SUPPLEMENTARY APPENDIX

Table of Contents

1.0 STUDY INVESTIGATORS

Austria (City)

- Florian W. Kiefer, MD, PhD; Vienna General Hospital (Vienna)

Belgium (City)

- Leen Antonio, MD; University Hospitals Leuven (Leuven)

Bulgaria (City)

- Ralitsa Robeva, MD; Medical University Sofia (Sofia)

Canada (City, Province)

- Syed A. Imran, MBBS; Nova Scotia Health Authority; Dalhousie University (Halifax, Nova Scotia)

Czech Republic (City)

- Jan Cap, MD, PhD; Charles University (Hradec Kralove)

France (City)

- Jerome Bertherat, MD, PhD; Hôpital Cochin (Paris)
- Olivier Chabre, MD, PhD; Université J. Fourier (Grenoble)
- Justine Cristante, MD, PhD; Université J. Fourier (Grenoble)
- Delphine Drui, MD; CHU de Nantes - Hôpital Guillaume-et-René-Laennec (Saint-Herblain)
- Patrice Rodien, MD; Claire Briet, MD, PhD; Frederic Illous, MD; Department of Endocrinology, Diabetology and Nutrition, Endo-ERN Center for Rare Endocrine Conditions, CHU d'Angers and Laboratoire MITOVASC, Université d'Angers (Angers)
- Philippe Touraine, MD; Hôpitaux Universitaires Pitié Salpétrière Charles Foix, Sorbonne University (Paris)
- Jacques Young, MD, PhD; University Paris Saclay (Paris)

Germany (City)

- Jörg Bojunga, MD; Frankfurt University Hospital (Frankfurt)
- Nicole Reisch, MD; Medizinische Klinik und Poliklinik IV, Endocrinology, LMU Klinikum München (Munich)

Greece (City)

- Zoe Efstathiadou, MD, PhD; ”Hippokration” General hospital of Thessaloniki (Thessaloniki)
- Gregory Kaltsas, MD, FRCP; Kapodistrian University of Athens, Medical School (Athens)
- Stylianos Tsagarakis, MD, PhD, FRCP; Evangelismos Hospital (Athens)

Israel (City)

- Alina German, MD; Emek Medical Center (Afula)

Italy (City)

- Gilberta Giacchetti, MD; University of Ancona (Ancona)

- Andrea M. Isidori, MD; Dario De Alcubierre, MD; Davide Ferrari, MD; Riccardo Pofi, MD, PhD; Sapienza University (Rome)
- Mario Maggi, MD; University of Florence (Florence)
- Uberto Pagotto, MD, PhD; University of Bologna (Bologna)
- Luca Persani; MD, PhD; University of Milan (Milan)
- Rosario Pivonello, MD, PhD; Chiara Simeoli, PhD; Angelica La Rocca; Erminio Massimo Crescenzo; University “Federico II” of Naples (Naples)
- Gianni Russo, MD; Marta D’Amico, MD; Marianna Rita Stancampiano, MD; Carmen Bucolo; S. Raffaele Hospital (Milan)
- Carla Scaroni, MD; University of Padova (Padova)

Poland (City)

- Lucyna Bednarek-Papierska, MD; Medical Academy of Warsaw (Warsaw)
- Marek Ruchala, MD; Karol Marcinkowski Medical University (Poznan)
- Grzegorz Sokolowski, MD; Centrum Nowoczesnych Terapii “Dobry Lekarz” (Krakow)

Portugal (City)

- Duarte Pignatelli, MD; University of Porto (Porto)

Spain (City)

- Miguel Angel Mangas Cruz, MD, PhD; Hospital Universitario Virgen del Rocio (Sevilla)
- Manuel Luque Ramirez, PhD, MD, MBA; Hospital Universitario Ramon y Cajal (Madrid)

Sweden (City)

- Henrik Falhammar, MD, PhD, FRACP; Martin Larsson, MD, PhD; Ana Maria Cristina Carasel; Karolinska University Hospital (Stockholm)

United Kingdom (City)

- Safwaan Adam, MRCP, PhD; The Christie NHS Foundation Trust (Manchester)
- Aled Rees, MRCP, PhD, FRCP, MAcadMEd; Cardiff University (Cardiff)
- Umasuthan Srirangalingam, MB ChB, FRCP, PhD; Matilde Calanchini, MD, PhD; Aiyappa Biddanda, MBBS; University College London Hospitals (London)

United States (City, State)

- Richard J. Auchus, MD; Xin He, MD, MBA (supported in part through grant T32DK007245 from the National Institutes of Diabetes, Digestive and Kidney Diseases); University of Michigan Medical School (Ann Arbor, MI)
- Irina Bancos, MD; Mayo Clinic (Rochester, MN)
- Laura J. Chalmers, DO; University of Oklahoma School of Community Medicine (Tulsa, OK)
- Patricia Y. Fechner, MD; Seattle Children’s Hospital (Seattle, WA)
- Eric Felner, MD, MS; Emory University School of Medicine (Atlanta, GA)

- Luigi R. Garibaldi, MD; UPMC Children's Hospital of Pittsburgh (Pittsburgh, PA)
- Oksana Hamidi, DO; Jessica Abramowitz, MD; Amy Shah, MD; Sasan Mirfakhraee, MD; UT Southwestern Medical Center (Dallas, TX)
- Erik A. Imel, MD, MS; Indiana University School of Medicine; James Whitcomb Riley Hospital for Children (Indianapolis, IN)
- Mimi S. Kim, MD, MSc; Children's Hospital Los Angeles (Los Angeles, CA)
- Kanthi Krishna, MD; UPMC Children's Hospital of Pittsburgh (Pittsburgh, PA)
- Chienying Liu, MD; University of California, San Francisco (San Francisco, CA)
- Karen Loechner, MD, PhD; Emory University School of Medicine (Atlanta, GA)
- Deborah Merke, MD, MS; Lee Ann Keener, CRNP [Department of Pediatrics, NIH Health Clinical Center]; Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health (Bethesda, MD)
- Alyson K. Myers, MD; Donald and Barbara Zucker School of Medicine at Hofstra/Northwell; Feinstein Institute (Great Neck, NY)
- Ron S. Newfield, MD; Rady Children's Hospital (San Diego, CA)
- Natalie J. Nokoff, MD; Children's Hospital Colorado (Aurora, CO)
- Maamoun Salam, MD; Washington University School of Medicine; Barnes-Jewish Hospital (St. Lous, MO)
- Kyriakie Sarafoglou, MD; University of Minnesota Masonic Children's Hospital (Minneapolis, MN)
- Maria G. Vogiatzi, MD; Children's Hospital of Philadelphia (Philadelphia, PA)
- Stuart A. Weinerman, MD; Donald and Barbara Zucker School of Medicine at Hofstra/Northwell; Feinstein Institute (Great Neck, NY)
- Selma Feldman Witchel, MD; UPMC Children's Hospital of Pittsburgh (Pittsburgh, PA)
- Mabel Yau, MD; Icahn School of Medicine at Mount Sinai; Mount Sinai Hospital; Beth Israel (New York, NY)

2.0 STUDY DESIGN

2.1 Participants

Key inclusion criteria:

