STATISTICAL ANALYSIS PLAN

Protocol No.:	ACP-101-302
Protocol Title:	A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group Study of Carbetocin Nasal Spray for the Treatment of Hyperphagia in Prader-Willi Syndrome
Drug:	Carbetocin nasal spray
Sponsor:	Acadia Pharmaceuticals Inc.
Version No. and Date	Version 3, 03 September 2025

SIGNATURE/APPROVAL PAGE

AUTHORS

See appended electronic signature page

PPD

Date

Associate Director, Biostatistics
Acadia Pharmaceuticals Inc.

APPROVERS

See appended electronic signature page

PPD, MD

Date

PPD, Clinical Development
Acadia Pharmaceuticals Inc.

See appended electronic signature page

PPD, PhD

Date

Executive Director, Biostatistics
Acadia Pharmaceuticals Inc.

TABLE OF CONTENTS

TABLE OF CONTENTS 3

LIST OF TABLES 6

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS 7

1. INTRODUCTION 9

2. OBJECTIVES AND ENDPOINTS 10
2.1 Primary Objective 10
2.1.1 Primary Endpoint 10
2.2 Secondary Objectives 10
2.2.1 Secondary Endpoints 10
2.3 Exploratory Objectives 10
2.3.1 Exploratory Endpoints 11
2.4 Safety Objective 11
2.4.1 Safety Endpoints 11
2.5 Pharmacokinetic Objectives 11
2.5.1 Pharmacokinetic Endpoints 11

3. STUDY DESIGN 12
3.1 General Study Design 12
3.2 Schedule of Assessments 13
3.3 Randomization 13
3.4 Blinding 13
3.5 Determination of Sample Size 13

4. ANALYSIS SETS 15

5. DATA HANDLING CONVENTIONS 16
5.1 General Data Reporting Conventions 16
5.2 Derived Efficacy Variables 16
5.2.1 Hyperphagia Questionnaire for Clinical Trials (HQ-CT) 17
5.2.2 Clinical Global Impression–Severity (CGI-S) for PWS, Clinical Global Impression–Change (CGI-C) for PWS, and Clinical Global Impression–Severity (CGI-S) for Hyperphagia in PWS 17
5.2.3 PWS Anxiousness and Distress Behaviors Questionnaire (PADQ) 17
5.2.4 Food Safe Zone 18
5.2.5 Zarit Burden Interview (ZBI) 18

5.2.6 PWS-specific Healthcare Resource Use and Caregiver Burden Questionnaire (PWS-HRUQ) 19
5.3 Study Day 19
5.4 Baseline Definition 19
5.5 Analysis Visit Windows 19
5.5.1 Unscheduled Assessments 20
5.5.2 Multiple Measurements Within Visit Windows 21
5.6 Missing or Incomplete Date for Last Dose of Study Drug 21
5.7 Missing or Incomplete Dates for Prior or Concomitant Medications 21
5.8 Missing or Incomplete Date for Adverse Events 21
5.9 Missing Severity Assessment for Adverse Events 22
5.10 Missing Relationship to Study Drug for Adverse Events 22
5.11 Character Values of Clinical Laboratory Variables 22

6. SUBJECT DISPOSITION 23

7. PROTOCOL DEVIATIONS 24

8. DEMOGRAPHICS AND BASELINE CHARACTERISTICS 25

9. MEDICAL HISTORY 26

10. EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE 27
10.1 Exposure to Study drug 27
10.2 Measurement of Treatment Compliance 27

11. PRIOR, CONCOMITANT, AND POST-TREATMENT MEDICATION 28

12. EFFICACY ANALYSES 29
12.1 Efficacy Variables 29
12.2 Adjustment for Covariates 29
12.3 Handling of Missing Data 30
12.4 Multiple Comparisons / Multiplicity 30
12.5 Examination of Subgroups 30

13. METHODS OF EFFICACY ANALYSES 31
13.1 Primary Efficacy Endpoints 31
13.1.1 Primary Analysis 31
13.1.2 Sensitivity Analyses 32
13.2 Secondary Efficacy Endpoints 36
13.2.1 CGI-S for PWS Score – Change From Baseline at Week 12 36
13.2.2 CGI-C for PWS Score at Week 12 36

13.2.3 Percentage of Subjects With Treatment Response (Defined as Improvement From Baseline ≥8 Points) Based on the HQ-CT at Week 12 36
13.2.4 CGI-S for hyperphagia in PWS Score – Change From Baseline at Week 12 36
13.3 Exploratory Efficacy Analyses 37
14. SAFETY ANALYSES 39
14.1 Adverse Events 39
14.2 Clinical Laboratory Variables 40
14.3 Vital Signs, Height, Weight, and BMI 44
14.4 Electrocardiogram 45
14.5 Physical Examination 47
14.6 Nasal Assessment 47
14.7 Other Safety Endpoints 47
15. CLINICAL PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES 48
16. INTERIM ANALYSIS 49
17. DATA AND SAFETY MONITORING BOARD 50
18. COMPUTER METHODS 51
19. CHANGES TO ANALYSES SPECIFIED IN PROTOCOL 52
20. REFERENCES 53
21. APPENDICES 54
Appendix A: Schedule of Assessments 54
Appendix B: Summary of Version Changes 57

LIST OF TABLES

Table 1 Visit Windows for Analyses of HQ-CT, PADQ, Food Safe Zone, and ZBI 20
Table 2 ECG and PK Analysis Visit Windows 20
Table 3 Other Analysis Visit Windows 20
Table 4 CCI 34
Table 5 PCI Criteria: Hematology and Chemistry 43
Table 6 PCI Criteria: Urinalysis 44
Table 7 PCI Criteria: Vital Signs and Weight 45
Table 8 PCI Criteria: ECG 46

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AE	adverse event
BMI	body mass index
CI	confidence interval
CGI-C	Clinical Global Impression–Change
CGI-S	Clinical Global Impression–Severity
DSMB	data and safety monitoring board
E-R	exposure-response
eCRF	electronic case report form
ECG	electrocardiogram
EOT	end of treatment
FAS	Full Analysis Set
HQ-CT	Hyperphagia Questionnaire for Clinical Trials
ICE	intercurrent event
LSM	least squares mean
MedDRA	Medical Dictionary for Regulatory Activities
MAR	missing at random
MI	multiple imputation
MMRM	mixed-effects model for repeated measures
MNAR	missing not at random
PADQ	PWS Anxiousness and Distress Behaviors Questionnaire
PCI	potentially clinically important
PK	pharmacokinetic(s)
PWS	Prader-Willi syndrome
PWS-HRUQ	PWS-specific Healthcare Resource Use and Caregiver Burden Questionnaire
Q1	quartile 1 (25^{th} percentile)
Q3	quartile 3 (75^{th} percentile)
SAE	serious adverse event
SI	Système International

SOC	system organ class
TEAE	treatment-emergent adverse event
ZBI	Zarit Burden Interview

1. INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the statistical analyses of efficacy and safety data as described in the study protocol amendment 2 GLOBAL, dated 24 February 2025.

Specifications for tables, figures, and listings are contained in a separate document. Statistical analyses for population pharmacokinetic (PK) and PK modeling will be presented in a separate report and therefore will not be included in this SAP.

This plan should be read in conjunction with the study protocol and the electronic case report forms (eCRFs).

2. OBJECTIVES AND ENDPOINTS

2.1 Primary Objective

The primary objective of this study is to investigate the efficacy of treatment with carbetocin nasal spray versus placebo on hyperphagia in Prader-Willi syndrome (PWS).

2.1.1 Primary Endpoint

The primary endpoint is:

- Hyperphagia Questionnaire for Clinical Trials (HQ-CT) score – change from Baseline at Week 12

2.2 Secondary Objectives

The secondary objectives are to investigate the efficacy of treatment with carbetocin versus placebo on:

- PWS
- Hyperphagia in PWS

2.2.1 Secondary Endpoints

The secondary endpoints for this study are:

- Clinical Global Impression–Severity (CGI-S) for PWS score – change from Baseline at Week 12
- Clinical Global Impression–Change (CGI-C) for PWS score at Week 12
- Percentage of subjects with treatment response (defined as improvement from Baseline ≥8 points) based on the HQ-CT at Week 12
- CGI-S for hyperphagia in PWS score – change from Baseline at Week 12

2.3 Exploratory Objectives

The exploratory objectives for this study are to investigate the benefit of treatment with carbetocin versus placebo on:

- Anxiousness in PWS
- Environmental controls
- Caregiver burden
- Healthcare utilization

2.3.1 Exploratory Endpoints

The exploratory endpoints for this study are:

- PWS Anxiousness and Distress Behaviors Questionnaire (PADQ) score – change from Baseline at Week 12
- Food Safe Zone score – change from Baseline at Week 12
- Zarit Burden Interview (ZBI) score – change from Baseline at Week 12
- PWS-specific Healthcare Resource Use and Caregiver Burden Questionnaire (PWS-HRUQ) – change from Baseline at Week 12

2.4 Safety Objective

The safety objective of this study is to investigate the safety and tolerability of treatment with carbetocin versus placebo.

