Statistical Analysis Plan

Phase 2b – Core Period

Title:	**A Phase 2b/3 study to evaluate the safety, tolerability, and effects of livoletide (AZP-531), an unacylated ghrelin analog, on food-related behaviors in patients with Prader-Willi syndrome**
Protocol Number:	**AZP01-CLI-003**
Product:	**AZP-531**
Sponsor:	**Millendo Therapeutics SAS** **15G Chemin du Saquin** **Ecully 69130, France**
Date:	**23-Mar-2020**
Version:	**Final 1.1**
Prepared by:	**Excelsus Statistics, Inc.** **4278 Av. Hampton** **Montreal (Qc), H4A 2K9** **Canada**

Signature Page

Title: A Phase 2b/3 study to evaluate the safety, tolerability, and effects of livoletide (AZP-531), an unacylated ghrelin analog, on food-related behaviors in patients with Prader-Willi syndrome

Protocol Number: AZP01-CLI-003

Prepared by:

Josée Morin, MSc
Senior Biostatistician
Excelsus Statistics

Signature

23MAR2020
Date

Approved by:

Diane Potvin, MSc
President
Excelsus Statistics

Signature

23 MAR 2020
Date

Vivian Lin, MD
Vice President
Clinical Development
Millendo Therapeutics

Signature

23MAR2020
Date

Modification History

Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
Draft 1.0	05-APR-2018	Josée Morin	N/A
Final 1.0	22-Nov-2019	Diane Potvin	Updated according to FDA review and protocol updates
Final 1.1	23-Mar-2020	Diane Potvin	Updated according to some of the FDA comments and the ZEPHYR (AZP01-CLI-003) protocol amendment. As additional change, the gatekeeper strategy was removed for the key secondary endpoints and Holm's method was proposed to control the overall type 1 error for the two comparisons on the primary endpoint.

Table of Contents

Abbreviations

AE	Adverse Event
AG	Acylated Ghrelin
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
BW	Body Weight
CGI	Clinical Global Impression
CGI-I	Clinical Global Impression-Improvement
CGI-S	Clinical Global Impression of Severity
CgGIS-H	Caregiver Global Impression of Severity – Hyperphagia
CGIS-H	Clinician Global Impression of Severity – Hyperphagia
CgGIC-H	Caregiver Global Impression of Change – Hyperphagia
CGIC-H	Clinician Global Impression of Change - Hyperphagia
CRO	Contractual Research Organization
CSR	Clinical Study Report
DBC-P	Developmental Behavior Checklist for Pediatric patients
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EPG	Elevated Postprandial Glucose
EQ-5D-5L	European Quality of Life Five Dimension Five Level Scale
FAS	Full Analysis Set
HOMA-IR	Homeostatic Model Assessment for Insulin Resistance
HQ-CT	Hyperphagia Questionnaire for Clinical Trials
IGF-1	Insulin-like growth factor 1
IGT	Impaired Glucose Tolerance
IQ	Intelligence Quotient
ITT	Intent To Treat
IWRS	Interactive Web Response System
kg	Kilogram
KR	Kenward-Rogers
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model Repeated Measures
MNAR	Missing-Not-At-Random
NDA	New Drug Application
NRS	Numerical Rating Scales
PK	Pharmacokinetic
PP	Per Protocol
PPMS	Pattern-Mixture Models
PT	Preferred Term
PWS	Prader-Willi Syndrome

REML	Restricted Maximum Likelihood
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SI	Système International
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TOEPH	Toeplitz
T2D	Type 2 Diabetes Mellitus
UAG	Unacylated Ghrelin
WC	Waist Circumference
WHO-DD	World Health Organization-Drug Dictionary
ZBI	Zarit Burden Interview

1.0 INTRODUCTION

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for Millendo Therapeutics clinical protocol AZP01-CLI-003 (version 1.3 dated March 23, 2020) and subsequent amendments. The protocol includes two parts: (i) a Phase 2b dose-response study (3-month double-blind, placebo-controlled Core Period and a 9-month Extension Period) and (ii) a Phase 3 study (6-month double-blind, placebo-controlled Core Period and a 6-month Extension Period). This document refers only to the Phase 2b Core Period. Statistical analyses pertaining to the extension period and to the Phase 3 part are presented in separate documents.

The investigational product is livoletide, formerly known as AZP-531.

This SAP has been developed prior to database lock and final analysis for the Phase 2b Core Period. All final analyses will be conducted after the clinical trial data are entered into the database, any discrepancies in the data are resolved, and the database is locked (closed to changes).

The derivation of the responder threshold for the Hyperphagia Questionnaire for Clinical Trials (HQ-CT) and analyses of the measurement properties of the HQ-CT are not covered in this SAP but will be described in a stand-alone Psychometric Analysis Plan.

The pharmacokinetic (PK) endpoints and related statistical analyses will be described in detail in a separate Pharmacokinetic Analysis Plan.

2.0 STUDY OBJECTIVES

2.1 Primary Objective

Core Period

- To demonstrate the efficacy of 3-month treatment with livoletide as compared to placebo for reducing caregiver-observed food-related behavior as assessed by the HQ-CT.

2.2 Secondary Objectives

Core Period

- To demonstrate the efficacy of 3-month treatment with livoletide as compared to placebo for:
    - Reducing total body fat mass in overweight/obese patients with Prader-Willi Syndrome (PWS);

- Reducing waist circumference (WC) in overweight/obese patients with PWS;
- Reducing body weight (BW) in overweight/obese patients with PWS;

where
Overweight/obese is defined as:
For patients ≥ 18 years of age: Body mass index (BMI) ≥ 27 kg/m^2
For patients 4-17 years of age: BMI ≥ 90th percentile for the same age and sex

2.3 Additional Assessments

Core Period

- To evaluate the effect of livoletide as compared to placebo on:
  - Hyperphagia severity as assessed by the Caregiver and Clinical Global Impression of Severity - Hyperphagia scales (CgGIS-H and CGIS-H, respectively);
  - Hyperphagia global impression of change as assessed by the Caregiver and Clinical Global Impression of Change - Hyperphagia scales (CgGIC-H and CGIC-H, respectively);
  - Clinical global impression of change as assessed by the Clinical Global Impression of Improvement (CGI-I) scale;
  - Clinical global impression of severity as assessed by the Clinical Global Impression of - Severity (CGI-S) scale;
  - Fasting (for approximately 8 hours) and postprandial glucose and insulin (profile);
  - HOMA-IR;
  - HbA1c;
  - Lipids;
  - Lean body mass;
  - Patient-reported appetite following breakfast as assessed by a Numeric Rating Scale (NRS);
  - Patients' Non-Food-Related behavior as assessed by the Developmental Behavior Checklist 2-Parent/Carer version (DBC2-P);
  - Patients' Quality of Life (QoL) as assessed by the Pediatric Quality of Life inventory^{TM} (PedsQL^{TM}) 4.0 Generic Core Scales;
  - Caregivers' disease burden as assessed by the Zarit Burden Interview (ZBI);
  - Health state utilities as assessed by European Quality of Life Five Dimension Five Level Scale (EQ-5D-5L) Self-complete version and Proxy version 1;
  - BMI and BMI Z-score (the latter for patients 4-17 years of age only);
- To evaluate the measurement properties of the HQ-CT in the study population (test-retest reliability, construct validity and ability to detect change);
- To derive the responder threshold for meaningful change in the HQ-CT;

- To compare the percentage of HQ-CT responders between groups, where HQ-CT responders will be defined as per the stand-alone Psychometric Analysis Plan.

2.4 Safety Objectives

To assess the safety and tolerability of livoletide as compared to placebo over a 3-month treatment period.

3.0 STUDY DESIGN

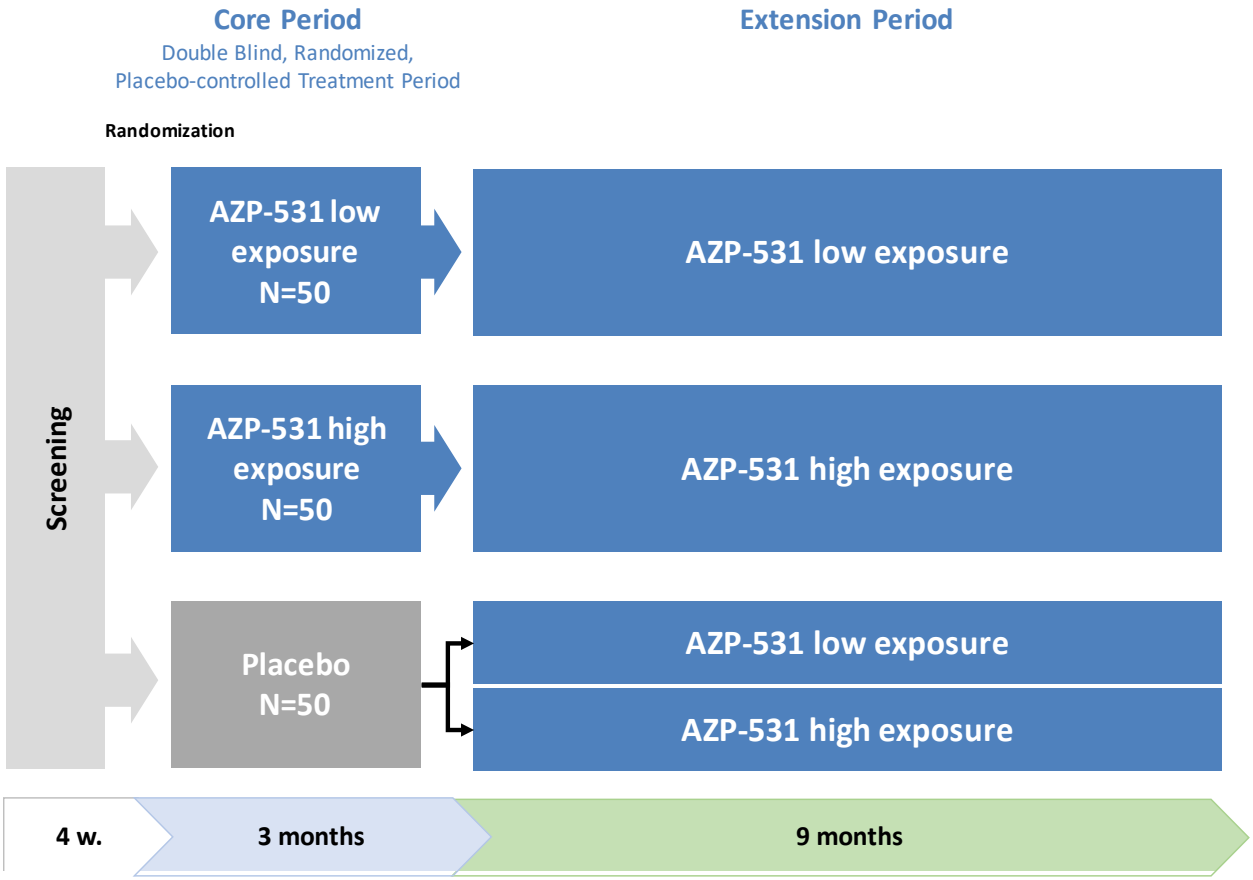
3.1 General Description

The Phase 2b part of the study will include adult and pediatric patients (4-65 years of age) with genetically confirmed PWS and evidence of hyperphagia as judged by the investigator and HQ-CT total score ≥ 10. For Phase 2b, a total of approximately 150 patients 8 to 65 years of age will need to be randomized. In addition to this cohort of 150 patients, a separate cohort of patients 4 to 7 years of age will also be randomized. Phase 2b will include a Core Period and an Extension Period. The patients will be randomized in approximately 40 sites in Europe, Australia and North America.