- Provided written informed consent.
- Female or male at least 18 years of age.
- Have a confirmed diagnosis of classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency (21OHD) based on standard medically accepted criteria such as elevated 17-hydroxyprogesterone (17OHP) level, confirmed CYP21A2 genotype, positive newborn screening with confirmatory second-tier testing, or cosyntropin stimulation.
- Chronic treatment with a supraphysiological glucocorticoid (GC) regimen, defined as >13 mg/m^2/day in hydrocortisone equivalents (HCe) adjusted for body surface area (BSA), that was stable for ≥1 month prior to screening and included any of the following orally administered GCs: hydrocortisone, prednisone, prednisolone, methylprednisolone, and/or dexamethasone.
    - Conversion factors for HCe were as follows: methylprednisolone, prednisolone, prednisone (4x); dexamethasone (60x)
- If treated with fludrocortisone, stable dose for ≥1 month prior to screening with an upright plasma renin activity (PRA) at screening that was not greater than upper limit of normal (ULN) usual sodium intake. If PRA >ULN, systolic blood pressure >100 mmHg, without orthostatic hypotension, and with serum sodium and potassium in the normal range.
    - Participants taking fludrocortisone were strongly encouraged to maintain their dose during the study. All dose changes were documented.
- For female participants of childbearing potential, consistent use of contraception from screening until final study visit or 30 days after the last dose of study drug, whichever is longer. For female participants not of childbearing potential:
    - Postmenopausal, defined as no menses for 12 months without an alternative medical cause and confirmed by elevated follicle-stimulating hormone (FSH) consistent with a postmenopausal range
    - Permanent sterilization procedure, such as hysterectomy, bilateral salpingectomy, or bilateral oophorectomy

Key exclusion criteria:

- Known or suspected diagnosis of any other form of classic CAH.
- History of bilateral adrenalectomy, hypopituitarism, or other condition requiring chronic therapy with oral GCs, or required chronic therapy with inhaled GCs that might interfere with study end points (per investigator judgment).
- Evidence of GC overtreatment during screening.
- Increased risk of developing adrenal crisis (per investigator judgment).

- Clinically significant medical condition or chronic disease (e.g., history of neurological, hepatic, renal, cardiovascular, gastrointestinal, significant malabsorption, hematologic, pulmonary, psychiatric, or endocrine disease [excluding CAH]) that would preclude participant or completion of the study or might confound interpretation of study outcome (per investigator judgment).
- History of malignancy, unless successfully treated with curative intent and considered to be cured.
- Known history of clinically concerning cardiac arrhythmia (including long QT syndrome) or prolongation of screening (pre-treatment) QT interval, corrected for heart rate using Fridericia's correction (QTcF).
- Known sensitivity (i.e., hypersensitivity) or allergy to any corticotropin-releasing hormone (CRH) receptor antagonist.
- Evidence of chronic renal or liver disease.
- Clinically significant hematologic or coagulation abnormalities at screening.
- Used any active investigational drug for another clinical trial within 30 days or 5 half-lives (whichever is longer) before screening, or plan to use such drug during the CAHtalyst study.
- Use of any of the following excluded concomitant medications (unless discontinuation during the study was possible):
    - Orally administered GCs for indications other than CAH
    - Strong inducers of CYP3A4 or CYP2B6 except topically administered medications
    - Medications that affect cortisol or GC metabolism (e.g., phenytoin, mitotane, phenobarbital, strong CYP3A4 inhibitors such as ketoconazole, clarithromycin, cholestyramine, certain antivirals) except topically administered medications
    - Aromatase inhibitors (e.g., anastrozole, letrozole, testolactone)
- Current dependence or abuse of any of the following: alcohol, controlled substances, non-prescribed use of prescription drugs, nicotine, or caffeine.
- Significant risk of suicidal or violent behavior.
- Blood loss ≥550 mL or blood/plasma donation within 8 weeks before baseline (day 1).
- Pregnancy or lactation (in females).
- Not capable of adhering to the protocol requirements (per investigator judgment).

2.2 Randomization and Trial Interventions

2.2.1 Randomization and Blinding

Eligible individuals were randomized 2:1 to crinecerfont or placebo at baseline using an interactive response technology. Randomization was stratified by total daily GC regimen (<20 mg/m^2/day or ≥20 mg/m^2/day in HCe, adjusted for BSA), GC type (hydrocortisone alone; prednisone, prednisolone, methylprednisolone, with or without hydrocortisone; dexamethasone, with or without another GC), and sex.

Blinding was maintained except for safety reasons. All participants were encouraged to complete follow-up even if an unblinding occurred. The participant, investigator, and all study center personnel remained blinded to the participant's randomized treatment assignment through database lock.

An independent Data Monitoring Committee (DMC) periodically reviewed ongoing unblinded clinical and safety data to ensure the safety and well-being of study participants and also reviewed results from the unblinded interim analysis (Section 4.4).

2.2.2 Glucocorticoid Reduction

From weeks 4 to 12, participants underwent a down titration (in 4 or fewer steps) to a target dose of 8 to 10 mg/m^2/day HCe. Based on the starting and target GC doses, the investigator had a detailed schedule for the dose reduction that would occur between weeks 4 and 12. Details of the participant's baseline GC regimen and dose reduction schedule were submitted to the Sponsor for review prior to randomization.

If the participant experienced any following signs or symptoms during GC reduction, the dose was returned to the previous tolerated dose: unexplained hyponatremia, orthostatic hypotension, severe nausea or vomiting, unacceptable symptoms of hyperandrogenism (e.g., hirsutism, acne, amenorrhea). Before GC reduction was stopped for orthostatic hypotension, volume status was optimized (e.g., with additional dietary salt, salt tablets, intravenous saline). GC dose reductions proceeded if androstenedione increased, provided that the increase was tolerated by the participant.

At week 12, based on each participant's hormone levels and clinical assessments, investigators determined which GC dose and regimen was most appropriate for continuation past week 12, based on tolerability and ability to maintain androgen control (i.e., androstenedione ≤120% of baseline or ≤ULN for age and sex).

2.2.3 Glucocorticoid Optimization

After weeks 12, 16, and 20, investigators reviewed laboratory results from the preceding study visit and determined whether the GC dose or regimen required adjustment in order to maintain androstenedione control.

2.2.4 Stress Dosing

Guidelines for GC reduction or optimization were followed except when a participant required stress or "sick-day" dosing (Table S1). If GC dosing was adjusted due to sick-day guidelines, the participant resumed their GC regimen for ≥3 days before their next scheduled hormone panel assessment.

3.0 ASSESMENTS AND ENDPOINTS

Analyses of the primary and key secondary end points were controlled for multiplicity (Section 4.2). Missing data for the primary and key secondary efficacy end points were imputed using a regression-based multiple imputation method which assumes data are missing at random (Section 4.3).

3.1 Primary Efficacy End Point

The primary efficacy end point was the percent change from baseline in GC daily dose (mg/m^2/day in HCe, adjusted for BSA) at week 24 while androstenedione was controlled (at $\leq$120% of baseline or $\leq$ULN for age and sex).

3.2 Key Secondary End Points

- Change from baseline in serum androstenedione at week 4
- Achievement of a reduction in GC daily dose to physiological levels ($\leq$11 mg/m^2/day) at week 24 while androstenedione was controlled
- Change from baseline in homeostatic model assessment of insulin resistance (HOMA-IR) at week 24 in fasting participants not on insulin
- Percent change from baseline in body weight at week 24
- Change from baseline in percent total fat mass at week 24, assessed using a whole-body scan

3.3 Secondary End Points

- Change from baseline in 17-hydroxyprogesterone (17OHP) at week 4
- Change from baseline in blood pressure at week 24
- Change from baseline in glucose tolerance at week 24, as measured by post-GTT (glucose tolerance test) load glucose levels
- Change from baseline in waist circumference at week 24
- Change from baseline in menstrual regularity at week 24 in premenopausal female participants not using hormonal or intrauterine device contraception
- Change from baseline in testicular adrenal rest tumor (TART) volume at week 24 in male participants, expressed as a percentage of total testicular volume

3.4 Exploratory Bone Marker End Points

Bone formation was assessed using changes in serum osteocalcin and bone-specific alkaline phosphatase. Serum C-terminal telopeptide (CTx) and urine N-terminal telopeptide (NTx, collected after the first morning void after an overnight fast) were used to assess bone resorption.

Changes from baseline in these measures at week 24 were defined as exploratory end points.