2.4.1 Safety Endpoints

The safety endpoints for this study are:

- Treatment-emergent adverse events (TEAEs)
- Serious adverse events (SAEs)
- Withdrawals due to adverse events (AEs)
- Potentially clinically important (PCI) changes in other safety assessments
- Device incidents, or device malfunctions

2.5 Pharmacokinetic Objectives

The PK objectives of this study are:

- To characterize the PK of carbetocin
- To assess the exposure-response relationship

2.5.1 Pharmacokinetic Endpoints

The PK endpoints for this study are:

- Plasma concentration of carbetocin
- Carbetocin PK parameters using the population PK approach
- Exposure-response (E-R) analysis methods

3. STUDY DESIGN

3.1 General Study Design

This study is a Phase 3, 12-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group study in subjects with PWS.

The study drug is carbetocin nasal spray or the matching placebo. Subjects will be dosed with carbetocin 3.2 mg or placebo, administered TID at mealtimes, if possible.

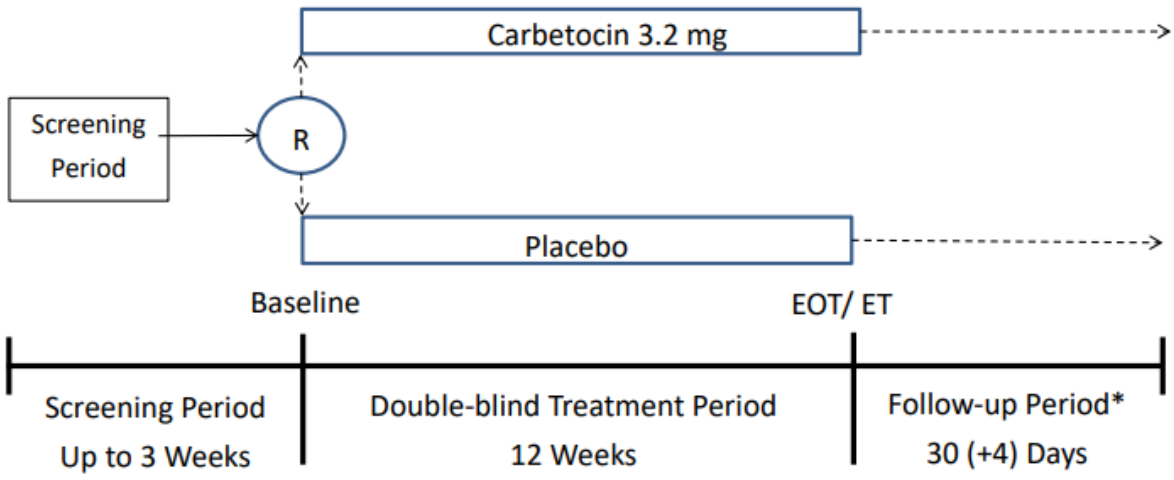
Subjects will be stratified according to age stratum (<12 years old, ≥12 to <17 years old, ≥17 years old) and by presence or absence of participation in a previous clinical study of carbetocin nasal spray. The Sponsor, subjects, caregivers, Investigators, and any other site or Sponsor personnel will be blinded to treatment assignment.

The duration of participation for individual study subjects will be approximately 19 weeks, consisting of a screening period of up to 3 weeks, a 12-week double-blind treatment period, and a safety follow-up period of 30 (+4) days. The study completion date is defined as the last date that subject data are collected, which includes the safety follow-up visit.

Approximately 42 global sites will participate in this study. Approximately 170 subjects will be randomized (85 subjects per treatment group).

The study design is summarized in Figure 1.

Figure 1 Schematic of Study Design



EOT=end of treatment; ET=early termination; R=Randomization

*If subjects continue into the open label extension (OLE) from the current study, they will not complete the follow-up visit and will roll over into the OLE.

3.2 Schedule of Assessments

The schedule of events and assessments is in Appendix A.

3.3 Randomization

Eligible subjects will be randomized into one of two treatment groups (carbetocin or placebo) in a 1:1 ratio using an interactive response technology (IRT) system. Randomization will be stratified by age group (<12 years old, ≥12 to <17 years old, ≥17 years old) and by presence or absence of participation in a previous clinical study of carbetocin nasal spray. The assignments will be based on a computer-generated randomization schedule.

3.4 Blinding

Treatment assignments will be blinded to all study subjects, caregivers, Investigators, site personnel, and Sponsor personnel. Blinding will be assured by restricting access of Investigators and Sponsor personnel and/or designee to the treatment codes and providing identical packaging for the carbetocin and placebo treatments. Data will be unblinded after database lock, unless specified otherwise.

In the event of a medical emergency where breaking the blind is required so that appropriate medical care can be provided, treatment assignment for the affected subject may be unblinded to the Investigator through the interactive response technology system. The date and rationale for unblinding must be clearly explained in source documentation. The Investigator should contact the Sponsor Medical Monitor promptly regarding any emergency unblinding. Once a subject's treatment assignment is disclosed to the Investigator, the subject should have study drug discontinued.

For the final analysis, the treatment codes for all subjects will be released to Acadia after all subjects have completed the study and the clinical database is locked.

3.5 Determination of Sample Size

The sample size calculation was performed for the primary efficacy endpoint as a single hypothesis test at an overall two-sided significance level of 0.05.

Assuming the common SD of change from Baseline at Week 12 in HQ-CT is 7 points, 80 subjects who complete the study per treatment group will provide 88% power to detect a difference in the mean change from Baseline at Week 12 in HQ-CT of 3.5 points between the carbetocin group and the placebo group at a significance level of 0.05 (two-sided t-test).

The difference of 3.5 points was selected based on findings from the previous Phase 3 study in IN carbetocin for hyperphagia in PWS (Study LV-101-3-01 [CARE-PWS]), where the MID, defined as one-half of the standard deviation of baseline scores on the outcome measures, was

2.6 points in HQ-CT. In this study, the observed between-group HQ-CT difference of -3.1 points for the 3.2 mg/dose group exceeded this threshold of clinical meaningfulness. For the current study, a difference of 3.5 points is selected to use as the basis on which to calculate sample size.

Adjusting for a potential discontinuation rate of up to 5%, approximately 170 subjects need to be randomized (85 subjects per treatment group).

4. ANALYSIS SETS

Randomized Analysis Set

The Randomized Analysis Set consists of all subjects who were randomized. Subjects will be classified according to the randomized treatment assignment.

Safety Analysis Set

The Safety Analysis Set consists of all randomized subjects who received at least one dose of study drug. The Safety Analysis Set will be used for the analyses of all safety endpoints based on the actual treatment received

Full Analysis Set

The Full Analysis Set (FAS) includes all randomized subjects who received at least one dose of study drug and who have a baseline value for the HQ-CT score. The FAS will be used for the analyses of all efficacy endpoints based on the randomized treatment assignment.

Pharmacokinetic (PK) Analysis Set

The PK Analysis Set includes all subjects in the Safety Analysis Set with at least one measurable carbetocin plasma concentration. Subjects will be classified according to the actual treatment received.

5. DATA HANDLING CONVENTIONS

.

5.1 General Data Reporting Conventions

Continuous variables will be summarized using the following descriptive statistics: number of subjects, mean, median, SD, Q1, Q3, minimum, and maximum. Unless specified otherwise, means, medians, and CIs will be presented to one more decimal place than the raw data, and the SDs and SEs will be presented to two more decimal places than the raw data. In general, the maximum number of decimal places is four and values will be truncated to four decimal places in situations where there are more than four decimal places. Wherever possible data will be decimal aligned.

Height (cm), weight (kg) and body mass index (BMI) (kg/m^2) will be presented with a maximum of one decimal place.

Categorical variables will be summarized by the number of subjects and the percent of subjects in each category; the number of subjects and the percentage of subjects with missing data will be summarized for demographic and baseline characteristics (if applicable). Categories with zero counts will be summarized but will not have the zero percentage displayed. For demographic summaries, percentages will be calculated by using the total number of subjects in the given treatment group as the denominator. Percentages will be presented with one decimal place.

Duration in months will be calculated as ([the number of days/365.25]*12).

Unless specified otherwise, all statistical tests will be two-sided hypothesis tests performed at the significance level of 5% for main effects and all CIs will be two-sided 95% CIs. P-values will generally be presented to four decimal places; values less than 0.0001 will be presented as <0.0001.

If methods based on the assumption that the data (of primary endpoint) are normally distributed are not adequate, analyses may be performed on transformed data or nonparametric analysis methods may be used as appropriate.

Clinical laboratory assessment values that are collected with “<” or “>” signs will generally be analyzed as the numerical value without the sign in tables and figures. In listings, these data will be reported as collected with the sign.

5.2 Derived Efficacy Variables

In general, assessment total scores and subscores will be derived within the analysis datasets. In the event that total scores and/or subscores are also collected on the eCRF, the derived values will be used for all analyses. Both the raw and derived scores will be presented in listings.

5.2.1 Hyperphagia Questionnaire for Clinical Trials (HQ-CT)

The HQ-CT is a nine-item questionnaire designed to be completed by caregivers of subjects with PWS. It is a revision of the 11-item Hyperphagia in Prader-Willi Syndrome Questionnaire–Responsiveness (HPWSQ-R) and has been further validated. The questionnaire uses a 5-point scale of 0 to 4 with the minimum score being 0 and the maximum score being 36. The Foundation for Prader-Willi Research has made the HQ-CT available for clinical studies in PWS, and it is the consensus instrument within the PWS research community for measuring observable behaviors that stem from subjects' excessive drive to eat. The HQ-CT should be completed by the same caregiver throughout the study.

