same manner as the primary efficacy outcome variable.
• To compare Chronocort to IRHC in terms of dose responders at 10 and 16 weeks after randomization.	• The outcome variables whether or not the participant is a dose responder at 08:00 hours at 10 and 16 weeks after randomization are compared between treatment arms by calculating the difference in proportion of participants responding. These outcome variables are to be analyzed in the same manner as the key secondary outcome variable.
• To compare Chronocort to IRHC in terms of total daily dose at 10 and 16 weeks after randomization.	• The outcome variables total daily dose at 10 and 16 weeks after randomization are compared between treatment arms by calculating the difference in mean total daily dose. These outcome variables are to be analyzed in the same manner as the second key secondary outcome variable.
• To compare Chronocort to IRHC in terms of biochemical control at 4, 10, 16 and 28 weeks after randomization.	• The outcome variables whether or not the participant is in biochemical control (provided total daily dose is not more than 30 mg) at 08:00 hours at 4, 10, 16 and 28 weeks after randomization are compared between treatment arms by calculating the difference in proportion of participants in control.
• To compare Chronocort to IRHC in terms of the impact on 17-OHP range.	• The difference in range as calculated as the difference between the 08:00 and 13:00 measurements of 17-OHP levels at 4, 10, 16 and 28 weeks after randomization and their changes from baseline will be summarized and compared between treatment arms.

• To compare Chronocort to IRHC in terms of the impact on A4 range.	• The difference in range as calculated as the difference between the 08:00 and 13:00 measurements of A4 levels at 4, 10, 16 and 28 weeks after randomization and their changes from baseline will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on mean 17-OHP and A4.	• The mean of the 08:00 and 13:00 measurements of 17-OHP levels and A4 levels at 4, 10, 16 and 28 weeks after randomization and their changes from baseline will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on glucocorticoid dose.	• The total daily glucocorticoid dose at 4, 10, 16 and 28 weeks after randomization will be summarized and compared between treatment arms. • The relationship between daily glucocorticoid dose and biochemical control at 28 weeks after randomization will be explored.
• To compare Chronocort to IRHC in terms of changes in menstrual regularity.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in menstrual regularity (only in pre-menopausal women without hysterectomy and not using hormonal contraception) will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on luteinizing hormone (LH) levels.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in LH levels (men only) will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on testicular adrenal rest tumors (TART) size.	• The change from baseline to 28 weeks of randomized treatment in size of TART (men only) will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on sperm count.	• The change from baseline to 28 weeks of randomized treatment in sperm count (men only) will be summarized and compared between treatment arms.

• To compare Chronocort to IRHC in terms of the impact on subjective hirsutism in female participants.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in subjective hirsutism (women only) will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on objective hirsutism in female participants.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in objective hirsutism using the Ferriman-Gallwey score (women only) will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on subjective acne in female participants.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in subjective acne using a visual analogue scale (VAS) (women only) will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on objective acne in female participants.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in objective acne using the Global Evaluation Acne (GEA) scale (women only) will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on glycated hemoglobin (HbA1c) levels.	• The change from screening to 4, 10, 16 and 28 weeks of randomized treatment in HbA1c levels will be summarized and compared between treatment arms.

• To compare Chronocort to IRHC in terms of the impact on waist circumference.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in waist circumference will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on body weight.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in body weight will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on quality of life (QoL) using the self-completed Medical Outcome Study 36-Item Short Form Health Survey (SF-36®) total score and the sub-domain of vitality.	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in QoL using the self-completed SF-36® total score and the sub-domain of vitality will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on QoL using the Multidimensional Assessment of Fatigue (MAF).	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in QoL using the MAF will be summarized and compared between treatment arms.
• To compare Chronocort to IRHC in terms of the impact on QoL using the 5-level Standardized Health Questionnaire (EQ-5D-5L™).	• The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in QoL using the EQ-5D-5L™ will be summarized and compared between treatment arms.
Compliance	
• To assess compliance over the study period.	• The percentage overall treatment compliance will be summarized.
Exploratory Efficacy	
• [redacted]	[redacted]

[redacted]	[redacted]
[redacted]	[redacted]
Safety	•
• To assess the safety and tolerability of Chronocort relative to IRHC.	• The incidence, nature, severity, relatedness, duration, outcome, seriousness and expectedness of treatment emergent adverse events (TEAEs) will be tabulated by treatment arm. AEs of special interest will additionally be tabulated separately, with particular note of adrenal crises.
• To assess the need for use of additional glucocorticoid doses (stress dosing rules).	• The use of medication from the stress dosing packs or use of any additional glucocorticoid treatment during the study will be tabulated by treatment arm.
• To evaluate the safety of Chronocort compared to IRHC by assessment of routine safety laboratory assessments, physical examination, vital signs, and electrocardiogram (ECG).	• Safety hematology, biochemistry, urinalysis, and vital signs conducted at each post-randomization visit and the physical examination and ECG assessments conducted after 28 weeks of randomized treatment, and their changes from baseline, will be summarized by treatment arm.

[1] This description aligns with the ICH E9 requirement for the definition of an Estimand.

Outcome variables are underlined

Note: The reference range for A4 is up to 150 ng/dL (5.2 nmol/L) for men and up to 200 ng/dL (7.0 nmol/L) for women (based on the luteal phase) (based on data from The Mayo Clinic).
The optimal range for 17-OHP in adults (post-pubertal males or females aged >16 years) is up to 1200 ng/dL (36.4 nmol/L).

3 SAMPLE SIZE

Sufficient participants will be screened to achieve approximately 50 participants randomly assigned to either Chronocort or IRHC (25 participants per treatment arm, regardless of baseline strata), or until a cut-off date of 30 April 2023, whichever is reached first. Individual study sites should not randomize more than 8 participants without first consulting the Sponsor. A sample size of 25 participants per treatment arm will provide at least 80% power to demonstrate non-inferiority of Chronocort to IRHC when the biochemical response rate in the IRHC arm is 40%, the non-inferiority margin is 15%, and the biochemical response rate in the Chronocort arm is 68% using a 1-sided continuity corrected z-test with unpooled variance and a 1-sided alpha of 2.5%. Sample size is calculated using PASS 2022 (v22.0.3).

Note: Basing the sample size calculation on a continuity corrected z-test provides a conservative estimate of sample size relative to the test, based on a logistic model adjusted for stratification factors, which will actually be used.

Note: Post-hoc analysis of study DIUR-005 showed that ~40% of participants receiving the control treatment (that is: their pre-randomization treatment) would have been down-titrated by the proposed rules for the current study at four weeks after randomization. Estimating therefore the response rate for the IRHC as 40% and expecting that no participants withdrawn completely from hydrocortisone treatment would qualify for down-titration, the estimated treatment benefit of IRHC is a gain in response rate of 40%. Setting the non-inferiority margin as minus 15% thus retains five eighths of the anticipated treatment benefit of IRHC.

4 RANDOMIZATION & UNBLINDING

This is a double-blinded study with 2 treatment arms. The blinded treatment to be taken by a participant will be assigned using an IRT on a 1:1 basis (Chronocort:IRHC). Before the study is initiated, the login information and instructions for use of the IRT will be provided to each site. The randomization list will be generated by the study statistician using a computer-generated central randomization list using balanced blocks. The site will access the IRT at the baseline visit (Day 1) and they will record the randomization number allocated to the participant on the applicable eCRF. Randomization will be balanced within site and stratified by the participant's fludrocortisone replacement status (requires fludrocortisone/does not require fludrocortisone).

Potential bias will be reduced by the following steps:

- Records will be kept of any participants screened, whether or not they enter the study. This will enable an assessment of the representative nature of the participants randomized.
- Both study treatments will be over-encapsulated so that they look the same, so both the participant and the Investigator will not know which treatment group the participant has been randomized to.
- The titration of each participant’s dose of study treatment will be the same in both arms and will be based on the adrenal insufficiency checklist and androgen sampling, thus ensuring that opportunities for optimization and control of androgens are the same in both arms.
- The dose titration decisions will be made by an independent blinded physician who will be unaware of the treatment allocation to each participant.
- The treatment packs for each participant will be pre-packed and packs issued by the IRT to ensure that each participant receives the correct number of active and placebo capsules.

After each visit, new supplies will be provided to the participant that contains the same number of capsules as the previous supplies (with placebo capsules masking any changes), regardless of whether the dose was titrated or not, thus ensuring the participant is not aware of any dose changes.

- The primary endpoint is an objective biochemical measure that is not subject to bias.
- Efforts will be made to ensure that all participants remain in the study and are evaluable.
- There will be rigorous monitoring of sites to ensure compliance with all study procedures.
- The statistical analysis plan (SAP) version 1.0 has been completed before screening of the first participant to the study. While it is not reasonable to assume that updates will not occur to the SAP after this point and prior to unblinding of the study, an initial record of planned analysis prior to any participant activity is provided.
- Treatment will be allocated using block randomization. The block size will be determined by the statistician and will not be communicated to the study team.
- To minimize bias whilst the study is ongoing, the study team will remain blinded to aggregated and individual data prior to database lock and any data reviews will use a dummy randomization list with generic labels.
- Participants and Investigators will be requested not to reveal details of the study on social media.

Participants will be allocated an identifying number sequentially per site at the screening visit (Visit 1) [redacted] When the participant is randomized at the baseline visit (Visit 2), a separate randomization number will be allocated, which will be used internally for the purposes of treatment allocation and stratum identification. The screening number will still be used for participant identification purposes, so there may be gaps in participant ID numbers for screen failures. Randomization numbers will be assigned sequentially within each stratum and site as soon as participants fully satisfy all of the inclusion and none of the exclusion criteria. Screening and randomization numbers once assigned to a participant must not be re-assigned to another participant. Re-screened participants should be assigned the same participant number as they were assigned during their initial screening visit.