3.5 Safety Assessments

3.5.1 Treatment-Emergent Adverse Events (TEAEs)

TEAEs included any untoward medical occurrence in a participant who received study treatment, regardless of whether the event was judged by the investigator to have a causal relationship with

the treatment. TEAEs included, but were not limited to abnormal laboratory findings, worsening of a condition that present at study entry, deterioration beyond what would be expected due to the primary illness, intercurrent illness, drug interaction, and suicidal ideation or behavior.

TEAEs were graded by the investigator for intensity (mild, moderate, severe) and relatedness to study drug (definite, possible, unlikely, not related).

Participants experiencing serious TEAE(s) or emergency situation were examined as quickly as possible by a physician, who did whatever was medically needed for their safety and well-being. Participants were followed until their serious TEAE(s) were resolved or until they were medically stabilized.

Serious TEAEs and pregnancies were reported within 24 hours of first knowledge of the event by any study personnel to Neurocrine's Drug Safety and Pharmacovigilance (DSPV) department. Investigators were requested to provide an assessment of the relationship to study drug at the time of this initial report.

3.5.2 Vital Signs and Physical Examinations

Vital signs were measured before blood sample collections at screening, baseline, and weeks 4, 6, 9, 12, 16, 20, and 24. These included blood pressure, pulse rate, respiratory rate, body temperature, and body weight.

Complete physical examinations were performed at screening, baseline, and weeks 4, 12, and 24. They included assessment of general appearance, skin and mucosae, head, eyes, ears, nose, throat, neck (including thyroid), lymph nodes, chest/lungs, cardiovascular, abdomen, extremities, musculoskeletal, and neurological system. Waist circumference was also assessed, using a measuring tape placed in a horizontal plane around the abdomen at the level of the iliac crest.

3.5.3 Electrocardiograms and Clinical Laboratory Tests

A standard 12-lead electrocardiogram (ECG) was administered at screening, baseline, week 4, and week 24. ECGs that were rated by investigators as "abnormal clinically significant" were recorded as a TEAE.

Routine laboratory tests (e.g., hematology, coagulation, clinical chemistry, urinalysis, thyroid, pregnancy [on Day 1 to confirm study eligibility]) were performed by a central laboratory, which provided instructions and supplies to clinical trial personnel before study initiation.

3.5.4 Brief Psychiatric Rating Scale (BPRS)

The BPRS was administered at screening, baseline, week 12, and week 24. It was designed to assess the severity of psychopathology in patients with schizophrenia and other psychotic disorders [Overall and Gorham 1988]. The BPRS includes 18 items (somatic concern, anxiety, emotional withdrawal, conceptual disorganization, guilt feelings, tension, mannerisms and posturing, grandiosity, depressive mood, hostility, suspiciousness, hallucinatory behaviors, motor retardation, uncooperativeness, unusual thought content, blunted affect, excitement, and disorientation), with each item rated on a scale of 1 (not present) to 7 (extremely severe) (total score range: 18 to 126).

3.5.5 Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS, a validated instrument used to prospectively assess suicidal ideation and behavior [Posner 2011], was administered throughout the treatment period. Versions have been designed for

use at screening/baseline and subsequent visits throughout the study. All versions of the C-SSRS include a series of questions related to suicidal ideation and suicidal behavior.

Responses of “yes” to one or more questions prompt additional questions that evaluate frequency and intensity of suicidal ideation and/or behavior. Individuals who had any suicidal behavior within the past year or suicidal ideation type 4 (active suicidal ideation with some intent to act, without specific plan) or type 5 (active suicidal ideation with specific plan and intent) in the 6 months before screening were excluded from the study.

4.0 STATISTICAL METHODS

4.1 Sample Size Determination

The protocol-specified sample size of 165 participants (55 placebo, 110 crinecerfont) was based on a power calculation for the primary end point and considerations for the size of the safety database. Based on a 2-sample t-test, an effect size of 0.75 with a sample size of at least 90 participants (30 placebo, 60 crinecerfont) was estimated to have greater >90% power to detect a treatment difference at a 0.05 level of significance. With the full sample size of 165 participants, there was >90% power to detect an effect size as small as 0.55.

4.2 Multiple Comparisons and Multiplicity

A testing procedure was used to control the family-wise Type I error rate for the primary and key secondary end points (Figure S2).

- The primary end point was tested using a significance level of 0.05.
- If the primary end point was statistically significant, then the first two key secondary end points were tested using an unequal allocation of significance levels. The change from baseline in androstenedione at week 4 was tested using a significance level of 0.02. The achievement of a reduction in GC daily dose to physiological levels (≤11 mg/m^2/day in HCe) was tested using a significance level of 0.03.
- The remaining key secondary end points were tested collectively using the Holm procedure [Holm 1979]. The significance level for this family of end points was dependent on whether the first two key secondary end points were statistically significant [Wiens 2003]. If both were statistically significant, then the significance level was 0.05. If only one was statistically significant, then the significance level for that end point was used for the family of end points. If neither was significant, then the family of end points was not tested for statistical significance.
- Secondary and exploratory end points were not controlled for multiplicity.

4.3 Statistical Models and Imputation Procedures

Analyses were performed using SAS® software, version 9.4 or later. Unless otherwise noted, efficacy data from the double-blind, placebo-controlled period (DBPC: baseline [day 1] to week 24) were summarized in all randomized participants, regardless of adherence to study drug treatment. Safety data from the DBPC were analyzed with descriptive statistics in all randomized participants who took ≥1 dose of study drug.

Missing data for the primary and key secondary efficacy end points were imputed using a regression-based multiple imputation method which assumes data are missing at random.

Continuous efficacy end points were analyzed using an analysis of covariance (ANCOVA) model, which included the change (or percent change) from baseline to the respective postbaseline visit where the end point was being evaluated. The model included the relevant baseline value as a covariate, treatment group, and stratification factors used in the randomization. Values of the stratification factors entered during randomization via the interactive response technology were used in the primary analysis methods.

Categorical efficacy end points were analyzed using the Cochran-Mantel-Haenszel (CMH) test. The CMH test included the stratification factors used in randomization.

In general, baseline was defined as the last assessment prior to study drug dosing on Day 1. Additional details for the primary and key secondary efficacy end points are described below.

4.3.1 Primary End Point

The primary efficacy end point was the percent change from baseline in GC daily dose (mg/m^2/day in HCe, adjusted for BSA) at week 24 while androstenedione was controlled (at ≤120% of baseline or ≤ULN for age and sex).

- Serum androstenedione was collected at baseline (day 1) and week 24 before the morning GC dose (pre-GC dose) and 2 hours after the morning dose (post-GC dose).
- The post-GC dose androstenedione values at day 1 was considered the baseline value for assessment of androstenedione control at week 24 (which also used post-GC dose value).
- If the baseline androstenedione value was missing, then the last post-GC dose androstenedione value collected prior to day 1 was used as the baseline.
- If the post-GC dose androstenedione value for week 24 was missing, then the pre-GC doses from baseline and week 24 were used to determine androstenedione control.
- If both pairs of pre- and post-GC androstenedione values at baseline and week 24 were incomplete, then the androstenedione values for the purpose of the primary end point (GC dose reduction at week 24 with androstenedione control) was considered missing and was multiply imputed using the methods described below.

The least-squares mean treatment difference (LSMD) between treatment groups (crinecerfont – placebo) was estimated using an ANCOVA model, where the values of the covariates were set at overall means.

- If a participant experienced a *decrease* in GC daily dose but did not maintain androstenedione control at week 24, their GC daily dose end point was set to 0% change from baseline.
- If a participant discontinued study drug but remained in the study through week 24, their observed GC daily dose and androstenedione value at week 24 were used in the analysis.
- If a participant was on study drug through week 24 but was missing GC dose and/or androstenedione values at that visit, their end point was multiply imputed from participants within the same treatment group who did not have missing data at week 24.
- If a participant in the crinecerfont treatment group discontinued from study drug and was missing GC dose and/or androstenedione values at week 24, their end point was multiply imputed using data from retrieved dropouts. If there was an insufficient number of retrieved dropouts, the participant's end point was imputed using observed data from participants in the placebo treatment group.
    - Retrieved dropouts were defined as randomized participants who discontinued study drug prior to time of end point collection but remained in the study and had an evaluable measurement for that end point. For end points with >1 data component (e.g., GC reduction with androstenedione control), the participant needed to have data for all of the components.
- If a medication was marked as "ongoing" at the time of a participant's discontinuation from study, or if they were lost to follow-up, the stop date for the medication was imputed to the participant's end of study date.