The HQ-CT will be administered to the caregiver by a rater using standardized prompts. The Food Safe Zone should be administered immediately before administration of the HQ-CT.

A higher score on the HQ-CT indicates greater severity of hyperphagia. A 7.7-point improvement (decrease) on the HQ-CT has been identified as a threshold for a within-patient responder definition that can be used to assess for a clinically meaningful response (RTI Health Solutions 2016).

5.2.2 Clinical Global Impression–Severity (CGI-S) for PWS, Clinical Global Impression–Change (CGI-C) for PWS, and Clinical Global Impression–Severity (CGI-S) for Hyperphagia in PWS

The CGI-S is a rating scale that records a clinician's global impression of the current severity of illness on a seven-point scale, using a range of responses from 1 (normal) to 7 (among the most severely ill subjects) (Guy 1976). The CGI-C is a similar rating scale that records a clinician's global impression of change in severity of illness, using a range of responses from 1 (very much improved) to 7 (very much worse).

The scales are to be administered by the Investigator or a designated rater (trained and qualified). For a given subject, the same person should perform this task throughout the study.

5.2.3 PWS Anxiousness and Distress Behaviors Questionnaire (PADQ)

The PADQ is a caregiver-reported instrument designed to capture the observable signs of anxiousness and distress that are common among subjects with PWS (Cotter et al. 2023). It was developed in collaboration with RTI-Health Solutions (RTI-HS), the Foundation for Prader-Willi Research, and a working group of three experienced researchers who are also parents of individuals with PWS. The PADQ score is computed based on responses to the first 14 items of the measure. Specifically, item responses are summed to yield scores ranging from 0 to 56, with higher scores indicating greater levels of anxiousness and psychological distress. Item 15 is included in the PADQ as an overall summary of caregivers' observations, but this item is not

included in the score. The PADQ has also been known as the PWS Anxiety and Distress Behaviors Questionnaire and the PWS Anxiety and Distress Questionnaire.

The PADQ should be completed by the same caregiver throughout the study.

5.2.4 Food Safe Zone

This caregiver rated questionnaire measures food-related restrictions or controls (e.g., where access to food has been limited by locking or other measures).There are 20 items, with each item rated as 1 (none of the time), 2 (some of the time), 3 (most/all of the time). Higher scores indicate a greater degree of caregiver control/oversight.

For a given subject, the same caregiver should perform this task throughout the study. The Food Safe Zone should be administered immediately before administration of the HQ-CT.

5.2.5 Zarit Burden Interview (ZBI)

The ZBI was designed to assess the stress experienced by caregivers of patients with dementia (Zarit et al. 1980). The ZBI will be administered as an interview. The interview consists of 22 statements reflecting how people sometimes feel when taking care of another person. The statements are phrased as questions for the family member caregiver to indicate how often they feel the way described in the statement. Responses are Never, Rarely, Sometimes, Quite Frequently, and Nearly Always. When the caregiver is not a family member, this scale will not be completed.

The ZBI should be completed by the same caregiver throughout the study.

A higher score on the ZBI indicates greater severity of stress experienced by the caregiver.

5.2.6 PWS-specific Healthcare Resource Use and Caregiver Burden Questionnaire (PWS-HRUQ)

A PWS specific caregiver reported questionnaire, PWS-HRUQ, was developed to assess the amount and type of healthcare resources utilized by the subject such as outpatient hospitalization, emergency rooms visits, physician/general practitioner visits, and allied health professional visits (e.g., nutritionist visits). Additionally, the number of inpatient hospitalizations, if any, and associated length of stay will be assessed.

Furthermore, the impact of PWS on subjects' classroom productivity will be assessed. The PWS-HRUQ also includes questions related to the caregiver's employment status and their work productivity related to absenteeism (i.e., amount of time absent from work) and presenteeism (i.e., amount of unproductive time at work).

All items in the PWS-HRUQ have a 2-week recall period and will be assessed at Week 0 (Baseline) and Week 12 (study endpoint). Analysis of this questionnaire will be outlined in a separate health economics analysis plan.

5.3 Study Day

If the date of assessment occurs on or after the first dose date, then study day will be calculated as (date of assessment – date of first dose) + 1. If the date of assessment occurs prior to the first dose date, then study day will be calculated as (date of assessment – date of first dose). There is no study Day 0.

Day 1 refers to the first day of dosing.

5.4 Baseline Definition

For efficacy variables, Baseline value is defined as the most recent nonmissing value collected prior to or on the day of administration of the first dose of study drug.

For other variables, Baseline value is defined as the most recent nonmissing value collected prior to the administration of the first dose of study drug.

5.5 Analysis Visit Windows

The assessments of HQ-CT, PADQ, Food Safe Zone, and ZBI will be summarized by analysis visit as presented in Table 1.

Table 1 Visit Windows for Analyses of HQ-CT, PADQ, Food Safe Zone, and ZBI

Analysis Visit	Target study Day	Study Day interval
Baseline	1	≤1
Week 2	15	2 – 22
Week 4	29	23 – 43
Week 8	57	44 – 71
Week 12	85	72 – 115

Abbreviations: HQ-CT=Hyperphagia Questionnaire for Clinical Trials; PADQ=PWS Anxiousness and Distress Behaviors Questionnaire; ZBI=Zarit Burden Interview

Electrocardiogram (ECG) and PK assessments will be summarized by analysis visit as presented in Table 2.

Table 2 ECG and PK Analysis Visit Windows

Analysis visit	Target study Day	Study Day interval
Baseline[a]	1	≤1
Day 1 post first dose[b]	1	1
Week 2	15	2 – 36
Week 8	57	37 – 71
Week 12	85	72 – 115

Abbreviations: ECG=electrocardiogram; PK=pharmacokinetic
[a] Assessments collected before first dose time on Day 1 are mapped to Baseline.
[b] Assessments collected after first dose on Day 1 are mapped to Day 1 post first dose.

Other efficacy and safety assessments will be summarized by analysis visit as presented in Table 3.

Table 3 Other Analysis Visit Windows

Analysis visit	Target study Day	Study Day interval
Baseline[a,b]	1	≤1
Week 2[c]	15	1 – 36
Week 8	57	37 – 71
Week 12	85	72 – 115

[a] Assessments (except CGI-S) collected before first dose time on Day 1 are mapped to Baseline.
[b] CGI-S assessments collected on Day 1 are mapped to Baseline.
[c] Assessments collected after first dose on Day 1 are mapped to Week 2.

5.5.1 Unscheduled Assessments

Both scheduled and unscheduled assessments, including the assessments at early termination (ET) visits, will be included for planned timepoint analyses based on the above analysis visit windowing rules. All assessments will be presented in data listings.

5.5.2 Multiple Measurements Within Visit Windows

In the event that more than one assessment falls within a given window, the assessment closest to the target study day will be selected for the by-visit analysis. If two assessments are equidistant from the target study day, then the chronologically last assessment will be used.

For safety analyses where the extreme values should be selected (e.g., overall post-Baseline minimum, overall post-Baseline maximum, and PCI values), all nonmissing post-Baseline values should be considered, regardless of whether the value is selected for the by-visit summaries. All assessments will be presented in data listings.

5.6 Missing or Incomplete Date for Last Dose of Study Drug

For subjects with completely missing last dose date, the last dose date will be imputed by the last expected dosing date, defined as the earliest of the following dates: last drug kit dispense date + scheduled dosing interval per protocol, end of treatment (EOT)/ET date and the return date of the last dispensed drug kit.

For subjects with partial missing last dose date, the imputation will be compared against the last expected dosing date as defined above. Detailed algorithms will be documented in a separate programming specifications document. The missing or incomplete dates will be displayed in the data listings as reported on the eCRF rather than the imputed dates.

5.7 Missing or Incomplete Dates for Prior or Concomitant Medications

Missing or incomplete medication start or stop dates will be imputed for the purpose of determining whether the medication is taken concomitantly or not (see Section 11 for definition). When the chronological order of medication use relative to the study drug treatment period is unclear due to missing or incomplete date(s), the medication will be considered as concomitant. The imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates as captured on the eCRF will be displayed in the data listings.

5.8 Missing or Incomplete Date for Adverse Events

Missing or incomplete AE start dates will be imputed for the purpose of determining whether the AEs are treatment-emergent or not (see Section 14.1 for definition). When the chronological order of an AE onset relative to the study drug treatment period is unclear due to missing or incomplete date(s), the AE will be considered as treatment-emergent. The imputation algorithms will be detailed in the analysis dataset specification document. The missing or incomplete dates captured on the eCRF will be displayed in the data listings.

5.9 Missing Severity Assessment for Adverse Events

If the severity is missing for an AE starting on or after the date of the first dose of study drug, then a severity of "severe" will be assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be used in data listings.

5.10 Missing Relationship to Study Drug for Adverse Events

If the relationship to study drug is missing for an AE starting on or after the date of the first dose of study drug, a causality of "related" will be assigned. The imputed values for relationship to study drug will be used for incidence summaries, while the actual values will be presented in data listings.