The Phase 2b Core Period will consist of a 3-month, randomized, double-blind, placebo-controlled treatment period. After a screening period (up to 4 weeks), approximately one hundred and fifty (150) patients 8 to 65 years of age will be randomized in a 1:1:1 ratio to one of the 3 arms (livoletide 60 µg/kg - low exposure [8 mg/mL], livoletide 120 µg/kg - high exposure [16 mg/mL], or placebo). In addition to this cohort of 150 patients, a separate cohort of patients 4 to 7 years of age will be randomized similarly (1:1:1 ratio). Randomization will be stratified based on age (age ≥4 and < 8 years, ≥8 and <18 years vs ≥18 years) and BMI (patients ≥ 18 years of age: BMI <27 kg/m^2 vs BMI ≥ 27 kg/m^2; patients 4-17 years of age: BMI <90th percentile vs BMI ≥ 90th percentile for the same age and sex). A minimum of 25 overweight/obese patients will be randomized per treatment group.

Patients randomized to livoletide for the Core Period will remain on the randomized dose (i.e., 60 µg/kg or 120 µg/kg) during the Extension Period. Patients randomized to placebo for the Core Period will cross-over to livoletide 60 µg/kg or 120 µg/kg in a randomized (1:1) manner. The Interactive Response Technology (IRT) will manage treatment assignment so that the double-blind status can remain intact at the start of the Extension Period.

Figure 1: Phase 2b Study Design



Abbreviations: AZP-531=livoletide; N=number of patients; pts=patients; w=weeks

3.2 Schedule of Events

Schedule of events is presented in Table 1.

Table 1: Schedule of Events – Phase 2b Core Period

Study Stage	Phase 2b Core Period				
	Screening[1]	Double-blind Treatment Period			
Visit#	V1	Randomization[2] V2	V3	V4	End of Core Period /ET Discontinuation V5
Study Week(s)	**-4 to -1**	**1**	**5**	**9**	**13**
Treatment Duration (weeks/months)	**-**	**-**	**4/1**	**8/2**	**12/3**
Study Day(s)	**-28 to -1**	**1**	**29**	**57**	**85**
Visit Window (±X days)	**0**	**0**	**3**	**3**	**3**
Informed consent and assign a study specific patient number	X				
Demography	X				
Medical history[3]	X				
Disease history[4]	X				
Inclusion/exclusion criteria	X				
IQ test[5]	X				
Concomitant medications	X	X	X	X	X
Physical examination[6]	X (C)	X (C)	X (A)	X (A)	X (C)
Vital signs[7]	X	X	X	X	X

[1] The screening period is up to 4 weeks.

[2] The randomization visit will occur a minimum of 2 weeks after the screening visit.

[3] Complete medical history will be obtained at screening and any clinically significant changes from the screening visit will be reported as AEs.

[4] A confirmed genetic diagnosis of PWS. Documentation of PWS subtype (chromosome 15 micro-deletion *versus* non-deletion) is also required for the study. If the PWS subtype is not known, a sample for testing may be obtained and the patient may continue on to be enrolled into the study if he or she meets all the other inclusion criteria and none of the exclusion criteria.

[5] Collect IQ scores and dates of completion if available in patient's medical record.

[6] "C" for complete physical examination, "A" for abbreviated physical examination. See Section **Error! Reference source not found.** of the protocol for assessments to be performed during a physical examination.

[7] Vital signs (respiration rate, pulse rate, and blood pressure) will be measured sitting after the patient has rested comfortably for 5 minutes. See Section **Error! Reference source not found.** of the protocol.

Study Stage	Phase 2b Core Period				
	Screening[1]	Double-blind Treatment Period			
Visit#	V1	Randomization[2] V2	V3	V4	End of Core Period /ET Discontinuation V5
Study Week(s)	**-4 to -1**	**1**	**5**	**9**	**13**
Treatment Duration (weeks/months)	**-**	**-**	**4/1**	**8/2**	**12/3**
Study Day(s)	**-28 to -1**	**1**	**29**	**57**	**85**
Visit Window (±X days)	**0**	**0**	**3**	**3**	**3**
WC, BW, and BMI[8]	X	X	X	X	X
Height[9]	X	X	X	X	X
12-lead ECG[10]	X	X[10]			X
DXA (total body fat mass, lean body mass, bone mineral density)[11]		X			X
Pregnancy test[12]	X	X	X	X	X
Safety laboratory evaluations in the fasting condition[13]	X	X	X		X
Blood samples for PK analysis[14]			X		
Blood samples for HbA1c	X	X			X
Blood samples for glucose and insulin profiles[15]		X	X		X
Blood samples for AG and UAG profiles[15]		X	X		X

8 Before breakfast. See Section **Error! Reference source not found.** and Section **Error! Reference source not found.** of the protocol.

9 Before breakfast. See Section **Error! Reference source not found.** of the protocol.

10 See Section **Error! Reference source not found.** of the protocol.

11 DXA can be obtained within 2 weeks prior to randomization (Visit 2) and the End of Core Period visit (Visit 5). See Section **Error! Reference source not found.** and Section **Error! Reference source not found.** of the protocol.

12 Pregnancy tests will be performed for women of childbearing potential only. See Section **Error! Reference source not found.** of the protocol.

13 Blood samples will be collected after fasting for approximately 8 hours. Clinical serum chemistries, hematology, coagulation, and urinalysis parameters are described in Section **Error! Reference source not found.** of the protocol. Clinical laboratory evaluations will be performed centrally.

14 PK samples (3-timepoint sampling) will be collected at the same time as glucose/insulin/AG/UAG profiles as outlined in the Pharmacokinetic Schedule of Activities provided in Section **Error! Reference source not found.** of the protocol.

15 Blood samples for glucose, insulin, AG, and UAG profiles will be collected after fasting for approximately 8 hours (all patients) and 30 minutes and 180 minutes post-start of an isocaloric breakfast (only patients 8 to 65 years).

Study Stage	Phase 2b Core Period				
	Screening[1]	Double-blind Treatment Period			
Visit#	V1	Randomization[2] V2	V3	V4	End of Core Period /ET Discontinuation V5
Study Week(s)	**-4 to -1**	**1**	**5**	**9**	**13**
Treatment Duration (weeks/months)	**-**	**-**	**4/1**	**8/2**	**12/3**
Study Day(s)	**-28 to -1**	**1**	**29**	**57**	**85**
Visit Window (±X days)	**0**	**0**	**3**	**3**	**3**
Blood samples for IGF-1		X			X
Blood samples for ADA analysis		X	X		X
Review and completion of questionnaires on eDevice [16]	X	X	X	X	X
Isocaloric breakfast taken on site and report any food not consumed [17]		X	X		X
Appetite-NRS[17]	X	X	X		X
HQ-CT	X	X	X	X	X
Quality of Life: PedsQL™ Parent-Proxy Report		X			X
CgGIS-H		X	X	X	X
CGIS-H		X	X	X	X
CgGIC-H			X	X	X
CGIC-H			X	X	X
CGI-I			X	X	X
CGI-S		X	X	X	X
Caregiver Disease Burden: ZBI		X			X
EQ-5D-5L Self-complete version and Proxy version 1		X			X

16 The eDevice will be used to collect data from all questionnaires and the appetite-NRS. Data from the questionnaires and appetite-NRS will be completed on the eDevice while on site, either by the caregiver, the patient (for the NRS), or by the investigator. The eDevice will not be dispensed to the patients.

17 Appetite-NRS should be administered (only to patients 8 to 65 years) before (after fasting for approximately 8 hours), at the end, and 120 minutes post-start of an isocaloric breakfast, except at screening visit, where NRS will be administer for training and acclimation purposes (no profile). Food not consumed when administered the isocaloric breakfast needs to be recorded. One-site consumption of isocaloric breakfast is not required for patients 4 to 7 years old.

Study Stage	Phase 2b Core Period				
	Screening[1]	Double-blind Treatment Period			
Visit#	V1	Randomization[2] V2	V3	V4	End of Core Period /ET Discontinuation V5
Study Week(s)	**-4 to -1**	**1**	**5**	**9**	**13**
Treatment Duration (weeks/months)	**-**	**-**	**4/1**	**8/2**	**12/3**
Study Day(s)	**-28 to -1**	**1**	**29**	**57**	**85**
Visit Window (±X days)	**0**	**0**	**3**	**3**	**3**
DBC2-P		X			X
Review/record AEs	X----------				----------X
Randomization via IRT[18]		X			
Dispensation of glucometer kit[19]		X			
Dispensation of study drug		X	X	X	X
Drug accountability			X	X	X
On-site study drug administration (during treatment period only)		X	X	X	X

[18] Randomization via IRT should be performed on the first day of the double-blind treatment period.
[19] Glucometer will be provided to patients with T2D if necessary (i.e. if they do not have one).

4.0 PLANNED ANALYSES

4.1 Data Monitoring Committee (DMC)

Safety data review will be performed at a regular basis during the trial by an external Data Monitoring Committee (DMC) operating independently of the Sponsor to make recommendations for the conduct of the study based on safety data. The DMC will operate under the rules of an approved charter defining the roles and responsibilities of its members.

4.2 Interim Analysis

No interim analysis is planned in the Core Period of the Phase 2b study.

4.3 Final Analysis

All analyses per age cohort (8-65 years old and 4-7 years old as described in section 3.1) will be performed following the signature of the analysis plan and the lock of the database for that age cohort for the Phase 2b Core Period of the study. Once database lock is performed for all age cohorts, statistical analyses will also be performed on all age cohorts combined. For more details, please refer to section 6.2.