As described above, participants missing GC dose or androstenedione at week 24 had their data imputed using a regression-based multiple imputation method assuming missing at random. The following steps were taken:

- Participants with a missing GC dose, androstenedione, or both at week 24 had their data multiply imputed using the fully conditional specifications (FCS) method.
- This method was implemented in SAS® 9.4 PROC MI using a series of conditional models (one for each incomplete variable at week 24). First, the missing week 24 post-dose androstenedione values was imputed using a regression-based multiple imputation model and the model included baseline androstenedione (post-GC dose), age, and sex (as specified in the electronic case report form [eCRF]).
- The subsequent FCS regression model imputed the week 24 GC total daily doses (in mg/m^2/day) and the model included baseline GC dose (mg/m^2/day), baseline androstenedione (post-GC dose), week 24 androstenedione (post-GC dose), and sex (as specified in the eCRF).
- Note that a minimum of 0 was set for imputed week 24 post-GC dose androstenedione values and the week 24 GC daily dose (mg/m^2/day) as it is biologically implausible for either of these values to be less than 0.
- For participants missing androstenedione at week 24, androstenedione "control" was determined based on the imputed week 24 androstenedione value and the percent change from baseline to week 24 in GC total daily dose was set to 0 if the imputed androstenedione value did not indicate "control."
- The number of imputed datasets created from the steps above was 100.
- Note that these PROC MI steps were done separately depending on the participant's treatment group and study drug status at the time of the end point (as described above).
- The multiply-imputed datasets generated from the steps above were set back together prior to the final analysis steps.
- Each of the 100 datasets were analyzed using the ANCOVA model described above. The results of these analyses were combined using PROC MIANALYZE.

4.3.2 First and Second Key Secondary End Points

The first key secondary end point was change from baseline in serum androstenedione at week 4.

Baseline and week 4 values were analyzed using samples that were taken prior to participants' morning GC doses (pre-GC dose).

- For participants on study drug who were missing serum androstenedione values at week 4, their missing data were imputed using a regression-based multiple imputation model based on data from participants from the same treatment group who had non-missing data. The model included the baseline value for the end point and sex.
- For participants in the crinecerfont group who were not on study drug and missing serum androstenedione values at week 4, missing data were imputed using a regression-based imputation model based on data from retrieved dropouts. If there were an insufficient number of retrieved

dropouts, the end point was imputed from participants in the placebo group. For participants in the placebo group, missing data were imputed using non-missing data in the placebo group. The model included the baseline value for the end point and sex.

- See Section 4.3.3 for additional information regarding the multiple imputation method.
    - The second key secondary end point was achievement of physiological GC dose (≤11 mg/m^2/day in HCe, adjusted for BSA) at week 24 while androstenedione was controlled
- If GC dose and/or androstenedione value were missing at week 24, the multiply-imputed datasets from the primary end point were used to determine achievement of this end point.
- For the second key secondary end point, the Wilson-Hilferty transformation [Ratitch 2013] was applied to the CMH test statistic to standardize the resulting normal variable before combining the CMH results via PROC MIANALYZE.

4.3.3 Additional Key Secondary End Points

Additional key secondary end points were the changes from baseline in HOMA-IR and percent total fat mass at week 24, and the percent change from baseline in body weight at week 24

- Participants taking insulin at baseline or not fasting at the time of the assessment were excluded from the HOMA-IR analysis.
- Percent total fat mass was calculated as follows: (total fat mass [g] / total body mass [g]) x 100.

A multiple imputation procedure was used for participants with missing HOMA-IR, body weight, or fat mass data at week 24. The following steps were taken to multiply-impute participants with missing data for these key secondary end points as well as the key secondary end point of change from baseline in androstenedione at week 4:

- Using SAS® 9.4 PROC MI, missing data for participants on study drug without the end point at the timepoint of interest were imputed using a regression-based multiple imputation model based on data from participants with non-missing end point data within the same treatment group. The model included the baseline value for the end point and sex.
- For participants in the crinecerfont group off study drug without the end point, missing data were imputed using a regression-based imputation model based on data from retrieved dropouts. If there were an insufficient amount of retrieved dropouts, the end point was imputed from participants in the placebo treatment group.
- For participants in the placebo group, missing data were imputed using non-missing data from participants in the placebo group. The model included the baseline value for the end point and sex.
- The number of imputed datasets created from the steps above for each end point was 100. After the multiply-imputed datasets were created, each of the datasets was analyzed using the ANCOVA model described above. The results of these analyses were combined using PROC MIANALYZE.
- Note: In the case of fat mass data, some participants were missing a baseline value. For this analysis, a series of FCS model statements were created within PROC MI where the baseline value was first imputed based on age and sex and then the week 24 fat mass values were imputed based on baseline fat mass and sex.

4.4 Interim Analysis

A planned interim analysis, unblinded only to the DMC, was conducted on the primary end point using the "promising zone" method [Mehta and Pocock, 2011] when 76 (46%) participants completed their week 24 assessments. The promising zone method has been shown to not inflate the type I error. This interim analysis included an unblinded sample-size re-estimation to mitigate against uncertainty of sample size assumptions as well as a nonbinding assessment of futility. The potential outcomes of the interim analysis, based on the 'zone' or level of the treatment effect and conditional power, were stop for futility, continue to the planned sample size, or increase the sample size to a maximum of 210 participants.

The interim analysis results were prepared by an Independent Statistical Center in January 2023 with a review of the results by the DMC. As a result of this meeting, the DMC recommended to continue the study as planned with no increase to sample size and the study was continued without modification.

The study Sponsor (except for the DMC), participants, and investigators all remained blinded during the interim analysis.

4.5 Sensitivity Analyses

4.5.1 Primary End Point

Sensitivity analyses for the primary end point (percent change from baseline in GC daily dose at week 24 while androstenedione was controlled) included a tipping-point analysis and complete-case analysis.

The tipping-point sensitivity analysis was performed to assess the robustness of the missingness assumptions. This analysis was based on "delta adjustments", a commonly used approach to assess the impact of missing data in clinical trials [O'Kelly and Ratitch 2014]. Imputed week 24 GC daily dose values (for participants in either treatment group with missing data) were delta-adjusted with a range of either penalties (for the crinecerfont group) or improvements (for the placebo group) until the treatment difference at week 24 was no longer statistically significant. Per feedback, the initial plan to apply penalties to just the crinecerfont group was later updated to apply penalties to the crinecerfont group while simultaneously applying improvements to the placebo group in what is called a 2-dimensional tipping point analysis. Results of this analysis were as follows:

- The tipping point ($P \geq 0.05$) was reached when the combined delta shift was 43 mg/m^2/day or greater (i.e., when the delta shift for crinecerfont was 23 mg/m^2/day or higher and the delta shift for placebo was -20 mg/m^2/day or lower)
- These values represent substantially large improvements (placebo) and penalties (crinecerfont) in order to tip the result into being non-significant; therefore, the tipping point analysis confirmed the robustness of the missing data assumptions and the primary analysis of the primary end point.

The complete-case sensitivity analysis was performed to test the robustness of the multiple imputation procedure. The ANCOVA analysis for the primary end point was repeated using only participants who had primary end point data at week 24. Those with missing data were excluded from the analysis. Results of this analysis were as follows:

- The LS mean percent change from baseline (±SEM) in GC daily dose at week 24 was -28.2% (±2.4) for the crinecerfont group and -11.0% (±3.2) for the placebo group.
- The LS mean difference (crinecerfont – placebo) was -17.2% (95% CI: -23.8, -10.6).

