NCT #: NCT05098509

A Phase 2/3, Randomized, Double-Blind, Placebo-Controlled Study of RAD011 (Cannabidiol Oral Solution) for the Treatment of Patients with Prader-Willi Syndrome

Protocol Number: SCOUT-015
Protocol Version: 5.0
Protocol Date: 10 March 2022

STATISTICAL ANALYSIS PLAN

Version 1.0

The effective date of the document is the last date of approval of the document

A Phase 2/3, Randomized, Double-Blind, Placebo-Controlled Study of RAD011 (Cannabidiol Oral Solution) for the Treatment of Patients with Prader-Willi Syndrome

STATISTICAL ANALYSIS PLAN
Version 1.0

Author: [redacted]
[redacted] CTI
[redacted]

CTI Biostatistics Reviewer: [redacted]
[redacted] CTI
[redacted]

Sponsor Approval: [redacted]
[redacted]
Radius Health, Inc.
[redacted]

Table of Contents

LIST OF ABBREVIATIONS

Abbreviation	Definition
ABC	Aberrant Behavior Checklist
ABC-I	Aberrant Behavior Checklist - Irritability subscale
AE	Adverse event
AESI	Adverse event of Special Interest
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
ATC	Anatomical, Therapeutic, and Chemical
AUS	Australia
BMI	Body mass index
CBD	Cannabidiol
CaGI-C	Caregiver Global Impression of Change
CaGI-S	Caregiver Global Impression of Severity
CGI-C	Clinician Global Impression of Change
CGI-S	Clinical Global Impression of Change –Severity Scale
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CS	Compound symmetry
CSH	Heterogeneous compound symmetry
C-SSRS	Columbia-Suicide Severity Rating Scale
DEXA	Dual-energy X-ray absorptiometry
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	Electronic case report form
EDC	Electronic data capture
ESS-CHAD	Sleepiness Scale for Children and Adolescents
EU	Europe
FCS	Fully conditional specification
GH	Growth hormone

Abbreviation	Definition
HbA1c	Glycosylated hemoglobin
HQ-CT	Hyperphagia Questionnaire for Clinical Trials
IP	Investigational product
IRT	Interactive Response Technology
ITT	Intent-to-Treat
LS	Least-squares
MAR	Missing at random
MCID	Minimal Clinically Important Difference
MCT	Medium Chain Triglycerides
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
mITT	Modified Intent-to-Treat
MMRM	Mixed Model for Repeated Measures
MNAR	Missing not at random
NA	North America
PP	Per Protocol
PK	Pharmacokinetics
PMM	Predictive means matching
PT	Preferred term
PWS	Prader-Willi Syndrome
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
ULN	Upper limit of normal
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) is based on Protocol SCOUT-015 V5.0, titled "A Phase 2/3, Randomized, Double-Blind, Placebo-Controlled Study of RAD011 (Cannabidiol Oral Solution) for the Treatment of Patients with Prader-Willi Syndrome." See the study protocol for full details.

This document details the statistical methods planned to perform the Data Monitoring Committee (DMC) analysis for Phase 2 part of the study, and the final analysis at the end of Phase 3 part of the study. The DMC will review the safety and tolerability data and recommend 1 or 2 RAD011 dose level(s) for evaluation in the Phase 3 part of the study.

2. OBJECTIVES AND ENDPOINTS

2.1 Objectives

2.1.1 Primary Objective

Phase 2

The primary objective of the Phase 2 part of this study is to assess the safety and tolerability of multiple doses levels of RAD011 in order to select 1 or 2 dose level(s) to further evaluate in the Phase 3 part of the study.

Phase 3

The primary objective* of the Phase 3 part of this study is to assess the effect of RAD011 on hyperphagia-related behavior in patients with Prader-Willi Syndrome (PWS).

*The Phase 3 part of the study will include all patients from the Phase 2 part of the study who were treated with dose level(s) selected for development in the Phase 3 part of the study.

2.1.2 Secondary Objectives

Phase 3

The secondary objectives* for the Phase 3 part of this study are to assess the following:

- Effect of RAD011 on irritability
- Effect of RAD011 on Clinician Global Impression of Change (CGI-C) in Hyperphagia
- Effect of RAD011 on Clinician Global Impression of Severity (CGI-S) in Hyperphagia
- Safety and tolerability of RAD011.

*The Phase 3 part of the study will include all patients from the Phase 2 part of the study who were treated with dose level(s) selected for development in the Phase 3 part of the study.

2.1.3 Other Objectives

Phase 3

The other objectives* for the Phase 3 part of this study are to access the following:

- Effect of RAD011 on Caregiver Global Impression of Change (CaGI-C) in Hyperphagia
- Effect of RAD011 on Caregiver Global Impression of Severity (CaGI-S) of Hyperphagia

- Effect of RAD011 on Caregiver Global Impression of Change (CaGI-C) in Irritability
- Effect of RAD011 on overall behavior
- Effect of RAD011 on sleep
- Effect of RAD011 on body mass index (BMI) and weight
- Effect of RAD011 on skin-picking behavior
- Effect of RAD011 on total muscle/fat composition (performed at selected US sites).

*The Phase 3 part of the study will include all patients from the Phase 2 part of the study who were treated with dose level(s) selected for development in the Phase 3 part of the study.

2.2 Endpoints

2.2.1 Primary Efficacy Endpoint

- Change in Hyperphagia Questionnaire for Clinical Trials (HQ-CT) scores from Baseline through End of Study/Week 34 Visit for RAD011 compared to placebo

2.2.2 Secondary Efficacy Endpoints

The secondary efficacy endpoints include the following:

- Change in PWS-associated Irritability from Baseline through End of Study/Week 34 Visit using the Aberrant Behavior Checklist (ABC) questionnaire – Irritability subscale (ABC-I) for RAD011 compared to placebo
- Change in hyperphagia as defined by the CGI-C in Hyperphagia through End of Study/Week 34 Visit
- Change in severity of hyperphagia as defined by the CGI-S response from Baseline through End of Study/Week 34 Visit

2.2.3 Other Endpoints

The other endpoints include the following:

- Change in hyperphagia as defined by the CaGI-C in Hyperphagia response through End of Study/Week 34 Visit for RAD011 compared to placebo
- Change in CaGI-S of Hyperphagia from Baseline through End of Study/Week 34 Visit
- Change in irritability as defined by the CaGI-C in Irritability through End of Study/Week 34 Visit
- Change in overall behavior from Baseline through End of Study/Week 34 Visit using the ABC questionnaire subscales for RAD011 compared to placebo
- Change in sleep from Baseline through End of Study/Week 34 Visit using the Epworth Sleepiness Scale for Children and Adolescents (ESS-CHAD) for RAD011 compared to placebo
- Change in Weight and BMI from Baseline through End of Study/Week 34 Visit for RAD011 compared to placebo

- Change in skin picking behavior using the Skin Picking Questionnaire from Baseline through End of Study/Week 34 for RAD011 compared to placebo
- Change in muscle/fat composition using a dual-energy X-ray absorptiometry (DEXA) scan from Randomization through End of Study/Week 34 visit for RAD011 compared to placebo

2.2.4 Pharmacokinetic (PK) Endpoints

Throughout the study, plasma samples will be collected to establish the PK profile of cannabidiol (CBD) and 2 metabolites, 7-OH-CBD and 7-COOH-CBD in patients with PWS. Furthermore, an exposure response analysis using CBD and 7-OH-CBD concentrations as well as a population PK model will be derived from the data. Data for the PK Analyses will be collected by ACM, samples analyzed by World Wide Clinical Trials (WWCT) and reported by Allucent. Allucent will provide details of the PK analyses in a separate analysis plan.

2.2.5 Safety Endpoints

The safety endpoints include:

- Treatment-emergent adverse events (TEAEs)
- Vital signs (blood pressure, heart rate, body temperature, respiratory rate, oxygen saturation)
- Electrocardiogram (ECG)
- Columbia-Suicide Severity Rating Scale (C-SSRS)
- Laboratory tests (chemistry, hematology, and urinalysis)

3. INVESTIGATIONAL PLAN

3.1 Study Design

This is a seamless Phase 2/3, double blind, randomized, placebo-controlled clinical study in patients diagnosed with PWS. The study will consist of six similar periods for the Phase 2 and Phase 3 parts of the study:

- Screening Period (3 weeks): Review of eligibility prior to proceeding to the Tolerability Period
- Tolerability Period (6 weeks): Determination of randomization eligibility
- Dose Escalation Period (3 weeks)
- Maintenance Period (24 weeks)
- Taper Period (2 weeks)
- Follow-Up Period (2 weeks).

Patients who do not meet criteria for randomization eligibility will not be randomized to study SCOUT-015. All patients participating in the Maintenance Period may be offered participation in the long-term extension study (SCOUT-016). Patients who do not enroll in the SCOUT-016 study

will have their investigational product (RAD011 or placebo) tapered over 2 weeks (14 days), followed by a 2-week (14-day) safety follow-up.

After approximately 45 patients complete 4 weeks of the Maintenance Period of the Phase 2 part of the study, the DMC will meet to review safety and tolerability data and to recommend 1 or 2 RAD011 dose level(s) for evaluation in the Phase 3 part of the study.

Screening Period (3 weeks, Visit 1 [Weeks -3 to 0]):

Phase 2 and Phase 3 Parts of the study

All patients with a signed informed consent/assent will undergo Screening assessments. Screening assessments will include a full medical history and physical exam, a review of documentation of PWS diagnosis, an ECG, laboratory assessments, and questionnaires (including the HQ-CT) as described in Appendix A Part 1. To be eligible for the Tolerability Period, patients must meet the following criterion:

- Mean HQ-CT score ≥13 obtained from HQ-CT scores as defined in Protocol Section 6.4.1. The EDC system will automatically calculate a mean HQ-CT score from Visit 1 and Visit 2. Patients must have a mean HQ-CT score ≥13 to be eligible to continue in the Tolerability Period. Patients with a mean score <13 will not be eligible to continue in the Tolerability Period and will be discontinued from the study.

Patients not meeting Inclusion/Exclusion criteria defined in Protocol Section 6.1 and 6.2 for entry in the study will not be eligible for the study and may not be screened again. Caregivers will be required either to assist the patient or complete the questionnaires on behalf of the patient. The same caregiver should be available throughout the study to help in completing the questionnaires for purposes of consistency.

Tolerability Period (6 weeks, Visits 2 to 5 [Weeks 1 to 7])

Phase 2 and Phase 3 of the study

Patients who meet the eligibility criteria will undergo a 6-week Tolerability Period. During this Period, all patients will receive placebo at a dose of 0.1 mL/kg/day to assess tolerability of the formulation without active CBD and give patients time to adjust to study procedures and future Investigational Product (IP) administration. Patients unable to tolerate the 0.1 mL/kg/day dose level will be withdrawn from the study.