5.0 ANALYSIS SETS

5.1 Full Analysis Set [FAS]

The Full Analysis Set (FAS) will include all randomized patients. The initial randomization will be preserved for this analysis set in order to comply with the Intent-to-Treat (ITT) principle.

5.2 Per Protocol Analysis Set [PP]

The Per Protocol Set (PP) will include all randomized patients who:

- Completed their first 3 months of treatment.
- Have non-missing observations at Baseline and Day 85 for the primary efficacy endpoint (HQ-CT total score).
- Have a baseline value for the primary efficacy endpoint (HQ-CT total score) ≥ 10.
- Have no major CSR-reportable protocol deviations defined as those that may have impacted the primary efficacy assessment.

Prior to the database lock of the Core Period, major CSR-reportable protocol deviations will be defined and reviewed by Millendo Therapeutics and the CRO in charge of the study follow-up in

a blinded fashion. Patients with major CSR-reportable protocol deviations will be excluded from the PP set.

5.3 Safety Analysis Set [SAF]

The Safety Analysis Set (SAF) will include all randomized patients who received at least one dose of the study drug. The treatment actually received by the subject will be used in this set.

6.0 STATISTICAL CONSIDERATIONS

6.1 Sample Size Determination

In the completed Phase 2a study (2 weeks of treatment), the change from baseline to Day 14 in the 9-item score of the HQ was -4.3 units for livoletide and -1.6 units for placebo, with a difference between groups of 2.7 units.

For the Phase 2b Core Period (N=150, patients 8 to 65 years of age), a between group difference of 4 units in the change from baseline in the HQ-CT score has been selected for sample size calculation based on the following:

1. It is considered reasonable to achieve this difference as a 3 months treatment period is expected to produce a higher effect as compared to short-term treatment (2 weeks).
2. It is in line with the HQ-CT distribution-based estimates of meaningful change observed in a landmark Phase 3 clinical trial[1] that included 107 obese individuals with PWS, 12 to 65 years of age (half-standard deviation [SD]=3.3, standard error of the mean [SEM]=4.4). In addition, the distribution-based estimates were found to be approximately the magnitude of the difference between the mean change of the "A little better" and "No change" in the CGIC subgroups.

Assuming a 2-sided parametric test, alpha of 2.5% (Bonferroni adjustment for 2 comparisons) and a SD of 5, forty-one (41) patients per group will allow to detect a difference between groups of at least 4 units in the primary endpoint 9-item score of the HQ-CT. This is considered a conservative approach since the higher the difference between groups, the lower the sample size.

Assuming a drop-out rate at 3 months of about 18%, a total of 50 patients (8 to 65 years of age) per group will need to be randomized with a minimum of 25 overweight/obese patients per treatment group.

There is no formal sample size calculation done for patients in the 4 to 7 years of age cohort since analysis of this cohort is supportive (i.e. they will not be included in the Phase 2b primary endpoint analysis).

6.2 Database Lock and Unblinding Procedure

During Phase 2b, there will be two database lock events: one at the conclusion of the Core Period and one at the end of the Extension Period, therefore requiring two unblinding procedures. Also, since the cohort of patients 4 to 7 years of age will take longer to recruit than the cohort of patients 8 to 65 years of age, there will be a rolling lock at the end of both periods (i.e., a first lock for the cohort of patients 8 to 65 years of age; and a second lock for all patients, to include the cohort of patients 4 to 7 years of age).

Per age cohort, once data checks have been resolved, and the analysis sets and the SAP have been finalized and signed, the clinical database related to the Phase 2b Core Period will be locked in order to proceed with the statistical analyses. As the cohort of the 4- to 7-year-old patients will take longer to recruit, the statistical analyses will be performed first with the 8- to 65-year-old patients and then with the cohort of the 4- to 7-year-old patients and all age cohorts combined.

6.3 Reference Start Date and Analysis Day

Analysis Day will be calculated from the reference start date and will be used to show start/stop day or study day of assessments and events.

Reference start date is defined as the day of the first dose of study drug (i.e., Day 1 of the Core Period).

- If the date of the event is on or after the reference date then:

  Analysis Day = (date of event – reference date) + 1.

- If the date of the event is prior to the reference date then:

  Analysis Day = (date of event – reference date).

In the situation where the event date is partial or missing, Analysis Day and any corresponding durations will appear partial or missing in the listings.

6.4 Baseline

The baseline is defined for each parameter as in Table 2 below:

Table 2: Baseline Definitions

Endpoint	V1	V2	Baseline Definitions	Additional Information
HQ-CT	X	X	Mean of V1 and V2	Include all unscheduled assessments between V1 and V2 in the Mean.
Total Body Fat Mass		X	V2 value	
Waist Circumference	X	X	Mean of V1 and V2	Include all unscheduled assessments between V1 and V2 in the Mean.
Body Weight	X	X	Mean of V1 and V2	Include all unscheduled assessments between V1 and V2 in the Mean.
CgGIS-H		X	V2	
CGIS-H		X	V2	
CgGIC-H			N/A	
CGIC-H			N/A	
CGI-I			N/A	
CGI-S		X	V2	
Fasting & Postprandial Glucose		X	V2 (time point specific)	Take the nominal time point into account: after fasting (all patients) and 30 minutes and 180 minutes post-start of an isocaloric breakfast (only patients 8 to 65 years of age).
Insulin		X	V2 (time point specific)	Take the nominal time point into account: after fasting (all patients) and 30 minutes and 180 minutes post-start of an isocaloric breakfast (only patients 8 to 65 years of age).

HOMA-IR		X	V2	
HbA1c	X	X	Mean of V1 and V2	Include all unscheduled assessments between V1 and V2 in the Mean.
Lipids	X	X	Mean of V1 and V2	Include all unscheduled assessments between V1 and V2 in the Mean.
Lean Body Mass		X	V2	Include values from 3 weeks before to one week after V2. including unscheduled assessments.
Appetite-NRS	X	X	V2	
DBC2-P		X	V2	
PedsQL		X	V2	
ZBI		X	V2	
EQ-5D-5L		X	V2	
BMI	X	X	Mean of V1 and V2	Include all unscheduled assessments between V1 and V2 in the Mean.
Vital Signs	X	X	Last **pre-dose** assessment on or prior to V2	Including unscheduled assessments.
Height	X	X	Last assessment on or prior to V2	Including unscheduled assessments.
PE	X	X	Last assessment on or prior to V2	Including unscheduled assessments.
Bone Mineral Density		X	Last assessment on or prior to V2	Include values from 3 weeks before to one week after V2, including unscheduled assessments.
ECG	X	X	Last **pre-dose** assessment on or prior to V2	Including unscheduled assessments.

Safety Labs	X	X	Last **pre-dose** assessment on or prior to V2	Including unscheduled assessments.
AG, UAG, AG/UAG		X	V2 (time point specific)	Take the nominal time point into account: after fasting (all patients) and 30 minutes and 180 minutes post-start of an isocaloric breakfast (only patients 8 to 65 years of age).
IGF-1		X	Last **pre-dose** assessment on or prior to V2	
ADA		X	Last **pre-dose** assessment on or prior to V2	

Adverse events (AEs) commencing on the reference start date will be considered to be treatment-emergent adverse events (TEAEs). Medications taken during the treatment core period will be considered to be concomitant medications.

6.5 Retests, Unscheduled Visits and Early Termination Data

Unscheduled and retest measurements will be provided in data listings. Those values might be used in the summary statistics for pre-dose measurements (see section 6.4), but they will not be used in by-visit summaries for post-dose measurements, unless there are no scheduled data available for a planned visit. If only unscheduled data are available for a planned visit (as per the schedule of assessments, using the statistical windowing), those data will be used in by-visit summaries for post-dose measurements as well.

For early termination visits, data will be mapped using the windowing conventions presented in section 6.7.

6.6 Definitions

- BMI (kg/ m^2) = weight (kg)/ height $(m)^2$
- Overweight/obese patients
    - patients ≥18 years of age: BMI ≥27 kg/m2
    - patients 4-17 years of age: BMI ≥90th percentile for the same age and sex
- Non overweight/obese patients

- patients ≥18 years of age: BMI <27 kg/m2
- patients 4-17 years of age: BMI <90th percentile for the same age and sex

- T2D patients are defined as those with any ongoing medical history with MedDRA preferred term = "diabetes mellitus" or "diabetes mellitus inadequate control" or "diabetes with hyperosmolarity" or "insulin resistant diabetes" or "insulin-requiring type 2 diabetes mellitus" or "type 2 diabetes mellitus", since patients with type 1 diabetes mellitus are excluded from the study.

- Prediabetic patients are defined as those not having ongoing medical history of T2D (as defined above), and meeting any of the following: HbA1c 5.7-6.4% at V1 and V2, fasting plasma glucose at V2 5.6-6.9 mmol/L, ongoing medical history with preferred term = "glucose tolerance impaired" or "impaired fasting glucose".

- Dyslipidemia at baseline is defined as (LDL-c >= 3.0 mmol/liter at V1 and V2) and/or (triglycerides >= 2.0 mmol/liter at V1 and V2).
- Elevated postprandial glucose (EPG) is defined as 3-hr post-start of breakfast glucose concentration ≥ 8.0 mmol/L (144 mg/dL) at V2.

6.7 Windowing Conventions

The following table describes assignment of visit windows for purposes of analysis for the Phase 2b Core Period. Windowing will be applied to the data prior to any missing data calculations. These windows will be used to report efficacy variables, vital signs, 12-lead ECG and safety laboratory evaluations. The windowing will also be applied to flags used for study completion (e.g., completion of the 3-month Core Period).

Table 3: Analysis Visit Windows

Visit/Study Week	Target Day	Visit Window As per protocol	Visit Window Statistical Analyses
Baseline	Day 1	Not Applicable	See section 6.4.
Visit 3 / Week 5	Day 29	Day 26 – Day 32	Day 2 – Day 43
Visit 4 / Week 9	Day 57	Day 54 – Day 60	Day 44 – Day 71
Visit 5 / Week 13	Day 85	Day 82 – 88	Day 72 – Day 103 [a, b]

[a] This should be prior to the first dose of study drug in the Extension Period, except for DXA.
[b] For DXA, measures taken within 3 weeks prior to or within 7 days after the V5 date will be considered as V5 values.

If more than one evaluation falls into the same time window, the evaluation nearest to the target day will be considered as the analysis value. If two evaluations are equally distant from the target day, the evaluation occurring after the target day will be retained.

6.8 Software Version

All analysis will be performed using SAS® software version 9.4 or higher.