- The results of the complete case analysis were consistent with the primary analysis of the primary end point with similar main estimates.

4.5.2 First Key Secondary End Point

Sensitivity analyses for the first key secondary end point (change from baseline in serum androstenedione at week 4) were the same as those used for the primary end point.

Results of the tipping-point analysis were as follows:

- The tipping point (P≥0.05) was reached when the combined delta shift was 3381 ng/dL (118 nmol/L) or greater (i.e., when the delta shift for crinecerfont was 1662 ng/dL [58 nmol/L] or higher and the delta shift for placebo was -1719 ng/dL [-60 nmol/L] or lower).
- These values represent substantially large improvements (placebo) and penalties (crinecerfont) in order to tip the result into being non-significant, therefore the tipping point analysis confirmed the robustness of the missing data assumptions and the primary analysis of the first key secondary end point.

Results of the complete-case analysis were as follows:

- The LS mean (±SEM) change from baseline in serum androstenedione at week 4 was -299.3 (±37.9) ng/dL (-10.5 [±1.3 nmol/L) for the crinecerfont group and 40.7 (±51.2) ng/dL (1.42 [±1.79] nmol/L) for the placebo group.
- The LS mean difference (crinecerfont – placebo) of -340 ng/dL (11.9 nmol/L) (95% CI: -453, -227).
- The results of the complete case analysis were consistent with the primary analysis of the first key secondary end point with similar main estimates.

4.5.3 Second Key Secondary End Point

A tipping-point analysis was not performed for the second key secondary end point (achievement of physiological GC dose with androstenedione control). Due to a small amount of missing data, it was not possible to “tip” the result to not statistically significant. Therefore, a highly conservative method was implemented in which participants in the crinecerfont group who were missing data were counted as not achieving the end point, while those missing data in the placebo group were counted as having achieved the end point. Results of this analysis were as follows:

- Seven participants had missing data (4 crinecerfont, 3 placebo) for this end point.
- When participants in the crinecerfont group were counted as not achieving the end point *and* those in the placebo group were counted as achieving the end point, the results were consistent with the primary analysis.

A complete-case analysis was performed to test the robustness of the multiple imputation procedure. Inferential statistics from the complete-case analysis were similar to those in the primary analysis of this end point.

5.0 FIGURES

Figure S1. Study Design

For the first 4 weeks of the double-blind, placebo-controlled (DBPC) period, GC dosing remained unchanged. From weeks 4 to 12, GC dosing was reduced in 4 or fewer steps to reach a target dose of 8-10 mg/m^2/day (in HCe) by week 12. A safety follow-up (SFU) was conducted by telephone within 1 week after each GC dose reduction. From weeks 12 to 24, GC doses were adjusted as necessary to reach the lowest tolerated dose needed to achieve androgen control.

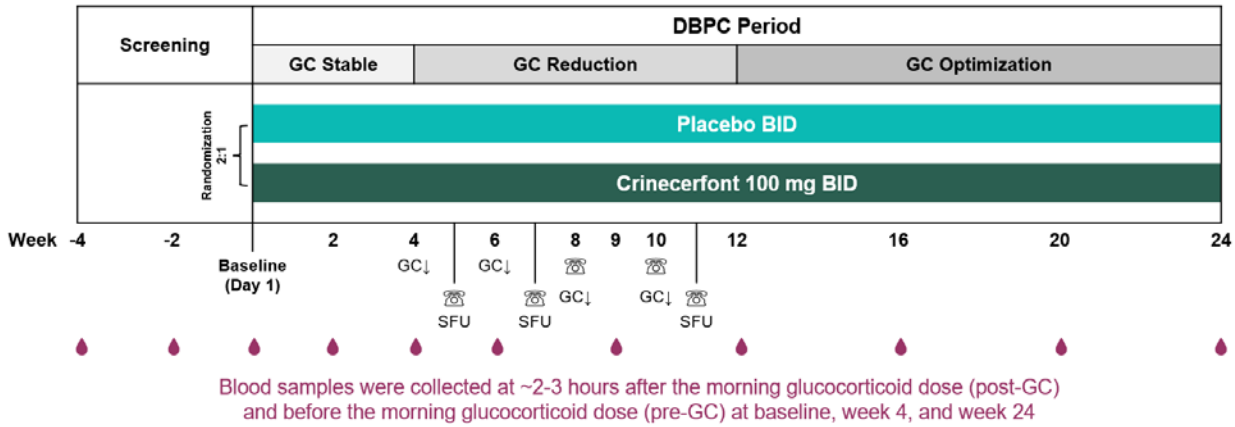
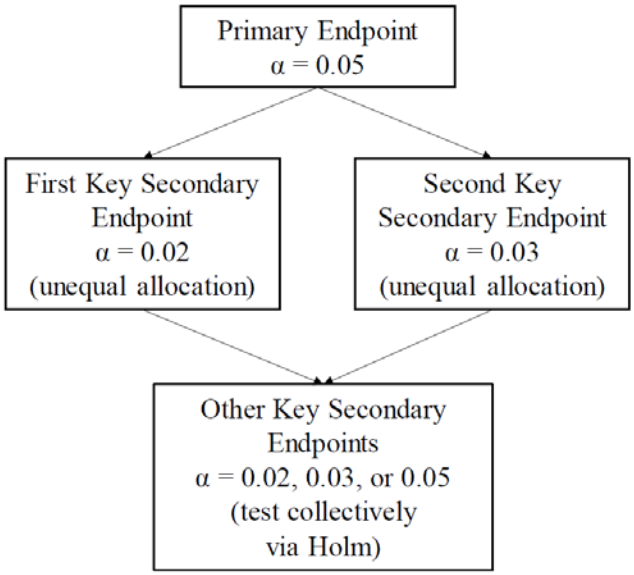


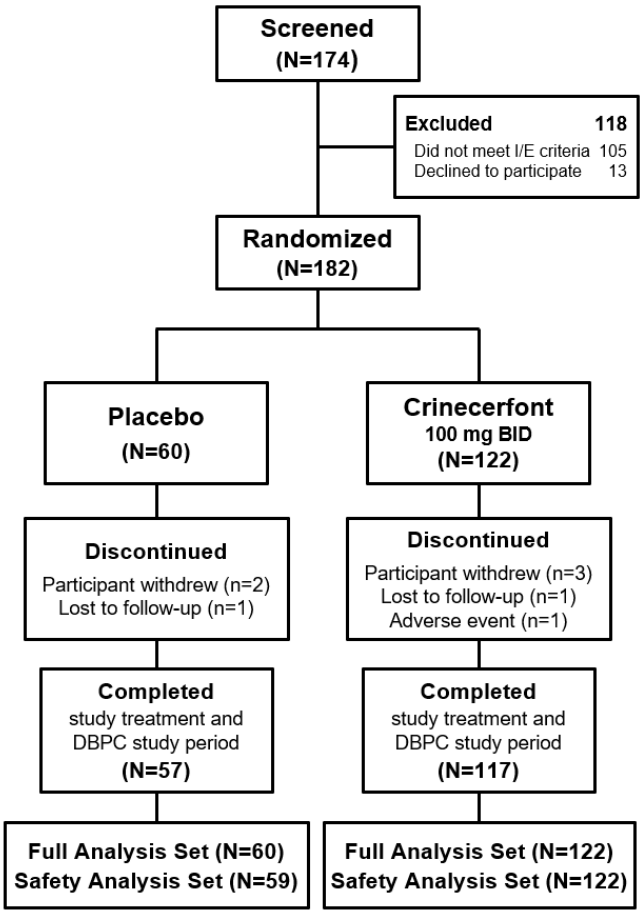
Figure S2. Procedure to Control for Multiple Comparisons



All significance levels were 2-sided. Secondary and exploratory end points were not adjusted for multiplicity.