5.11 Character Values of Clinical Laboratory Variables

If the reported value of a clinical laboratory variable cannot be used in a statistical analysis due to, for example, a character string reported for a numeric variable, an appropriately determined coded value may be used in the statistical analysis. The coding algorithms will be detailed in the analysis dataset specification document. The actual values as reported in the database will be presented in data listings.

6. SUBJECT DISPOSITION

For subjects who participate in the screening phase but are not randomized (screen failures), their demographics information (including age, sex, and primary race), screen failure reasons (the specific inclusion/exclusion criterion [or criteria] not met) and protocol version will be listed. If a subject is rescreened, then the rescreening subject ID and the final enrollment status (whether eventually enrolled) will also be displayed in this listing. In addition, the frequency of reported screen failure reasons will be summarized. Note that a subject may be deemed ineligible for multiple inclusion/exclusion criteria and may be allowed to rescreen with the permission of the Medical Monitor, provided the screen failure was due to a temporary condition that subsequently resolved.

The number of sites that screened at least 1 subject, number of sites that randomized at least 1 subject, number of subjects screened, and number of unique subjects screened will be tabulated. In addition, the number of subjects enrolled at each site will also be tabulated by analysis set and by treatment group and overall.

The number and percentage of subjects who completed the study, discontinued early, and the reason for discontinuation will be summarized by treatment group using the Randomized and Full Analysis Sets.

7. PROTOCOL DEVIATIONS

Protocol deviations will be reviewed periodically over the course of the study. The review process, definition of the deviation categories, and the classification of a deviation as major or minor are detailed in the protocol deviation management plan.

A summary of the number and percentage of subjects with major protocol deviations for each deviation category will be presented by treatment group for the Randomized Analysis Set.

A data listing of all protocol deviations will be provided.

8. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Demographics and baseline characteristics will be summarized by treatment group and overall for the Randomized Analysis Set and the FAS using descriptive statistics. Variables include, but are not limited to, age, age group (<12 years old, ≥12 to <17 years old, ≥17 years old), sex, race, ethnicity, height, weight, BMI and participation in a previous clinical study of carbetocin nasal spray (presence or absence).

Race will also be categorized by White versus Non-White. The reported age reflects a subject's age at the informed consent date.

Baseline characteristics will be summarized by treatment group. HQ-CT, CGI-S for PWS, CGI-S for hyperphagia, PADQ, Food Safe Zone, and ZBI will be summarized using descriptive statistics. Additionally, CGI-S for PWS category (1-7) and CGI-S for hyperphagia category (1-7) will be summarized by number and percentage of subjects.

9. MEDICAL HISTORY

Medical history reported terms will be coded with Medical Dictionary for Regulatory Activities (MedDRA), Version 28.0 or newer. The subject incidence will be summarized for each system organ class (SOC) and preferred term by treatment group and overall for the Safety Analysis Set. A subject will be counted only once per SOC or per preferred term for the summary.

A listing of the SOC, preferred term, verbatim for the medical history condition/event, start and stop dates (when available), and an indicator for whether or not the condition is ongoing will be provided.

PWS History

Age at PWS diagnosis and genetic subtype of PWS will be summarized.

10. EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE

Extent of exposure and treatment compliance will be summarized as both continuous variables and categorical variables by treatment group for Safety Analysis set.

10.1 Exposure to Study drug

Duration of exposure to study drug will be calculated for each subject as (last dose date – first dose date + 1). The number and percentage of subjects within each of the following exposure levels in terms of duration of exposure will also be tabulated: <2 weeks (1 to 13 days), 2 to <4 weeks (14 to 27 days), 4 to <6 weeks (28 to 41 days), 6 to <8 weeks (42 to 55 days), 8 to <10 weeks (56 to 69 days), 10 to <12 weeks (70 to 83 days), and ≥12 weeks (84 days or longer).

10.2 Measurement of Treatment Compliance

Study drug compliance will be calculated based on the drug accountability data as collected on the eCRF. The study drug compliance will be calculated as (the total number of vials actually used divided by the total number of vials expected to be used)*100.

The total number of vials actually taken will be calculated as (total vials dispensed – total vials unused and returned).

Treatment compliance will be summarized as a categorical variable. The number and percentage of subjects within each of the following compliance levels will be tabulated: <80%, 80% to 120% and >120%.

If study drug return data is incomplete, diary data may be used to estimate the number of vials used.

11. PRIOR, CONCOMITANT, AND POST-TREATMENT MEDICATION

Prior medication is defined as any medication with a stop date prior to the date of the first dose of study drug. Concomitant medication is defined as any medication that is ongoing at the first dose of study drug or with a start date between the dates of the first dose and last dose of study drug, inclusive. Post-treatment medication is defined as any medication with a start date after the date of the last dose of study drug. Medications will be coded using WHO Drug Global Dictionary March 2025 or newer version.

The number and percentage of subjects taking prior, concomitant, and post-treatment medications will be tabulated separately by each drug class (anatomical/therapeutic/chemical [ATC] Level 3) and medication preferred term, treatment group and overall for the Safety Analysis Set. Multiple medication usage by a subject in the same category will be counted only once. Listings of the prior, concomitant, and post-treatment medications will also be provided.

Note that not all medications can be coded down to level 3 (e.g., KRILL OIL). If ATC level 3 is not available, then ATC level 2 will be used instead. If ATC level 2 is also not available, then ATC level 1 will be used instead.

12. EFFICACY ANALYSES

Unless otherwise specified, all efficacy analyses will be performed using the planned treatment assignments based on the randomization schedule for the FAS.

All statistical tests will be two-sided hypothesis tests performed at the 5% level of significance. In addition, all CIs will be two-sided 95% CIs, unless stated otherwise.

12.1 Efficacy Variables

Primary Efficacy Endpoint

The primary efficacy endpoint is:

- HQ-CT score – change from Baseline at Week 12

Secondary Efficacy Endpoints

The secondary endpoints for this study are:

- CGI-S for PWS score – change from Baseline at Week 12
- CGI-C for PWS score at Week 12
- Percentage of subjects with treatment response (defined as improvement from baseline ≥8 points) based on the HQ-CT at Week 12
- CGI-S for hyperphagia in PWS score – change from Baseline at Week 12

Exploratory Efficacy Endpoints

The exploratory endpoints for this study are:

- PADQ score – change from Baseline at Week 12
- Food Safe Zone score – change from Baseline at Week 12
- ZBI score – change from Baseline at Week 12
- PWS-HRUQ score – change from Baseline at Week 12

12.2 Adjustment for Covariates

The Baseline HQ-CT score will be included as a covariate for the analysis of change from baseline in the HQ-CT score using the mixed-effects model for repeated measures (MMRM). The Baseline HQ-CT score will also be included as a covariate for the analysis of the proportion of treatment responders (defined as improvement ≥8 points from Baseline) using logistic regression.

The Baseline CGI-S for PWS score will be included as a covariate for the analysis of change from Baseline in the CGI-S for PWS score using the MMRM method. The Baseline CGI-S for

PWS score will be included as a covariate for the analysis of CGI-C for PWS score using the MMRM method. The Baseline CGI-S for hyperphagia in PWS score will be included as a covariate for the analysis of change from baseline in the CGI-S for hyperphagia in PWS score using the MMRM method.

Unless otherwise specified, sex and the randomization stratification variables of age group (<12 years old, ≥12 to <17 years old, ≥17 years old) and participation in a previous carbetocin nasal spray study (presence or absence) will also be included as fixed effects in the MMRM and logistic regression methods.

12.3 Handling of Missing Data

The primary analysis of

- HQ-CT score – change from Baseline at Week 12 (primary efficacy endpoint)

will be performed assuming missing at random (MAR) using the direct likelihood-based MMRMs and missing scores will not be imputed. The MMRM method is unbiased under the MAR assumptions in the estimation of treatment effect that could have been observed if all subjects had continued on treatment for the full study duration (EMA 2010).

12.4 Multiple Comparisons / Multiplicity

Multiplicity adjustment is not applicable in this study.

12.5 Examination of Subgroups

Treatment effect will be examined with respect to the subgroups of sex, age group (<12 years old, ≥12 to <17 years old, ≥17 years old), participation in a previous carbetocin nasal spray study (presence or absence), genetic subtype (deletion type 1, deletion type 2, deletion (type unknown), uniparental disomy (UPD), imprinting defect (ID), unknown), HQ-CT Baseline score (>22 or ≤22), and PADQ Baseline score (in upper three quartiles or in lowest quartile). If a subgroup size is too small, categories may be pooled and footnoted appropriately.

13. METHODS OF EFFICACY ANALYSES

13.1 Primary Efficacy Endpoints

13.1.1 Primary Analysis

Estimand

Target population: Male or female subjects 5 to 30 years of age, inclusive, with PWS, as defined by the inclusion/exclusion criteria of the study.

Variable (primary endpoint): Change from Baseline in HQ-CT score at Week 12.

Treatment condition: carbetocin 3.2 mg TID and placebo.

CCI

Population-level summary: Difference of means between carbetocin 3.2 mg TID and placebo in the change from Baseline in HQ-CT score at Week 12.

Hypothesis

Let Δ ($\boldsymbol{\mu}_{CARB} - \boldsymbol{\mu}_{PBO}$) be the difference in the mean change from Baseline at Week 12 in the HQ-CT score between the active treatment ($\boldsymbol{\mu}_{CARB}$) and placebo ($\boldsymbol{\mu}_{PBO}$) groups.