5 PLANNED ANALYSES

The SAP supersedes the protocol in case of any differences; a list of any changes from the analyses presented in the protocol is provided in Section 8. The SAP provides a robust plan of the analyses to be conducted for the trial; however, it is recognized that no SAP prepared in advance of the data can be absolutely definitive, and the final Clinical Study Report (CSR) may

contain additional tables or statistical tests if warranted by the data obtained. The justification for any such additional analyses will be fully documented in the final CSR.

The SAP describes all the analyses and summaries that will be performed on the total participant populations. In addition, all analyses and summaries will be repeated on the subgroup of French participants only.

5.1 Analysis Sets

For purposes of analysis, the following populations are defined:

Population	Description
Screened Set	All participants who sign the ICF.
Safety Analysis Set (SAF)	All participants who received, either as part of the run-in or post-randomization periods, at least 1 dose of Chronocort or IRHC. Participants will be analyzed according to the treatment they actually received. The SAF is a subset of the Screened Set.
Full Analysis Set (FAS)	All participants with CAH who were randomized into the study and who received at least 1 dose of post randomization study treatment (Chronocort or IRHC). Participants in the FAS will be analyzed according to randomized treatment. The FAS is a subset of the SAF.
Efficacy Evaluable Set (EES)	All participants in the FAS who had Visit 2 and Visit 6 (EOS) 17-OHP and A4 measurements taken at 08:00 hours, and who have no critical protocol deviations that affect the assessment of efficacy. The EES is a subset of the FAS.

Participants excluded from the analysis sets and the reason for their exclusion will be listed, and further details regarding exclusion decisions will be documented during the Data Review Meeting.

5.2 Derived Data

This section describes the derivations required for statistical analysis. Unless otherwise stated, variables derived in the source data will not be re-calculated.

5.2.1 Race

Where more than one race category has been selected for a participant, these race categories will be combined into a single category labeled "Multiple Race" in the summary tables. The listings will reflect the original selected categories.

5.2.2 Baseline

Baseline is defined as the last non-missing value (either scheduled, unscheduled or repeat) before the participant receives the first dose of study drug following randomization. Generally, this will be the assessment performed at the Baseline visit (V2).

For androgens / biochemical control status and difference / average of 8:00 and 13:00 17-OHP and A4 results, baseline will be defined as value obtained at Baseline visit (V2).

5.2.3 Duration / Study Day / Time

Study day will be calculated as the number of days from first dose of *study drug* after randomization

- date of event – date of first dose of study drug + 1, for events on or after first dose
- date of event – date of first dose of study drug, for events before first dose

As the protocol baseline is at Day 1 (i.e. there is no Day 0), any derivation of target day for a scheduled visit and/or applicable visit windows per protocol is in line with this approach. Durations of events related to an anchor point of start of the event will be calculated in a similar manner, replacing date of first dose of study drug with date of start of the event.

5.2.4 Quality of Life: SF-36, MAF, EQ-5D

QoL questionnaires (SF-36®, MAF, and EQ-5D-5L™) will be completed by the participants at Visits 2, 3, 4, 5 and 6 through an electronic data collection tool. No missing data will be allowed by the system and therefore no missing data imputation method will be needed.

The scores for the SF-36 questionnaire for each domain and total will be calculated using the Optum software. The total scores for MAF and EQ-5D will be derived programmatically as per their respective user guides. See Section 11.1 for further details of derivation.

5.2.5 Participants Recruited before Implementation of Protocol Version 6.0

The week 28 visit was not scheduled prior to protocol version 6.0. Therefore some participants recruited before the implementation of protocol version 6.0 may be missing the week 28 assessment. In such cases, should the week 34 assessment be available, this will be used for the week 28 assessment. Therefore, where a week 34 assessment is available, no imputation for the missing week 28 assessment will occur. Furthermore, for participants recruited before the implementation of protocol version 6.0, if such a participant does not have a week 28 assessment, but their end of study is between week 16 and week 34, then the end of study is mapped to week 28.

Note, the following conventions will apply for the efficacy and safety analyses.

For efficacy data, all summaries will be presented by all available scheduled visits. In these summaries, week 28 and week 34 will be presented separately. In addition, a further summary, and where applicable analysis, will be performed for week 28 where week 34 assessments are mapped to week 28 as described above. Note, this will be the main analysis for inference.

For safety data, all summaries will be presented by all available scheduled visits. No mapping of week 34 to week 28 will occur.

5.2.6 Conventions for Missing and Partial Dates

All rules explained below for partial / missing dates will be followed unless contradicted by any other data recorded on the electronic Case Report Form (eCRF).

All dates presented in the individual participant listings will be presented as recorded on the eCRF (i.e., not completed as per the below rules).

5.2.7 Missing / Partial Start / Stop Date of Adverse Events (AE) and Concomitant Medications

Missing and partial start and stop date will be imputed for analysis purposes as follows.

Partial or missing stop date will be imputed as follows:

If the stop date is completely missing and the event has resolved, or the participant has stopped taking the concomitant medication, the stop date will be imputed as the date of the participant's last clinic visit in the study.

- If only the year is known, the stop date will be imputed as "31-Dec" of that year or as the date of the participant's last clinic visit in the study if in the same year.
- If the month and year are known, the stop date will be imputed as the last day of that month unless the stop date corresponds to the same month as the participant's last clinic visit in which case the date of participant's last clinic visit in the study will be used instead.

Missing start date will be imputed as follows:

- If the stop date occurs on or after the start of study drug or the event / concomitant medication is ongoing, the start date will be imputed as the date of the first dose of study drug.
- If the stop date occurs before the start of study drug, the start date of the event / concomitant medication will be imputed as the participant's screening date or the stop date of the event / concomitant medication whichever is the earlier.

Partial start date (year present, but month and day missing)

- If the stop date occurs on or after the start of study drug or the event / concomitant medication is ongoing, and the year is the same as the year of first dosing the start date will be imputed as the date of the first dose of study drug. If the year is different from the year of first dosing "01-Jan" will be used.
- If the stop date occurs before the start of study drug, the start date of the event / concomitant medication will be imputed as the "01-Jan" of the same year.

Partial start date (month and year present, but day missing)

- If the stop date occurs on or after the start of study drug or the event / concomitant medication is ongoing, the start date will be imputed as the first day of the same month and year unless this partial start date is in same month as the first dose of study drug in which case the date of first dose of study drug will be used.
- If the stop date occurs before the start of study drug, the start date will be imputed as the first day of the month and year of the partial stop date.

If the start time is missing it will be imputed only in the case where the start date of the concomitant medication / event corresponds to the date of the first dose of study drug. The time will be imputed as the same time as the first dose of study drug. In all other cases the time will not be imputed.

5.2.8 Missing Last Dates of Study Drug Dosing

If the date of last dose of study drug is completely missing, then the date of last dose of study drug will be taken for analysis purposes as the date when the participant would have run out of study drug assuming full compliance from the date the study drug was last dispensed or the date of participant's last clinic visit in the study or early withdrawal or death whichever the earlier.

If only the month and year of the last dose was recorded, then the date of last dosing will be taken for analysis purposes as the date the participant would have run out of study drug assuming full compliance from the date the study drug was last dispensed, the last day of the month of the recorded last dose or the date of participant's last clinic visit in the study or early withdrawal or death whichever the earlier.

5.2.9 Missing Diagnosis Dates

If the month and year are present but the day is missing, for the purpose of calculating duration of disease at baseline/randomization, the diagnosis date will be set to first day of the relevant month. If only the year is recorded the diagnosis date will be set as "01-Jan" for that year.

5.2.10 Exposure to Study Drug

The following derivations will be implemented with respect to study drug exposure:

- Total treatment duration (days) = (date of last dose of study drug – date of first dose of study drug)+1
- Total glucocorticoid exposure (days) = sum of all periods of glucocorticoid treatment, where each period is calculated as (date of last dose of glucocorticoid in period of treatment – date of start of glucocorticoid treatment during period) +1

5.2.11 Inexact Values

In the case where a variable is recorded as "> x", "≥ x", "< x" or "≤ x", a value of x will be taken for analysis purposes.

5.2.12 Electrocardiogram (ECG) Data

For ECG data recorded on continuous scales, if more than one value is recorded at a time point, the mean value rounded to the integer will be presented. For overall interpretation if more than one value is recorded, the most severe (worst case) of the respective readings will be taken.

5.2.13 Visit Windows

There are no plans to derive visit windows, and visits will be used in the analyses as reported on the eCRF. For participants prematurely withdrawing from the study, the early withdrawal visit will be remapped to the next scheduled visit after the last dose. Note, for participants with an end of study visit between weeks 28 and 34, the end of study visit will be remapped to week 28 for the analysis combining the week 28 and 34 results, see Section 5.2.5. Note, for participants recruited before implementation of protocol version 6.0 (see Section 5.2.5), week 34 assessments used in lieu of week 28 assessments will be used for analysis for the EES, as will any assessments from an end of study visit between week 28 and week 34.

5.2.14 Unscheduled Visits

Only scheduled post-baseline values will be tabulated. Post-baseline repeat / unscheduled assessments will in general be disregarded, although unscheduled visits will be included in the derivation of shifts to "worst" post-baseline value. All post-baseline assessments (scheduled and unscheduled) will be listed in the relevant appendices to the CSR.