Figure S3. Study Flow

Of 300 adults who were screened for study entry, the most common reasons for ineligibility were androstenedione or 17OHP below the lower limit of normal based on sex and age (n=65) and not being treated with supraphysiological GC doses (n=26). Of 182 participants who met the eligibility criteria and were randomized (2:1) to crinecerfont or placebo, 176 (96.7%) completed the double-blind placebo-controlled period, including 2 participants (1 in each treatment arm) who withdrew after completing all week 24 assessments and contributed data to the primary endpoint. One participant in the placebo group who did not receive study drug withdrew and was excluded from the safety analysis set. One participant in the crinecerfont group discontinued during the double-blind period due to an adverse event of gastric ulcer, which was judged by the investigator as unlikely related to study therapy.



6.0 TABLES

Table S1. Stress Dosing Instructions

Type of Illness	Instruction
Minor illness or low-grade fever	Take hydrocortisone 10 mg 3x/day in addition to usual GC regimen
Major illness or high-grade fever	Take hydrocortisone 20 mg 3x/day in addition to usual GC regimen; increase fluid intake with frequent ingestion of simple and complex carbohydrates
All illnesses	30 g of simple carbohydrates (1 cup juice or regular soda) for lethargy; increase fluids. Hospital/physician evaluation for lethargy and decreased oral intake and urine output.
Vomiting	Repeat the dose if patient vomits within 1 hour of medication intake. If patient vomits again, administer intramuscular injection of hydrocortisone 100 mg. If unable to tolerate fluids, call emergency services for evaluation following intramuscular injection.

Table S2. Hormone Reference Ranges

Lab Analyte: Conversion Factor	Reference Ranges: Conventional Units, SI Units
Androstenedione: ng/dL to nmol/L (x 0.0349)	Male adult: • 18–30 years: 50–220 ng/dL; 1.75–7.68 nmol/L • 31–50 years: 40–190 ng/dL; 1.40–6.63 nmol/L • 51–60 years: 50–220 ng/dL; 1.75–7.68 nmol/L Female adult: • Mid Follicular: 51–213 ng/dL; 1.78–7.43 nmol/L • Surge: 73–230 ng/dL; 2.55–8.03 nmol/L • Mid Luteal: 73–184 ng/dL; 2.55–6.42 nmol/L • Postmenopausal: 20–75 ng/dL; 0.70–2.62 nmol/L
17–hydroxyprogesterone: ng/dL to nmol/L (x 0.0303)	Male adult: • 18–30 years: 32–307 ng/dL; 0.97–9.30 nmol/L • 31–40 years: 42–196 ng/dL; 1.27–5.94 nmol/L • 41–50 years: 33–195 ng/dL; 1.00–5.91 nmol/L • 51–60 years: 37–129 ng/dL; 1.12–3.91 nmol/L • ≥61 years: not established Female adult: • Pre-menopausal mid-follicular: 23–102 ng/dL; 0.70–3.09 nmol/L • Pre-menopausal surge: 67–349 ng/dL; 2.03–10.57 nmol/L • Pre-menopausal mid-luteal: 139–431 ng/dL; 4.21–13.06 nmol/L • Post-menopausal phase: ≤45 ng/dL; 1.36 nmol/L
Adrenocorticotropic hormone: pg/mL to pmol/L (x 0.22)	Male and female adults: 7.2–63.3 pg/mL; 1.6–13.9 pmol/L
Follicle stimulating hormone: mIU/mL to IU/L (x 1.0)	Male adult: 1.6–11.0 mIU/mL (IU/L) Female adult: 1.8–20.4 mIU/mL (IU/L)
Luteinizing hormone: mIU/mL to IU/L (x 1.0)	Male adult: 1.7–8.6 mIU/mL (IU/L) Female adult: 1.0–95.6 mIU/mL (IU/L)
Testosterone: ng/dL to nmol/L (x 0.0347)	Male adult: 250–1100 ng/dL; 8.68–38.2 nmol/L Female adult: 2–45 ng/dL; 0.07–1.6 nmol/L

Table S3. Demographics and Baseline Characteristics (Randomized Population)

	All Participants (N=182)	Crinecerfont (N=122)	Placebo (N=60)
Demographics			
Age, mean (SD), years	30.8 (9.9)	31.3 (9.8)	29.8 (10.2)
Sex, n (%)			
Male	92 (50.5)	61 (50.0)	31 (51.7)
Female	90 (49.5)	61 (50.0)	29 (48.3)
Race, n (%)			
White	164 (90.1)	107 (87.7)	57 (95.0)
Asian	5 (2.7)	5 (4.1)	0
Black or African American	2 (1.1)	2 (1.6)	0
Native American / Alaska Native	1 (0.5)	1 (0.8)	0
Multiple (not specified)	2 (1.1)	1 (0.8)	1 (1.7)
Other (not specified)	3 (1.6)	2 (1.6)	1 (1.7)
Not collected	5 (2.7)	4 (3.3)	1 (1.7)
Ethnicity, n (%)			
Hispanic or Latino	14 (7.7)	6 (4.9)	8 (13.3)
Not Hispanic or Latino	168 (92.3)	116 (95.1)	52 (86.7)
Geographic region, n (%)			
United States	69 (37.9)	48 (39.3)	21 (35.0)
Rest of the world[a]	113 (62.1)	74 (60.7)	39 (65.0)
CAH medications at baseline			
GC daily dose in HCe, mean (SD), mg/day	32.3 (9.3)	32.4 (9.2)	32.1 (9.5)
Adjusted for BSA, mean (SD), mg/m^2/day	17.6 (4.9)	17.5 (4.5)	17.9 (5.5)
≥20 mg/m^2/day, n (%)	39 (21.4)	25 (20.5)	14 (23.3)
GC type, n (%)			
Hydrocortisone (HC) alone	106 (58.2)	71 (58.2)	35 (58.3)
Prednisone, prednisolone, or methylprednisolone, with or without HC	53 (29.1)	34 (27.9)	19 (31.7)
Dexamethasone, with or without another GC	23 (12.6)	17 (13.9)	6 (10.0)
Fludrocortisone, n (%)	157 (86.3)	107 (87.7)	50 (83.3)
Physical exams and tests at baseline			
Body weight, mean (SD), kg	79.3 (18.3)	80.8 (17.8)	76.2 (18.9)
BMI, mean (SD), kg/m^2	29.8 (7.0)	30.1 (6.9)	29.0 (7.1)
BMI ≥30, n (%), kg/m^2	85 (46.7)	58 (47.5)	27 (45.0)
BSA, mean (SD), m^2	1.8 (0.2)	1.9 (0.2)	1.8 (0.2)
Percent total fat mass, mean (SD)[b]	35.7 (9.2)	36.3 (9.0)	34.6 (9.5)
Waist circumference, mean (SD), cm[b]	94.8 (15.7)	96.1 (14.8)	92.1 (17.2)
HOMA-IR, mean (SD)[b]	3.2 (2.8)	3.2 (2.7)	3.1 (3.1)
Plasma glucose 2 hours post 75-g glucose load, mean (SD), mg/dL[c]	98.6 (25.6)	99.5 (26.1)	96.9 (24.7)

	All Participants (N=182)	Crinecerfont (N=122)	Placebo (N=60)
Blood pressure, mean (SD), mmHg			
Systolic	122.8 (12.6)	124.1 (12.3)	120.3 (13.0)
Diastolic	76.5 (9.7)	77.6 (9.1)	74.4 (10.5)
TARTs in male participants, n (%)[e]	53 (66.3)	35 (66.0)	18 (66.7)

a Includes Canada, Europe, and Israel.

b Number of participants with missing physical exams or tests: percent total fat mass (18 crinecerfont, 7 placebo); waist circumference (2 crinecerfont), HOMA-IR (4 crinecerfont, 2 placebo). Percent total fat mass corresponds to mean (SD) total fat mass in kg as follows: all participants, 28.9 (12.0); crinecerfont, 29.4 (11.3); placebo, 27.8 (13.3).

c In 172 participants (117 crinecerfont, 55 placebo) without diabetes mellitus.

d Based on blood samples collected before the morning GC dose. Normal ranges and conversion factors for conventional units to standard international units are in Appendix Table S2.

e In 80 male participants (53 crinecerfont, 27 placebo) with ultrasound data.