6.9 Concept of Estimands and General Strategy for Analyses

As per ICH E9 Addendum[2], an estimand defines in detail what needs to be estimated to address a specific scientific question of interest. In this study, the specific question or objective is to demonstrate the efficacy of a 3-month (or 6-month) treatment with livoletide as compared to placebo for reducing caregiver-observed food-related behavior. As described in Figure 2 below, a description of an estimand includes four attributes:

A. the population, that is, the patients targeted by the scientific question;

B. the variable (or endpoint), to be obtained for each patient, that is required to address the scientific question;

C. the specification of how to account for intercurrent events to reflect the scientific question of interest.

D. the population-level summary for the variable which provides, as required, a basis for a comparison between treatment conditions.

The main estimand in this study will follow a treatment policy strategy. As per protocol section 7.2.1, a patient might continue in the study, but off drug (thus adhering to the Treatment policy strategy). Because it is not anticipated that patients will be followed for very long once they stop study drug, since patients typically do not want to be “off drug on study”, we had identified the main intercurrent event as discontinuation of study treatment. Since the patient will most likely not be followed-up after treatment discontinuation, this would result in missing data for the main estimand efficacy assessment at 3 months (for the phase 2b part) or 6 months (for the phase 3 part).

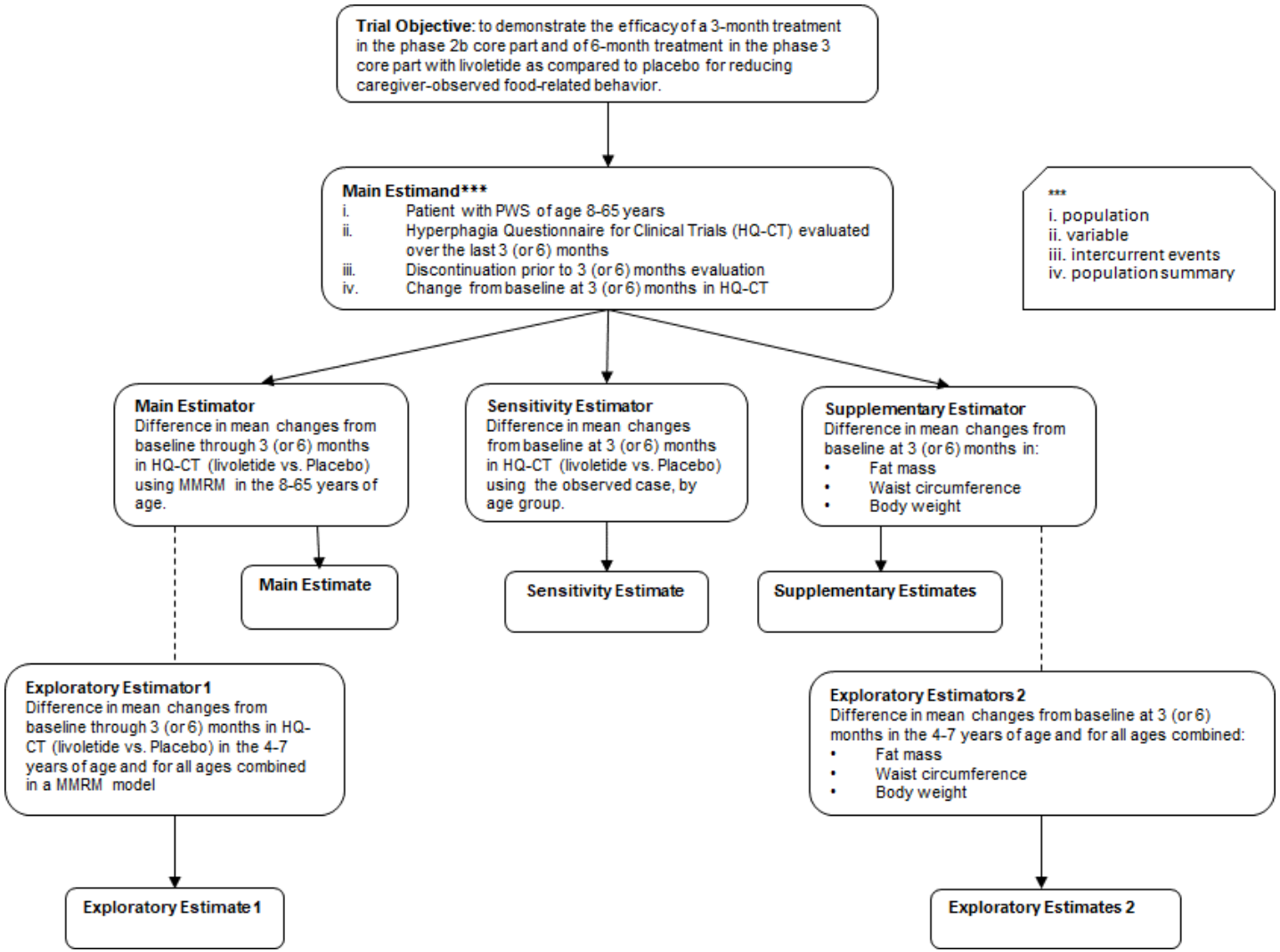
The main estimand will be based on the 8-65 years of age cohort.

The main estimator of the main estimand will be based on the cohort of patients 8-65 years of age using a mixed model repeated measures (MMRM) approach to handle the missing observations. Sensitivity analyses will be performed on observed cases at 3 months (for the phase 2b part) or 6 months (for the phase 3 part) to evaluate the robustness of the missing observations strategy using the MMRM approach in the cohort of patients 8-65 years of age. If missing observations are not at random or if the drop-out rate is more than 10%, additional imputation methods for missing observations will be performed by using a pattern-mixture models (PMMs) with PROC MI in SAS. PMMs are a straightforward and transparent way of modeling a wide variety of assumptions about missing data, and are particularly suited to implementing missing-not-at-random (MNAR) assumptions.

Supplementary analyses will consist of the analyses of the key secondary endpoints in a manner similar to the main estimator of the main estimand. They will also consist of the statistical analyses of all additional study assessments as described in section 2.3. These supplementary analyses will all be supportive in providing additional insights into the understanding of the treatment effect.

Exploratory analyses for the main estimand will present the results in the cohort of patients 4-7 years of age separately, as well as combined with the patients 8-65 years of age in a same MMRM model.

Figure 2: General concept of estimands in phase 2b and phase 3 studies



6.10 General Presentation

All data presentations are limited to the Core Period, which includes the scheduled Core Period visits and unscheduled visits that occur prior to the end of the Core Period/withdrawal during the Core Period. AEs are limited to only those that start during the Core Period, and concomitant medications are limited to only those that started prior to the end of the Core Period (i.e., start date <= end of Core Period data). The first study drug administration refers to the first administration during the Core Period. The last drug administration refers to the last administration during the Core Period. The last contact date refers to the last contact during the Core Period.

All continuous variables will be summarized by presenting the number of patients, mean, standard deviation, median, minimum, and maximum. Categorical variables (including level of scales) will be presented as frequencies and percentages. Summary tables will be presented by treatment group (60 µg/kg, 120 µg/kg and Placebo), visit, and time point, when applicable.

An overall alpha of 5% will be used. Details related to multiplicity are presented in Section 6.13.

6.11 Multicenter Studies

This study will be conducted at multiple sites worldwide, at approximately 40 sites in Europe, Australia and USA. To obtain a sufficient and better-balanced number of patients among study sites, pooling of sites will be applied prior to performing the statistical analyses and unblinding of treatments. Sites without at least 5 patients in the FAS will be incorporated into pooled sites as described below. European sites will be pooled together and USA sites will be pooled together.

<u>Sites in USA</u>

1) Sites with fewer than 5 patients in the FAS will be pooled with the geographically closest sites in a common geographical area (i.e., state). Within a geographical area:

    a. Sites with fewer than 5 patients will be ordered from lowest to highest in terms of number of FAS patients. In case of ties, the ordering for tied sites will be determined according to the site identification number (from smallest to largest).

b. Sites will be combined beginning at the smallest until the resulting pooled site contains at least 5 FAS patients. The sites pooled in this way will be considered as a single site in the statistical analyses.

c. The process described above will resume for the remaining sites not meeting the criterion of having at least 5 FAS patients. If the final set of pooled sites does not meet the criterion of having at least 5 FAS patients, the final set will be pooled with the preceding pooled site.

In the case where fewer than 5 FAS patients are available in one geographical area, the patients will be pooled with the closest geographical area (i.e., with the closest state). If there is more than one closest geographical area, the one with the smallest site identification number will be chosen.

Sites in Europe (for each Country) and Australia

1) Sites with fewer than 5 patients will be ordered from lowest to highest in terms of number of FAS patients. In case of ties, the ordering for tied sites will be determined according to the site identification number (from smallest to largest).

2) Sites will be combined beginning at the smallest until the resulting pooled site contains at least 5 FAS patients. The sites pooled in this way will be considered as a single site in the statistical analyses.

3) The process described above will resume for the remaining sites not meeting the criterion of having at least 5 FAS patients. If the final set of pooled sites does not meet the criterion of having at least 5 FAS patients, the final set will be pooled with the preceding pooled site.

In the case where fewer than 5 FAS patients are available in one country, the patients will be pooled as described in points 1) a to c (with countries instead of states).

6.12 Missing data

Unless otherwise specified, missing safety data will not be imputed. Missing efficacy endpoints will be handled as described in Section 13 of this analysis plan.

See **Error! Reference source not found.**Appendix 1 for handling of completely or partially missing dates for prior and concomitant medications and adverse events. In general, the following rules will be used:

Event Start Date Imputation

- Imputation of event end date should be done before imputation of event start date.
- Completely missing: Impute to the first study drug administration date.
- Missing day and month: Impute to January 1st, unless year is the same as year of first study drug administration then impute to the first study drug administration date.
- Missing day: Impute to the 1st of the month, unless month and year are the same as month and year of first study drug administration then impute to the first study drug administration date.
- If imputed event start date is after event end date (imputed or not), set the event start date to the event end date.

Event End Date Imputation

- Completely missing (and not flagged as "ongoing"): Impute to the last contact date.
- Missing day and month: Impute to December 31st, unless year is the same as last contact date then impute to the last contact date.
- Missing day: Impute to the last day of the month, unless year and month are the same as year and month of last contact date, then impute to the last contact date.

6.13 Multiple Comparisons/Multiplicity

In order to control for the overall type 1 error at 5% for the primary endpoint, Holm's method will be used for the two comparisons of interest (i.e., livoletide 60 µg/kg vs. placebo and livoletide 120 µg/kg vs. placebo). Using Holm's method, the comparison of interest with the smallest p-value will first be tested at alpha of 2.5% (2-sided). If not statistically significant, both comparisons will be declared not statistically significant. If statistically significant, that comparison will be declared statistically significant and the second comparison of interest (with the larger p-value) will be tested at alpha of 5% (2-sided).