Patients/caregivers will also complete serial HQ-CT questionnaires at 2-week intervals. At the end of the 6-week Tolerability Period and prior to randomization, patient randomization eligibility based on tolerability and HQ-CT scores will be evaluated. To be eligible for randomization, patients must meet both of the following criteria:

- Mean HQ-CT score ≥13 obtained from prior HQ-CT scores as defined in Protocol Section 6.4.2. The EDC system will automatically calculate a mean HQ-CT score from Visit 3, Visit 4 and Visit 5. Patients must have a mean HQ-CT score ≥13 to be eligible to be randomized and proceed to the Dose Escalation Period. Patients with a mean score <13 will not be eligible for randomization and will not have the opportunity to participate in the long-term extension study (SCOUT-016).
- Decrease from the Tolerability Period to the Screening Period in HQ-CT score ≤7 points as defined in Protocol Section 6.4.2. The EDC system will automatically calculate the change in

HQ-CT score from the Tolerability Period to the Screening Period. This change is calculated as:

Mean HQ-CT score (Visit 3, Visit 4, Visit 5) – Mean HQ-CT score (Visit 1, Visit 2)

Patients with a decrease HQ-CT score ≤7 will be eligible for randomization. Patients with a decrease in HQ-CT score >7 will not be eligible for randomization and unable to proceed to the Dose Escalation Period.

Randomization

Phase 2 of the study

Following the Tolerability Period, patients meeting randomization eligibility criteria will be randomized to 1 of 6 groups (3 active to 3 placebo with the ratio of 2:2:2:1:1:1): low dose (10 mg/kg/day [0.1 mL/kg/day]) RAD011, mid dose (20 mg/kg/day [0.2 mL/kg/day]) RAD011, high dose (40 mg/kg/day [0.4 mL/kg/day]) RAD011, low volume (0.1 mL/kg/day) placebo, mid volume (0.2 mL/kg/day) placebo, or high volume (0.4 mL/kg/day) placebo divided in two daily doses taken with food approximately 12 hours apart.

Phase 3 of the study

Following review of the Phase 2 safety and tolerability data by the DMC, 1 or 2 dose level(s) will be recommended for further development and assessment in Phase 3 of the study.

Dose Escalation Period (3 weeks, Visits 5 to 7 [Weeks 7 to 9])

Phase 2 of the study

After completing the Screening and Tolerability Periods, all randomized patients will be initiated on the IP 10 mg/kg/day (0.1 mL/kg/day) for a period of 1 week at Visit 5 (Week 7). Patients randomized to low dose or low volume IP (10 mg/kg/day) will continue at a dose of 10 mg/kg/day (0.1 mL/kg/day) until the end of Visit 7 (Week 9). Patients randomized to mid dose or mid volume IP (20 mg/kg/day) will have their dose increased to 20 mg/kg/day (0.2 mL/kg/day) at Visit 6 (Week 8) and will continue at that dose until the end of Visit 7 (Week 9). Patients randomized to high dose or high volume IP (40 mg/kg/day) will have their dose increased to 20 mg/kg/day (0.2 mL/kg/day) at Visit 6 (Week 8) and to 40 mg/kg/day (0.4 mL/kg/day) at Visit 7 (Week 9).

The Investigator, or designee, will contact patients/caregivers by phone call, video call, email, or other remote means of communication at Visits 6 (Week 8) and 7 (Week 9) to determine if the patient's dose may be escalated to the next dose level. Dose reductions will not be allowed during the Dose Escalation Period. Patients unable to tolerate their assigned dose level(s) during the Dose Escalation Period will be withdrawn from the study.

Phase 3 of the study

After completing the Screening and Tolerability Periods, all patients will be initiated on 10 mg/kg/day IP (0.1 mL/kg/day) for a period of 1 week at Visit 5 (Week 7). Dose escalation will occur as in the Phase 2 of the study, according to the DMC-recommended Phase 3 dose level(s).

The Investigator, or designee, will contact patients/caregivers by phone call, video call, email, or other remote means of communication at Visits 6 (Week 8) and 7 (Week 9) to determine if the patient's dose may be escalated to the next dose level. Dose reductions will not be allowed during the Dose Escalation Period. If a patient cannot be escalated to the next dose level, the patient will

be discontinued from the study.

Details of the dose escalation schema for Phase 2 and Phase 3 of the study are provided in Table 1:

Table 1. Dose Escalation Regimen

Group for Analysis	**Dosing Group** Target dose	**Visit 5** Week 7	**Visit 6** Week 8	**Visit 7** Week 9
RAD011 low dose	RAD011 low dose (10 mg/kg/day = 0.1 mL/kg/day)	10 mg/kg/day 0.1 mL/kg/day	10 mg/kg/day 0.1 mL/kg/day	10 mg/kg/day 0.1 mL/kg/day
RAD011 mid dose	RAD011 mid dose (20 mg/kg/day = 0.2 mL/kg/day)	10 mg/kg/day 0.1 mL/kg/day	20 mg/kg/day 0.2 mL/kg/day	20 mg/kg/day 0.2 mL/kg/day
RAD011 high dose	RAD011 high dose (40 mg/kg/day = 0.4 mL/kg/day)	10 mg/kg/day 0.1 mL/kg/day	20 mg/kg/day 0.2 mL/kg/day	40 mg/kg/day 0.4 mL/kg/day
Placebo	Placebo low volume (0.1 mL/kg/day)	0.1 mL/kg/day	0.1 mL/kg/day	0.1 mL/kg/day
	Placebo mid volume (0.2 mL/kg/day)	0.1 mL/kg/day	0.2 mL/kg/day	0.2 mL/kg/day
	Placebo high volume (0.4 mL/kg/day)	0.1 mL/kg/day	0.2 mL/kg/day	0.4 mL/kg/day

Maintenance Period (24 weeks, Visits 8 to 14 [Weeks 10 to 34])

Phase 2 and Phase 3 of the study

After completing the 3-week Dose Escalation Period, patients will enter the 24-week Maintenance Period.

Some patients may have their dose adjusted based on tolerability during the Maintenance Period.

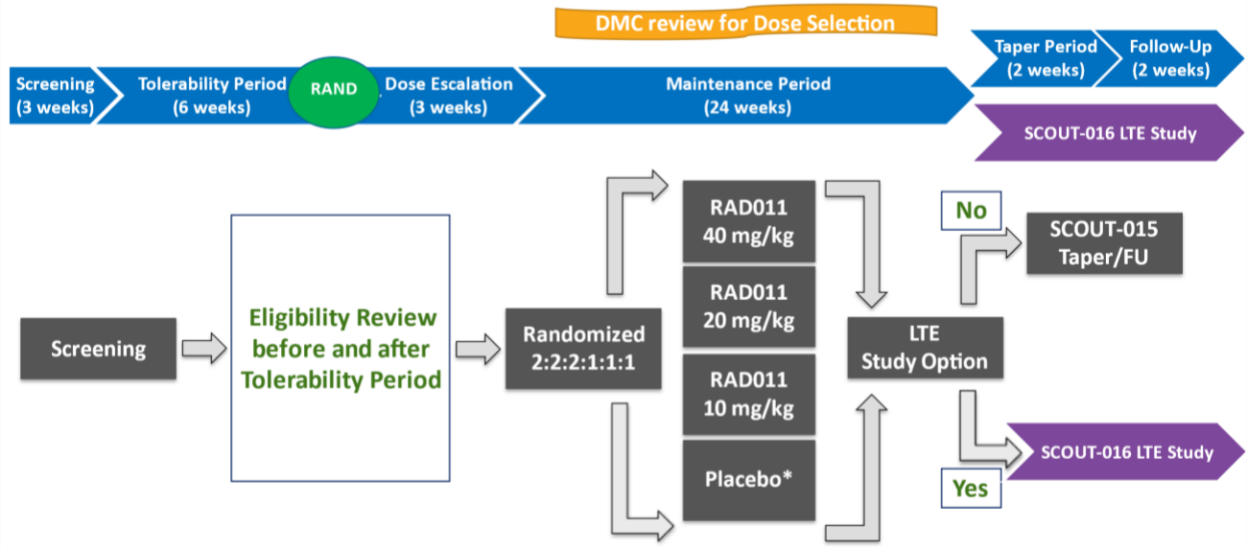
After completion of the Maintenance Period, patients may be offered the opportunity to enroll in the long-term extension study (SCOUT-016). Patients who do not elect to enroll will be tapered off the IP.

Taper Period (14 days, Visit 14 [Week 34]) and Follow-up Period (14 days, Visits 15 and 16 [Weeks 36 and 38])

Patients who do not enroll in the long-term extension study (SCOUT-016) will have their IP tapered off over 14 days according to the following schedule: the IP should be decreased by 50% during the first 7 days, followed by an additional 50% during the following 7 days, and then discontinued. The taper schedule may be modified at the discretion of the Investigator but cannot be extended to a total period lasting more than 14 days. A final follow up will be conducted 2 weeks after the last dose of IP for patients who are tapered off.

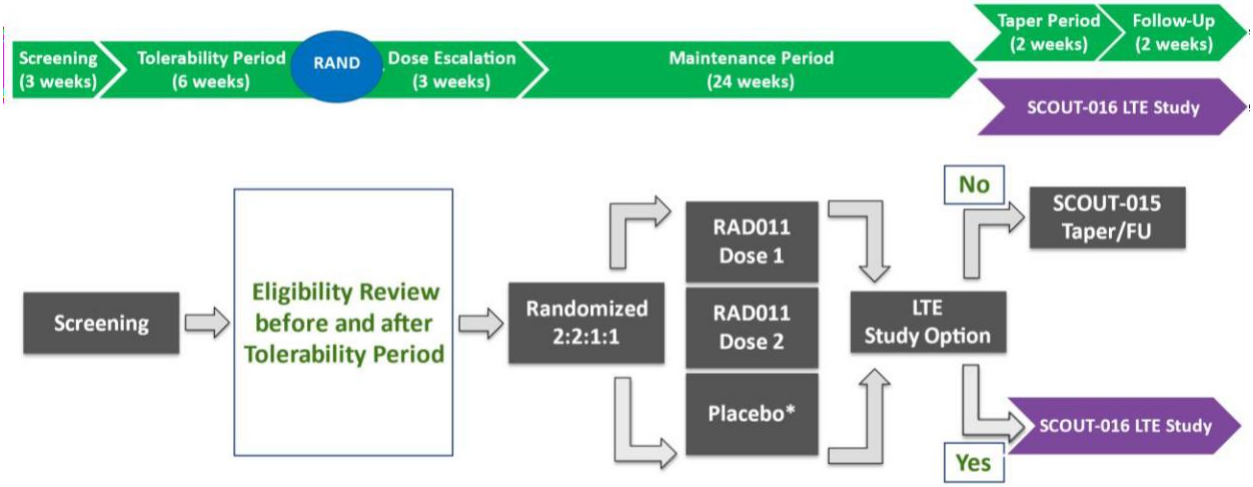
A schematic representation of the Phase 2 and Phase 3 of the study is provided in the figures below:

Study Design Schematic – Phase 2 of the Study



* Patients assigned to placebo will include low, mid, and high volume placebo to match active drug and will be combined for statistical analysis.

Study Design Schematic – Phase 3 of the Study



Abbreviations: FU = follow-up; LTE = long-term extension; RAND = randomization.
* Patients assigned to placebo will be further broken down to achieve balanced exposure between placebo and RAD011 patients. Patients in the placebo group (matching volume) will be combined for statistical analysis. Final doses are to be determined based on the DMC decision. Schematic shows the scenario where 2 doses are chosen for Phase 3.