Abbreviations: BMI, body mass index; BSA, body surface area; CAH, congenital adrenal hyperplasia; GC, glucocorticoid; HCe, hydrocortisone equivalents (conversion factors: 4x for methylprednisolone, prednisolone, prednisone; 60x for dexamethasone); HOMA-IR, homeostatic model assessment for insulin resistance; SD, standard deviation; TART, testicular adrenal rest tumor.

Table S4. Hormones and Hormone Precursors at Baseline (Randomized Population)

	All Participants (N=182)			Crinecerfont (N=122)			Placebo (N=60)		
	n	Mean (SD)	Median (Min, Max)	n	Mean (SD)	Median (Min, Max)	n	Mean (SD)	Median (Min, Max)
Androstenedione, ng/dL[b]	180	620 (729)	387 (16, 5815)	121	635 (796)	417 (41, 5815)	59	590 (572)	369 (16, 2364)
17OHP, ng/dL[b]	179	9467 (8829)	7207 (53, 49299)	121	9314 (8560)	7507 (92, 49299)	58	9787 (9435)	6295 (53, 44466)
ACTH, pg/mL[b]	181	264 (317)	160 (2.7, 2000)	122	261 (286)	177 (2.7, 1758)	59	269 (375)	136 (2.7, 2000)
Testosterone (female participants), ng/dL[c]	85	86 (85)	57.0 (10, 548)	56	87 (88)	60.0 (10, 548)	29	84 (81)	56.0 (12, 327)
Testosterone (male participants), ng/dL[g]	88	444 (193)	416 (2, 1008)	60	450 (204)	413 (86, 1008)	28	433 (171)	423 (2, 838)
A4:T ratio (male participants)[d]	88	2.2 (2.4)	1.2 (0.1, 11.6)	60	2.1 (2.3)	1.3 (0.1, 9.6)	28	2.3 (2.7)	1.0 (0.2, 11.6)
FSH (male participants), mIU/mL[d]	90	6.3 (9.4)	4.0 (0.3, 75.2)	60	5.7 (6.2)	4.0 (0.3, 36.5)	30	7.5 (13.9)	3.9 (0.3, 75.2)
LH (male participants), mIU/mL[d]	90	4.6 (5.0)	3.3 (0.3, 39.3)	60	4.1 (3.2)	3.1 (0.3, 15.1)	30	5.5 (7.3)	3.8 (0.3, 39.3)

[a] Based on blood samples collected before the morning GC dose. Normal ranges and conversion factors for conventional units to standard international units are in Appendix Table S2.

[b] Number of participants with missing hormone or hormone precursor assessments at baseline: androstenedione (1 crinecerfont, 1 placebo); 17OHP (1 crinecerfont, 2 placebo); ACTH (1 placebo).

[c] Number of female participants with missing testosterone assessment at baseline (5 crinecerfont).

[d] Number of male participants with hormone assessments at baseline: testosterone (1 crinecerfont, 3 placebo); A4:T ratio (1 crinecerfont, 3 placebo); FSH (1 crinecerfont, 1 placebo); LH (1 crinecerfont, 1 placebo).

Abbreviations: 17OHP, 17-hydroxyprogesterone; A4:T, androstenedione-to-testosterone; ACTH, adrenocorticotropic hormone; FSH, follicle-stimulating hormone; LH, luteinizing hormone; SD, standard deviation.

Table S5. Medical Conditions of Interest (Randomized Population)

The number and percentage of participants with self-reported medical conditions of interest (indicated by a “yes” response at study entry) are presented below.

Condition – no. (%)[a]	All Participants (N=182)	Crinecerfont (N=122)	Placebo (N=60)
Irregular menstrual cycles (female participants)[b]	35 (53.0)	22 (48.9)	13 (61.9)
TARTs (male participants)[c]	44 (47.8)	30 (49.2)	14 (45.2)
Hirsutism (female participants)[c]	42 (46.7)	25 (41.0)	17 (58.6)
Acne (female participants)[c]	21 (23.3)	14 (23.0)	7 (24.1)
Anxiety	30 (16.5)	19 (15.6)	11 (18.3)
Osteopenia	28 (15.4)	22 (18.0)	6 (10.0)
Depression	26 (14.3)	17 (13.9)	9 (15.0)
Hypertension	18 (9.9)	16 (13.1)	2 (3.3)
Hyperlipidemia	18 (9.9)	12 (9.8)	6 (10.0)
Osteoporosis	9 (4.9)	6 (4.9)	3 (5.0)
Impaired glucose tolerance	5 (2.7)	3 (2.5)	2 (3.3)
Hypertriglyceridemia	5 (2.7)	4 (3.3)	1 (1.7)
Diabetes mellitus	3 (1.6)	1 (0.8)	2 (3.3)

[a] Self-reported by ≥3% of participants in either treatment group.

[b] In 66 female participants of child-bearing potential who were not on hormonal or intrauterine contraception (45 crinecerfont, 21 placebo).

[c] Participants included for analyses were as follows: 92 males (61 crinecerfont, 31 placebo); 90 females (61 crinecerfont, 29 placebo).

Abbreviation: TART, testicular adrenal rest tumor.

Table S6. Secondary and Exploratory End Point Results

	Crinecerfont		Placebo		LSM Difference (95% CI)
	N	LSM CFB (SEM)	N	LSM CFB (SEM)	
Secondary end points (at week 24)[a]					
Serum 17OHP (at week 4), ng/dL[b]	117	-5594 (533)	55	-156 (725)	-5838 (-7441, -4235)
Systolic blood pressure, mmHg	118	-3.1 (1.0)	57	-2.0 (1.4)	-1.1 (-4.2, 1.9)
Diastolic blood pressure, mmHg	118	-2.2 (0.8)	57	-0.9 (1.1)	-1.3 (-3.8, 1.2)
Waist circumference, cm	115	-0.2 (0.6)	57	0.4 (0.8)	-0.6 (-2.4, 1.2)
Plasma glucose 2 hours post 75-g glucose load, mg/dL	113	-0.9 (2.4)	54	-0.6 (3.2)	-0.3 (-7.2, 6.7)
Menstrual cycle regularity, n (%)[c]					
In female participants who were regular at baseline	12	9 (75.0)	6	2 (33.3)	Not applicable
In female participants who were irregular at baseline	9	7 (77.8)	4	3 (75.0)	Not applicable
TART volume in male participants, n (%)[d]	29	0.1 (2.4)	15	-3.1 (3.2)	3.2 (-4.2, 10.6)
Exploratory end points (at week 24)[a]					
Serum bone-specific alkaline phosphatase, µg/L	116	1.6 (0.3)	58	1.0 (0.4)	0.6 (-0.3, 1.5)
Serum CTx I, ng/L	116	110.3 (18.9)	58	22.2 (25.0)	88.1 (33.3, 142.9)
Urine NTx, nmol BCE/mmol creatinine	117	12.3 (2.6)	57	7.1 (3.5)	5.2 (-2.5, 12.9)
Serum osteocalcin, µg/L	116	8.2 (0.9)	58	3.3 (1.2)	4.9 (2.3, 7.4)

[a] End points are presented as least-squares mean (LSM) change from baseline (CFB) at week 24 with standard of the mean (SEM), unless indicated otherwise. Where applicable, the LSM difference between treatment groups (crinecerfont – placebo) is noted, along with the 95% confidence interval (CI). P-values for secondary and exploratory end points were considered nominal and are not shown. 95% CIs these end points were not adjusted for multiplicity, and the intervals may not be used in place of hypothesis testing.