5.2.15 Randomization Strata

Randomization of participants to treatment arm is stratified by fludrocortisone treatment and where applicable, the analyses will be stratified by baseline biochemical control status.

5.3 General Considerations

5.3.1 Tables and Listings

All data listings, summaries, figures and statistical analyses will be generated using SAS [SAS Institute Inc., Cary, NC, 27513, USA] version 9.4 or higher[1].

Summaries will be presented by treatment group (Chronocort vs IRHC).

Overall columns are to be included within the table shells as follows:

Disposition	Treatment group and overall
Demography and Baseline Characteristics	Treatment group and overall
Prior, Concomitant, Subsequent and Rescue Medications	Treatment group and overall
Exposure to Study Drug	Treatment group only

Efficacy	Treatment group only
AEs	Treatment group and overall
Other safety	Treatment group only

Listings will be sorted in the following order: treatment group, participant, parameter, and visit unless otherwise stated. All data will be listed; participants who were not randomized will be displayed after the randomized treatment groups as "Not randomized to study treatment".

Continuous variables will be summarized by the number of non-missing observations, mean, median, standard deviation, quantiles and minimum and maximum. For tabulations of changes from baseline data for efficacy endpoints, the lower and upper 95% confidence limits for the mean for the individual treatments will be presented.

Categorical variables will be summarized by presenting the frequency and percent. Percentages will be based on the number of non-missing observations or the participant population unless otherwise specified. For each variable, all categories will be shown. Zero frequencies (but not the percent) within a category will be presented.

5.3.2 Decimal Places

Decimal places for derived data described in Section 5.2 will be determined by the scale of measurement unless otherwise stated. No decimal places will be displayed if the smallest calculated value is ≥ 100; 1 decimal place will be displayed when the smallest value is within the interval (10, 100), with 10 being inclusive; 2 decimal places will be displayed when the smallest value is within (1, 10), with 1 being inclusive; and so on for even smaller scales of measurement.

Derived data where it is known in advance the result will be an integer, for example day, month, year, number of days and total scores (for rating scales) will be presented with zero decimal places.

Means, medians and percentiles will be displayed to one more decimal place than the data**,** dispersion statistics (e.g. standard deviation) will have two more decimal places, and the minimum and maximum will be displayed to the same number of decimal places as reported in the raw data. Percentages will be displayed with one decimal place.

P-values will be quoted to 3 decimal places. P-values < 0.001 will be presented as "p<0.001". Where this value is less than 0.05 (two-sided test), 0.025 (one-sided test), 0.01 or 0.001, attention will be drawn to this fact using the conventional "*", "**" or "***" annotation, respectively. These conventions will be footnoted in the table or figure shell.

5.4 Participant Disposition

Participant disposition will be summarized descriptively as follows:

- The number and percentage of participants achieving the following criteria will be summarized by treatment group and overall for the Screened Set:
  - Entering the study or failed screening, with reasons
  - Multiple screened and a) entering the study or b) persistent failed screening
  - Received IRHC in the run-in period but were not randomized, with reasons
  - Randomized
  - Included and excluded from each analysis set, with reasons for exclusion
  - Completed or prematurely discontinued treatment, with reasons (including the number of participants with premature discontinuation of study treatment, that continued study assessments)
  - Completed or prematurely discontinued study, with reasons
- As a separate summary, the number of participants included in each analysis set will be summarized for each region, country and study site by treatment group and overall for the Screened Set.
- The number of participants who complete each phase of the study and the number of participants present at each scheduled visit will be summarized by treatment group and overall for the FAS.

Participant disposition will be listed for the Screened Set. The reasons for exclusion from each of the analysis sets will also be listed.

5.5 Protocol Deviations

A critical protocol deviation is a protocol deviation (PD) with the potential to compromise the efficacy analysis. Critical PDs may include, but are not limited to:

- Randomization criteria violations.
- Inclusion/exclusion criteria violations.
- Inadequate compliance with study drug.
- Prohibited medications taken.
- Significant deviations from the study drug administration schedule.
- Other protocol deviations that could affect participants' efficacy outcomes.

A Data Review Meeting will be convened after the data has been cleaned and the database soft-locked but before final database lock and analysis has commenced. As part of the review, a determination will be made on whether PDs are classed as 'critical', 'major', 'minor' or not a PD

at all. The review will also document the allocation of participants to analysis sets, among other items.

A listing of all PDs defined as major and/or critical will be presented. Participants experiencing at least one major and/or critical PD, and the reasons for deviation will be summarized by treatment group and overall.

5.6 Demographic and Other Baseline Characteristics

Demographic and other baseline variables will be summarized by treatment group, split by baseline biochemical control status, using the FAS. The comparability of treatment groups with respect to participant demographics and baseline characteristics will be assessed in a descriptive manner; no formal statistical testing will be performed.

Standard continuous or categorical variable summaries will be presented by randomized treatment group for the following variables. Participant listings will also be presented.

- Demographic data (e.g. age, sex, race, ethnicity, height (cm), weight (kg), BMI (kg/m^2), fludrocortisone replacement status [as stratified for randomization, and actual if different], baseline biochemical control)
- Disease history (e.g. duration since diagnosis, hospitalization and number of adrenal crises within the 12 months prior to randomization, previous CAH medication)
- Medical history by MedDRA System Organ Class (SOC) and Preferred Term (PT)

All variables collected only at baseline or prior to baseline will be summarized here. Variables that are collected pre and post-baseline (e.g. vital signs, labs, physical exam, ECGs and efficacy) will be presented at Screening and/or Baseline in the by-visit tables within their respective analysis section (i.e. under efficacy or safety).

5.7 Medical History

Separate tabulations of previous and ongoing conditions at screening will be presented by treatment group and overall for the FAS. Conditions will be presented according to Medical Dictionary for Regulatory Activities (MedDRA, version 20.0 or later) primary system organ class and preferred term. The dictionary version used will be displayed in the footnote of the presented data displays.

5.8 Other Baseline Characteristics

Pregnancy test performed at screening, baseline and at other scheduled visits for females of childbearing potential will be listed only.

5.9 Covid-19 Visit Impact

At each visit, using the FAS, a summary will be provided of any impact on the visit due to Covid-19. The number of participants with an impact will be summarized, as well as what the impact was (missed visit, interrupted and/or out of window in person visit at site, remote visit, other).

5.10 Exposure to Study Drug

Extent of exposure will be presented using total treatment duration (days) by treatment group for the SAF.

The number and percentage of participants experiencing a study drug dose change (either increase or decrease) during the study, plus reasons for change, will be summarized by treatment group for the SAF. Total glucocorticoid exposure during the study will also be summarized by treatment group for the SAF.

5.11 Efficacy Analyses

Analyses will be stratified or adjusted, as appropriate, on the basis of (i) the participant's fludrocortisone replacement status (requires fludrocortisone/does not require fludrocortisone) and (ii) whether they are in biochemical control at baseline or not (based on the androgen results from the baseline visit [Visit 2]).

5.11.1 Statistical Hypotheses

5.11.1.1 Primary Efficacy Endpoint: Biochemical Response after 28 Weeks of Randomized Study Treatment

A participant is a biochemical responder if:

i) they are in biochemical control at 08:00 hours after 28 weeks of randomized treatment (Visit 6)
where 'biochemical control' is defined as a 17-OHP concentration equal to or less than the upper limit of optimal control and a A4 concentration equal to or less than the upper limit of the reference range - if either or both these conditions are not met the participant is not in biochemical control, **AND**

ii) their total daily dose after 28 weeks of randomized treatment is not more than 25 mg (if the participant was in biochemical control at baseline) and not more than 30 mg (if the participant was not in biochemical control at baseline).

The primary null hypothesis is that the response rate in the Chronocort arm is worse than or equal to the response rate in the IRHC arm minus 15 percentage points. The primary alternate hypothesis is that the response rate in the Chronocort arm is better than the response rate in the IRHC arm minus 15 percentage points. Biochemical non-inferiority of Chronocort to IRHC will be declared if the 95% CI for the difference in response rates between the 2 treatment arms (Chronocort minus IRHC) is wholly above minus 15 percentage points.

5.11.1.2 First Key Secondary Efficacy Endpoint: Dose Response after 28 Weeks of Randomized Study Treatment

A participant is a dose responder if:

i) their total daily dose after 28 weeks of randomized treatment is not more than 25 mg, **AND**

ii) they are in biochemical control at 08:00 hours after 28 weeks of randomized treatment (Visit 6)
where 'biochemical control' is defined as a 17-OHP concentration equal to or less than the upper limit of optimal control and a A4 concentration equal to or less than the upper limit of the reference range - if either or both these conditions are not met the participant is not in biochemical control.

The first key secondary null hypothesis is that the response rate in the Chronocort arm is worse than or equal to the response rate in the IRHC arm. The first key secondary alternate hypothesis is that the response rate in the Chronocort arm is better than the response rate in the IRHC arm.

Dose superiority of Chronocort to IRHC will be declared if the two-sided 95% CI for the difference in response rates between the 2 treatment arms (Chronocort minus IRHC) is wholly above zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared under the primary efficacy objective.

5.11.1.3 Second Key Secondary Efficacy Endpoint: Mean Total Daily Dose of Hydrocortisone after 28 Weeks of Randomized Study Treatment

The second key secondary null hypothesis is that the mean total daily dose of hydrocortisone after 28 weeks of randomized treatment in the Chronocort arm is higher than or equal to the mean total daily dose after 28 weeks of randomized treatment in the IRHC arm. The second key secondary alternate hypothesis is that the mean total daily dose after 28 weeks of randomized treatment in the Chronocort arm is lower than the mean total daily dose after 28 weeks of randomized treatment in the IRHC arm.