3.2 Treatment

3.2.1 Randomization Scheme and Treatment Arm Assignment

Phase 2 of the study

Patients who meet eligibility criteria will be randomized in a 2:2:2:1:1:1 manner to receive double-blind treatment with IP, defined as either low dose (10 mg/kg/day, [0.1 mL/kg/day]) RAD011, mid dose (20 mg/kg/day, [0.2 mL/kg/day]) RAD011, high dose (40 mg/kg/day [0.4 mL/kg/day])

RAD011, low volume (0.1 mL/kg/day) placebo, mid volume (0.2 mL/kg/day) placebo, or high volume (0.4 mL/kg/day) placebo using a centralized randomization system. Randomization will be stratified according to current use of Growth Hormone (GH) treatment at the time of randomization (yes/no) and age (≥16 or <16 years old).

A central randomization system, Medidata Rave Randomization & Trial Supply Management (RTSM)®, will be used for this study.

Phase 3 of the study

Following review by the DMC, 1 or 2 dose levels will be recommended for continued development and assessment in Phase 3 of the study. Therefore, following the Tolerability Period, patients meeting randomization eligibility criteria will be randomized using a centralized randomization system RTSM in a way to obtain a balanced number of patients on RAD011 and placebo to receive double-blind treatment with IP at dose levels recommended by the DMC.

If 1 dose is recommended for continued development and assessment in Phase 3 of the study, randomization will be 1:1 (RAD011: placebo).

If 2 doses are recommended for continued development and assessment in the Phase 3 part of the study, randomization will be 2:2:1:1 (RAD011 dose 1: RAD011 dose 2: placebo: placebo).

Patients aged 8 to <12 years may be allowed to participate in the Phase 3 part of the study and will be randomized in a way to obtain a balanced number of patients on RAD011 (at dose levels recommended by the DMC) and placebo. Patients aged 8 to <12 will not be included in the primary efficacy analysis.

3.2.2 Blinding

Both Investigators and patients will be blinded to the randomization assigned treatment. During the Tolerability Period, all patients will receive 0.1 mL/kg/day of placebo. To maintain the blind, multiple placebo groups were created to mimic the RAD011 groups: low volume (0.1 mL/kg/day), mid volume (0.2 mL/kg/day), and high volume (0.4 mL/kg/day). During the dose escalation period, the volume of blinded therapy for the mid volume and high volume placebo groups will increase in the same manner as the mid dose and high dose RAD001 groups, further ensuring blinding between RAD011 and placebo.

The matching (appearance, smell and taste) placebo and RAD011 product will utilize the same medium chain triglycerides (MCT) vehicle.

The assigned unblinded statistician will have access to unblinded data for DMC meeting purposes.

Breaking the treatment blind for a patient should be done only in the event of a medical emergency where the identity of the IP is necessary to appropriately treat the patient. The Investigator may unblind the treatment received by the patient through the interactive response technology (IRT). The IRT will automatically document and record any such unblinding and notify the Sponsor Medical Monitor of the unblinding event. If possible, the Sponsor and/or designee should be contacted prior to unblinding of the treatment. The study monitor will not be apprised of the actual treatment assignment.

3.2.3 Dosing Schedule

The IP (RAD011 or placebo) will be administered orally twice daily with food, approximately 12

hours apart. Patients must take the IP with food throughout the study. All patients will receive a dose on a mg/kg/day basis, and dose reductions will be allowed per the details below.

1. During Dose Escalation Period, dose reductions for tolerability are not allowed. Patients unable to tolerate their assigned dose level(s) during the Dose Escalation Period will be withdrawn from the study.
2. In the event that a patient does not tolerate their assigned dose level on a mg/kg/day basis during the Maintenance Period, the Investigator may decrease the patient's dose by 25% at any time during the Maintenance Period after consultation with the Sponsor as shown in Table 2. Patients undergoing a dose reduction may not have their dose increased back to their originally assigned dose. If a second dose reduction is required for tolerability during the Maintenance Period, the patient should be discontinued from the study in consultation with the Sponsor.

Table 2. Dose Level (mg/kg/day) Reduction for Tolerability during the Maintenance Period

Dose Prior to Performing Dose Reduction	Dose Following Dose Reduction (25% Decrease)
10 mg/kg/day 0.1 mL/kg/day	7.5 mg/kg/day 0.075 mL/kg/day
20 mg/kg/day 0.2 mL/kg/day	15 mg/kg/day 0.15 mL/kg/day
40 mg/kg/day 0.4 mL/kg/day	30 mg/kg/day 0.3 mL/kg/day

During the Maintenance Period, all patients will be dosed using their actual body weight at Visit 5 (Week 7), and this dose will be maintained through the course of the study.

3.2.4 Patient Withdrawal and Study Termination

The Investigator should withdraw the patient from the study if they believe that participation is no longer in the patient's best interest, or in the event of study termination by the Sponsor. If the patient elects to withdraw from the study, reason(s) for withdrawal should be documented. The Investigator will discuss with the patient procedures for withdrawal and any additional care or treatment alternatives available at that time. In addition, patients unable to tolerate the 0.1 mL/kg/day dose level during the Tolerability Period or their assigned dose level(s) during the Dose Escalation Period will be withdrawn from the study.

If a patient withdraws consent/assent from the study, the patient should be encouraged to complete all End of Study Visit assessments (Visit 14, Week 34) completed as described in the Schedule of Assessments (Appendix A, Part 2). At time of withdrawal of consent/assent, the patient will be withdrawn from all additional study procedures and follow-up.

Data collected until patient withdrawal or lost to follow-up will be retained and included in the analysis of the study in accordance with local rules and regulations.

4. GENERAL CONSIDERATIONS FOR DATA ANALYSIS

Unless otherwise specified, continuous variables will be summarized using descriptive statistics (the number of non-missing observations [n], mean, standard deviation [SD], median, minimum, and maximum). Categorical variables will be summarized by showing the number and percentage of patients within each category.

Summary results will be provided for the RAD011 and placebo groups. The low, mid, and high volume placebo groups will be combined for summary analysis for the purposes of the DMC analysis. If 2 doses are selected for the Phase 3 part of the study, the 2 corresponding volume placebo groups will be combined for all Phase 3 statistical analyses and the other placebo group will be presented in its own column. All tabulations will be based on pooled data across centers.

If the DMC recommends two RAD011 doses for Phase 3 of the study, all statistical comparisons will be performed using two-sided tests at an $\alpha = 0.025$ significance level for each RAD011 dose group and placebo comparison. If the DMC recommends one RAD011 dose for Phase 3 of the study, all statistical comparisons will be performed using two-sided tests at an $\alpha = 0.05$ significance level.

All p-values will be displayed to four decimal places, with p-values less than 0.0001 presented as '<0.0001'. Minimum and maximum values will be reported with the same number of decimal places as in the units of collection with 3 decimals being maximum value; the mean will be presented with 1 decimal place more and the standard deviation 2 decimal places more than the units of collection.

All analyses will be performed using SAS for Windows statistical software, version 9.4 or higher (SAS, Cary, NC), except where other software may be deemed more appropriate.

CTI will perform all efficacy and safety analyses described in this SAP. PK samples will be collected by ACM, samples analyzed by WWCT, and analyzed and reported by Allucent. The PK analysis methods will be covered in a separate analysis plan provided by Allucent.

Data with qualifiers (e.g., "<") will be listed with the qualifier but summarized without the qualifier, unless otherwise specified.

Patient data will be listed, sorted by treatment group (RAD011 followed by placebo), investigational center, and patient number. Demographic listing will also include Study Part enrolled (Phase 2/3). When applicable, listings will be additionally sorted by nominal visit and assessment date/time.

4.1 Data Quality Control Procedures

Once all the source verification is complete, all queries are resolved, and the database has been updated appropriately, the database will be locked and made available to CTI Biostatistics for final analysis.

Safety data will be pulled by CTI Biostatistics for the DMC analysis at a time when source verification and query resolution is ongoing, however, all critical data needed for DMC analysis will be entered and fully source data verified for this data-cut.

4.2 Analysis Populations

The following populations will be defined for analysis:

1. Intent-to-Treat (ITT) Population: the ITT Population will include all randomized patients.
2. Modified ITT (mITT) Population: the mITT Population will include all patients aged ≥ 12 years who were randomized, received at least 1 dose of IP, had at least 1 post-randomization HQ-CT questionnaire completed, and were randomized to receive the dose level(s) recommended for development in Phase 3 part of the study, including patients in the Phase 2 part of the study who received the chosen dose level(s).
3. Safety Population: the Safety Population will include all randomized patients who were treated with at least 1 dose of IP.
4. Per Protocol (PP) Population: the PP Population will include all patients aged ≥12 years who were randomized, received at least 1 dose of IP, completed Week 34 without significant protocol deviations that were prospectively defined to impact efficacy and received the dose level(s) recommended for development in the Phase 3 part of the study, including patients in Phase 2 of the study who received the chosen dose level(s).

   Prior to database lock and unblinding, patients who had a significant protocol deviation leading to exclusion from the PP population will be identified in a blinded manner by the study team.
5. PK Evaluable Population: the PK Evaluable Population will include all patients who received at least 1 dose of IP and underwent at least 1 PK sample collection.

4.3 Assessment Windows

Data will be summarized by nominal study visit recorded in the database except the early termination visit which will be re-assigned to an applicable visit for summary purpose.

To re-assign the early termination visit, the protocol specified visit windows will be extended to the midpoint between visits so that there is no gap between the visits (Table 3). If a re-assigned early termination visit falls into an existing scheduled visit where data already exists, the data collected from the scheduled visit will be used for the summary analysis. The re-mapped early termination visit will be presented in the summary tables only if there is no data already existing for the visit. The early termination visit recorded in the database together with the calculated study days will be displayed in the data listings.

Table 3. Early Termination Visit Re-mapping Windows

Visit Number	Study Time Point	Minimum Window Day	Maximum Window Day
5	Day 43	40	46
6	Day 50	47	53
7	Day 57	54	60
8	Day 64	61	81
9	Day 99	82	113
10	Day 127	114	141

11	Day 155	142	169
12	Day 183	170	193
13	Day 204	194	218
14	Day 232	219	235

The analysis of exposure and compliance by analysis period and the safety analyses on TEAEs based on the adverse event (AE) onset date during an analysis period are described as below in Table 4.

Table 4. Analysis Period

Analysis Period	Protocol Defined Study Windows (Start – Stop)	Window Start Time	Window Stop Time
Tolerability	Day 1 - Day 42	1st placebo dosing date	Date at Visit 5 - 1
Dose Escalation	Day 43 - Day 63	Date at Visit 5	Date at Visit 8 - 1
Maintenance	Day 64 - Day 232	Date at Visit 8	Date at Visit 14
Taper*	Day 233 - Day 246	Date at Visit 14 + 1	Date at Visit 15
Follow-Up*	Day 247 - Day 260	Date at Visit 15+1	Date at Visit 16

* Taper and Follow-up periods only apply to the patients who do not participate in long term extension study (SCOUT-016).

4.4 Handling of Dropouts or Missing Data

Missing data on demographic, baseline information and safety will be treated as missing. Patient-level listings will present data as reported. Missing or partially missing dates that are required for date-dependent definitions (e.g., treatment-emergent adverse events, concomitant medications) will be assumed to be the most conservative date possible. See Appendix B for the details of imputation rules.

Details of any missing data imputation for primary and secondary efficacy analyses will be specified in Section 6.1 and 6.2. No imputations will be performed for the other efficacy endpoints.