[b] Based on pre-morning glucocorticoid dose samples. Normal ranges and conversion factors for conventional units to standard international units are in Table S2.

[c] Analyzed in female participants of childbearing potential who were not on hormonal or intrauterine contraception. Regularity at baseline was based on participant self-report. For the purposes of this analysis, a menstrual cycle was defined as 2 consecutive calendar days with any amount of flow. Regularity was defined as a menstrual cycle every 21-35 days.

[d] TART volume is expressed as the percentage of total testicular volume.

Abbreviation: 17OHP, 17-hydroxyprogesterone; CFB, change from baseline; CI, confidence interval; CTx I, collagen c-telopeptide Type I; LSM, least-squares mean; NTx, N-terminal telopeptide; TART, testicular adrenal rest tumor.

Table S7. Observed Mean Values for Key End Points

	At Week 4 (End of GC Stable Period)				At Week 24 (End of Double-Blind Treatment)			
	Crinecerfont		Placebo		Crinecerfont		Placebo	
	n	Mean (SEM)	n	Mean (SEM)	n	Mean (SEM)	n	Mean (SEM)
Androstenedione, ng/dL[a]	115	316 (62.5)	55	624 (92.7)	116	607 (73.8)	57	974 (108)
17-hydroxyprogesterone, ng/dL[a]	115	3012 (452)	54	9228 (1172)	116	8744 (645)	56	15368 (1295)
Glucocorticoid dose, mg/m^2/day[b]	117	17.4 (0.4)	56	18.5 (0.8)	118	10.7 (0.4)	57	13.7 (0.6)

[a] Based on pre-morning glucocorticoid dose samples. Normal ranges and conversion factors for conventional units to standard international units are in Table S2.

[b] Glucocorticoid doses are presented in hydrocortisone equivalents (conversion factors: 4x for methylprednisolone, prednisolone, prednisone; 60x for dexamethasone), adjusted for body surface area. Unlike the approach taken for the primary endpoint, these observed values for glucocorticoid doses were not set to 0 if the participant did not achieve androstenedione control at week 24 (i.e., had androstenedione less than or equal to 120% of baseline or the upper limit of normal for age and sex).

Abbreviations: GC, glucocorticoid; SEM, standard error of the mean.

Table S8. Representativeness of the Randomized Study Population

The information in this table is supported by the following:

- Endocrine Society treatment guidelines for CAH, which compile data from newborn screenings and national case registries [Speiser 2018]
- A review of mortality rates due to CAH and adrenal insufficiency [Lousada 2021]
- A worldwide estimate of CAH incidence based on newborn screening [Pang and Clark 1993]
- A systematic review and meta-analysis of ethnic and national differences [Navarro-Zambrana and Sheets 2023]
- A cross-sectional study of clinical characteristics of adult and pediatric patients from the National Institutes of Health [Finkelstein 2012]

Category	Details
Disease, problem, or condition under investigation	Classic congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency (21OHD)
Special considerations related to:	
Sex and gender	CAH affects an equal proportion of men and women. Fetuses with XX chromosomes may be born with varying degrees of genital ambiguity. In areas where newborn screening for CAH is not available, female patients with CAH may outnumber male patients due to higher mortality rates in male patients from delayed or incorrect diagnosis, as there is no genital ambiguity in male patients.[a]
Age	CAH is present at (and prior to) birth. Before the availability of glucocorticoid therapy 70 years ago, 80% of patients with CAH died of adrenal crisis within the first year of life.[b] Consequently, there are few elderly patients with CAH.
Race, ethnicity, and geography	CAH occurs in approximately 1:15,000 births worldwide.[a,c] Based on neonatal screening and national case registries, the worldwide incidence in most studies ranges from 1:14,000 to 1:18,000 births, but the incidences vary depending on whether or not newborn screening is available and on the genetic diversity with the population studied.[a,c,d] CAH is more prevalent in small, genetically isolated groups with a smaller gene pool, particularly in remote geographic regions.[a,c,d]
Other considerations: glucocorticoid treatment dose and type	Glucocorticoid treatment in an adult cohort of patients with classic CAH on hydrocortisone in an NIH registry was 16.4 ± 4.6 mg/m^2/day for those with salt wasting CAH and 17.7 ± 6.9 mg/m^2/day for those with simple virilizing CAH.[d] The proportions of patients on hydrocortisone, prednisone, and dexamethasone were generally similar.
Overall representativeness of this trial	The trial enrolled equal proportions of men and women, reflecting the expected distribution. Biologic sex was reported by the investigative site. Options were 'male' and 'female.' Gender was not reported. Ethnicity was reported as 'Hispanic or Latino' or 'Not Hispanic or Latino.' Race was reported as 'American Indian or Alaska Native,' 'Asian,' 'Black or African American,' 'White,' 'Native Hawaiian or Other Pacific Islander,' or 'Other.' The only Black or African American participants enrolled were in the US, comprising 3% of enrolled participants in the US (and 1% overall). Asian people comprised 6% of the enrolled participants in the US (and 3% overall), similar to the total population distribution of Asian people in the US. The investigative sites were located in the US, Canada, Europe, and Israel; thus, no participants were enrolled in Central or South America, Africa, or Asia. The mean age was approximately 30 years and the oldest enrolled participant was 58 years of age, reflecting the lack of elderly patients in the general population. The mean glucocorticoid dose of 17.6 ± 4.9 mg/m^2/day was similar to that reported in the NIH registry,[e] potentially due in part to the exclusion of patients with glucocorticoid overtreatment (based on hormone levels) and those with a glucocorticoid dose ≤13 mg/m^2/day from the study. The percentage of participants on hydrocortisone alone (58%) was higher and the percentage on dexamethasone (13%) was lower than reported in the NIH registry, possibly reflecting the exclusion of patients with glucocorticoid overtreatment. Coexisting conditions were consistent with epidemiologic and registry data from the US and Europe.

[a] Speiser et al, 2018. [b] Lousada et al, 2021. [c] Pang and Clark, 1993. [d] Navarro-Zambrana and Sheets, 2023. [e] Finkielstain et al, 2012. See References (Section 7.0) for full citations.

7.0 REFERENCES

Finkielstain GP, Kim MS, Sinaii N, et al. Clinical characteristics of a cohort of 244 patients with congenital adrenal hyperplasia. *J Clin Endocrinol Metab*. 2012;97(12):4429-38.

Holm S, A simple sequentially rejective multiple test procedure. *Scand J Stat.* 1979;6(2):65-70.

Lousada LM, Mendonca BB, Bachega TASS. Adrenal crisis and mortality rate in adrenal insufficiency and congenital adrenal hyperplasia. *Arch Endocrinol Metab*. 2021;65(4):488-94.

Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. *Statist Med*. 2011;30(28):3267-84.

Navarro-Zambrana AN and Sheets SR. Ethnic and national differences in congenital adrenal hyperplasia incidence: a systematic review and meta-analysis. *Horm Res Paediatr*. 2023;96(3):249-58.

O'Kelly M and Ratitch B, *Clinical Trials with Missing Data: A Guide for Practitioners*. West Sussex, United Kingdom: John Wiley & Sons; 2014.

Overall JE and Gorham DR, The Brief Psychiatric Rating Scale (BPRS): recent developments in ascertainment and scaling. *Psychopharmacol Bull.* 1988;24(1):97-9.

Pang S and Clark A. Newborn screening, prenatal diagnosis, and prenatal treatment of congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Trends Endocrinol Metab*. 1990;1(6):300-7.

Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 2011;168(12):1266-77

Ratitch B, Lipkovich I, O'Kelly M. Combining results from multiply imputed categorical data. *PharmaSUG* 2013; Paper No. SP03:8-11.

Speiser PW, Arlt W, Auchus RJ, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab 2018*;103(11): 4043-88.

Wiens BL. A fixed sequence Bonferroni procedure for testing multiple end points. *Pharmaceut Statist.* 2003;2(3):211-5.