With Holm's method, the study will be considered positive in terms of efficacy, if at least one of the comparisons is declared statistically significant. This strategy, which is a step-down version of the Bonferroni test, will control the overall type 1 error to no more than 5%.

The key secondary and other efficacy parameters will be tested at a nominal alpha of 5% for all comparisons, without control for multiplicity of testing.

6.14 Examination of Subgroups

For the cohort of patients 8-65 years of age, the following subgroups will be assessed in regards to the primary efficacy endpoint:

- Age (8-17 and ≥18 years of age)
- BMI (not overweight/obese, overweight/obese)

In addition, analyses of blood glucose, insulin, HOMA-IR and HbA1c will be performed on patients with elevated postprandial glucose (EPG) and also based on prediabetic/T2D patients. Lipid data will be presented for patients with dyslipidemia at baseline.

6.15 Examination of Cohort of Patients 4-7 Years of Age

The cohort of patients 4-7 years of age will be presented separately as well as combined with the cohort of patients 8-65 years of age.

7.0 DISPOSITION AND WITHDRAWALS

All patients who provide informed consent will be accounted for. The number of patients randomized, treated, completing and discontinuing from the Phase 2b Core Period as well as the number of patients in each analysis population will be summarized by treatment group. For patients who did not complete the Phase 2b Core Period, the reasons for withdrawal will be presented. The reasons for which a patient is excluded from the analysis populations will be presented. The number of patients screened for inclusion will also be presented. Reasons for screen failures will be reported.

Summaries will be presented for the patients age 8-65 years of age, for the cohort of patients 4-7 years of age and for all patients (4-65 years of age) combined.

8.0 PROTOCOL DEVIATIONS

Clinical study report (CSR)-reportable-protocol deviations (as identified by the Sponsor prior to the database lock) will be summarized for all randomized patients by treatment group, and all CSR-reportable protocol deviations will be presented in a by-subject listing. Summaries will be presented for the cohort of patients 8-65 years of age, for the cohort of patients 4-7 years of age and for all patients (4-65 years of age) combined.

9.0 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Baseline measures and demographic information will be summarized by treatment group for FAS, Safety, and PP sets. The demographic and baseline characteristics include, but are not

limited to: sex, age, age group (4-7, 8-17 and ≥ 18 years), genetic subtype (chromosome 15 micro-deletion, maternal uniparental disomy, imprinting center defect, non-deletion), body weight, height, body mass index (BMI), BMI category (not overweight/obese or overweight/obese), waist circumference, waist circumference by BMI category, total body fat mass, total body fat mass by BMI category, HQ-CT, CgGIS-H, CGIS-H, ZBI, use of growth hormone medication (yes/no), fasting AG, fasting UAG, AG/UAG ratio, HbA1c, fasting glucose, fasting insulin (for patients not on insulin at V2 (WHODrug ATC3=A10A)), prediabetes or T2D (yes/no), EPG (yes/no), dyslipidemia (yes/no) and HOMA-IR (for patients not on insulin at V2 (ATC3=A10A)).

Summaries will be presented for the patients 8-65 years of age, for the cohort of patients 4-7 years of age and for all patients (4-65 years of age) combined.

Individual patient data listings will be provided for all demographic and baseline characteristics. IQ will be listed only.

10.0 SURGICAL AND MEDICAL HISTORY

Medical history will be presented for the Safety analysis set. Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) dictionary, version 22.0. A summary table of medical history will be presented by treatment group and all treatments combined. Summaries will be presented for the patients 8-65 years of age, for the cohort of patients 4-7 years of age and for all patients (4-65 years of age) combined. A by-patient listing of all medical history will be presented.

11.0 PRIOR AND CONCOMITANT MEDICATIONS

Prior medications are defined as any medications started prior to the first dose of study drug. Concomitant medications are defined as any medications taken during the treatment period of the study, including those that started prior to the first dose date of study drug and continued past that date. All medications will be coded using the World Health Organization Drug Dictionary (WHO DD), March 2019.

Incidence of concomitant medications will be tabulated by treatment group and all treatments combined by Anatomical Therapeutic Chemical (ATC) Classification System levels 1 to 4. Summaries will be presented for the patients 8-65 years of age, for the cohort of patients 4-7 years of age and for all patients (4-65 years of age) combined. A by-patient listing of all

medications will be presented with an indication of whether the medication is considered concomitant or not and prior or not.

12.0 STUDY DRUG EXPOSURE AND COMPLIANCE

The total extent of exposure, number of injections and compliance during the treatment period will be summarized by treatment group for the FAS, Safety, and PP sets. Summaries will be presented for the patients 8-65 years of age, for the cohort of patients 4-7 years of age and for all patients (4-65 years of age) combined. A by-patient listing of all exposure and compliance data will be presented.

12.1 Derivations

The number of injections performed is calculated as:

> Date of last dose of study drug – Date of first dose of study drug – any missed doses + 1

The total extent of exposure (expressed as a number of days) will be defined as:

> Date of last dose of study drug - Date of first dose of study drug + 1

Compliance (as a percentage) will be calculated as follows:

$$100 * (\text{Number of injections performed}) / (\text{Total extent of exposure (in days)})$$

13.0 EFFICACY OUTCOMES

Efficacy analyses will be performed on the FAS as main analyses and on the PP set as supportive analyses (when applicable).

Actual change and percentage change (when applicable) from baseline results for all variables, including each item of the HQ-CT, will be summarized by treatment group and visit, using descriptive statistics, including n, arithmetic mean, standard deviation (SD), 95% confidence interval to the mean, median, minimum, Q1, Q3, and maximum values. Descriptive summaries will be presented for the patients 8-65 years of age, for the cohort of patients 4-7 years of age and for all patients (4-65 years of age) combined.

For each inferential analysis, the distribution of efficacy residuals will be visually examined to determine whether substantial departures from normality and homogeneity of variance are apparent. If the data are inconsistent with the assumption of normality and homogeneity of

variance, alternative transformation of the dependent variable or non-parametric models may be presented instead.

Individual patient data listings will be provided for all efficacy outcomes.

13.1 Primary Efficacy Variable

13.1.1 Primary Efficacy Variable & Derivation

The primary efficacy endpoint is the change from baseline to the end of the 3-month Core Period (i.e., V5/Week 13) in HQ-CT total score. The HQ-CT is derived from the Hyperphagia Questionnaire and is a 9-item instrument that has been developed and validated for use in clinical trials.

All items on the HQ-CT are rated on a five-point scale (0 = not a problem to 4 = severe and/or frequent problem) and are summed to generate a total score from 0 to 36. The HQ-CT questionnaire is completed at screening, randomization (V2), V3, V4 and V5 during the Core Period.

13.1.2 Missing Data Methods for Primary Efficacy Variable

Missing observations for the primary endpoint will be addressed by using the mixed-effect model repeated measures (MMRM) approach with an unstructured (UN) variance-covariance pattern across the visits. Details are presented in section 13.1.3.

If missing observations are not at random, additional imputation methods for missing observations will be performed by using PROC MI in SAS.

13.1.3 Primary Analysis of Primary Efficacy Variable

The primary analysis for the primary efficacy endpoint will be done using an MMRM analysis on the change from baseline in HQ-CT total score within the first 3 months in the cohort of patients 8-65 years of age. The model will include treatment, visit, stratification variables (age group 8-17 years, $\geq$ 18 years), BMI category (overweight/obese, not overweight/obese), and treatment-by-visit interaction as fixed effects and screening HQ-CT total score and V2 minus screening HQ-CT total scores as covariates. An unstructured covariance matrix will be used to model the correlation. The method of estimation will be REML and the method for approximating the degrees of freedom of the denominator will be Kenward-Rogers (KR). If convergence issues arise, the heterogeneous Toeplitz (TOEPH) structure will be used. If TOEPH is used, a sensitivity analysis of the MMRM results will be performed by requesting the empirical sandwich estimator (option EMPIRICAL) and by allowing SAS to use the default between-within DF estimation method for this sensitivity analysis. Contrasts will be defined

in the mixed model to test for treatment differences at different time points, but the primary comparison will be done at 3 months (i.e., V5/Week 13). This analysis will be performed on both FAS (primary analysis) and PP.

As a supportive statistical analysis, MMRM on the change from baseline in HQ-CT total score will be done with treatment, visit, stratification variables, pooled sites, treatment by visit interaction, pooled sites-by-visit interaction, pooled sites-by-treatment interaction and pooled sites-by-treatment-by-visit interaction as fixed effects and baseline HQ-CT total score and screening HQ-CT total score as covariates. An unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. If pooled sites-by-visit interaction, pooled sites-by-treatment interaction or pooled sites-by-treatment-by-visit interaction are not statistically significant ($p<0.05$), they will be removed from the model. Contrasts will be defined in the mixed model to test for treatment differences at different time points. This analysis will be performed on FAS only.

As exploratory statistical analyses, every item of the HQ-CT global score will be analyzed as described in the first paragraph, but by using each item as dependent variable in a separate statistical model. These analyses will be performed on FAS only.

13.1.4 Sensitivity Analyses

13.1.4.1 Analysis of Covariance at 3 Months

As a sensitivity analysis, an analysis of covariance (ANCOVA) on the change-from-baseline HQ-CT total score at 3 months will be performed (with no imputation of missing values) in the cohort of patients 8-65 years of age. The model will include treatment and the stratification variables as fixed effects, and screening HQ-CT total score and V2 minus screening HQ-CT total scores as covariates. The treatment differences in least-squares means and corresponding 95% CIs will be presented. This analysis will be performed on FAS only.

13.1.4.2 Subgroup Analyses

In addition to the primary analysis which will present the effect of treatment with Age (8-65 years) and BMI combined, subgroup analyses will be performed to explore the treatment effect within each level of age group and each level of BMI category.

A sensitivity analysis will be performed on the change from baseline in HQ-CT total score within the first 3-month treatment period with subgroups of age (age 8 to 17 and age ≥18 years) included in the statistical model. The MMRM model will be used with treatment, age group, BMI category, visit, and the interactions treatment-by-age group, treatment-by-visit and treatment-by-visit-by-age group as fixed effects and screening and V2 minus screening HQ-

CT total scores as covariates. Contrasts will be defined to test for treatment differences for each level of age group and visit. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented by visit for all ages combined.