Superiority of Chronocort to IRHC (with respect to the total daily dose after 28 weeks of randomized treatment) will be declared if the two-sided 95% CI for the difference in means between the 2 treatment arms (Chronocort minus IRHC) is wholly below zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared under the primary efficacy objective, and dose superiority of Chronocort to IRHC has been declared under the first key secondary efficacy objective.

5.11.1.4 Overview of Analyses for Efficacy Endpoints

Analysis	Statistical Analysis Methods
Primary analysis	Non-inferiority: a 95% CI will be calculated for the difference (Chronocort minus IRHC) in biochemical responder rates. Biochemical non-inferiority will be declared if the 95% CI lies wholly above minus 15%. The primary analysis will be conducted using the FAS.
First key secondary analysis	Superiority analysis: a 95% CI will be calculated for the difference (Chronocort minus IRHC) in dose responder rates. Dose superiority will be declared if the 95% CI lies wholly above zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared under the primary efficacy objective. The first key secondary analysis will be conducted using the FAS.
Second key secondary analysis	Superiority analysis: a 95% CI will be calculated for the difference (Chronocort minus IRHC) in mean total daily dose after 28 weeks of randomized treatment. Superiority with respect to total daily dose after 28 weeks of randomized treatment will be declared if the 95% CI lies wholly below zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared under the primary efficacy objective and dose superiority of Chronocort to IRHC has been declared under the first key secondary objective. The second key secondary analysis will be conducted using the FAS.
Other analyses of the primary outcome variable and of the key secondary outcome variables (classed as secondary outcome analyses)	Analyses of the primary outcome variable other than the primary analysis and analyses of the key secondary outcome variables other than the key secondary analyses, such as: • analysis at other timepoints • assessment of the effect of baseline and disease characterizing variables, including stratification variables • analysis using the EES and subgroups. are deemed to be supplementary analyses and will be performed using the same statistical methodology as will be used for the primary analysis and key secondary analyses. Calculated p-values will be indicative rather than substantive.

Other secondary outcome variables and exploratory outcome variables	Analysis of any other secondary outcome variables and of exploratory outcome variables will be performed using the standard statistical methodology. Calculated p-values will be indicative rather than substantive.

5.11.2 Estimands

5.11.2.1 Primary Efficacy Endpoint

Treatment Conditions of Interest	Chronocort vs IRHC
Population	Participants Aged 16 Years and Over with Congenital Adrenal Hyperplasia
Assessment	17-OHP and A4 plasma concentration obtained from the 08:00 hours blood sample at Visit 6 (EOS) and daily dose of hydrocortisone
Outcome Variable	The proportion of participants in each treatment arm who are biochemical responders at 08:00 hours at Visit 6 (EOS). See Section 5.11.1.1 for details on criteria for biochemical response
Population-Level Summary	The difference (Chronocort arm minus IRHC arm) in the proportion of participants who are biochemical responders at 08:00 hours at Visit 6 (EOS)
Analysis	The difference (Chronocort arm minus IRHC arm) in the proportion of participants who are biochemical responders is estimated using logistic regression (Ge, 2011) with treatment, fludrocortisone use at baseline, biochemical control at baseline, and the interaction of fludrocortisone use and biochemical control at baseline as explanatory variables.
Intercurrent Events (ICEs)	(i) Participants who withdraw from treatment before Visit 6 (EOS) or who do not meet the criteria for being a biochemical responder are deemed not to be a biochemical responder. (ii) Participants receiving rescue medication (under 'stress dosing rules') will continue in the study with unchanged procedures, except that Visit 6 (EOS) will be delayed if necessary, to 5 days after the use of rescue medication
Strategies to Handle ICEs	(i) Discontinuation of study treatment before Visit 6 (EOS), for any reason *Composite strategy* (ii) Receipt of rescue medication under 'stress-dosing rules' *Treatment policy strategy*

Where under the treatment policy strategy, available data occurring on or after the ICE will be analyzed as observed, and under the composite strategy, any observed biochemical response data

after the ICE will be considered as biochemical non-response. Additionally, any missing Visit 6 biochemical response data will be multiple imputed per Section 5.11.3.1.

Interpretation: Biochemical non-inferiority of Chronocort to IRHC will be declared if the 95% CI for the difference in proportions lies wholly above minus 15 percentage points. Confirmatory testing will occur on the primary estimand.

Choice of Non-Inferiority Margin: See Section 3 for details on the choice of the NI margin.

5.11.2.2 First Key Secondary Efficacy Endpoint

Treatment Conditions of Interest	Chronocort vs IRHC
Population	Participants Aged 16 Years and Over with Congenital Adrenal Hyperplasia
Assessment	17-OHP and A4 plasma concentration obtained from the 08:00 hours blood sample at Visit 6 (EOS) and daily dose of hydrocortisone
Outcome Variable	The proportion of participants in each treatment arm who are dose responders at 08:00 hours at Visit 6 (EOS). See Section 5.11.1.2 for details on criteria for dose response.
Population-Level Summary	The difference (Chronocort arm minus IRHC arm) in the proportion of participants who are dose responders at 08:00 hours at Visit 6 (EOS)
Analysis	The difference (Chronocort arm minus IRHC arm) in the proportion of participants who are dose responders is estimated using logistic regression (Ge, 2011), with treatment, fludrocortisone use at baseline, biochemical control at baseline, and the interaction of fludrocortisone use and biochemical control at baseline as explanatory variables.
ICEs	(i) Participants who withdraw from treatment before Visit 6 (EOS) or who do not meet the criteria for being a dose responder are deemed not to be a dose responder. (ii) Participants receiving rescue medication (under 'stress dosing rules') will continue in the study with unchanged procedures, except that Visit 7 will be delayed if necessary, to 5 days after the use of rescue medication
Strategies to Handle ICEs	(i) Discontinuation of study treatment before Visit 6 (EOS), for any reason *Composite strategy* (ii) Receipt of rescue medication under 'stress-dosing rules' *Treatment policy strategy*

Note: Handling strategies for ICEs will be the same as for the primary efficacy endpoint.

Interpretation: Dose superiority of Chronocort to IRHC will be declared if the 95% CI for the difference in proportions lies wholly above zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared.

5.11.2.3 Second Key Secondary Efficacy Endpoint

Treatment Conditions of Interest	Chronocort vs IRHC
Population	Participants Aged 16 Years and Over with Congenital Adrenal Hyperplasia
Assessment	Total daily dose of hydrocortisone at Visit 6 (EOS)
Outcome Variable	The mean total daily dose in each treatment arm at Visit 6 (EOS).
Population-Level Summary	The difference (Chronocort arm minus IRHC arm) in the mean total daily dose of hydrocortisone at Visit 6 (EOS)
Analysis	The mean difference (Chronocort arm minus IRHC arm) in total daily dose after 28 weeks of randomized treatment is estimated using a repeated measures model on timepoints 10, 16 and 28 weeks after randomization with treatment, timepoint, treatment by timepoint interaction, fludrocortisone use at baseline, biochemical control at baseline, and the interaction of fludrocortisone use and biochemical control at baseline as explanatory variables.
ICEs	(i) Participants who withdraw from treatment before Visit 6 (EOS) will have their total daily dose imputed. (ii) Participants receiving rescue medication (under 'stress dosing rules') will continue in the study with unchanged procedures, except that Visit 6 (EOS) will be delayed if necessary, to 5 days after the use of rescue medication
Strategies to Handle ICEs	(i) Discontinuation of study treatment before Visit 6 (EOS), for any reason *Treatment policy strategy* (ii) Receipt of rescue medication under 'stress-dosing rules' *Treatment policy strategy*

Where under the treatment policy strategy, available data occurring on or after the ICE will be analyzed as observed. Data missing at 28 weeks (or any other visit) will be imputed as per Section 5.11.3.3.

Interpretation: Superiority of Chronocort to IRHC with respect to total daily dose after 28 weeks of randomized treatment will be declared if the 95% CI for the difference in means lies wholly below zero, provided that biochemical non-inferiority of Chronocort to IRHC has been declared under the primary efficacy objective and dose superiority of Chronocort to IRHC has been declared under the first key secondary efficacy objective.

5.11.3 Estimand Related Data Handling

5.11.3.1 Primary Efficacy Endpoint – Biochemical Response at Visit 6 (EOS)

For the imputation of biochemical response where a composite variable strategy is employed, any biochemical response data occurring on or after the applicable ICE will be considered as biochemical non-response. This imputation will occur first, before any imputation of missing data occurs.

For the imputation of missing data, the binary response status will be multiple imputed. The number of imputations will be 50, representing the set of 50 plausible values obtained from a prediction model based primarily on observed data (i.e. applying a MAR approach in general). The fully conditional specification (FCS) method will be used for dealing with arbitrary non-monotone missing data patterns. The FCS is based on an iterative algorithm; at each iteration and for each variable of the prediction model, there is a prediction step and an imputation step. As imputations will be for binary components, the models used for prediction and imputation will be logistic regression models.

The full MI approach will comprise of 4 steps, as follows:

- Prediction Step: the current (iteration) values of the observed and imputed values are used to derive the predictive distribution of the missing values under MAR conditions
- Imputation Step: updated imputations are generated by draws from the predictive distribution defined by the updated logistic regression model
- Analysis Step: The analysis as described in Section 5.11.4.1 is conducted for each of the 50 imputations
- Combining Step: Rubin's rule (SAS PROC MIANALYZE) is used to combine the 50 point estimates and standard errors, in order to provide a single point estimate and confidence interval that can be compared to the respective Non-inferiority delta in the null hypotheses described in Section 5.11.1 to provide rejection or acceptance.