The null hypothesis is: $\Delta = 0$.

The alternative hypothesis is: $\Delta \neq 0$.

Primary Estimator

The hypothesis testing will be performed for the FAS using the direct likelihood MMRM method assuming missing data are MAR. The MMRM for the HQ-CT analysis will include effects for treatment group, sex, age group (<12 years old, ≥12 to <17 years old, ≥17 years old), participation in a previous clinical study of carbetocin nasal spray (presence or absence), visit, treatment-by-visit interaction as fixed effects, Baseline HQ-CT score, and the Baseline HQ-CT score-by-visit interaction as covariates and subjects as random effect. An unstructured covariance matrix will be used and the Kenward-Roger approximation will be used to adjust the denominator degrees of freedom.

The following is sample SAS code for the MMRM:

CCI

Summary statistics for the HQ-CT score (observed and change from Baseline), least squares means (LSMs), the between-group difference in LSM with the corresponding 95% CI, p-value, and the effect size (Cohen's *d*) will be reported for the FAS.

CCI

LSM ± SE from the MMRM models over time for the change from Baseline values by treatment group will be displayed in line plots for the FAS.

13.1.2 Sensitivity Analyses

The following sensitivity analyses of the primary efficacy endpoint are planned to account for ICEs. Additional sensitivity analyses may be performed as appropriate.

13.1.2.1 CCI

CCI

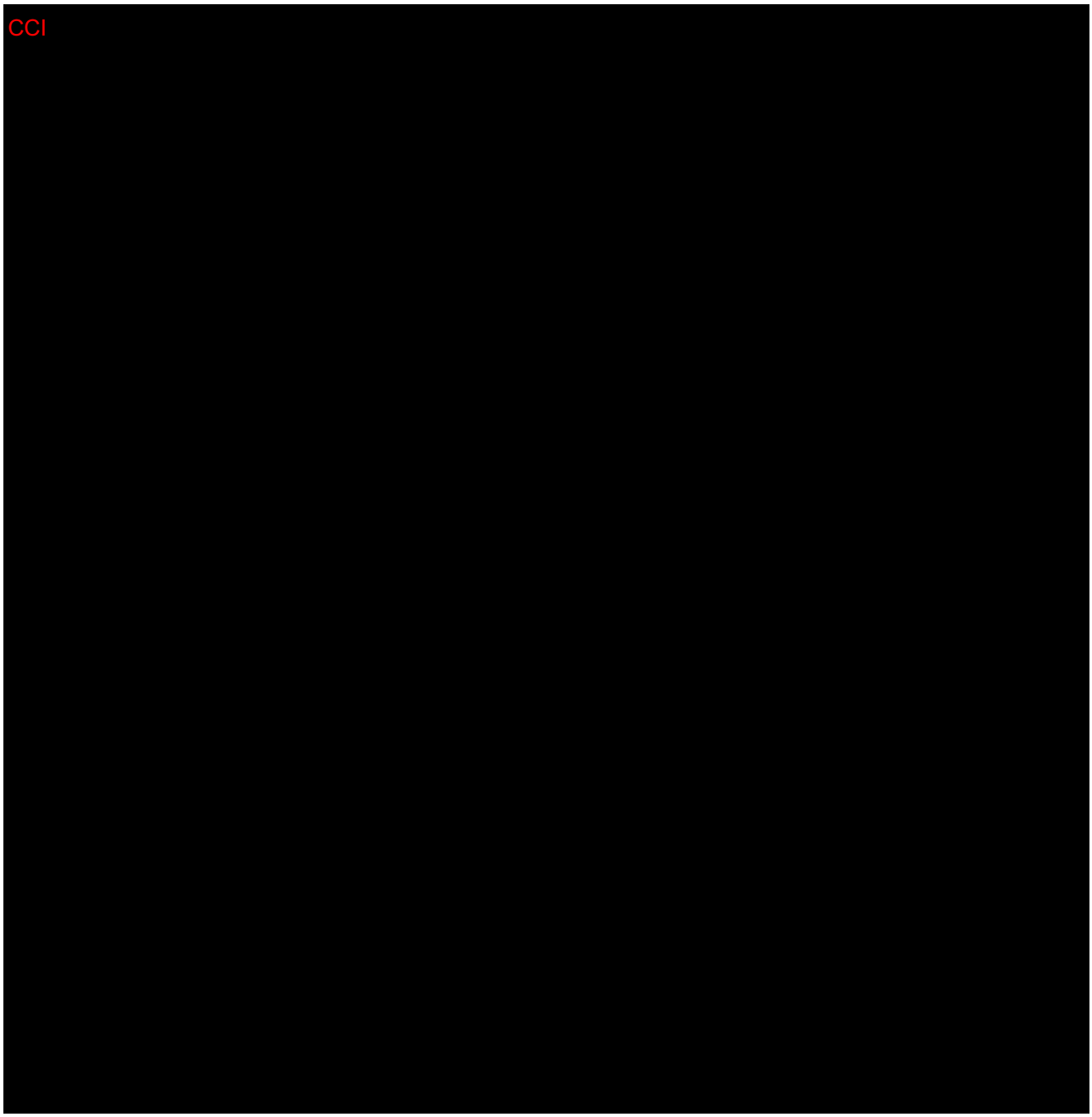
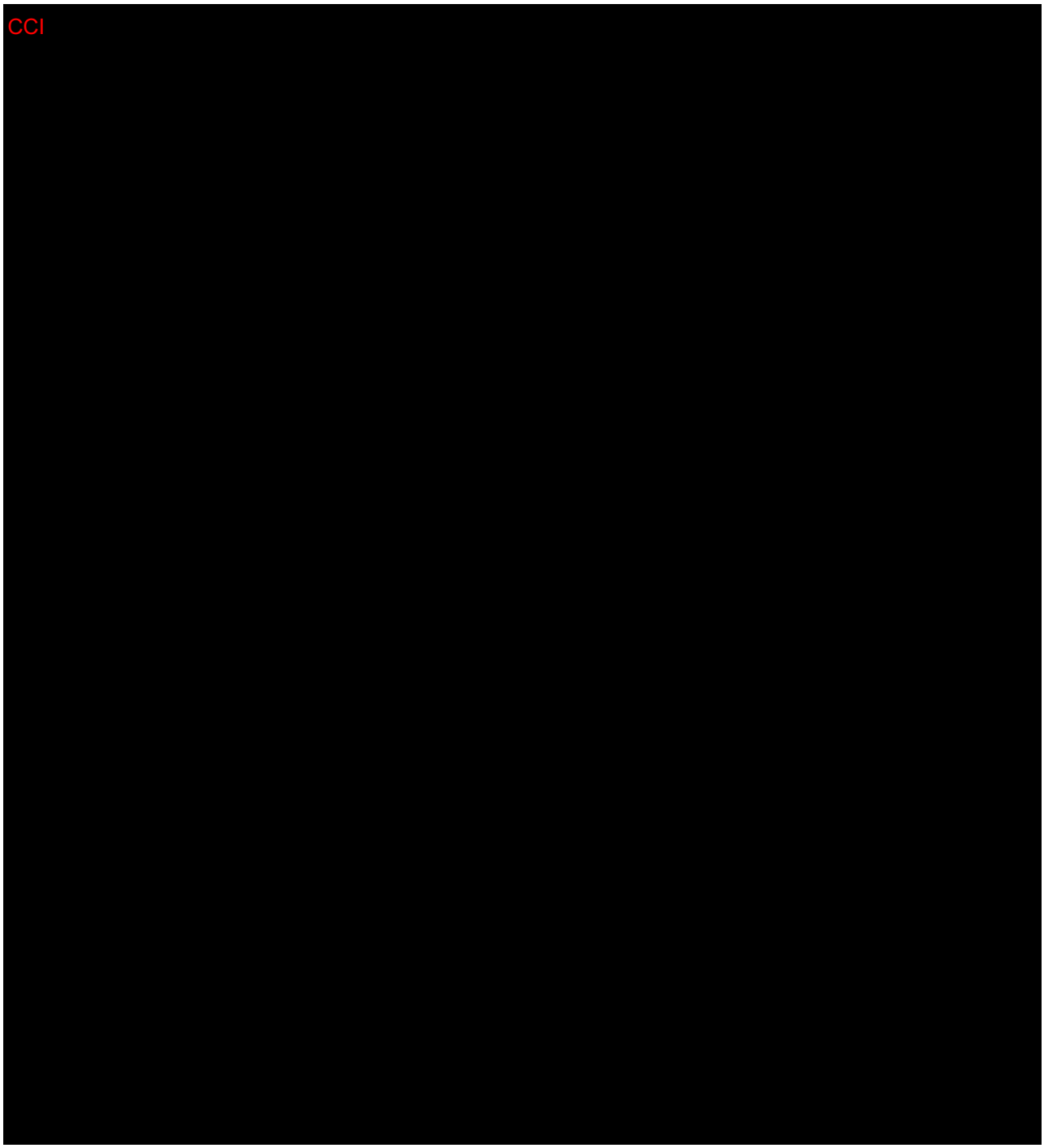

CCI

Table 4 CCI

CCI


CCI

13.2 Secondary Efficacy Endpoints

Analyses for the secondary efficacy endpoints in this section will be performed for the FAS using the observed values.