Similar analyses will be done on the change from baseline in HQ-CT total score within the first 3-month treatment period with subgroups of BMI (not overweight/obese and. overweight/obese) included in the statistical model. The MMRM model will be used with treatment, age group, BMI category, visit, and the interactions treatment-by-BMI category, treatment-by-visit and treatment-by-visit-by-BMI category as fixed effects and screening and V2 minus screening HQ-CT total scores as covariates. Contrasts will be defined to test for treatment differences for each level of BMI category and visit. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented by visit for all BMI combined.

Results will also be presented graphically (Forest Plot) using the FAS.

13.1.4.3 Analysis of Cohort of Patients 4-7 Years of Age and All Age Groups Combined

In addition to descriptively summarizing the HQ-CT scores for the cohort of patients 4 to 7 years of age (See section 13.0), the analysis of section 13.1.4.2 will be repeated with age group separated in three categories: age 4-7 years, age 8 to 17 years, and age ≥18 years. The MMRM model will be used with treatment, age group, BMI category, visit, and the interactions treatment-by-age group, treatment-by-visit and treatment-by-visit-by-age group as fixed effects and screening and V2 minus screening HQ-CT total scores as covariates. Contrasts will be defined to test for treatment differences for each level of age and visit. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented by visit for all age groups combined. These analyses will be done on FAS.

13.2 Secondary Efficacy Variables

13.2.1 Secondary Efficacy Variables & Derivations

The key secondary efficacy endpoints include the percentage change from baseline to the end of the 3-month Core Period in total body fat mass, the change from baseline to the end of the 3-month Core Period in waist circumference and the percentage change from baseline to the end of the 3-month Core Period in body weight. The key secondary efficacy endpoints are

assessed in overweight/obese patients as main analyses and in not overweight/obese patients and on all patients as supportive analyses.

Before breakfast, patients will be weighed clothed (underwear, light gown or light clothing only), without footwear (no shoes or sandals) or heavy jewelry, using a calibrated scale. The same scale should be used throughout the study if possible. The conditions under which patients are weighed should be kept consistent if possible. Total body fat mass will be assessed by Dual Energy X-ray Absorptiometry (DXA). Waist circumference and body weight are measured at screening, randomization (V2), V3, V4 and V5 during the Core Period and fat mass is measured at V2 and V5 during the Core Period.

13.2.2 Missing Data Methods for Secondary Efficacy Variables

Missing observations for the secondary endpoints of body weight and waist circumference will be handled the same way as for the primary endpoint (i.e., using the MMRM approach with an unstructured (UN) variance-covariance pattern across the visits). Details are presented in section 13.1.3. For total body fat mass, since this parameter is only measured at randomization and at week 13, no missing observations will be imputed.

13.2.3 Analysis of Secondary Efficacy Variables

The primary and supportive analyses described in section 13.1.3 will be performed on the change from baseline in waist circumference and percentage change from baseline in body weight based on overweight/obese patients. The stratification variable related to the BMI will not be included in the different statistical models.

For the percentage change from baseline in total body fat mass, an ANCOVA on the percentage change from baseline in total body fat mass at 3 months will be performed (with no imputation of missing values). The model will include treatment and the stratification variable age group (age 8 to 17 years and age ≥18 years) as fixed effects and total body fat mass at baseline as covariate. The treatment differences in least-squares means and corresponding 95% CIs will be presented. The analysis will be done on both FAS and PP. A supportive analysis including pooled sites and pooled sites by treatment will also be performed. The pooled sites-by-treatment interaction will be removed from the statistical model if not statistically significant ($p<0.05$). The analysis will be done on FAS in overweight/obese patients only.

A second supportive ANCOVA will be done on the change from baseline in total body fat mass using a similar model as for the percentage change-from-baseline in total body fat mass. The analysis will be done on FAS.

The statistical analyses described above will be done in overweight/obese patients as main analyses and in not overweight/obese patients and on all patients as exploratory analyses (in FAS only). The analyses on all patients will include BMI category as covariate.

13.2.4 Analysis of the Cohort of Patients 4-7 Years of Age and All Age Groups Combined

In addition to descriptively summarizing the key secondary efficacy variables for the cohort of patients 4-7 years of age (See section 13.0), the following statistical analyses will be performed:

- The sensitivity analysis including the 4- to 7-year-old age group described in section 13.1.4.3 will be performed on the change from baseline in waist circumference and percentage change from baseline in body weight based on overweight/obese patients. The stratification variable related to the BMI will not be included in the different statistical models.
- For the percentage change from baseline and the change from baseline in total body fat mass, an ANCOVA at 3 months will be performed (with no imputation of missing values). The model will include treatment and the stratification variable age group (age 4-7 years, age 8-17 years, and age ≥18 years) as fixed effects, the interaction age group-by-treatment and total body fat mass at baseline as covariate. Contrasts will be defined to test for treatment differences for each level of age. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented for all age groups combined.

The statistical analyses described above will be done in overweight/obese patients as main analyses and in not overweight/obese patients and on all patients as exploratory analyses (in FAS only). The analyses on all patients will include BMI category as covariate.

13.3 Additional Assessments

13.3.1 Additional Assessments & Derivations

The additional assessments include the following:

Change from baseline to the end of the 3-Month Core Period in:

- Caregiver and Clinician Global Impression of Severity - Hyperphagia (CgGIS-H and CGIS-H) scores

- Clinical Global Impression of Severity (CGI-S) score
- Fasting/postprandial glucose and insulin
- HOMA-IR
- HbA1c
- Lipids (total cholesterol, triglycerides, low density lipoprotein-cholesterol (LDL-c), high density lipoprotein-cholesterol (HDL-c))
- Lean body mass
- Appetite-Numeric Rating Scale (NRS) scores
- Developmental Behavior Checklist Parent/Carer (DBC-P) score
- PedsQL^{TM} Parent-Proxy age-appropriate Reports scores
- Zarit Burden Interview (ZBI) scores
- Health state utilities as assessed by EQ-5D-5L
- BMI
- BMI Z-score (patients <18 years of age only)

Actual end of the 3-Month Core Period assessment in:

- Caregiver and Clinician Global Impression of Change - Hyperphagia (CgGIC-H and CGIC-H) scores
- Clinical Global Impression of General Improvement (CGI-I) score

In addition, the percentage of HQ-CT responders at 3 months as defined by anchor-based and distribution-based methods will be determined as specified in the Psychometric and Responder Analysis Plan and results will be presented in the corresponding report, which will be included in an appendix to the clinical study report.

Glucose and insulin are collected after fasting for approximately 8 hours (all patients) and 30 minutes and 180 minutes post-start of an isocaloric breakfast (only patients 8 to 65 years of age) at V2, V3 and V5.

HOMA-IR is defined as fasting insulin (µU/L) x fasting plasma glucose (nmol/L)/22.5 (obtained at V2; calculated only for patients not currently on insulin (ATC3=A10A)). HOMA-IR will be derived at V2, V3 and V5.

HbA1c is measured at screening, V2 and V5.

Lipids are measured at screening, V2, V3 and V5.

Lean body mass is measured at V2 and V5.

Clinical Global Impression scales have been developed as research tools for clinical practice and have been since extensively used in clinical trials. They evaluate either severity of illness or clinical global improvement. Scales for severity of hyperphagia (CGIS-H, CgGIS-H and CGI-S) and hyperphagia improvement (CGIC-H, CgGIC-H and CGI-I) have been adapted from these tools. CGIS-H, CgGIS-H and CGI-S are measured at V2, V3, V4 and V5. CGIC-H, CgGIC-H and CGI-I are measured at V3, V4 and V5.

Patients are asked to rate appetite/prospective food consumption using 6-point NRS ("nothing at all" to "A very very large amount") specifically designed for the PWS patient population. NRS is measured at screening (only to patients 8 to 65 years), V2, V3 and V5. NRS will be measured before (after fasting for approximately 8 hours), at the end, and 120 minutes post-start of an isocaloric breakfast, except at screening visit, where NRS will be administer for training and acclimation purposes (no profile).

The DBC-P is a 96-item instrument completed by parents or other primary caregivers, used for the assessment of behavioral and emotional problems in people with developmental and intellectual disabilities. For each item that does describe the patient, number 2 would be circled if the item is very true or often true. Number 1 would be circled if the item is somewhat or sometimes true. If the item is not true, then 0 would be circled. If the patient is unable to behave in the way referred to in an item, 0 will be selected. All items for DPC-P are summed to generate a total score. DBC-P is measured at V2 and V5.

The PedsQL^{TM} generic core scale consists of 23 items related to physical functioning (eight items), emotional functioning (five items), social functioning (five items) and school/work/studies functioning (five items). All items are rated on a five-point scale (0 = Never to 4 = Almost Always) and are summed to generate a total score from 0 to 92.

The ZBI assess the level of burden experienced by the principal caregiver of study participants. It is a 20-item questionnaire rated on a five-point scale (0 = Never to 4 = Nearly Always) and summed to generate a total score from 0 to 80. PedsQL^{TM} is measured at V2 and V5.

The EQ-5D-5L consists of 2 pages: the EQ-5D-5L descriptive system and the EQ Visual Analogue scale (EQ VAS). The descriptive system comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). For each dimension, respondents select which statement best describes health on that day from a possible 5 options

of increasing levels of severity (no problems, slight problems, moderate problems, severe problems, and extreme problems). This will result in a 1-digit number which will express the level selected for that dimension: '1' (no problems); '2' (slight problems); '3' (moderate problems); '4' (severe problems) and '5' (extreme problems). A unique EQ-5D health state is referred to by a 5-digit code. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples.

In addition to the descriptive system, respondents also assess their health on the day of assessment on a visual analogue scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). EQ-5D-5L is measured at V2 and V5.

The BMI Z-scores, also called BMI standard deviation scores, are measures of relative weight adjusted for child age and sex. BMI Z-score will be determined based on the age, sex, BMI, and the CDC reference standard[3] for 4- to 17-year-old patients. BMI is measured at screening, V2, V3, V4 and V5. BMI Z-score will be derived at those visits as well.

13.3.2 Analysis of Additional Assessments

13.3.2.1 Analysis of the Cohort of Patients 8-65 Years of Age

CgGIS-H/CGIS-H and CGI-S scores

Analysis will be done using a mixed model repeated measures analysis (MMRM) on the change-from-baseline in CgGIS-H/CGIS-H and CGI-S scores. The model will include treatment, visit, age group (8-17 years, ≥18 years), BMI category, and treatment by visit interaction as fixed effects and baseline result as a covariate. An unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined in the mixed model to test for treatment differences at different time points. The differences in least-squares means and corresponding 95% CIs will be presented. Analysis will be performed on the FAS set.