4.5 Multiple Comparisons

A hierarchical, serial gatekeeper testing strategy will be used to protect the overall two-sided significance level of $\alpha = 0.05$ when testing the primary and secondary efficacy outcome measures.

DMC Recommendation of two RAD011 Doses

The primary efficacy analysis will be performed using two-sided tests at an $\alpha = 0.025$ significance level for each RAD011 dose group and placebo comparison (RAD011 low dose vs. placebo and RAD011 high dose vs. placebo) so overall Type 1 error is maintained at two-sided 0.05 level.

The primary endpoint for each RAD011 dose group and placebo comparison will serve as a

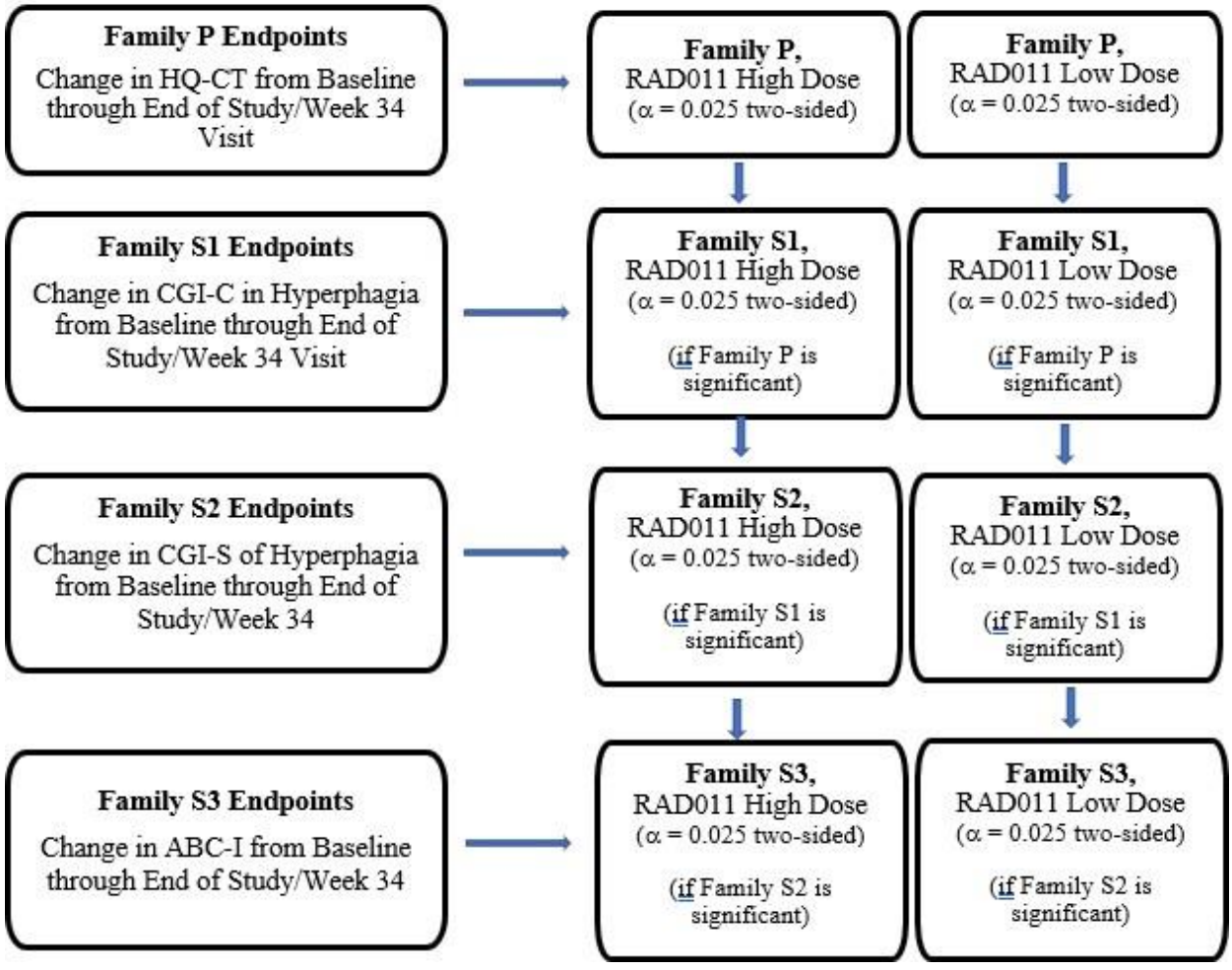
gatekeeper for the secondary efficacy endpoints with the same dose. Thus, if and only if the change in HQ-CT score from Baseline to End of Study/Week 34 visit is significantly different from placebo in a RAD011 dose group, then statistical significance will also be assessed for the secondary efficacy endpoints for the same dose group comparison.

The secondary endpoints will be tested sequentially in a hierarchical order as specified below to protect the overall significance level at 5% (two-sided):

1. Change in hyperphagia as defined by the CGI-C in Hyperphagia response through End of Study/Week 34 Visit
2. Change in severity of hyperphagia as defined by the CGI-S of Hyperphagia response from Baseline through End of Study/Week 34 Visit
3. Change in ABC-I from Baseline through End of Study/Week 34 Visit

Error! Reference source not found. presents the complete gatekeeping strategy for this study. Each family from the secondary endpoints will be tested at $\alpha = 0.025$ level (two-sided) provided all analyses in the preceding families were significant within the same dose.

Figure 2 Serial Gatekeeper Strategy



If the DMC recommends one RAD011 dose for Phase 3 of the study, the same gatekeeper strategy

described above applies to both primary and secondary efficacy analyses using the two-sided $\alpha = 0.05$ significance level for the comparison between RAD011 and placebo.

No adjustment for multiplicity will be performed for testing the other endpoints.

4.6 Data Derivations and Transformations

The following derivations will be used in this study:

The day of first IP administration in Tolerability Period at Visit 2 is defined as Day 1. All other study days will be presented relative to Day 1.

Study Day:

- Date of assessment – date of first IP administration + 1 for assessments done on or after Day 1
- Date of assessment – date of first IP administration for assessments done prior to Day 1

Baseline Observation: Unless otherwise specified, the baseline value is the last non-missing value prior to the first IP administration.

Baseline HQ-CT is defined as the mean HQ-CT score from Visit 1 and Visit 2.

Baseline DEXA scan is defined as the assessment done at Visit 5.

Duration:

- Duration in days = end date – start date + 1

5. STUDY PATIENTS

5.1 Disposition of Patients

Patient disposition will be listed and summarized for all screened patients. The frequency counts and percentages of all patients in the ITT, mITT, Safety, and PP Population will be provided. Patient disposition including study completion status and reasons for early termination, together with whether they continue into the long-term extension study (SCOUT-016) will be tabulated by treatment group and overall.

5.2 Protocol Deviations

The number and percentage of patients in the ITT Population that have minor, major and significant protocol deviations, including the type of deviation within each category, will be tabulated for the treatment groups. Significant protocol deviations may include randomized patients that didn't meet I/E criteria, patients who developed withdrawal criteria but were not withdrawn, patients where the patient/site staff where unproperly unblinded, and received wrong treatment, are a subset of major deviations, and are defined as affecting efficacy or safety conclusions of the study. Significant protocol deviations will be identified prior to unbinding.

A listing of all protocol deviations will be provided.

5.3 Demographic and Baseline Characteristics

Demographic and Baseline characteristics will be listed and summarized descriptively for the ITT

Population. The demographic and baseline characteristics will consist of age and age group (≥16 years, ≥12 and ≤15 years, and ≥8 and ≤11 years), gender, race, ethnicity, region (NA, EU, AUS), baseline height, baseline weight, baseline Body Mass Index [BMI], calculated age-normalized baseline BMI z-scores for patients aged 8 to 19 years inclusively, growth hormone use (Yes/No) at randomization, and whether experienced any seizure during past 12 months (Yes/No).

The summary analysis will be repeated for the mITT, Per Protocol, and Safety populations. The analysis on Safety population will be based on the treatment actually received. If a patient was randomized to placebo but received an active kit by accident, then the patient's "actual" treatment would be the corresponding RAD011 volume (e.g. if randomized to mid volume placebo but receives active kit, then actual treatment is mid dose RAD011).

In addition, the seizure history, coagulation, and serology results will be presented in a listing.

5.4 Medical History

All medical conditions and surgical procedures will be classified by system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA v24.1 or later). The number and percentage of patients with each medical condition and surgical procedure will be presented for each SOC and PT for the ITT population.

5.5 Prior and Concomitant Medications

All prior and concomitant medications collected will be coded using the World Health Organization (WHO) Drug Dictionary version September 2021 B3 Global WHO-Drug or later. The number and percentage of patients in the ITT population using prior and concomitant medications will be tabulated by Anatomical, Therapeutic, and Chemical (ATC) level 2 and preferred name.

Prior medications will be summarized separately from concomitant medications.

Prior medications are defined as medications that ended prior to the date of first IP administration. Concomitant medications are defined as medications that started at any time but ended on or after the date of first IP administration, including those that are ongoing at study completion. In the case of a missing or partial end date/time, in order to determine whether a medication is prior or concomitant, the imputation rules described in Appendix B will be used. Note, the imputed date/time will only be used to determine whether a medication is prior or concomitant. The actual date/time reported on the eCRFs will be presented in the listings.

6. EFFICACY ANALYSES

6.1 Primary Efficacy Endpoint and Analysis

The primary efficacy endpoint for this study is the change in HQ- CT scores from Baseline through End of Study/Week 34 Visit for RAD011 compared to placebo in the mITT population based on the treatment policy estimand.

The HQ-CT questionnaire consists of 9 items, with a 2-week recall period. Each question is rated on a scale of 0 to 4, providing a maximum score of 36. The HQ-CT total score will be used to demonstrate improvements in hyperphagia-related behavior (extreme uncontrollable hunger) between RAD011 and placebo groups. An HQ-CT score of 13 is associated with moderate to

severe hyperphagia[1] (Fehnel 2015).

After approximately 45 patients complete 4 weeks of the Maintenance Period of the Phase 2 part of the study, the DMC will meet to review safety and tolerability data and to determine 1 or 2 RAD011 dose level(s) for evaluation in the Phase 3 part of the study. If DMC recommends 2 RAD011 doses, the primary comparison will be mean change in HQ-CT as calculated via least square means comparing RAD011 low dose vs. placebo and RAD011 high dose vs. placebo at Week 34. Low dose and high dose groups will be compared to placebo at the two-sided 0.025 level so overall type 1 error is maintained at two-sided 0.05 level. If the DMC recommendation is one RAD011 dose, the primary comparison will be mean change in HQ-CT as calculated via least square means comparing RAD011 vs. placebo at Week 34 at the two-sided 0.05 level.

Change in HQ- CT total score from Baseline through End of Study/Week 34 Visit will be evaluated by fitting a linear mixed model for repeated measures (MMRM). The model will include change from baseline in HQ-CT score as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline HQ-CT score as a covariate. Patient will be included as a random effect.