13.2.1 CGI-S for PWS Score – Change From Baseline at Week 12

The change from Baseline will be analyzed using the MMRM method similar to the primary analysis for the primary endpoint. The MMRM model will include effects for treatment group, sex, age group (<12 years old, ≥12 to <17 years old, ≥17 years old), participation in a previous clinical study of carbetocin nasal spray (presence or absence), visit, the treatment-by-visit interaction, Baseline score, and the Baseline score-by-visit interaction.

Summary statistics (observed and change from Baseline), LSMs, the between-treatment difference in LSM with the corresponding 95% CI, p-value, and the effect size (Cohen's *d*) will be presented at each post-Baseline visit.

13.2.2 CGI-C for PWS Score at Week 12

The score will be analyzed using the MMRM method similar to the primary analysis for the primary endpoint. The MMRM model will include effects for treatment group, sex, age group (<12 years old, ≥12 to <17 years old, ≥17 years old), participation in a previous clinical study of carbetocin nasal spray, visit, the treatment-by-visit interaction, CGI-S Baseline score, and the CGI-S Baseline score-by-visit interaction.

Summary statistics, LSMs, the between-treatment difference in LSM with the corresponding 95% CI, p-value, and the effect size (Cohen's *d*) will be presented at each post-Baseline visit.

13.2.3 Percentage of Subjects With Treatment Response (Defined as Improvement From Baseline ≥8 Points) Based on the HQ-CT at Week 12

The percentage of subjects with treatment response will be summarized by treatment group and analyzed using the logistic regression model adjusted by baseline HQ-CT, treatment group, sex, and randomization stratification factors of age grouping and participation in a previous clinical study of carbetocin nasal spray. Missing HQ-CT will be imputed as non-responder.

Response rates, odds ratio with corresponding 95% CI, and p-value will be presented for each post-Baseline visit.

13.2.4 CGI-S for hyperphagia in PWS Score – Change From Baseline at Week 12

The change from Baseline will be analyzed using the MMRM method similar to the primary analysis for the primary endpoint. The MMRM model will include effects for treatment group, sex, age group (<12 years old, ≥12 to <17 years old, ≥17 years old), participation in a previous

clinical study of carbetocin nasal spray visit, the treatment-by-visit interaction, Baseline score, and the Baseline score-by-visit interaction.

Summary statistics (observed and change from Baseline), LSMs, the between-treatment difference in LSM with the corresponding 95% CI, p-value, and the effect size (Cohen's *d*) will be presented at each post-Baseline visit.

13.3 Exploratory Efficacy Analyses

Analyses for the exploratory efficacy endpoints in this section will be performed for the FAS using the observed values.

PADQ, Food Safe Zone, and ZBI

The change from Baseline will be analyzed using the MMRM method similar to the primary analysis for the primary endpoint. The MMRM model will include effects for treatment group, sex, age group (<12 years old, ≥12 to <17 years old, ≥17 years old), participation in a previous clinical study of carbetocin nasal spray, visit, the treatment-by-visit interaction, Baseline score, and the Baseline score-by-visit interaction.

Summary statistics (observed and change from Baseline), LSMs, the between-treatment difference in LSM with the corresponding 95% CI, p-value, and the effect size (Cohen's *d*) will be presented at each post-Baseline visit.

PWS-HRUQ

Analyses of the PWS-HRUQ questionnaire will be outlined in a separate health economics analysis plan. A listing of PWS-HRUQ values will be provided.

PADQ Subgroup Analysis

The analyses of PADQ will be performed for the following subgroups:

- HQ-CT Baseline score (>22 or ≤22)
- PADQ Baseline score (in upper three quartiles or in lowest quartile)

Percentage of Subjects With Treatment Response (Defined as Improvement from Baseline ≥11 Points) Based on PADQ at Week 12

The percentage of subjects with treatment response will be summarized by treatment group and analyzed using the logistic regression model adjusted by baseline PADQ, treatment group, sex, and randomization stratification factors of age grouping and participation in a previous clinical study of carbetocin nasal spray.

Response rates, odds ratio with corresponding 95% CI, and p-value will be presented for each post-Baseline visit.

Question-Level Analysis of HQ-CT and PADQ

Question-level changes from Baseline may be analyzed using the MMRM method similar to the primary analysis for the primary endpoint, as needed.

14. SAFETY ANALYSES

All safety analyses will be performed using the actual treatment for the Safety Analysis Set.

14.1 Adverse Events

Adverse events will be coded using MedDRA dictionary, Version 28.0 or newer.

An AE (classified by preferred term) will be considered a TEAE if started after first dose administration and no later than last dose date + 30 days. AEs based on predose findings (e.g., clinically significantly abnormal vital signs, laboratory test results, or electrocardiogram parameters) will not be considered as TEAEs.

The event counts, the number, and percentage of subjects reporting TEAEs in each treatment group will be tabulated by SOC and preferred term; by SOC, preferred term, and maximum severity; and by SOC, preferred term, and relationship to study drug. If more than one TEAE occurs with the same preferred term for the same subject, then the subject will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by severity and by relationship to study drug. The display in these tables will be sorted alphabetically by SOC and then by descending subject frequency for the preferred terms (counts from carbetocin group) within each SOC.

The event counts, the number and percentage of subjects with any TEAEs will also be tabulated by preferred term without SOCs. This table will be sorted by descending subject frequency within carbetocin group.

Treatment-emergent adverse events (TEAEs), TEAEs leading to drug withdrawal, TEAEs related to study drug, TEAEs by maximum severity, fatal TEAEs, treatment-emergent SAEs, treatment-emergent SAEs related to study drug, treatment-emergent nonserious AEs, treatment-emergent nonserious AEs related to study drug, and TEAEs related to study drug leading to drug withdrawal will be summarized by SOC, preferred term, and treatment group. The tables will be sorted alphabetically by SOC and then by descending frequency within each SOC in carbetocin group. Device incidents or malfunctions will be summarized by treatment group. In addition, the incidence of fatal treatment-emergent AEs (i.e., events that cause death) will be summarized separately by SOC, preferred term and treatment group.

Common TEAEs are those that occur in at least 5% of subjects in either treatment group. Common TEAEs, common nonserious TEAEs, common related TEAEs and common related nonserious TEAEs will be summarized by SOC, preferred, term, and treatment group.

An AE listing by subject will display all events, including those which occur during screening, and will include the verbatim term in addition to the MedDRA SOC and preferred term. This listing will also include all relevant eCRF data associated with the event: date of onset, date

resolved, date of last dose, severity, frequency, outcome, relationship to study drug, and action taken with study drug. Separate listings will be presented for subjects with treatment-emergent SAEs, treatment-emergent nonserious AEs, related TEAEs, TEAEs leading to drug withdrawal, device incidents and malfunctions and fatal TEAEs (if any).

14.2 Clinical Laboratory Variables

It is possible that some test results may be collected from a local laboratory. Local laboratory results and the associated normal ranges will be converted to Système International (SI) units; local laboratory results, in SI units, will then be normalized if applicable to central lab ranges to be included in summary data analysis together with the central laboratory results. CCI

CCI

CCI . Summary of change from baseline, shift, and PCI analyses will be presented.

Clinical laboratory assessments are performed at Screening (Visit 1), Baseline (Week 0), Week 2, Week 8, and Week 12 (EOT)/ET.

- Clinical chemistry serum tests (CHEM):
    - Sodium (Na), potassium (K), carbon dioxide (CO_2), chloride (Cl), calcium (Ca), blood urea nitrogen (BUN), creatinine (CR)
    - Glucose
    - Albumin (ALB)
    - Total protein
    - Alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), total bilirubin (TBIL), lactate dehydrogenase (LDH)

    - Thyroid function tests
        - Thyroid stimulating hormone (TSH)
            - TSH should be performed only at Screening
            - Free T4 will be measured only if the TSH is abnormal

- Pregnancy tests
    - A serum pregnancy test should be performed at Screening for women of childbearing potential
    - A urine pregnancy test should be performed at all designated visits after Screening for women of childbearing potential
        - If urine cannot be obtained in women of childbearing potential, a serum pregnancy test should be done in its place
- Hematology tests:
    - Complete blood count (CBC) including:
        - White blood cell (WBC) count
        - Complete differential (relative and absolute)
        - Hematocrit (Hct), hemoglobin, red blood cells (RBC), platelets
- Urinalysis (UA):
    - Blood, RBCs, WBCs, protein, glucose, ketones, specific gravity, pH

Clinical laboratory values (in SI units) and the change from Baseline values will be summarized by treatment group at each post-Baseline visit using descriptive statistics. The overall minimum and maximum observed and change from Baseline values will also be summarized. For urinalysis with categorical results, the number and percentage of subjects will be tabulated by category at Baseline and each post-Baseline visit, and the denominator is the number of subjects with nonmissing values for the given parameter, visit and treatment group.