5.11.3.1.1 Prediction Step

The prediction step attempts to predict the biochemical response based on observed data that may have an influence on it (assuming MAR).

The prediction model will include the following variables, to attempt to estimate the missing value as closely as possible based on available characteristics.

Timepoint	Variables
Variable Used for Randomization (per IWRS)	• Fludrocortisone replacement status (requires fludrocortisone or does not require fludrocortisone)
Baseline Assessments	• Participant's assessment of biochemical control • Sex
Assessments at V4, V5, V6 and V7	• Participant's biochemical response status
Other	• Randomized study treatment (Chronocort or IRHC)

5.11.3.1.2 Multiple Imputation Step for All Missing Data (MAR)

Within the Bayesian framework, the task of imputing missing values is achieved by drawing random values from the posterior predictive distribution of the missing primary efficacy endpoint data (predicted by the logistic regression prediction model specified above). This posterior predictive distribution is a function of the observed data and regression parameters (or function of regression parameters) under MAR conditions.

When the last variable in the sequence of variables listed above in the prediction model has been imputed, the algorithm cycles again through each variable, repeating the chain of regression estimation and imputation draw steps. These cycles are repeated 200 times ("burn-in" iterations) and finally there will be 50 draws from the predictive distribution for missing values. SAS PROC MI will be used for this process.

Example SAS Code for Prediction and Imputation Steps:

[redacted]

The LOGISTIC option LIKELIHOOD=AUGMENT may be used when the maximum likelihood parameter estimates do not exist. This option adds observations in each response group to the likelihood function. Each added observation contributes the same weight, with the number of

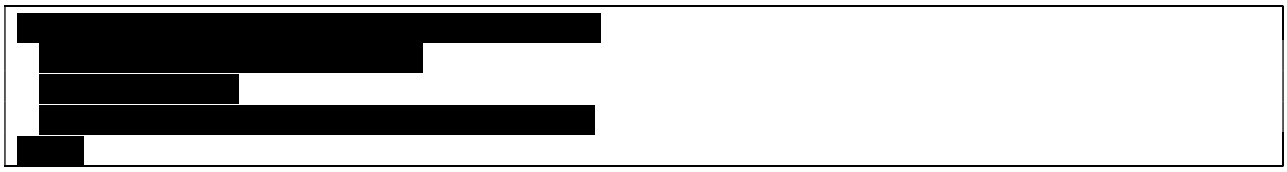
parameters in the logistic regression model as the total added weight. Existence of maximum likelihood estimates will be checked after general unblinding with the actual randomization.

5.11.3.1.3 Analysis Step

Each of the imputations will be analyzed using a logistic regression model, as described in Sections 5.11.4.1 and 11.2.

5.11.3.1.4 Combining Step

SAS Proc MIANALYZE will be used to combine the 50 point estimates for the common response probability difference and standard errors, allowing appropriate statistical inference as described below:



5.11.3.2 First Key Secondary Efficacy Endpoint – Dose Response at Visit 6

The approach for dose response will be similar to that described in Section 5.11.3.1 for biochemical response. The primary analysis conducted using MAR approach for imputation will be presented for this endpoint, similar to the primary efficacy endpoint.

5.11.3.3 Second Key Secondary Efficacy Endpoint – Total Daily Dose of Hydrocortisone at 28 weeks

For the second key secondary efficacy endpoint, where an ICE occurs, every effort will be taken to ascertain the total daily dose of hydrocortisone (or equivalent) at 28 weeks whether on study or not. Where hydrocortisone is not used, and instead an alternative treatment (such as prednisone, prednisolone or dexamethasone) is received, the dose of the treatment received will be converted to hydrocortisone equivalent as follows [Nebesio et al. 2016]:

- Dexamethasone x 80
- Prednisolone/one x 5

This could occur at pre Run-in Baseline, and for participants that discontinue study treatment.

Where missing data exists for total daily hydrocortisone dose after discontinuation, no explicit imputation will be done.

5.11.4 Analysis of the Primary Efficacy Endpoint

5.11.4.1 Main Analysis of the Primary Efficacy Endpoint

The main analysis of the primary endpoint is based on the proportion of participants achieving biochemical response at V6, with ICE handling per Section 5.11.2.1 and imputing missing data using the MI approach described in Section 5.11.3.1.

The primary estimand will be analyzed using a logistic regression model. The model will include the following factors: treatment arm (Chronocort versus IRHC, reference IRHC), the participant's fludrocortisone replacement status at baseline (requires fludrocortisone or does not require fludrocortisone) and whether they are in biochemical control at baseline or not (based on the androgen results from the baseline visit [Visit 2]), plus the interaction of these variables. The estimated response rate for each treatment arm, and the corresponding difference in rates along with the 2-sided 95% CI for the difference will be presented using the procedure described by Ge et al. (2011). See Section 11.2 for further details of the analysis.

1) In case issues related to convergence of the model or other issues that may be expected based on the frequency counts within the combinations of the factors involved lead to warnings (such as, for example, complete or quasi- separation, model less than full rank, etc.) and/or prevent the model from producing valid results (execution stops, some terms in the model not estimated, unusually large standard errors obtained, etc.), an initial assessment of the distribution of response/non-response outcomes will occur in order to ascertain the possible best next steps. However, a priori, in lieu of such knowledge, the following steps will be implemented in order specified below: Exclude the interaction term;
2) Exclude the interaction term and one of the main effects (fludrocortisone replacement status at baseline or biochemical control at baseline, whichever produces combinations with lower counts);
3) Exclude both factors and their interaction.

At each step, the model will be rerun and evaluated for continued occurrence of issues. The decisions made at each step of the model selection process will be fully documented.

17-OHP and A4 results will be listed, and results from visits outside the protocol specified visit window will be flagged.

5.11.4.2 Sensitivity and Supportive Analyses

The main statistical analysis assumes that the treatment effect is constant over the standard explanatory variables. Additional sensitivity analyses will investigate whether the treatment

effect varies with the explanatory variables: including, but not limited to, subgroup analyses by biochemical control status at baseline, fludrocortisone use at baseline, region or country as appropriate and pre-treatment corticosteroid. Separate summary tables will be presented for each subgroup and a measure of treatment-by-subgroup interaction together with confidence interval and P-value provided. Forest plots will additionally be presented to provide a visual comparison of the differences in treatment effect between subgroups.

All subgroup analyses will be performed as unstratified/unadjusted.

Summary tables split by biochemical control at baseline will also be presented for disposition, demography and baseline characteristics and exposure to study drug, (Participants with baseline 17-OHP equal to or below the upper limit for optimal control at 08:00 hours and baseline A4 equal to or below the upper limit of the reference range will be classified as 'in biochemical control'. Participants with baseline 17-OHP above the upper limit for optimal control and/or baseline A4 above the upper limit of the reference range will be classified as 'not in biochemical control'.)

Supportive analysis: The main statistical analysis will additionally be performed on the EES.

Other analyses: Any other analyses of 17-OHP and A4 concentrations and dose, not otherwise specified as a secondary analysis, are supplementary analyses. The number and percentage of participants in receipt of rescue medication overlapping with a scheduled visit, as well as the number and percentage of participants with moved Visit 6 assessment due to stress dosing rules will be summarized by treatment group, split by baseline biochemical control status.

All observed data for the primary endpoint will additionally be summarized descriptively by visit, with no imputation.

5.11.5 Key Secondary Efficacy Endpoints and Analyses

The high-level handling approaches for the secondary endpoints are as follows. All observed data for these endpoints will additionally be summarized descriptively by visit, with no imputation.

Endpoint	Analysis Set	Modelling Method	Data Handling
Percentage of Participants Dose Response after 28 Weeks of Randomized Study Treatment	FAS and EES	Logistic Regression	As per Primary efficacy endpoint.

Endpoint	Analysis Set	Modelling Method	Data Handling
Mean Total Daily Dose after 28 Weeks of Randomized Study Treatment	FAS and EES	MMRM	Treatment policy strategy for participants who stop treatment before visit 6. Treatment policy strategy for participants whose visit 6 is delayed because of administering rescue medication. Present all visits up to and including V6.

5.11.5.1 Percentage of Participants Achieving Dose Response after 28 Weeks of Randomized Study Treatment

The percentage of participants achieving dose response over the course of the study will be analyzed in a similar manner to the primary efficacy endpoint (logistic regression) on the FAS.

5.11.5.2 Mean Total Daily Dose after 28 Weeks of Randomized Study Treatment

Mean Total Daily Dose will be analyzed on the FAS using a repeated measures mixed model with factors for treatment arm (Chronocort versus IRHC, reference IRHC), the participant's fludrocortisone replacement status at baseline (requires fludrocortisone or does not require fludrocortisone) and whether they are in biochemical control at baseline or not (based on the androgen results from the baseline visit [Visit 2]) and their interaction, as well as the treatment-visit interaction. A restricted maximum likelihood method will be specified, along with an unstructured covariance matrix. The estimate of the least-squared means difference at each visit and the corresponding 95% CI for the difference will be presented. Example SAS code for the analysis is provided in Section 11.3.

Considerations for model adjustments specified in Section 5.11.4 also apply to the analysis of this key secondary efficacy endpoint.