An unstructured covariance matrix will be used to model the within-subject error. Parameters will be estimated using restricted maximum likelihood with the Kenward-Roger method for calculating the denominator degrees of freedom. In case the model will not converge with the unstructured covariance structure, the heterogeneous compound symmetry (CSH) will be used instead. Furthermore, if the model fails to converge with CSH, the compound symmetry (CS) structure will be used. In each case of CSH and CS, the sandwich estimator (SAS option: EMIPRICAL) will be used with the SAS option BETWITHIN for the denominator degrees of freedom. Data collected at Baseline (Visits 1 and 2), Weeks 10, 19, 27, and/or 34 will be included in the analysis. Baseline is defined as the mean HQ-CT score from Visit 1 and Visit 2.

For the primary endpoint analysis at Week 34, multiple imputation (MI) will be used to assign a value to those cases with missing data based on the missing at random (MAR) assumption. The SAS fully conditional specification (FCS) method with predictive means matching (PMM) as described in Berglund & Heeringa [2] will be used. This method uses all of a patient's known primary outcome measures at Baseline, Weeks 10, 19, 27, and/or 34 weeks to impute any missing values through Week 34. The model will include randomization treatment assignment and stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]). The SAS FCS REGPMM method will be used in the PROC MI. The MI procedure will be repeated 50 times and each derived imputation will be analyzed separately using the MMRM as described above, and then pooled with PROC MIANALYZE.

The output from the MIANALYZE procedure including the least-squares (LS) means and 97.5% or 95% confidence intervals (CIs) for each treatment group based on number of doses chosen for Phase 3, the treatment difference in the LS means, the 97.5% or 95% CI for the difference based on number of doses chosen for Phase 3, and the associated p-value will be reported on Weeks 10, 19, 27 and 34. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase 3, then 95% CIs will be presented.

6.1.1 Sensitivity Analyses for the Primary Efficacy Endpoint

The primary endpoint analysis will be repeated using the PP population if different from the mITT population.

Descriptive statistics on primary endpoint for patients aged 8 to <12 may be calculated if the sample size for this age group warrants a meaningful summary analysis. Otherwise, a listing for this age group would be sufficient.

To explore the uncertainty in parameter estimation due to missing data, the following sensitivity analyses will be performed using the mITT population:

- Repeat the MMRM analysis without imputation
- RAD011 patients who discontinued from the study due to an AE will be treated as missing not at random (MNAR) and the imputation for these patients' missing data will be based on the observed data in the placebo group. This can be done by assigning these patients to placebo in the SAS MI procedure so that the missing data imputed for these patients will be based on the placebo group. After generating a set of imputation datasets, these patients will be assigned back to RAD011 for the MMRM analysis.
- Repeat the primary endpoint analysis using the actual stratifications, if ≥10% are different from the randomized stratifications.
- An additional sensitivity analysis using the tipping-point approach will be conducted to assess the robustness of the primary analysis approach only if the primary analysis is significant. This is particularly important if a significant treatment effect is found from the primary analysis based on the MAR assumption. In which case, the validity of the result should be examined to explore the impact of MNAR on the outcome.

  Tipping point analysis is a means of exploring the influence of missingness on the overall conclusion from statistical inference by positing a wide spectrum of assumptions regarding the missingness mechanism (from less conservative to more conservative). The analysis finds a (tipping) point in this spectrum of assumptions, at which conclusions change from being favorable to the experimental treatment to being unfavorable. After such a tipping point is determined, clinical judgment can be applied as to the plausibility of the assumptions underlying this tipping point.

  The tipping point imputation analysis will use a specified sequence of shift parameters (k), which adjust the imputed values for observations in the active treatment group. This can be done in the SAS MI procedure with the MNAR statement using the imputed datasets from the MAR described above in Section 6.1. Each derived imputation from a shift parameter will be analyzed separately using the MMRM method, and then pooled with PROC MIANALYZE. A p-value for the treatment effect will be produced for each value of k. The smallest value of k where the treatment effect is no longer significant is identified as the tipping point.

6.1.2 Subgroup Analyses for the Primary Efficacy Endpoint

Subgroup analyses by randomized age group stratification (>=16; <16 years), randomized GH group stratification (Yes/No), gender, Race (White; non-White), Region (NA, EU, AUS), Ethnicity (Hispanic or Latino; Not Hispanic or Latino), patients with DEXA performed and BMI

group (normal/overweight/obese based on age and race defined categorization in Table 5) will be performed on the primary efficacy endpoint using the method specified in Section 6.1. The imputed dataset created from the primary analysis will be used rather than re-imputing based on the subgroups. The results for Week 34 will be presented graphically using Forest Plot with LS Means and 97.5% or 95% CI based on number of doses chosen for Phase 3 on X-axis and subgroups on Y-axis. The BMI group will be defined based on the following categories in Table 5:

Table 5. BMI (kg/m^2) Categories and Ranges

Age (yrs)	Normal Weight (kg/m^2)	Overweight (kg/m^2)	Obese (kg/m^2)
≥20, Non-Asian	18.5-<25.0	25-<30.0	≥30
≥20, Asian	18.5-<23.0	23-<27.5	≥27.5
≤19	≥5th percentile and <85th percentile	≥85th percentile and <95th percentile	≥95th percentile

6.1.3 Clinical Meaningfulness of Primary Efficacy Endpoint

Separate plots of the cumulative distribution function of the change from baseline in HQ-CT to Week 34 for all patients in the mITT Population in each treatment group used in the primary efficacy analysis for the CaGI-S of Hyperphagia and CaGI-C in Hyperphagia at Week 34. The change from baseline will be calculated based on the data collected in the database without imputation. The minimal clinically important difference (MCID) for the change from baseline in HQ-CT at Week 34 will be examined using anchor-based methods for patients meeting the following criteria:

- Improved vs. Not Improved based on the CaGI-C in Hyperphagia where Improved will be defined as any patient with a CaGI-C in Hyperphagia response of 'Much better', 'Moderately better', or 'A little better' and Not Improved will be defined as a response of 'No difference', 'A little worse', 'Moderately worse', or 'Much worse'.
- No Severity vs. Any Severity based on the CaGI-S of Hyperphagia where No Severity will be defined as having a CaGI-S response of 'None' and Any Severity will be defined as having a response of 'Mild', 'Moderate', 'Severe', or 'Very Severe'.

The mean change from baseline in HQ-CT at Week 34 will be presented for all patients in each group (Improved/Not Improved or No Severity/Any Severity) along with the difference in means between the categories in each questionnaire. The primary assessment of clinical meaningfulness will be based on the mean change from baseline in HQ-CT for all patients in the Improved group (ie all patients with a score of 'Much better', 'Moderately better', or 'A littler better').

In addition, a responder analysis will be performed based on the cut-off point using the mean change from baseline in HQ-CT at Week 34 within the Improved group. The proportion of patients with change from baseline values at Week 34 ≥ the cut-off point in each RAD011 group will be compared to placebo using the CMH method, stratified by age category at randomization

[>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic will be reported.

6.2 Secondary Efficacy Endpoints and Analyses

The following secondary endpoints comparing RAD011 with placebo will be analyzed using a hierarchical testing, serial gatekeeping strategy as described above in Section 4.5 based on the mITT population. The fixed-sequence test procedure will be performed in the order specified as below:

- Change in hyperphagia as defined by the CGI-C in Hyperphagia through End of Study/Week 34 Visit
- Change in severity of hyperphagia as defined by the CGI-S of Hyperphagia response from Baseline through End of Study/Week 34 Visit
- Change in PWS-associated Irritability from Baseline through End of Study/Week 34 Visit using the ABC-I subscale for RAD011 compared to placebo

These three secondary efficacy endpoints will be tested for the same dose only if the primary efficacy endpoint for that dose meets statistical significance (α=0.025 if 2 RAD011 dose levels continue into Phase 3; α=0.05 if 1 RAD011 dose level continues into Phase 3), thus maintaining the overall significance level at two-sided 5%. If the first secondary endpoint for a dose level fails to meet significance (α=0.025 if 2 RAD011 dose levels; α=0.05 if 1 RAD011 dose level), then the subsequent analyses will be considered descriptive, and the statistical tests will provide only nominal p-values.

As a sensitivity analysis, each of the secondary endpoints will also be analyzed using the PP population, assuming there are differences between the mITT and PP populations.

Change in hyperphagia as defined by CGI-C in Hyperphagia through End of Study/Week 34 Visit

The CGI-C of hyperphagia is a single-item, clinician-rated measure, assessing the clinician's impression about changes in the patient's hyperphagia condition since the start of taking the study medication at the initiation of the Tolerability Period. The CGI-C of hyperphagia utilizes a 5-point response scale: 1=Much better; 2=A little better; 3=No change; 4=A little worse; 5=Much worse. Data will be collected at End of Study/Week 34 visit.

The analysis of CGI-C in hyperphagia assessed at Week 34 will be conducted using the Cochran-Mantel-Haenszel (CMH) method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic, together with the descriptive summary (frequency count and percentage) for each CGI-C response and for categories of Improved (response scales of 1 and 2), No change (response scale of 3), and Worse (response scales of 4 and 5) will be presented. Missing data will remain as missing, and no imputation is planned.

In addition, the 5-point response scale will be dichotomized into two categories (responder and non-responder). The responder category will include the response scales of 1 and 2, and the non-responder category will include the response scales of 3 to 5. The binary results will be analyzed using the CMH method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic will be

reported and will serve as the primary test of the change in CGI-C in Hyperphagia.

Change in CGI-S of Hyperphagia Scale from Baseline through End of Study/Week 34 Visit

The CGI-S of hyperphagia is a single-item, clinician-rated measure, assessing the clinician's impression of the severity of a patient's hyperphagia condition. The CGI-S of hyperphagia utilizes a 5-point response scale: 1=None; 2=Mild; 3=Moderate; 4=Severe; 5=Very Severe.

The analysis of change in CGI-S of hyperphagia from Baseline through Week 34 will be conducted using the CMH method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic at each post-Baseline visit (Weeks 10, 19, 27, and/or 34), together with the descriptive summary (frequency count and percentage) for each CGI-S response at each visit will be presented. Missing data will remain as missing, and no imputation is planned.

Change in ABC-I subscale score from Baseline through End of Study/Week 34 Visit

The ABC questionnaire is an informant-rated questionnaire assessing severity of behavioral symptoms and is one of the few empirically developed scales designed to measure psychiatric symptoms and behavioral disturbance exhibited by individuals with developmental disabilities (intellectual disability, autism spectrum disorder, cerebral palsy, epilepsy).

Overall, the ABC contains 58 items divided in five subscales: (a) irritability and agitation (15 items), (b) lethargy and social withdrawal (16 items), (c) stereotypic behavior (7 items), (d) hyperactivity and noncompliance (16 items), and (e) inappropriate speech (4 items).

Each item is scored as 0 (never a problem), 1 (slight problem), 2 (moderately serious problem), or 3 (severe problem). If a caregiver is required to assist the patient in completing the questionnaire, the same caregiver should be available throughout the study.

The ABC-I (irritability and agitation) subscale score will not be calculated if ≥4 items are missing in that subscale. In this case, the missing ABC-I subscale score will remain missing, and no imputation is planned for this analysis. If there are less than 4 missing items in the irritability subscale, then the subscale score will be calculated as follows: (# of items in the subscale/# of completed items) * the sum of the completed items for that subscale.

Change in ABC-I subscale score from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM. The model will include change from baseline in ABC-I subscale score as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline ABC-I subscale score as a covariate. Patient will be included as a random effect.