DBC-P, PedsQL™ and ZBI scores

Analysis will be done using an ANCOVA on the change-from-baseline in DBC-P, PedsQL™ and ZBI scores. The model will include treatment, age group (8-17 years, ≥18 years) and BMI category as fixed effects and baseline result as a covariate. The treatment differences in least-squares means and corresponding 95% CIs will be presented. Analysis will be performed on the FAS set.

Glucose (Fasting/Postprandial), Insulin (Fasting/Postprandial), HOMA-IR and HbA1c

The analysis will be done using a mixed model repeated measures analysis (MMRM) to model the change from baseline glucose/insulin/HOMA-IR at visits 3 and 5. The model will include treatment, time (before breakfast, and 30 min and 180 min post-start of breakfast), visit, age group (8-17 years, ≥18 years), BMI category, time-by-treatment, treatment-by-visit, time-by-visit and time-by-treatment-by-visit interactions as fixed effects and corresponding timepoint baseline glucose/insulin/HOMA-IR as covariate. Unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined in the mixed model to test for treatment differences at the different time points. The differences in least-squares means and corresponding 95% CIs will be presented.

For HbA1c, an ANCOVA on the change-from-baseline HbA1c at 3 months will be performed (with no imputation of missing values). The model will include treatment, age group (8-17 years, ≥18 years) and BMI category as fixed effects and screening and V2 minus screening HbA1c as covariates. The treatment differences in least-squares means and corresponding 95% CIs will be presented. Moreover, HbA1c will be descriptively presented by visit and treatment in terms of n (%) in the following categories: <5.7%, 5.7-5.9%, 6.0-6.4% and >6.4%. A Cochran-Mantel-Hansel test will be performed at month 3 to test the treatment effect, controlling for age group. A shift table from worst post-dose HbA1c category from baseline will also be descriptively presented.

All analyses will be performed on FAS.

In addition, the same models described above on glucose, insulin, HOMA-IR and HbA1c will be performed on patients with EPG, and on patients with prediabetes/T2D.

All analyses performed on insulin and HOMA-IR will not include patients who were on insulin (ATC3=A10A).

Analyses performed on glucose will also be performed without patients who were on insulin.

Lipids

The analysis will be done using a mixed model repeated measures analysis (MMRM) to model the percentage change and change from baseline in total cholesterol, triglycerides, LDL-c and HDL-c at visits 3 and 5. The model will include treatment, visit, age group (8-17 years, ≥18 years), BMI category, time-by-treatment, treatment-by-visit, time-by-visit and time-by-treatment-by-visit interactions as fixed effects and corresponding timepoint screening and V2 minus screening total cholesterol/triglycerides/LDL-c/HDL-c as covariates. Unstructured

covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined in the mixed model to test for treatment differences at the different time points. The differences in least-squares means and corresponding 95% CIs will be presented. The analyses will be done on all patients and on patients with dyslipidemia at baseline. All analyses will be performed on FAS.

Lean body mass

For the change from baseline in lean body mass, an ANCOVA on the change from baseline in lean body mass at 3 months will be performed (with no imputation of missing values). The model will include treatment and the stratification variable age group (age 8 to 17 years and age ≥18 years) as fixed effects and lean body mass at baseline as covariate. The treatment differences in least-squares means and corresponding 95% CIs will be presented. The analysis will be done on FAS.

These statistical analyses will be done in overweight/obese patients, in not overweight/obese patients and on all patients. The analyses on all patients will include BMI category as covariate.

Appetite-NRS

This analysis will be done using an MMRM analysis to model the change from baseline Appetite-NRS score at visits 3 and 5. The model will include treatment, time (end of breakfast and 120 min. post-start of breakfast), visit, age group (8-17 years, ≥18 years), BMI category, time-by-treatment, treatment-by-visit, time-by-visit and time-by-treatment-by-visit interactions as fixed effects and corresponding timepoint baseline (V2) Appetite-NRS score as covariate. An unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined in the mixed model to test for treatment differences at the different time points. The differences in least-squares means and corresponding 95% CIs will be presented. Analysis will be performed on the FAS set.

CgGIC-H/CGIC-H and CGI-I

An MMRM analysis will be used to model the post dose CgGIC-H/CGIC-H and CGI-I at visits 3, 4 and 5. The model will include treatment, visit, age group (8-17 years, ≥18 years) and BMI category as fixed effects, and treatment-by-visit interaction. Unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined in the mixed model to test for treatment differences at different time

points. The differences in least-squares means and corresponding 95% CIs will be presented. Analysis will be performed on the FAS set.

EQ-5D-5L and VAS

Frequency count and percentage for each level of each dimension will be presented by treatment group and visit. Actual and change from baseline results in EQ-5D-5L derived utility index and VAS will be summarized by treatment group and visit, using descriptive statistics, including n, arithmetic mean, standard deviation (SD), 95% confidence interval to the mean, median, minimum, Q1, Q3, and maximum values.

An ANCOVA will be performed on the change from baseline EQ-5D-5L utility index and VAS at 3-month. The model will include treatment, age group (8-17 years, ≥18 years), and BMI category as fixed effects and the baseline EQ-5D-5L utility index/visual analogue scale as covariate. The treatment differences in least-squares means and corresponding 95% CIs will be presented.

Analysis will be performed on the FAS set.

HQ-CT responders

The analysis of HQ-CT responders at 3 months will be described in the Psychometric and Responder Analysis plan and results presented in the corresponding report, which will be included as an appendix to the Clinical Study Report.

BMI and BMI Z-Score

BMI will be analysed for patients 18-65 years of age.

BMI Z-scores will be analysed for patients 8-17 years of age.

The analysis will be done using an MMRM analysis to model the percentage change from baseline in BMI and the change from baseline in BMI at visits 3, 4 and 5. The model will include treatment, visit, age group (8-17 years, ≥18 years) and BMI category as fixed effects, treatment-by-visit interaction and corresponding timepoint screening and V2 minus screening BMI as covariates. Unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined in the mixed model to test for treatment differences at the different time points. The differences in least-squares means and corresponding 95% CIs will be presented.

For change from baseline in BMI Z-score, a similar model will be used, but without the inclusion of age group in the model.

The number and proportion of patients with decreases of ≥5% vs. <5% and ≥10% vs. <10% in BMI will be presented descriptively by age group, visit and treatment. Cochran-Mantel-Hansel tests will be performed at each visit to test the treatment effect, controlling for age group.

The change-from-baseline BMI z-score will be descriptively presented, per visit, as n (%) in the following categories:

- >1.0
- 0.6 to 1.0
- 0.1 to 0.5
- -0.5 to -0.0
- -1.0 to -0.6
- <-1.0

As well as in the following categories for the BMI z-score at each visit:

- >4.0
- 3.1 to 4.0
- 2.1 to 3.0
- 1.1 to 2.0
- -1.0 to 1.0
- < -1.0

All analyses will be performed on FAS.

13.3.2.2 Analysis of the Cohort of Patients 4-7 Years of Age and All Age Groups Combined

Descriptive summaries for the additional assessments will be presented for the 4- to 7-year of age-old cohort (See section 13.0).

For CgGIS-H/CGIS-H and CGI-S, a sensitivity analysis including the 4- to 7-year-old cohort will be performed. The MMRM model will be used with treatment, age group (4-7, 8-17, ≥18 years), BMI category, visit, and the interactions treatment-by-age group, treatment-by-visit and treatment-by-visit-by-age group as fixed effects and baseline result as a covariate. Contrasts will be defined to test for treatment differences for each level of age and visit. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented by visit for all age groups combined.

For DBC-P, PedsQL™, ZBI scores, an ANCOVA on the change from baseline in DBC-P, PedsQL™ and ZBI at 3 months will be performed (with no imputation of missing values) in the 4- to 65-year-old patients. The model will include treatment and the stratification variable age group (4-7, 8-17, ≥18 years) and BMI category as fixed effect, the interaction treatment-by-age group, and the corresponding value at baseline as a covariate. Contrasts will be defined to test for treatment differences for each level of age. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented for all age groups combined.

For glucose (fasting), insulin (fasting) and HOMA-IR, the analysis will be done using an MMRM analysis to model the change from baseline glucose/insulin/HOMA-IR at visits 3 and 5. The model will include treatment, age group (4-7, 8-17, ≥18 years), BMI category and visit as fixed effects, the interactions treatment-by-age group, treatment-by-visit and treatment-by-visit-by-age group and corresponding baseline glucose/insulin/HOMA-IR as covariate. Unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined to test for treatment differences for each level of age group and visit. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented by visit for all age groups combined.

For HbA1c, an ANCOVA model similar to the one described for DBC-P, PedsQL™ and ZBI scores will be performed.

In addition, the same models described above on glucose, insulin, HOMA-IR and HbA1c will be performed on patients with EPG, and on patients with prediabetes/T2D.

All analyses performed on insulin and HOMA-IR will not include patients who were on insulin (ATC3=A10A) at V2.

Analyses performed on glucose will also be performed without patients who were on insulin.

For lipids, a similar model as described for glucose will be performed, but with screening and V2 minus screening scores as covariates. The analyses will be done on all patients and on patients with dyslipidemia at baseline.

For lean body mass, a similar model as described in section 13.2.4 for total body fat mass will be performed.

For CgGIC-H/CGIC-H and CGI-I, an MMRM analysis will be used to model the post dose CgGIC-H/CGIC-H and CGI-I at visits 3, 4 and 5. The model will include treatment, visit, age

group (4-7, 8-17, ≥18 years) and BMI category as fixed effects and the interactions treatment-by-age group, treatment-by-visit and treatment-by-visit-by-age group. Unstructured covariance will be used to model the correlation. If convergence issues arise, TOEPH structure will be used. Contrasts will be defined to test for treatment differences for each level of age group and visit. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented by visit for all age groups combined.

For EQ-5D-5L and VAS, descriptive statistics will be presented as described in section 13.3.2.1, for age group 4-7 years and for all ages combined. Moreover, the ANCOVA will be performed on the change from baseline EQ-5D-5L utility index and VAS at 3 months. The model will include treatment, age group (4-7, 8-17, ≥18 years) as fixed effects, treatment-by-age group interaction and the baseline EQ-5D-5L utility index/visual analogue scale as covariate. Contrasts will be defined to test for treatment differences for each level of age. The differences in least-squares means and corresponding 95% CIs will be presented. Additionally, the treatment differences in least-squares means and corresponding 95% CIs will be presented for all age groups combined.

HQ-CT responder analyses, as described in section 13.3.2.1, will be repeated for the age group of 4- to 7-year-olds and for all age groups combined.