5.11.5.3 Sensitivity and Supportive Analyses for the Key Secondary Efficacy Endpoints

Sensitivity analyses:

The main statistical analysis assumes that the treatment effect is constant over the standard explanatory variables. Additional sensitivity analyses will investigate whether the treatment effect varies with the explanatory variables: including, but not limited to, subgroup analyses by biochemical control status at baseline, fludrocortisone use at baseline, region or country as appropriate and pre-treatment corticosteroid. Separate summary tables will be presented for each subgroup and a measure of treatment-by-subgroup interaction together with confidence interval

and P-value provided. Forest plots will additionally be presented to provide a visual comparison of the differences in treatment effect between subgroups.

Supportive analysis: The main statistical analysis for each of the key secondary endpoints will additionally be performed on the EES.

5.11.6 Other Secondary Efficacy Endpoints

5.11.6.1 Standard Explanatory Variables for Other Secondary Efficacy Endpoints

The standard explanatory variables are (1) whether or not the participant was in biochemical control at baseline, (2) whether or not the participant was receiving fludrocortisone at baseline and (3) their interaction. Randomization of participants to treatment arm is stratified by fludrocortisone treatment and the analyses will include the stratification of biochemical control.

Analyses will be stratified or adjusted, as appropriate, on the basis of (i) the participant's fludrocortisone replacement status at baseline (requires fludrocortisone or does not require fludrocortisone) and (ii) whether they are in biochemical control at baseline or not (based on the androgen results from the baseline visit [Visit 2]). Participants with 17-OHP equal to or below the upper limit for optimal control at 08:00 hours and A4 equal to or below the upper limit of the reference range will be classified as ‘in biochemical control’. Participants with 17-OHP above the upper limit for optimal control and/or A4 above the upper limit of the reference range will be classified as ‘not in biochemical control’.

Considerations for model adjustments specified in Section 5.11.4 also apply to the analysis of other secondary efficacy endpoints.

5.11.6.2 Standard Methodology for Comparisons of Other Secondary Efficacy Endpoints

Unless otherwise stated, continuous outcome variables will be compared between treatment arms using analysis of covariance (ANCOVA) adjusting for the standard explanatory variables. Unless otherwise stated, a treatment policy approach will be assumed for all ICEs, with no imputation of missing data. If a non-normal distribution is anticipated or observed, then consideration will be given to using a logarithmic transformation or a non-parametric analysis. Analyses of change from baseline will always include the baseline value as a covariate. Example code is provided in Section 11.4.

Unless otherwise stated, binary outcome variables will be compared between treatment arms using logistic regression adjusted for the standard explanatory variables. Unless otherwise stated, a treatment policy approach will be assumed for all ICEs, with non-responder imputation of

missing data. The procedure described by Ge et al. (2011) will be used to estimate the standard error and resulting confidence interval around the covariate adjusted difference in proportions

5.11.6.3 Analyses of Other Secondary Efficacy Endpoints

Outcome Variables	Outline Analyses
Whether or not the participant is a biochemical responder at 4, 10 and 16 weeks after randomization.	The same analysis will be used as for the primary estimand of the primary outcome variable (except that at 4 weeks there has been no opportunity for a dose titration so the condition on dose is not used).
Whether or not the participant is a dose responder at 10 and 16 weeks after randomization.	The same analysis will be used as for the primary estimand of the first key secondary outcome variable.
Total daily dose at 10 and 16 weeks after randomization.	The same analysis will be used as for the primary estimand of the second key secondary outcome variable.
Whether or not the participant is in biochemical control (provided total daily dose is not greater than 30mg) at 4, 10, 16 and 28 weeks after randomization.	The same analysis will be used as for the primary estimand of the primary outcome variable.
The difference in range as calculated as the difference between the 08:00 and 13:00 17-OHP levels at 4, 10, 16 and 28 weeks after randomization.	Standard methodology for comparisons.
The difference in range as calculated as the difference between the 08:00 and 13:00 A4 levels at 4, 10, 16 and 28 weeks after randomization.	Standard methodology for comparisons.
The mean of the 08:00 and 13:00 levels for each androgen at 4, 10, 16 and 28 weeks after randomization.	Standard methodology for comparisons.
The total daily glucocorticoid dose at 4, 10, 16 and 28 weeks after randomization.	Standard methodology for comparisons. Total glucocorticoid daily dose will be calculated as mean total daily dose of study medication and other additional glucocorticoids over the time interval since previous visit. Additionally, the mean total daily dose by treatment will be calculated for participants in and not in biochemical control at the end of study visit.

Outcome Variables	Outline Analyses
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in menstrual regularity	Standard methodology for comparisons (only in pre-menopausal women without hysterectomy and not using hormonal contraception). Note, post treatment menstrual regularity will be derived from the menstrual cycle log using the following definitions: • Oligomenorrhoea is defined as fewer than 9 menstrual cycles per year or cycle length >35 days • Amenorrhoea is defined as absent menses for ≥3 months.
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in LH levels	Standard methodology for comparisons (men only)
The change from baseline to 28 weeks of randomized treatment in size of TART	Descriptive statistics only (men with TART at baseline). Percentage change from baseline will be summarized.
The change from baseline to 28 weeks of randomized treatment in sperm count	Standard methodology for comparisons (men only)
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in objective hirsutism using the Ferriman-Gallwey score	Standard methodology for comparisons (women only)
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in objective acne using the Global Evaluation Acne (GEA) scale	Standard methodology for comparisons (women only)
The change from screening to 4, 10, 16 and 28 weeks of randomized treatment in HbA1c levels	Standard methodology for comparisons
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in waist circumference	Standard methodology for comparisons
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in body weight	Standard methodology for comparisons
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in QoL using the self-completed SF-36® total score and the sub-domain of vitality	Standard methodology for comparisons
The change from baseline to 4, 10, 16 and 28 weeks of randomized treatment in QoL using the MAF	Standard methodology for comparisons
The change from baseline to weeks 4, 10, 16 and 28 in subjective hirsutism and acne using a participant-reported VAS.	Standard methodology for comparisons (women only).

Outcome Variables	Outline Analyses
The change from baseline to weeks 4, 10, 16 and 28 in QoL using the EQ-5D-5L™.	Standard methodology for comparisons.

[redacted]

5.11.8 Multiplicity

The results of the key secondary analysis (dose response) will only be considered substantive if the primary estimand of the primary analysis demonstrates non-inferiority, otherwise the results will be considered indicative. Similarly, the analysis of the second key secondary efficacy endpoint (total daily dose) will be considered substantive if the primary estimand of the primary efficacy endpoint and key first secondary endpoint demonstrate success.

All other analyses of secondary endpoints, analyses of exploratory endpoints and sensitivity and supportive analyses will be considered as descriptive evidence of efficacy and will be analyzed without any procedures to account for multiple comparisons.

5.12 Safety Analyses

The safety analyses will be presented by the treatment received for the Safety Analysis Set.

5.12.1 Summary

All safety analyses will be performed on the SAF.

Endpoint	Statistical Analysis Methods
AEs	AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 20.0 or later and summarized descriptively. Only TEAEs will be included in the main safety analysis. A TEAE is defined as any AE that has an onset on or after the first dose of study drug and before the last dose of study drug + 30 days or is any pre-existing AE that has worsened in severity on or after the first dose of study drug and before the last dose of study drug + 30 days. The incidence, nature, severity, relatedness, duration, outcome, seriousness and expectedness of TEAEs will be tabulated. AEs of special interest will be additionally separately tabulated.
Prior, concomitant and subsequent medications	Medications will be coded using WHO Drug dictionary and summarized descriptively using ATC level and preferred term. See Section 5.12.3 for further details.
Rescue medications	Rescue medications will be coded and summarized similarly to prior and concomitant medications. An additional summary of the use of rescue medication overlapping with each visit will be presented. See Section 5.12.4 for further details.
Need for additional glucocorticoid doses	Use of hydrocortisone from the stress dosing packs or use of any additional glucocorticoid treatment during the study will be summarized descriptively.
Other safety assessments	For laboratory parameters, vital signs, and ECGs, change from baseline will be calculated at each scheduled post-baseline visit and summary statistics will be presented both for absolute values (at baseline and each scheduled post-baseline visit) and changes from baseline by treatment arm, and visit. Shift tables of baseline to minimum and maximum on-treatment will be produced for the appropriate parameters displaying the number and percentage of participants in each of the categories (High, Low and Normal) relative to the category of their baseline assessment. Results considered clinically significant by the Investigator will be summarized at each visit. Physical examination results, and their change from baseline, will be summarized at each visit.

5.12.2 Adverse Events

A treatment emergent adverse event (TEAE) is defined as:

- Any AE that has an onset on or after the first dose of study drug and before last dose of study drug + 30 days.
- Any pre-existing AE that has worsened in severity on or after the first dose of study drug and before the last dose of study drug + 30 days.

A TEAE which starts or worsens on or after the first dose of IRHC in the run-in period but before the first dose of randomized study treatment is allocated to the run-in period. A TEAE which starts or worsens after the first dose of randomized study treatment is allocated to the treatment actually received during the randomized treatment period. A TEAE which starts or worsens in the run-in period and worsens in the randomized treatment period is allocated to both periods.

A treatment-related AE is defined as an AE that is recorded by the Investigator as definitely, probably, possibly or likely to be related to the study drug. If an AE has missing relationship it is assumed to be related to the study drug for analysis purposes. Maximum severity will be assumed for an AE with missing severity.

TEAEs will be summarized separately for the run-in period and after randomization. In the list of AE tables below, only the first 2 will be done for both the run-in period and after randomization. All the others will just be done after randomization.