An unstructured covariance matrix will be used to model the within-subject error. Parameters will be estimated using restricted maximum likelihood with the Kenward-Roger method for calculating the denominator degrees of freedom. In case the model will not converge with the unstructured covariance structure, the CSH will be used instead. Furthermore, if the model fails to converge with the CSH, the CS structure will be used. In each case of CSH and CS, the sandwich estimator (SAS option: EMIPRICAL) will be used with the SAS option BETWITHIN for the denominator degrees of freedom. ABC-I subscale scores collected at Baseline (Visit 1), Weeks 10, 19, 27, and/or 34 will be included in the analysis.

From this model, the LS means and CIs for each treatment group, the treatment difference in the

LS means, the CI for the difference, and the associated p-value will be reported on Weeks 10, 19, 27 and 34. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase 3, then 95% CIs will be presented.

A plot of the cumulative distribution function of the change from baseline in ABC-I to Week 34 for all patients in the mITT Population in each treatment group used in the primary efficacy analysis for the CaGI-C in Irritability at Week 34. The MCID will be examined using anchor-based methods based on the mean change from baseline in ABC-I at Week 34 will be compared for patients meeting the following criteria:

- Improved vs. Not Improved based on the CaGI-C in Irritability where Improved will be defined as any patient with a CaGI-C response of 'Much better', 'Moderately better', or 'A little better', and Not Improved will be defined as a response of 'No difference', 'A little worse', 'Moderately worse', or 'Much worse'.

The mean change from baseline in ABC-I to Week 34 will be presented for each group of patients (Improved/Not Improved) along with the difference in means between the categories. Additionally, the clinical meaningfulness will be assessed on a distribution-based approach in accordance with the ABC-2 Community/Residential Manual which states "we regard half-standard deviation changes on the ABC subscales to be meaningful".

6.3 Other Efficacy Endpoints and Analyses

The following other efficacy endpoints comparing RAD011 with placebo will be analyzed based on the mITT population. Missing data from all other efficacy endpoints will remain as missing and no imputations are planned.

Change in hyperphagia as defined by CaGI-C in Hyperphagia response through End of Study/Week 34 Visit

The CaGI-C is a single-item, caregiver-rated measure, assessing the caregiver's impression about changes in the patient's food-related behavior compared to before the beginning of the trial. The CaGI-C utilizes a 7-point response scale (1=Much better; 2=Moderately better; 3=A little better; 4=No difference; 5=A little worse; 6=Moderately worse; 7=Much worse). Data will be collected at End of Study/Week 34 visit.

The analysis of CaGI-C assessed at Week 34 will be conducted using the CMH method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic, together with the descriptive summary (frequency count and percentage) for each CaGI-C response and for categories of Improved (response scales of 1 to 3), No change (response scale of 4), and Worse (response scales of 5 to 7) will be presented.

In addition, the 7-point response scale will be dichotomized into two categories (Improved and Not Improved). The Improved category will include the response scales of 1 to 3; and the Not Improved category will include the response scales of 4 to 7. The binary results will be analyzed using the CMH method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic will be reported.

Change in CaGI-S of Hyperphagia from Baseline through End of Study/Week 34 Visit

The CaGI-S of hyperphagia is a single-item, caregiver-rated measure, assessing the caregiver's impression of the severity of a patient's hyperphagia condition. The CaGI-S of hyperphagia utilizes a 5-point response scale: 1=None; 2=Mild; 3=Moderate; 4=Severe; 5=Very Severe.

The analysis of change in CaGI-S of hyperphagia from baseline through Week 34 will be conducted using the CMH method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic at each post-Baseline visit (Weeks 10, 19, 27, and/or 34), together with the descriptive summary (frequency count and percentage) for each CaGI-S response at each visit will be presented.

Change in irritability as defined by CaGI-C in Irritability

The CaGI-C of irritability is a single-item, caregiver-rated measure, assessing the caregiver's impression about changes in the patient's irritability condition since the start of taking the study medication at the initiation of the Tolerability Period. The CaGI-C of irritability utilizes a 7-point response scale: 1=Much better; 2=Moderately better; 3=A little better; 4=No difference; 5=A little worse; 6=Moderately worse; 7=Much worse. Data will be collected at End of Study/Week 34 visit.

The analysis of CaGI-C in irritability assessed at Week 34 will be conducted using the CMH method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic, together with the descriptive summary (frequency count and percentage) for each CaGI-C response and for categories of Improved (response scales of 1 to 3), No change (response scale of 4), and Worse (response scales of 5 to 7) will be presented.

In addition, the 7-point response scale will be dichotomized into two categories (Improved and Not Improved). The Improved category will include the response scales of 1 to 3; and the Not Improved category will include the response scales of 4 to 7. The binary results will be analyzed using the CMH method, stratified by age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]. The p-value for row mean scores differ statistic will be reported.

Change in overall behavior from Baseline through End of Study/Week 34 Visit using the ABC questionnaire subscales

The ABC questionnaire was detailed in Section 6.2 above for the analysis of irritability subscale scores. The list below outlines the scenarios when the score cannot be calculated for the other subscales:

- For Lethargy/social withdrawal and Hyperactivity/Noncompliance subscales, the score will not be calculated if ≥4 items are missing in that subscale.
- For Stereotypic behavior subscale, the score will not be calculated if ≥3 items are missing in that subscale.
- For Inappropriate speech subscale, the score will not be calculated if ≥2 items are missing in that subscale.

In these cases, the missing ABC subscale scores will remain missing, and no imputation is planned for this analysis. If missing items are present in a subscale, but the score can be calculated, then the subscale score will be calculated as follows: (# of items in the subscale/# of completed items)

* the sum of the completed items for that subscale.

Change in each of the ABC subscale scores from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM. For each subscale, the model will include change from baseline in the subscale score as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline ABC subscale score as a covariate. Patient will be included as a random effect.

An unstructured covariance matrix will be used to model the within-subject error. Parameters will be estimated using restricted maximum likelihood with the Kenward-Roger method for calculating the denominator degrees of freedom. In case the model will not converge with the unstructured covariance structure, the CSH will be used instead. Furthermore, if the model fails to converge with CSH, the CS structure will be used. In each case of CSH and CS, the sandwich estimator (SAS option: EMIPRICAL) will be used with the SAS option BETWITHIN for the denominator degrees of freedom. ABC subscale scores collected at Baseline (Visit 1), Weeks 10, 19, 27, and/or 34 will be included in the analysis.

From this model, the LS means and CIs for each treatment group, the treatment difference in the LS means, the CI for the difference, and the associated p-value will be reported on Weeks 10, 19, 27 and 34. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase 3, then 95% CIs will be presented.

Change in sleep from Baseline through End of Study/Week 34 Visit using the ESS-CHAD

The ESS-CHAD is an 8-question scale with a 1-month recall period that has been validated in school-attending children and adolescents to evaluate daytime sleepiness. Each question is rated on a 0 to 3 scale, with a maximum score of 24, and a score above 10 is considered to represent mild excessive daytime sleepiness

Change in ESS-CHAD score from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM. The model will include change from baseline in ESS-CHAD score as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline ESS-CHAD score as a covariate. Patient will be included as a random effect.

An unstructured covariance matrix will be used to model the within-subject error. Parameters will be estimated using restricted maximum likelihood with the Kenward-Roger method for calculating the denominator degrees of freedom. In case the model will not converge with the unstructured covariance structure, the CSH will be used instead. Furthermore, if the model fails to converge with CSH, the CS structure will be used. In each case of CSH and CS, the sandwich estimator (SAS option: EMIPRICAL) will be used with the SAS option BETWITHIN for the denominator degrees of freedom. ESS-CHAD score collected at Baseline (Visit 2), Weeks 19 and/or 34 will be included in the analysis.

From this model, the LS means and CIs for each treatment group, the treatment difference in the LS means, the CI for the difference, and the associated p-value will be reported on Weeks 19 and 34. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase 3, then 95% CIs will be presented.

Change in weight and BMI from Baseline through End of Study/Week 34 Visit

Change in age-normalized BMI z-scores [3] will be used for patients aged 8 to 19 years inclusively, whereas percent change in weight will be used for patients aged 20 to 65 years.

The BMI Z-scores, also called BMI standard deviation scores, are measures of relative weight adjusted for child age and sex. BMI Z-score can be determined based on the age, sex, BMI, and the CDC reference standard. The CDC growth charts contain 10 smoothed percentiles (between the 3rd and 97th) of BMI for 24- to 240-month-old children and adolescents. These smoothed estimates were subsequently used to derive lambda (L, the power transformation to achieve normality), mu (M, mean or median) and sigma (S, coefficient of variation) parameters for Cole's LMS method [4, 5]. Estimates of these 3 parameters allow the BMI of any children to be expressed as a z-score and percentile relative to children of the same sex and age in the CDC growth charts. To calculate the z-score for any BMI value (BMIz) with the LMS method, the following formula is used:

$$BMIz = [(\frac{BMI}{M})^{L} - 1] \div (L \times S)$$

Values for the L, M, and S are available on the CDC website for each sex and month of age.

For BMI values at or above the 95th percentile, the BMI z-score will be re-calculated using the methods described in Rong Wei's paper [3], in which the half normal distributions are used to model the distribution of BMI ≥ 95th percentile. These values combined with the CDC BMI z-scores (below 95th percentile) will be used for the analysis.

For mITT patients aged ≥20 years, the percent change in weight from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM. The model will include percent change from baseline in weight as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variable (GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline weight as a covariate. Patient will be included as a random effect.

In addition, for mITT patients aged ≥18 years, the percent change in BMI from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM as described above for the analysis of percent change in weight.

For mITT patients aged <20 years, the change in age-normalized BMI z-scores from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM. The model will include change from baseline in BMI z-score as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variable (GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline BMI z-score as a covariate. Patient will be included as a random effect.

For both MMRM analyses, an unstructured covariance matrix will be used to model the within-subject error. Parameters will be estimated using restricted maximum likelihood with the Kenward-Roger method for calculating the denominator degrees of freedom. In case the model will not converge with the unstructured covariance structure, the CSH will be used instead. Furthermore, if the model fails to converge with CSH, the CS structure will be used. In each case of CSH and CS, the sandwich estimator (SAS option: EMIPRICAL) will be used with the SAS option BETWITHIN for the denominator degrees of freedom. Weight, BMI and BMI z-scores

collected at Baseline (Visit 2), Weeks 10, 19, 27 and/or 34 will be included in the analysis.

From this model, the LS means and CIs for each treatment group, the treatment difference in the LS means, the CI for the difference, and the associated p-value will be reported on Weeks 10, 19, 27 and 34. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase 3, then 95% CIs will be presented.

Change in skin picking behavior using the Skin Picking Questionnaire from Baseline through End of Study/Week 34 Visit

The Skin Picking questionnaire is a 6-item questionnaire of skin picking symptoms, with scores from 0 to 24. Each item utilizes a 4-point scale with 0 indicating none or minimal and 4 indicating extreme. A score of ≥7 is considered highly suspicious for a skin picking disorder.

Change in skin picking score from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM. The model will include change from baseline in skin picking score as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline skin picking score as a covariate. Patient will be included as a random effect.