For BMI Z-Score (calculated for patients 4-17 years of age), a similar MMRM model as described for glucose will be performed but with screening and V2 minus screening scores as covariates. The factor age group has two categories (4-7 years of age and 8-17 years of age).

14.0 SAFETY OUTCOMES

Safety analyses will be performed on the Safety Analysis Set. Individual patient data listings will be provided for all safety outcomes. The safety summaries will first be presented for the cohort of patients 8-65 years of age and will be repeated for the cohort of patients 4-7 years of age, as well as for all patients (4-65 years of age) combined.

14.1 Adverse Events

An adverse event (AE) is defined as any untoward medical occurrence associated with the use of a drug in humans, whether or not considered drug-related.

Only AEs during the Core Period are considered in this SAP. An AE will be considered as part of the Core Period if its start date is earlier than or the same as the last contact date in the Core Period.

All AEs, regardless of relationship to study drug, should be collected beginning from the time the subject signs the informed consent form until the last study visit or 30 days after the last dose of study drug, whichever is later. Any serious adverse event (SAE) judged by the Investigator to be related to the study treatment should be reported to the Sponsor regardless of the length of time that has passed since study completion. AEs in study patients include any change in the subject's condition. This includes symptoms, physical findings, or clinical syndromes. All AEs and SAEs must be followed to resolution or, if resolution is unlikely, to stabilization.

All AEs will be coded to body system and preferred term using the Medical Dictionary for Regulatory Activities (MedDRA) dictionary, version 22.0.

AEs will be classified as treatment-emergent AEs (TEAEs) or pre-treatment AEs. A TEAE is defined as any event that started on the same day as or after the first dose of study drug. A pre-treatment AE is defined as any event that started prior to the day of first dose of study drug (at or after informed consent date and prior to first dose date). An increase of intensity or frequency of a pre-treatment AE will also be considered as a TEAE. See Appendix 1 for handling of partial dates for AEs. In the case where it is not possible to define an AE as treatment emergent or not, the AE will be classified by the worst case, i.e., treatment emergent. An AE with an unknown/unreported onset date will also be counted as a treatment-emergent AE.

For general presentation of summary tables by System Organ Class (SOC) and Preferred Term (PT), a patient experiencing the same TEAE multiple times will be counted only once at each level of summarization. The summary tables will be sorted alphabetically by SOC and within each SOC the PT will be presented by decreasing order of total frequency.

An overall summary table of AEs will be presented. The number and percentage of patients who experienced TEAEs, related TEAEs, treatment-emergent SAEs, related treatment-emergent SAEs and TEAEs leading to treatment discontinuation will be presented. The number of events will also be presented.

All AEs will be presented in a listing, by treatment group and patient number, and will include the following information: patient identifier, SOC, PT, reported term, severity, seriousness, action taken, outcome, relationship, date of onset, duration, and end date. Separate listings will

be provided for SAEs, TEAEs leading to study drug discontinuation, and TEAEs leading to death.

14.1.1 All TEAEs

The number and percentage of patients who experience TEAEs and the number of events will be presented by SOC and by PT within SOC for each treatment group. TEAEs will also be presented by PT in descending order.

14.1.1.1 Severity

The number and percentage of patients who experience TEAEs will be presented by SOC, PT within SOC, and Severity (mild/moderate/severe) for each treatment group. If a patient experiences more than one TEAE with different severity within the same SOC/PT, only the worst case (worst severity) will be reported. A TEAE with an unknown severity will be considered as severe.

14.1.1.2 Relationship to Study Drug

Related AEs are defined as any AEs assessed by the Investigator as "Definitely Related", "Probably Related" or "Possibly Related". AEs assessed as "Unlikely Related" or "Unrelated" will be defined as not related. Adverse events with an unknown relationship will be considered as related.

The number and percentage of patients who experience TEAEs will be presented by System Organ Class (SOC), Preferred Term (PT) within SOC and Relationship for each treatment group. If a patient experiences more than one TEAE with different relationship within the same SOC/PT, only the TEAE with the greatest degree of relationship to study drug will be reported.

14.1.2 TEAEs Leading to Discontinuation of Study Drug

The number and percentage of patients who experience TEAEs leading to discontinuation of the study drug will be presented by SOC and by PT within SOC for each treatment group.

14.1.3 Serious Adverse Events

The number and percentage of patients who experience treatment-emergent SAEs will be presented by SOC and by PT within SOC for each treatment group. Treatment-emergent SAEs will also be summarized by Severity and Relationship, as described in sections 14.1.1.1 and 14.1.1.2 above.

14.2 Deaths

If any patients die during the study, the information will be presented in a data listing.

14.3 Vital Signs and Height

Descriptive statistics for each vital sign measurement and height will be presented by treatment group and available visit. Change from baseline to each available post-dose visit will also be summarized. These statistics will also be presented for all patients 8-17 years of age, patients 8-17 years of age who are not on growth hormone (WHODrug ATC code=H01AC), and patients 8-17 years of age who are on growth hormone (WHODrug ATC code=H01AC).

14.4 Physical Examination

Physical Exam will be summarized by visit in terms of number and percentage of patients with normal/abnormal (not clinically significant)/abnormal (clinically significant) results per body system and treatment group.

14.5 Bone Mineral Density

Descriptive statistics for bone mineral density measurements will be presented by treatment group and available visit. Change from baseline to each available post-dose visit will also be summarized.

14.6 12-Lead ECG

Descriptive statistics for ECG measurements will be presented by treatment group and available visit and time point. Change from baseline to each available post-dose visit and time point will also be summarized. Overall assessment of ECG will also be summarized by visit and time point in terms of number and percentages of patients with normal/abnormal (not clinically significant)/abnormal (clinically significant) results. For quantitative measures reported in triplicate, the mean of the triplicates will be reported in the tables. For qualitative assessments reported in triplicate, the worst assessment will be reported in the tables. Qualitative assessment will be presented for both central and local readings, but quantitative results will be presented only from central readings.

In addition, abnormal quantitative ECG measurements will be identified and summarized by visit, time point, and treatment group in accordance with the following predefined criteria:

- Absolute values for QT, QTc interval and QTcF will be classified as:

  > 450 msec

> 480 msec
> 500 msec

- Change from Baseline for QT interval, QTc interval and QTcF will be classified as:

  >30 msec increase from baseline
  >60 msec increase from baseline

14.7 Safety Laboratory Parameters

Results of all clinical laboratory assays (except urine pregnancy test) will be performed at central laboratories and will be included in the reporting of this study for Blood Chemistry, Coagulation, Hematology and Urinalysis. A list of laboratory assessments to be included in the outputs is included in Section 8.1.2 of the protocol (Version 1.2, dated 31 July 2019).

Presentations will use both Système International (SI) units and conventional units where these differ. Quantitative laboratory measurements reported as "< X", i.e., below the lower limit of quantification (BLQ), or "> X", i.e., above the upper limit of quantification (ULQ), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e., as "< X" or "> X" in the listings.

Descriptive statistics for each clinical laboratory test will be presented by treatment group and available visit. Change from baseline to each available post-dose visit will also be summarized.

According to the laboratory normal ranges, laboratory test results will be categorized as low (< lower normal limit), normal (within normal range), and high (> upper normal limit). Shift tables for selected laboratory values comparing the distributions of these three categories at baseline versus each available post-dose visit will be presented by treatment group.

The selected laboratory parameters are:

Chemistry

- Serum creatinine (<LLN, normal, >ULN, missing, total)
- AST, ALT (Low or Normal, >1 – 3x ULN, >3 – 5x ULN, >5x ULN, missing, total)
- Total bilirubin (Low or Normal, >1 – 1.5x ULN, >1.5 – 2x ULN, >2x ULN, missing, total)
- Alkaline phosphatase (Low or Normal, >1 – 1.5x ULN, >1.5 – 2x ULN, >2x ULN, Missing, total)

Listings of all clinical laboratory data for each patient will be provided with values outside the normal ranges indicated as well as the change from baseline. A listing of all abnormal laboratory findings will also be provided.

The blood levels for AG, UAG and AG/UAG ratio will be presented descriptively by visit and treatment group. Change from baseline to each available post-dose visit will also be summarized.

Results from anti-livoletide antibodies (ADA: Anti-Drug Antibodies) testing will also be summarized if data are available.

14.8 IGF-1 Evaluation

IGF-1 levels and change in IGF-1 levels will be summarized in patients 8-17 years of age who are:

- On growth hormone (WHODrug ATC code=H01AC), and for which the growth hormone dose did not change between V1 and V5.
- Not on growth hormone.
- For all patients 8-17 years of age.

15.0 Changes from Protocol

- Some subgroup analyses were added: analysis on HbA1c and HOMA-IR in patients with elevated postprandial glucose (EPG), analyses on blood glucose, insulin, HOMA-IR and HbA1c based on prediabetes/T2D patients (as per the information in the medical history), and analysis on lipids in patients with dyslipidemia at baseline.
- IGF-1 evaluation was added in section 14.8 in order to investigate the effect of livoletide on IGF-1 levels in patients on growth hormone.

16.0 REFERENCES

1. McCandless SE, Yanovski JA, Miller J, Fu C, Bird LM, Salehi P, Chan CL, Stafford D, Abuzzahab MJ, Viskochil D, Barlow SE, Angulo M, Myers SE, Whitman BY, Styne D, Roof E, Dykens EM, Scheimann AO, Malloy J, Zhuang D, Taylor K, Hughes TE, Kim DD, Butler MG. Effects of MetAP2 inhibition on hyperphagia and body weight in Prader-Willi syndrome: A randomized, double-blind, placebo-controlled trial. Diabetes Obes Metab. 2017 Dec;19(12):1751-1761.

2. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Final version. 20 November 2019. ICH Expert Working Group.
3. 2000 CDC Growth Charts for the United States: Methods and Development. Data from the National Health Examination Surveys and the National Health and Nutrition Examination Surveys. National Center for Health Statistics and Division of Health Examination Statistics. May 2002.

APPENDIX 1

Algorithm for Treatment Emergence of Adverse Events

If part of the date/time of an AE is missing, but the existing parts allow determination of timing of AE onset/end relative to start and stop date of the study drug, then the AE will be classified (treatment-emergent or not treatment-emergent) per review of the existing parts of the date/time field. If timing of the AE onset/end relative to start and stop date of the study drug cannot be made, then the AE will be assumed to be a treatment-emergent adverse event (TEAE).

Algorithm for Concomitant Medications

For the purposes of determining whether a medication was taken during the treatment period of the study, any missing or partial start or stop dates for which a definitive determination cannot be made will result in that medication being assumed to be concomitant.