The following tables will be presented for AEs by treatment group:

- Overall incidence and the number of AEs, Serious Adverse Events (SAEs), Treatment related TEAEs, TEAEs leading to withdrawal of study drug, TEAEs leading to early withdrawal from study, TEAEs leading to study drug dose increase or decrease, TEAEs leading to study drug dose interruption, TEAEs leading to use of rescue medication, Treatment Emergent AEs of Special Interest (TEAESIs), non-treatment emergent AEs occurring > 30 days after last dose of study drug, and AEs leading to death – run-in period and after randomization
- TEAEs by system organ class (SOC) and preferred term (PT), incidence and number of events – run-in period and after randomization
- Treatment related TEAEs by SOC and PT, incidence and number of events – after randomization only
- Serious TEAEs by SOC and PT, incidence and number of events – after randomization only
- TEAEs by SOC, PT and maximum severity, incidence – after randomization only

- Treatment related TEAEs by SOC, PT and maximum severity, incidence – after randomization only
- TEAEs by SOC, PT and relationship, incidence – after randomization only
- TEAEs leading to withdrawal of study drug by SOC and PT, incidence – after randomization only
- TEAEs leading to study drug dose increase or decrease by SOC and PT, incidence – after randomization only
- TEAEs leading to study drug dose interruption by SOC and PT, incidence – after randomization only
- TEAEs leading to use of rescue medication by SOC and PT, incidence – after randomization only
- TEAEs leading to early withdrawal from study by SOC and PT, incidence – after randomization only
- TEAESIs (including events of unexpected benefit, adrenal crises, signs and symptoms of adrenal insufficiency and over-treatment) by SOC and PT, incidence – after randomization only
- Non-treatment emergent AEs occurring > 30 days after last dose of study drug by SOC and PT, incidence – after randomization only
- Listing of Serious TEAEs (presented in the Table section of the appendices)
- Listing of Deaths (presented in the Table section of the appendices).

In counting the number of AEs reported, a continuous event (i.e. reported more than once and which did not cease), will be counted only once; non-continuous AE reported several times by the same participant will be counted as multiple events.

For incidence summaries by SOC and PT, a participant will be counted only once at the SOC level and once at each PT level within the SOC level, even if the participant experienced more than one AE within a specific SOC and PT. For summaries by SOC, PT, and severity, a participant is typically counted once at each severity level for which the event occurred at the SOC level and at each severity level for which the event occurred for each unique PT within that SOC level. Therefore, participants may contribute to multiple severity levels within a PT or SOC.

All AEs will be listed.

5.12.3 Prior, Concomitant and Subsequent Medication

Medications/procedures will be coded using the World Health Organization Drug Dictionary (WHODrug). The dictionary version used will be displayed in the footnote of the presented data displays.

5.12.3.1 Prior and Concomitant Medications

Separate tabulations will be produced for prior and concomitant medications presented by treatment group and overall, for the SAF.

Prior medications are defined as all medications with stop date after the signing of the informed consent form but before first dose of randomized study drug. Concomitant medications are defined as medications that either start after the first dose of randomized study drug, or started before the first dose of randomized study drug and continue to be taken after the first dose of randomized study drug. For details on imputation of missing/partial start and stop dates of prior/concomitant medications, see Section 5.2.7.

Participant incidence will be tabulated for concomitant medications by Anatomic Therapeutic Class (ATC) level 3 and preferred term for the SAF, by treatment group and overall. Participants will be counted only once for each ATC or preferred term in the event that they have multiple records of the same ATC or preferred term in the database. Prior medications will only be included in data listings.

Prior and concomitant medications will be listed with the occurrence (prior versus concomitant) identified.

Medications for CAH taken in the 26 weeks before the start of the run-in period will be separately listed and summarized.

5.12.3.2 Subsequent Medication

Medications that start after the last dose of randomized study drug but before 28 weeks after randomization are considered subsequent medications. Subsequent medications will be analyzed in the same way as concomitant medications (using ATC and preferred term, by treatment group and overall for the SAF), and will be included in the above listings of prior/concomitant/subsequent.

5.12.4 Rescue Medication

Medication to be used when the 'stress dosing rules' are implemented will be supplied by the study site as part of an emergency treatment pack. The content will typically include (according to local practice):

- A one-week supply of 10 mg oral hydrocortisone tablets (to allow dosage of up to 60 mg daily)
- Hydrocortisone for injection plus syringes and needles
- The standard information guidance regarding 'stress dosing rules'.

Participants receiving rescue medication will be presented in a data listing. The extent of receipt of rescue medication in relation to efficacy assessments will be covered under Section 5.11.4.2. The Data Review will document further details of all medications regarded as rescue, and their timing.

Medications defined as rescue will be analyzed in the same way as concomitant medications (using ATC and preferred term, by treatment group and overall, for the SAF); however the number of separate uses (similar to presentation of the number of events for the AEs) will also be included. Rescue medications will be listed separately to prior, concomitant, and subsequent medications.

5.12.5 Need for Additional Glucocorticoid Doses

Use of hydrocortisone from the stress dosing packs or use of any additional glucocorticoid treatment during the study will be summarized descriptively.

5.12.6 Laboratory Data

Descriptive statistics of the observed values and change from baseline (continuous data) will be presented by treatment group and visit for each hematology, urinalysis, and serum chemistry parameter. Each measurement (continuous data) will be classed as below, within, or above normal range, based on ranges supplied by the laboratory used. Shift tables in relation to the normal range from baseline to end of study/early withdrawal visit, and worst high/low on-treatment during the study will be presented. For the worst high/low shifts, all lab assessments will be considered (both scheduled and any unscheduled).

A listing of any clinically significant laboratory measurements recorded throughout the study will be presented.

5.12.7 Vital Signs

Descriptive statistics for observed values and changes from baseline in the following vital signs will be presented by treatment group and visit:

- Systolic blood pressure (mmHg)
- Diastolic blood pressure (mmHg)
- Pulse rate (bpm)
- Respiration rate (breath / min)
- Body temperature (degrees Celsius)

5.12.8 Electrocardiogram Data

Descriptive statistics for observed values and changes from baseline in the following ECG variables will be tabulated at each follow-up:

- Heart rate (bpm)
- PR interval (ms)
- RR interval (ms)
- QRS complex (ms)
- QT interval (ms)
- QTc interval (ms)

In addition, shift tables relating to the overall interpretation (Normal, Abnormal NCS, and Abnormal CS) from baseline to end of study/early withdrawal visit will be presented by treatment group.

5.12.9 Physical Examination

The body systems within the physical examination data at the end of the study will be summarized by treatment group (Normal; Abnormal NCS, Abnormal CS). Changes from baseline will also be tabulated. Details of clinically significant findings will be listed.

6 INTERIM ANALYSIS

No formal interim analysis is planned.

7 DATA SAFETY MONITORING BOARD ANALYSIS

An independent DSMB will meet on a regular basis during the study to review safety data. The DSMB will operate in accordance with a charter, which is a separate controlled document.

Safety data will be generated at regular intervals for review by the independent DSMB (see Section 9.6 of the study protocol). Details of these DSMB outputs will be specified in the Kick-Off Meeting slides and minutes and the DSMB Charter. Any outputs provided to the DSMB will be unblinded (a separate statistical team will provide the unblinded outputs to the DSMB). Outputs provided to other team members to remain in a blinded capacity will use a dummy randomization list with generic labels.

During the course of the study, if unblinded analyses indicate that AEs are occurring more frequently than anticipated, the DSMB will notify the Sponsor, who will notify regulatory authorities expeditiously as appropriate.

8 CHANGES TO PLANNED PROTOCOL ANALYSIS

- Section 2: Analysis of Biochemical Response Rate at Week 4 was specified in the Protocol but not including a criterion for dose as the participant is still receiving the initial dose at Week 4. This analysis has been removed as it is identical to the analysis of Biochemical Control Rate at Week 4.

- No comparative analysis will be performed for the endpoint The change from baseline to 28 weeks of randomized treatment in size of TART (men only). Only summaries including percentage change from baseline will be provided.

- No compliance calculation will be performed. The extent by which participants did not follow the instructions for the return of unused and empty medication containers rendered calculations of compliance uninterpretable.

9 CHANGES TO PLANNED ANALYSIS FROM SAP VERSION 1.0

- In line with protocol amendment 6.0, the primary analysis timepoint has been updated from week 52 to week 28
- Sensitivity analyses of primary efficacy endpoint and first key secondary endpoint of MNAR multiple imputation and tipping point analysis have been removed.
- Analysis of second key secondary endpoint will now be mixed model repeated measures analysis. Analysis using MAR multiple imputation has been removed.
- Description of steps to be implemented in case of issues with statistical models, including potential exclusion of interaction term and/or other factors, was added.
- It was specified that all subgroup analyses will be performed as unstratified/unadjusted.
- Additional considerations for baseline derivation were added.

- The compliance calculation and analysis was removed.
- Updates to match to protocol version 7.0.

10 REFERENCES

Brand, J. P. L. (1999). "Development, Implementation, and Evaluation of Multiple Imputation Strategies for the Statistical Analysis of Incomplete Data Sets." Ph.D. thesis, Erasmus University.

[redacted]

Drano B et al. Development and evaluation of a Global Acne Severity Scale.

Ge M, Durham LK, Meyer RD, et al. Covariate-adjusted difference in proportions from clinical trials using logistic regression and weighted risk differences. Drug Inf J 2011;45:481–93.

(GEA Scale) suitable for France and Europe Journal of the European Academy of Dermatology and Venereology 2011, 25, 43–48.