An unstructured covariance matrix will be used to model the within-subject error. Parameters will be estimated using restricted maximum likelihood with the Kenward-Roger method for calculating the denominator degrees of freedom. In case the model will not converge with the unstructured covariance structure, the CSH will be used instead. Furthermore, if the model fails to converge with CSH, the CS structure will be used. In each case of CSH and CS, the sandwich estimator (SAS option: EMIPRICAL) will be used with the SAS option BETWITHIN for the denominator degrees of freedom. Skin picking score collected at Baseline (Visit 2), Weeks 19 and/or 34 will be included in the analysis.

From this model, the LS means and CIs for each treatment group, the treatment difference in the LS means, the CI for the difference, and the associated p-value will be reported on Weeks 19 and 34. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase 3, then 95% CIs will be presented.

Change in muscle/fat composition using a DEXA scan from Randomization through End of Study/Week 34 Visit

The assessment of change in muscle/fat composition using DEXA scan will be performed at a selected number of US sites only. Results of the DEXA scan will be reported as total mass, fat mass, lean mass, percentage body fat, and bone mineral content. Data will be collected at randomization (Visit 5) and End of Study/Week 34 visit.

The analysis of change from baseline in each DEXA scan parameter at Week 34 will be performed using an analysis of covariance (ANCOVA) model. The model will include change from baseline in each DEXA scan parameter as the response variable; randomization treatment assignment (as a class variable) and stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]) as fixed effects; DEXA scan parameter at baseline as a covariate. From this model, the LS means and CIs for each treatment group, the treatment difference in the LS means, the CI for the difference, and the associated p-value will be presented. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase

3, then 95% CIs will be presented.

Change in Appetite Hormones from Baseline through End of Study/Week 34 Visit

Appetite hormones including insulin and leptin as well as HbA1c and insulin resistance will be collected at the time points presented in Appendix A. Insulin resistance will be calculated as insulin (mU/L) x glucose (nmol/L)/22.5. Change in each parameter from Baseline through End of Study/Week 34 Visit will be evaluated by using MMRM. The model will include change from baseline in each parameter as the response variable; randomization treatment assignment (as a class variable), visit (as a class variable), stratification variables (age category at randomization [>=16 or <16 years] and GH treatment at randomization [yes/no]), and treatment-by-visit interaction as fixed effects; baseline parameter value as a covariate. Patient will be included as a random effect.

An unstructured covariance matrix will be used to model the within-subject error. Parameters will be estimated using restricted maximum likelihood with the Kenward-Roger method for calculating the denominator degrees of freedom. In case the model will not converge with the unstructured covariance structure, the CSH will be used instead. Furthermore, if the model fails to converge with CSH, the CS structure will be used. In each case of CSH and CS, the sandwich estimator (SAS option: EMIPRICAL) will be used with the SAS option BETWITHIN for the denominator degrees of freedom. Appetite hormones collected at Baseline (Visit 1), Weeks 10, 27, and/or 34 will be included in the analysis.

From this model, the LS means and CIs for each treatment group, the treatment difference in the LS means, the CI for the difference, and the associated p-value will be reported on Weeks 10, 27, and 34. If 2 doses are chosen for Phase 3, then 97.5% CIs will be presented. If 1 dose is chosen for Phase 3, then 95% CIs will be presented.

7. SAFETY ANALYSIS

Safety assessments will include assessment of AEs, C-SSRS, vital signs, ECGs, and laboratory assessments. All safety analyses will be conducted using the Safety Population. No formal hypothesis testing will be performed to compare differences between treatment groups.

7.1 Extent of Exposure and Compliance

The duration of exposure (days), total dose administered (mg/kg) will be summarized by period: Tolerability, Dose Escalation, Maintenance, and Taper periods. Additionally, the number of patients with a dose level reduction during Maintenance Period will be presented.

Duration of exposure will be calculated as:

> For each period (defined in Table 4 for start and stop dates): Duration of exposure (days) = The last study drug administration date - The first study drug administration date – 1

Total dose administered (mg/kg) will be calculated as:

> For each period: Total dose administered (mg/kg) = [(sum of dispensed weights in grams – sum of returned weights in grams)* (2800/69.5)]/patient weight in kg
>
> Note: 1) The dispensed weight and returned weight are based on the “Study Drug Accountability” eCRF for the total initial kit weight and returned kit weight respectively.

Bottles that are not returned will be assumed to have not been used. 2) For calculation during the Tolerability period, the patient weight (kg) will be based on the assessment on Visit 2; for all other periods, the patient weight (kg) will be based on the assessment on Visit 5.

In addition, the treatment compliance (%) will be calculated as follows and summarized for the Maintenance Period. Additionally, number of patients with treatment compliance in each category of <80%, ≥80% to ≤120%, and >120% will be presented.

Treatment compliance (%) = ([Total dose administered (mg/kg)]/(Total Expected Dose (mg/kg)) x 100.

Note:

1) Total dose administered (mg/kg) = [(sum of dispensed weights in grams – sum of returned weights in grams) * (2800/69.5)]/ patient weight in kg
2) Total expected dose (mg/kg) = Treatment duration in days * Dose Level of 10/20/40 on patient level, where treatment duration is based on the patient actual treatment days in the Maintenance period. If a patient has a dose reduction, then the 25% dose reduction will be considered for the expected dose calculations by using Total expected dose (mg/kg) = Treatment duration in the Maintenance period prior to dose reduction * dose Level of 10/20/40 plus remaining treatment duration during the Maintenance Period * reduced dose level; the start date of dose reduction is the visit date when the first dose reduction is recorded as "Yes" on the Study Drug Dispensation eCRF page during the maintenance period.

Details of IP administration based on the information collected on the "Study Drug Accountability" eCRF page will be listed and will include: the kit number, date dispensed, total initial kit weight (g), date of kit returned, returned kit weight (g), number of bottles returned, and any comments.

A separate listing will be provided to include treatment start and stop date, the duration of exposure (days), total dose administered (mg/kg), the visit the dose level was reduced during Maintenance Period, and treatment compliance for each treatment period.

7.2 Adverse Events

An AE is any untoward medical occurrence in a clinical investigation patient administered a pharmaceutical product, which does not necessarily have a causal relationship with the treatment. This includes any newly occurring event which occurs during the study, having been absent at Screening, or any previous condition, if present at Screening, that appears to worsen (i.e., increased in severity or frequency) after the administration of study drug.

All AEs will be collected from the time of consent/assent until the following time points:

- For patients who are not randomized: until time of screen failure
- For randomized patients:
    - For patients continuing in the long-term extension study SCOUT-016: through Visit 14 (Week 34) inclusively
    - For patients entering the follow-up period with IP tapering: through Visit 16 (Week

38).

Investigators are not obligated to actively seek AEs or serious adverse events (SAEs) after the participant has concluded study participation. However, if the investigator learns of any SAE, including a death, at any time after a participant has completed the study or 30 days after the last dose of IP, and he/she considers the event to be reasonably related to the study intervention, the investigator must promptly report the SAE.

7.2.1 Treatment-emergent Adverse Events

A treatment-emergent adverse event (TEAE) is defined as an AE that first occurs or worsens in intensity after the first dose of study drug. For adverse events with missing or partially missing start/stop dates, the imputation of missing dates will be performed as described in Appendix B.

7.2.2 Adverse Event Intensity

For both SAEs and non-SAEs, the Investigator must determine both the intensity of the event and the relationship of the event to study drug administration.

Intensity for each AE will be defined according to the following criteria as noted in Table 6. If the intensity of an AE changes within a day, the maximum intensity should be recorded. If the intensity changes over a longer period of time, the changes should be recorded as separate events (having separate onset and stop dates for each intensity).

Table 6: Intensity Classification and Definition of Adverse Events

Intensity	Definition
Mild	Awareness of sign or symptom, but easily tolerated
Moderate	Discomfort enough to cause interference with normal daily activities
Severe	Inability to perform normal daily activities

7.2.3 Adverse Event Relationship to Study Drug

Every effort should be made by the Investigator to assess the relationship (causality) of the AE, if any, to the IP.

Investigators should use their knowledge of the patient, the circumstances surrounding the event, and an evaluation of any potential alternative causes to determine whether or not an AE is considered to be related to study drug, indicating on the eCRF that an AE is either Related or Not Related.

7.2.4 Serious Adverse Events

An SAE is defined as any adverse experience that suggests a significant hazard, contraindication, side effect, or untoward medical occurrence that, in the view of the Investigator or Sponsor, results in any of the following criteria:

1. Death
2. Is immediately life threatening
3. Requires in-patient hospitalization or prolongation of existing hospitalization

4. Results in persistent or significant disability or incapacity
5. Results in a congenital abnormality or birth defect
6. Other medically important event

7.2.5 Adverse Events of Special Interest (AESI)

An AESI is an AE that is designated to be of special medical or scientific interest to the Sponsor. All AEs within the category of AESIs should be reported promptly, following the reporting procedures for abnormal laboratory, AEs or SAEs as appropriate.

- Suicide ideation and behavior AESI

Included in this AESI are AEs following administration of RAD011 that are assessed by the Investigator to be any of the following: suicidal ideation, suicidal behavior, suicide attempt, and completed suicide.

- Hepatic dysfunction AESI

Included in this AESI are abnormal laboratory findings or AEs suggestive of hepatic dysfunction following administration of RAD011 such as greater or moderate elevation of liver enzymes (AST, ALT, ALP, gamma-glutamyl transferase [GGT]) and bilirubin; or AEs of right upper quadrant abdominal pain, chronic fatigue, loss of appetite, unexplained nausea and vomiting, yellowing of skin or eyes, itching, dark urine, and light-colored stools.

The preferred terms for AESI are listed in Appendix C.

7.2.6 Adverse Event Summaries

All AEs (serious and non-serious) occurring after the first dose of study drug administration and before the end of study, regardless of relationship to study drug, will be included and classified by SOC and PT using MedDRA, v24.1 or later.

For treatment-emergent AEs (TEAEs), the following will be summarized and presented for the Safety Population:

i. An overall summary of TEAEs will be provided for treatment period AEs defined as AEs with onset dates during Dose Escalation, Maintenance, Taper, and Follow-up periods. The table will include:
    a. the number and percentage of patients experiencing a TEAE
    b. the number and percentage of patients experiencing a TEAE by strongest relationship to study drug
    c. the number and percentage of patients experiencing a TEAE by maximum severity
    d. the number and percentage of patients experiencing a SAE during the study
    e. the number and percentage of patients experiencing a treatment-emergent SAE (TESAE)
    f. the number and percentage of patients experiencing a TEAE leading to study drug discontinuation
    g. the number of percentages of patients experiencing an AE leading to death during

the study

h. the number and percentage of patients experiencing a TEAE leading to death

i. the number of percentage of patients experiencing an AESI

ii. the number and percentage of patients and number of events for patients experiencing a TEAE by SOC and PT for the treatment period and by period (Tolerability, Dose Escalation, Maintenance, Taper, and Follow-up)

iii. the number and percentage of patients and number of events for patients experiencing a related TEAE by SOC and PT for Dose Escalation and Maintenance periods each separately

iv. the number and percentage of patients experiencing a TEAE by SOC, PT and the strongest relationship to study drug for Dose Escalation and Maintenance periods each separately

v. the number and percentage of patients experiencing a TEAE by SOC and PT and the maximum severity for Dose Escalation and Maintenance periods each separately

vi. the number and percentage of patients experiencing a TESAE by SOC and PT for the treatment period, Dose Escalation, and Maintenance periods each separately

vii. the number and percentage of patients experiencing a TEAE leading to study drug discontinuation by SOC and PT for the treatment period, Dose Escalation, and Maintenance periods each separately

viii. the number and percentage of patients experiencing an AESI by SOC and PT for the treatment period

In the overall summary of TEAEs table (i), besides tabulating the number and percentage of patients, the total number of events will also be provided. If a patient has repeated episodes of a particular TEAE, all events will be counted in the summary table. In the remaining summary tables, the incidence of TEAEs will be calculated by dividing the number of patients who have experienced the event by the total number of patients in the Safety Population. Thus, the incidence of TEAEs is shown in terms of the total number of patients and not in terms of the total number of events. If a patient has repeated events of a particular TEAE, only the most severe episode, or the event with the strongest causal relationship to study drug, will be counted in the summary tables. A patient with more than one type of TEAE in a particular SOC or PT will be counted only once in the total of patients experiencing TEAEs in that particular SOC or PT.