Laboratory values will also be summarized in shift tables by treatment group, to determine the number and percentage of subjects with values classified as below, within, and above normal ranges at each post-Baseline visit relative to the same classification at the Baseline visit. For the by-visit shift summary, the denominator is the number of subjects with nonmissing values at Baseline and the given visit for the given parameter and treatment group. For the shift to the overall post-Baseline minimum or maximum, all post-Baseline values will be considered, including unscheduled and out of window values and the denominator is the number of subjects

with nonmissing Baseline value and at least one post-Baseline value for the given parameter and treatment group.

Clinical laboratory values are PCI if they meet either the low or high PCI criteria listed in Table 5 and Table 6. The number and percentage of subjects with post-Baseline PCI values for each of the categories in Table 5 and Table 6 will be summarized by treatment group for selected parameters. For the overall post-Baseline summaries of PCI values, all post-Baseline values will be considered, including unscheduled and out of window values. Subjects with multiple PCI values for a given parameter will be counted only once for that parameter. For the overall post-Baseline summary, the numerator of the percentage is the number of subjects with at least one post-Baseline PCI laboratory value for the given parameter and treatment group, and the denominator is the number of subjects with at least one post-Baseline laboratory value for the given parameter and treatment group.

Table 5 PCI Criteria: Hematology and Chemistry

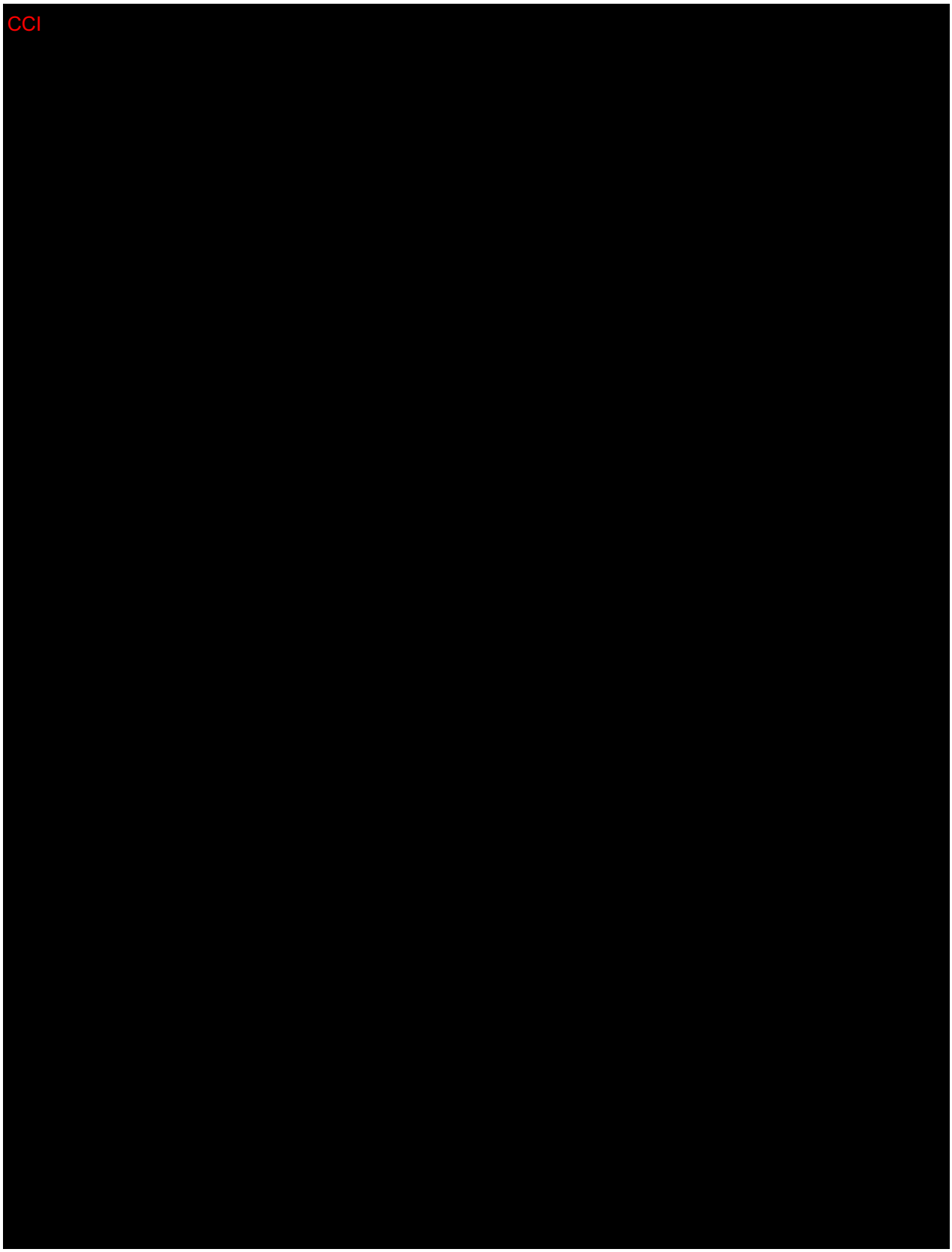


Table 6 PCI Criteria: Urinalysis

CCI

Clinical laboratory data will be displayed in data listings with date and study day of collection. All units will be displayed according to SI conventions for units. Out of range

values will be flagged in the data listings (i.e., "L" or "H"). A separate listing will be provided for a subset of the chemistry, hematology, and urinalysis analytes with values classified as PCI.

The latest version of the reference ranges provided by the central laboratory will be used for all analysis; PCI status and categorical labs (low, normal, high) and shifts will be derived from the latest version of the main reference range. Both the reference ranges (main and alternative) and the corresponding categorical assessment will be listed.

The pregnancy results (positive or negative) will be presented in a listing.

14.3 Vital Signs, Height, Weight, and BMI

Vital signs will be collected throughout the study at every visit in the clinic; height will be measured at Baseline and Week 12 (EOT)/ET; weight will be measured at Screening, Baseline, Week 2, Week 8 and Week 12 (EOT)/ET. Observed vital signs, height, weight, and BMI and the changes from Baseline at each post-Baseline visit will be summarized by treatment group using descriptive statistics.

Vital sign and weight values will be considered PCI if they meet the criteria listed in Table 7. The number and percentage of subjects with post-Baseline vital signs that are PCI will be summarized by treatment group at each post-Baseline visit and for overall post-Baseline. For the overall post-Baseline summaries, all post-Baseline values will be considered, including unscheduled and out of window values. Subjects with multiple PCI values for a given parameter will be counted only once for that parameter. For the by-visit summary, the numerator for the percentage is the number of subjects with a post-Baseline PCI vital sign/weight for the given parameter, visit and treatment group, and the denominator is the number of subjects with nonmissing values for the given parameter, visit and treatment group. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least one post-Baseline PCI vital sign for the given parameter and treatment group, and the denominator is the number of subjects with at least one post-Baseline vital sign/weight for the given parameter

and treatment group. A listing of overall and of subjects with any PCI vital sign or weight values will be provided.

Table 7 PCI Criteria: Vital Signs and Weight



14.4 Electrocardiogram

12-lead ECGs are collected throughout the study at every visit in the clinic.

ECG readings will occur both before the first dose is given (predose) and after the first dose is given (postdose). These two assessments will be summarized separately. If multiple ECG assessments are performed at any other visit (or for the predose or first postdose ECG), the average of all available values for that planned ECG will be considered as one assessment for the summaries, with the averages rounded to integer values.

Observed (average) of ECG variables (e.g., heart rate, PR interval, QRS interval, QT interval, corrected QT interval using Bazett's correction method [QTcB] and corrected QT interval using Fridericia's correction method [QTcF]) and the changes from Baseline at each assessment time point will be summarized by visit and treatment group.

Observed (average) of QTcB and QTcF will also be categorized into the following categories (ms) and the number and percentage of subjects in each category will be summarized by treatment group at each visit and for the overall post-Baseline maximum:

- Observed: ≤450, 451 to ≤480, 481 to ≤500, and >500; >450; >480
- Change from Baseline: ≤10, 11 to 30, 31 to 60, and >60; >30

For cardiologist's interpretations, the number and percentage of subjects with ECG results that are determined as normal or abnormal will be summarized at scheduled visits. The overall post-baseline worst interpretation will also be summarized (i.e., if a subject has one or more post-baseline ECG results that is/are considered as abnormal, this subject will be counted in the abnormal category). Cardiologist's interpretations will also be summarized in a shift table. The shifts from Baseline to overall post-Baseline worst interpretation will also be presented. For the by-visit shift summary, the denominator is the number of subjects with nonmissing cardiologist's interpretation at Baseline and the given visit for the given treatment group. For the summaries of shift from Baseline to the overall post-Baseline worst interpretation, the denominator is the number of subjects with nonmissing Baseline and at least one post-Baseline cardiologist's interpretation for the given treatment group.

Electrocardiogram variable average values will be considered PCI if they meet the criteria listed in Table 8. The number and percentage of subjects with post-baseline PCI values will be summarized by treatment group at each post-Baseline visit and for overall post-Baseline. For the by-visit summary, the numerator for the percentage is the number of subjects with a post-Baseline PCI ECG for the given parameter, visit and treatment group, and the denominator is the number of subjects with nonmissing values for the given parameter, visit and treatment group. For the overall post-Baseline summary, the numerator for the percentage is the number of subjects with at least one post-Baseline PCI ECG for the given parameter and treatment group, and the denominator is the number of subjects with at least one post-Baseline ECG value for the given parameter and treatment group. A listing of overall and of all subjects with any PCI ECG values will be provided.