Hatch R. et al. Hirsutism: Implications, etiology, and management AJOG 1981 VOLUME 140, ISSUE 7, P815-830.

Nebesio et al. Differential effects of hydrocortisone, prednisone, and dexamethasone on hormonal and pharmacokinetic profiles: a pilot study in children with congenital adrenal hyperplasia. International Journal of Pediatric Endocrinology (2016) 2016:17. DOI 10.1186/s13633-016-0035-5

Ratitch, B., and O'Kelly, M. (2011). "Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures." In Proceedings of PharmaSUG 2011 (Pharmaceutical Industry SAS Users Group). Paper SP04. Cary, NC: SAS Institute Inc.

Rubin, D. B. (1976). "Inference and Missing Data." Biometrika 63:581–592.

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

Van Buuren, S. (2007). "Multiple Imputation of Discrete and Continuous Data by Fully Conditional Specification." Statistical Methods in Medical Research 16:219–242.

11 APPENDICES

11.1 Quality of Life Derivation

11.1.1 SF-36 Scoring Information

Derivation of summary scores, including appropriate handling of missing items, will be performed using the scoring software provided by the owner of the scale (Optum software, owned by QualityMetric®). Details of the specific derivations are held by QualityMetric.

The SF-36 (v2.0) consists of 36 items, as follows:

- *1:In general, would you say your health is*
- *2:Compared to one year ago, how would you rate your health in general now*
- *3A:Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports*
- *3B:Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf*
- *3C:Lifting or carrying groceries*
- *3D:Climbing several flights of stairs*
- *3E:Climbing one flight of stairs*
- *3F:Bending, kneeling, or stooping*
- *3G:Walking more than a mile*
- *3H:Walking several hundred yards*
- *3I:Walking one hundred yards*
- *3J:Bathing or dressing yourself*
- *4A:As a result of your physical health, cut down on the amount of time you spent on work or other activities*
- *4B:As a result of your physical health, accomplished less than you would like*
- *4C:As a result of your physical health, were limited in the kind of work or other activities*
- *4D:As a result of your physical health, had difficulty performing the work or other activities (for example, it took extra effort)*
- *5A:As a result of any emotional problems, cut down the amount of time you spent on work or other activities*
- *5B:As a result of any emotional problems, accomplished less than you would like*
- *5C:As a result of any emotional problems, did work or other activities less carefully than usual*
- *6:During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups*
- *7:How much bodily pain have you had during the past 4 weeks*
- *8:During the past 4 weeks how much did pain interfere with your normal work (including both work outside the home and housework)*

- *9A:How much of the time during the past 4 weeks: did you feel full of life*
- *9B:How much of the time during the past 4 weeks: have you been very nervous*
- *9C:How much of the time during the past 4 weeks: have you felt so down in the dumps that nothing could cheer you up*
- *9D:How much of the time during the past 4 weeks: have you felt calm and peaceful*
- *9E:How much of the time during the past 4 weeks: did you have a lot of energy*
- *9F:How much of the time during the past 4 weeks: have you felt downhearted and depressed*
- *9G:How much of the time during the past 4 weeks: did you feel worn out*
- *9H:How much of the time during the past 4 weeks: have you been happy*
- *9I:How much of the time during the past 4 weeks: did you feel tired*
- *10:During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting your friends, relatives, etc.)*
- *11A:I seem to get sick a little easier than other people*
- *11B:I am as healthy as anybody I know*
- *11C:I expect my health to get worse*
- *11D:My health is excellent*

From the responses to these questions, the physical component score (PCS) and mental component score (MCS) are derived from the following subscores:

Scale	**Item**	**Coding**	**Range of Raw Score**	**Normalization***
Physical Functioning	Items 3a to 3j	Yes, limited a lot = 1 Yes, limited a little = 2 No, not limited = 3	10 (Worst) – 30 (Best)	(S-10)/20x100
Role-Physical	Items 4a to 4d	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	4 (Worst) – 20 (Best)	(S-4)/16x100
Bodily Pain	Item 7	None = 6 Very Mild = 5.4 Mild = 4.2 Moderate = 3.1 Severe = 2.2 Very Severe = 1	2 (Worst) – 12 (Best)	(S-2)/10x100
	Item 8 if both items 7 and 8 are answered	Not at all and Item 7 is "None" = 6 Not at all and Item 7 is not "None" = 5		

Scale	Item	Coding	Range of Raw Score	Normalization*
		A little bit = 4 Moderately = 3 Quite a bit = 2 Extremely = 1		
	Item 8 if item 7 is not answered	Not at all = 6 A little bit = 4.75 Moderately = 3.5 Quite a bit = 2.25 Extremely = 1		
General Health	Item 1	Excellent = 5 Very good = 4.4 Good = 3.4 Fair = 2 Poor = 1	5 (Worst) – 25 (Best)	(S-5)/20x100
	Items 11a and 11c	Definitely true = 1 Mostly true = 2 Don't know = 3 Mostly false = 4 Definitely false = 5		
	Items 11b and 11d	Definitely true = 5 Mostly true = 4 Don't know = 3 Mostly false = 2 Definitely false = 1		
Vitality	Items 9a and 9e	All of the time = 5 Most of the time = 4 Some of the time = 3 A little of the time = 2 None of the time = 1	4 (Worst) – 20 (Best)	(S-4)/16x100
	Items 9g and 9i	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5		
Social Functioning	Item 6	Not at all = 5 Slightly = 4 Moderately = 3	2 (Worst) – 10 (Best)	(S-2)/8x100

Scale	Item	Coding	Range of Raw Score	Normalization*
		Quite a bit = 2 Extremely = 1		
	Item 10	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5		
Role-Emotional	Items 5a to 5c	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	3 (Worst) – 15 (Best)	(S-3)/12x100
Mental Health	Items 9b, 9c and 9f	All of the time = 1 Most of the time = 2 Some of the time = 3 A little of the time = 4 None of the time = 5	5 (Worst) – 25 (Best)	(S-5)/20x100
	Items 9d and 9h	All of the time = 5 Most of the time = 4 Some of the time = 3 A little of the time = 2 None of the time = 1		

Note: * S = raw score = sum of item scores after coding

Each scale score is then standardized based on the latest US population figures (held by Optum) to produce the following standardized Scale Scores:

- Physical Functioning
- Role-physical
- Bodily pain
- General health
- Vitality
- Social functioning
- Role-emotional
- Mental heath

Finally, the SF-36 PCS and MCS transformed scores are derived. Additionally, norm-based transformation of the 8 scale scores will occur.

11.1.2 Multidimensional Assessment of Fatigue (MAF)

The items of the MAF will be scored as follows (per American College of Rheumatology guidance):

- Items 1, 4 – 14: 10pt scale from 1 = 'not at all' to 10 = 'a great deal'
- Item 2: 10pt scale from 1 = 'mild' to 10 = 'severe'
- Item 3: 10pt scale from 1 = 'no distress' to 10 = 'a great deal of distress'
- Timing items 15 and 16: Categorical responses: 1 to 4
- To calculate the Global Fatigue Index (GFI), convert item 15 to a 0 - 10 scale by multiplying each score by 2.5 and then sum items 1, 2, and 3, and average 4 - 14, and newly scored item 15. Do not assign a score to items 4 - 14 if the respondent gave a response of "do not do any activity for reasons other than fatigue." If a respondent selects "no fatigue" on item 1, assign a zero to items 2 - 16. Item 16 is not included in the GFI.

A higher score indicates more severe fatigue, fatigue distress, or impact on activities of daily living.

11.1.3 EQ-5D

The EQ-5D is a standardized PRO for use as a general measure of health outcome. The EQ-5D has 5 questions used to create a descriptive "health state" and a VAS to capture a patient's overall health rating. The 5 questions and derivation of the total score are as follows. Additionally, the 5 questions will be mapped to an index value using the crosswalk link function based on the published EQ-5D-5L Crosswalk Index Value Set.

Question	Coding
Mobility (EQ5D0201)	I have no problems walking = 1 I have slight problems walking = 2 I have moderate problems walking' = 3 I have severe problems walking = 4 I am unable to walk = 5
Self-Care (EQ5D0202)	I have no problems washing and dressing myself = 1 I have slight problems washing and dressing myself = 2 I have moderate problems washing and dressing myself = 3 I have severe problems washing and dressing myself = 4 I am unable to wash and dress myself = 5

Question	Coding
Usual Activities (EQ5D0203)	I have no problems doing my usual activities = 1 I have slight problems doing my usual activities = 2 I have moderate problems doing my usual activities = 3 I have severe problems doing my usual activities = 4 I am unable to do my usual activities = 5
Pain/Discomfort (EQ5D0204)	I have no pain or discomfort = 1 I have slight pain or discomfort = 2 I have moderate pain or discomfort = 3 I have severe pain or discomfort = 4 I have extreme pain or discomfort = 5
Anxiety/Depression (EQ5D0205)	I am not anxious or depressed = 1 I am slightly anxious or depressed = 2 I am moderately anxious or depressed = 3 I am severely anxious or depressed = 4 I am extremely anxious or depressed = 5
EQ-5D Total Score	Determine 'health state' which for example if all questions were answered '1' then it would be 11111 per participant per visit where: • EQ5D0201 outcome x 10000 • EQ5D0202 outcome x 1000 • EQ5D0203 outcome x 100 • EQ5D0204 outcome x 10 • EQ5D0205 outcome x 1
EQ-5D Index	Corresponding index value from the US country value set from the Euroquol list, based on the 'health state'. The crosswalk value set can be found here.

[redacted]

[redacted]