Table 8 PCI Criteria: ECG



14.5 Physical Examination

Physical examinations are performed throughout the study at every visit in the clinic. Physical examination results (normal, abnormal, and not done) will be summarized in a frequency table by treatment group, body system and visit. A listing of overall physical examination data will be presented.

14.6 Nasal Assessment

Nasal assessments are performed throughout the study at every visit in the clinic. Nasal assessment results (grade 0 = no abnormal findings to grade 4 = nasal septum perforation) will be summarized in a frequency table by treatment group and visit. A listing of nasal assessment data will be presented.

14.7 Other Safety Endpoints

There are no other safety endpoints in this study.

15. CLINICAL PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

Plasma concentration data for carbetocin will be listed based on the Safety Analysis Set. Plasma concentration data for the carbetocin-treated group will be summarized using descriptive statistics based on the PK Analysis Set. Carbetocin plasma concentration data will remain blinded until the unblinding of the clinical database at the end of the study. Concentrations that are below the limit of quantitation (BLQ) will be displayed as "BLQ" in the data listings and imputed as 0 for computing summary statistics.

Additionally, population PK analyses will be performed to further characterize the PK profile. The E-R relationship of carbetocin using measures of safety and efficacy parameters will be assessed. The details of the population PK and E-R relationship analyses will be presented in a separate report in accordance with a separate data analysis plan.

16. INTERIM ANALYSIS

No interim analysis is planned in this study.

17. DATA AND SAFETY MONITORING BOARD

An external data and safety monitoring board (DSMB) will review ongoing safety data, including data on AEs, SAEs, and safety laboratory data. The DSMB will be independent of the Sponsor and the Investigators and will be empowered to recommend whether the study should be terminated due to safety concerns, continue as planned, or continue with modification.

The DSMB membership, activities, responsibilities, and frequency of meetings will be described separately in the DSMB charter.

While the DSMB will be making recommendations regarding future conduct of the study, the Sponsor retains final decision-making authority on all aspects of the study.

18. COMPUTER METHODS

Statistical analyses will be performed using Version 9.4 (or newer) of SAS® on a qualified and validated environment.

Validation and quality control of the tables, listings, and figures containing the results of the statistical analyses will follow appropriate standard operating procedures.

19. CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

No changes are made to the analyses specified in the protocol.

20. REFERENCES

Cotter SP, Schwartz L, Strong TV, Bender RH, Fehnel SE. The Prader-Willi Syndrome Anxiousness and Distress Behaviors Questionnaire: development and psychometric validation. *Value Health*. 2023;26(2):243-250.

European Medicines Agency, Committee for Medical Products for Human Use. *Guideline on Missing Data in Confirmatory Clinical Trials*. EMA/CPMP/EWP/1776/99 Rev. 1. European Medicines Agency; 2010.

Guy W. Clinical Global Impressions. In: *Early Clinical Drug Evaluation (ECDEU) Assessment Manual for Psychopharmacology, revised*. US Department of Health, Education, and Welfare Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, National Institutes of Mental Health Psychopharmacology Research Branch, Division of Extramural Research Programs;1976:218-222.

International Council for Harmonisation. *Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials* ICH E9(R1). International Council for Harmonisation; November 2019.

RTI Health Solutions. *Psychometric Evaluation of the Hyperphagia Questionnaire for Clinical Trials.* 24 March 2016. RTI-HS Project Number 0304043.

Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist*. 1980;20(6):649-655.

21. APPENDICES

Appendix A: Schedule of Assessments

Period	Screening	Double-blind Treatment Period					Safety Follow-up[c]
Visit Week	Week -3 to 0	Baseline 0	2	4	8	12 (EOT)/ET	30 Day follow-up[c]
Visit Number	1	2[a]	3[b]	4[b]	5[b]	6[b]	
Visit window (days)		N/A	±3	±3	±4	+4	+4
Visit Type	Clinic	Clinic	Clinic	Remote	Clinic	Clinic	Telephone or Video Teleconference
Informed consent	X						
Caregiver consent	X						
Inclusion/exclusion criteria	X	X					
Medical history and demographics	X						
Prader-Willi syndrome history	X						
Confirm documentation of mutation causing PWS[d]	X						
Physical examination including nasal assessment	X	X	X		X	X	
12-lead electrocardiogram (ECG)[e]	X	X	X		X	X	
Clinical laboratory tests (hematology, chemistry, urinalysis)	X	X	X		X	X	
TSH (and free T4, if necessary)	X						
Vital signs	X	X	X		X	X	
Height		X				X	
Weight	X	X	X		X	X	
Pregnancy test[f]	X	X	X		X	X	
Randomization		X					

Blood samples for pharmacokinetics		X	X		X	X	
Hyperphagia Questionnaire for Clinical Trials (HQ-CT)[g]	X	X	X	X	X	X	
Clinical Global Impression–Severity (CGI-S) for PWS	X	X	X		X	X	
Clinical Global Impression–Change (CGI-C) for PWS			X		X	X	
Clinical Global Impression–Severity (CGI-S) for hyperphagia in PWS	X	X	X		X	X	
PWS Anxiousness and Distress Behaviors Questionnaire (PADQ)	X	X	X	X	X	X	
Food Safe Zone[g]	X	X	X	X	X	X	
Zarit Burden Interview (ZBI)		X	X	X	X	X	
PWS-specific Healthcare Resource Use and Caregiver Burden Questionnaire (PWS-HRUQ)		X				X	
Adverse event review	X	X	X	X	X	X	X
Device incident review		X	X	X	X	X	
Concomitant medications review	X	X	X	X	X	X	X
Study drug dispensed[h]		X					
Study drug return[h]			X		X	X	
Study drug compliance diary			X		X	X	
Study drug accountability[h]			X		X	X	
Optional caregiver interview[i] (US only)						X	

Abbreviations: EOT=end of treatment; ET=early termination; PWS=Prader-Willi syndrome; TSH=thyroid stimulating hormone

a The Baseline visit assessments must occur prior to first dose of the double-blind study drug. Administration of the first dose and postdose ECG, and PK samples collection may be done on the same day or on the following day.

b Timing of postbaseline visits will be calculated from the first day of dosing (Day 1) (e.g., the Week 2 visit will occur 2 weeks [±3 days] after the first day of dosing).

c The Safety Follow-up visit will assess adverse events and concomitant medications and occur 30 (+ 4) days after the last dose of study drug. Subjects who roll over into the OLE study will not have the safety follow-up visit.

d If documentation of mutation is not complete, genomic testing for the mutation causing PWS should be conducted per the Investigator's standard practice.

e The (predose) ECG at the Baseline visit will be completed prior to administration of the first dose of study drug. The postdose ECG will be completed 20 minutes (within a 10 to 40 minute window) after the first dose of study drug is administered.

f A serum pregnancy test should be performed at Screening for women of childbearing potential. A urine pregnancy test should be performed at all designated visits after Screening for women of childbearing potential. If urine cannot be obtained in women of childbearing potential, a serum pregnancy test should be done in its place.

g The primary assessment, the HQ-CT, will be administered to the caregiver by a rater using standardized prompts. Caregiver scales should be rated at the clinic whenever possible, however they may be done remotely if the primary caregiver is unable to attend the clinic visit. The Food Safe Zone should be administered immediately before administration of the HQ-CT.

h Study drug will be dispensed at the site during the Baseline visit. Thereafter study drug may be shipped directly to the subject (where permitted). Study drug shipment, return, and accountability will be performed in accordance with the Pharmacy Drug Distribution Manual.

i The caregiver interview will be completed by 3rd-party trained moderators within 2 weeks after an end-of-treatment or early termination visit.

Appendix B: Summary of Version Changes

Version No:	Document history Description of update	Author(s)	Version date
1.0	Original plan.	PPD PPD	2024-11-21
2.0	Main Updates: 1) Updated baseline definition for efficacy variables; 2) Modified language in section 13.1.1 per FDA comments; 3) Added subgroup analyses for HQ-CT based on genetic subtype, Baseline HQ-CT score and Baseline PADQ score; 4) Added exploratory subgroup analyses for PADQ based on Baseline HQ-CT score and based on Baseline PADQ score; 5) Added exploratory PADQ Responder analysis; 6) Added question-level analysis of HQ-CT and PADQ; 7) Updated language throughout text to align with protocol amendment 2 GLOBAL.	PPD	2025-05-30
3.0	Main Updates: 1) Add statement on normality assumption check in Section 5.1 per FDA comment; 2) Add SAS code for multiple imputations in Section 13.1.2.1 per FDA comment; 3) Add Two-way Delta Adjustment Imputation Method, in Section	PPD PPD	2025-09-03

	13.1.2.1 per FDA comment; 4) Week 2 analysis window definition updated in Section 5.5.		

Signature Page for ACP-101-302 Statistical Analysis Plan

Approval w eSig Task	PPD Biostatistics 05-Sep-2025 15:31:54 GMT+0000

Approval w eSig Task	PPD Biostatistics 05-Sep-2025 15:57:38 GMT+0000

Approval w eSig Task	PPD Clinical 05-Sep-2025 16:01:19 GMT+0000

Signature Page for ACP-101-302 Statistical Analysis Plan