All occurrences of all AEs and SAEs will be listed for each patient, grouped by treatment group. The listing will contain the following information: treatment group, verbatim term, SOC, PT, severity, relationship to study drug, date and day of onset, date and day of resolution, action taken with regard to study drug, the outcome, whether the event was an SAE, whether it led to study discontinuation, and whether it is a TEAE. Listings will be sorted by treatment group, patient identification number, onset date, SOC, and PT. If the onset date is completely missing, then these events will be presented first. If the onset date is missing a month or a day, then these events will be presented before any complete dates.

A separate listing will be presented for AESI.

Percent of patients by AE Onset Date in each period experiencing an AE with a PT with ≥10% incidence will be presented in a line graph. The X axis will represent the AE onset period in the

Tolerability, Dose Escalation, Maintenance, Taper, and Follow-up periods. The Y axis will represent the percentage of patients within each onset period by treatment group. Each treatment group will be distinguished by different color.

7.3 Clinical Laboratory Assessments

Clinical laboratory tests will be obtained at the time points presented in Appendix A. For continuous laboratory variables summary statistics of the actual values and change from Baseline values will be presented by treatment group at each treatment visit separated by panel (chemistry, hematology, and urinalysis). In addition, summary statistics for the actual value and change from Baseline are to be presented by treatment group for the last visit, minimum, and maximum post-Baseline values during the defined treatment period. Shifts from Baseline for the last visit, minimum, and maximum post-Baseline values during the defined treatment period (Visit 5 through Visit 16) according to normal range criteria will also be presented. Repeated or unscheduled laboratory assessments during the study will not be presented in by-visit summaries, but will be considered when determining the last visit, minimum, and maximum post-Baseline values during the defined treatment period.

Marked laboratory abnormalities are detailed in Appendix D for selected hematology parameters. The number and percentage of patients with treatment-emergent marked abnormalities (Low/High) will be presented in a summary table for each lab parameter specified in Appendix D by visit and treatment group. Summary for the last visit including unscheduled assessments will also be presented.

The number of patients meeting the criteria for Hy's Law along with number of patients meeting the individual lab thresholds (ALT ≥ 3×Upper limit of normal [ULN], AST ≥ 3×ULN, or total bilirubin ≥ 2×ULN) will be summarized. An eDISH plot will be generated for maximum total bilirubin (Y-axis) vs. maximum ALT or AST (x-axis). The criteria of Hy's Law are as follows:

- (ALT or AST ≥ 3×ULN) and total bilirubin ≥ 2×ULN

To meet the above criteria, a patient must experience the elevation in total bilirubin and ALT or AST at the same visit. For example, a patient who experiences a >=2 x ULN elevation of total bilirubin at one visit and a 3 x ULN elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's Law criteria. A patient with ALT and AST values missing or a patient with total bilirubin value missing has not fulfilled the Hy's Law criteria

All laboratory data from all (scheduled or unscheduled) visits will be listed.

7.4 Vital Signs

Descriptive summaries of the vital signs (both actual and change from baseline values) including systolic and diastolic blood pressure, heart rate, respiratory rate, oxygen saturation, and temperature will be prepared for each study treatment by visit.

Unscheduled visits will be excluded from the summaries but will be included in the data listings.

Marked abnormalities for vital signs are shown as below in Table 7.

Table 7 Marked Vital Signs Abnormalities

Parameter	Abnormality Criteria: Low	Abnormality Criteria: High
Systolic blood pressure (mmHg)	<=90 and >=20 decrease from baseline	>=180 and >=20 increase from baseline
Diastolic blood pressure (mmHg)	<=50 and >=15 decrease from baseline	>=105 and >=15 increase from baseline
Heart rate (beats/min)	<=50 and >=15 decrease from baseline	>=120 and >=15 increase from baseline

The number and percentage of patients with treatment-emergent marked abnormalities (Low/High) for systolic/diastolic blood pressure and heart rate will be presented in a summary table by visit and treatment group. Summary for the last visit including unscheduled assessments will also be presented.

7.5 ECG

Electrocardiogram measurements (RR interval, PR interval, QRS duration, QT interval, and QTcF interval) will be summarized by presenting descriptive statistics of actual data and change from baseline values at each visit. The findings (normal; abnormal, not clinically significant; Abnormal, clinically significant) will be included in the data listing.

A summary table will be presented for the number and percentage of patients with QTcF >450 (ms), >480 (ms), >500 (ms), as well as with change from baseline in QTcF >30 ms and >60 ms by visit. The summary for change from baseline in QTcF >30 ms and >60 ms will also be presented for any post-baseline visit.

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

The C-SSRS is a broadly validated scale aimed at assessing suicidal risk in a variety of children and adult populations. The scale includes suggested questions to solicit the type of information needed to determine if suicidal ideation and/or behavior occurred. The first time the C-SSRS questionnaire is administered, on Visit 5 (Week 7), the Baseline/Screening questionnaire [6] should be used, and the "Since Last Visit" questionnaire should be used for all subsequent assessments.

The standard C-SSRS scale should be used for patients aged 6 and over and the Children's C-SSRS scale should be used in patients aged less than 6 years. However, the Children's version may be used in patients aged 6 years or older when the Investigator or designee determines that the patient is unable to understand the first couple of questions in the standard version. The same C-SSRS scale should be used for a given patient throughout the study.

The C-SSRS outcomes include the following categories and have binary responses (yes/no) [7]:

1. Wish to be Dead
2. Non-specific Active Suicidal Thoughts
3. Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
4. Active Suicidal Ideation with Some Intent to Act, without Specific Plan

5. Active Suicidal Ideation with Specific Plan and Intent
6. Preparatory acts or Behavior
7. Aborted Attempt
8. Interrupted Attempt
9. Actual Attempt (non-fatal)
10. Completed Suicide
11. Self-Injurious Behavior (Non-suicidal and/or intent unknown)

The following outcome is a numerical score to assess the intensity of ideation:

- Suicidal Ideation Score: The maximum suicidal ideation category (1-5 on the C-SSRS, with 1 being the least severe and 5 being the most severe) present at the assessment. A score of 0 will be assigned if no ideation is present.

The number and percentage of patients with an event in each of the 11 categories listed above will be presented by visit.

A shift table will be provided for change in C-SSRS categories from Baseline (Tolerability) to the End of Study/Week 34 by visit. The C-SSRS categories for this shift table will include: No suicidal ideation or behavior, Suicidal Ideation, and Suicidal Behavior. Suicidal Ideation includes any one of the five suicidal ideation events (Categories 1-5 as shown above). Suicidal behavior includes any one of the six suicidal behavior events (Categories 6-11 as shown above). Each patient will be counted only once in one of the 3 categories, so patients who have both Suicidal Ideation and Suicidal Behavior are included in the Suicidal Behavior category.

In addition, a separate shift table for change in maximum C-SSRS suicidal ideation scores (0-5) from Baseline to End of Study/Week 34 by visit will also be presented.

A by-patient data listing will be provided.

7.7 Other Safety Measures

Data listings will be provided for food diary and patient impression of seizures (No change; Became better; Became worse). In addition, a descriptive summary of patient impression of seizures will be provided.

8. INTERIM ANALYSES

After approximately 45 patients complete 4 weeks of the Maintenance Period of the Phase 2 part of the study, an interim analysis for DMC will be performed. The DMC will meet to review the safety and tolerability data and recommend 1 or 2 RAD011 dose level(s) for evaluation in the Phase 3 part of the study.

The DMC will also be available to assist in the review of aggregate safety events as requested by the Sponsor at any time during the study.

9. SAMPLE SIZE AND POWER CALCULATIONS

The DMC will recommend one RAD011 Dose or two RAD011 Doses after Phase 2 part of the study. The sample size and power calculation will vary based on the DMC recommendation.

DMC Recommendation of Two RAD011 Doses

For the primary efficacy analysis, a total of 150 patients (50 patients per treatment group) aged 12 to 65 years will provide 90% power to compare each dose group vs. placebo with a two-sided alpha of 0.025. The study is powered to detect a difference in the HQ-CT from Baseline to Week 34 of 5 points with a standard deviation of 7 points. The 0.05 Type 1 error is split equally between groups (low dose and high dose). The low and high volume placebo groups will be combined for all comparisons to placebo. Assuming a 15% post-randomization dropout rate, approximately 180 patients (~60 patients per treatment group) will be evaluated for the primary efficacy analysis.

Thus, if the DMC recommends two RAD011 doses for Phase 3 of the study, approximately a total of 191 patients aged 12 to 65 years (approximate target of 200 patients) will be randomized in this study. The 191 reflects 176 patients required for primary efficacy analysis plus 15 patients randomized in the Phase 2 part of the study to the dose not selected for Phase 3 of the study.

Assuming a 20% pre-randomization eligibility failure rate (including HQ-CT eligibility criteria), approximately 240 patients will be enrolled into the Screening Period. An additional 10 patients per treatment group aged 8 to <12 years may be allowed to participate in Phase 3 of the study and will be randomized in a way to obtain a balanced number of patients on RAD011 (at dose levels recommended by the DMC) and placebo. Patients aged 8 to <12 will not be included in the primary efficacy analysis.

DMC Recommendation of One RAD011 Dose

For the primary efficacy analysis, a total of 100 patients (50 patients per treatment group) aged 12 to 65 years will provide 94% power to compare active vs. placebo with a two-sided alpha of 0.05. The study is powered to detect a difference in the HQ-CT from Baseline to Week 34 of 5 points with a standard deviation of 7 points. Assuming a 15% post-randomization dropout rate, approximately 118 patients (~59 patients per treatment group) will be evaluated for the primary efficacy analysis.

Thus, if the DMC recommends one RAD011 dose for Phase 3 of the study, approximately 148 patients aged 12 to 65 years (target approximately 150 patients) will be randomized in this study.

The 148 reflects 118 patients required for primary efficacy analysis plus 30 patients randomized in the Phase 2 part of the study to the two doses not selected for Phase 3 of the study.

Assuming a 20% pre-randomization eligibility failure rate (including HQ-CT eligibility criteria), approximately 185 patients will be enrolled into the Screening Period. An additional 10 patients per treatment group aged 8 to <12 years may be allowed to participate in Phase 3 of the study and will be randomized in a way to obtain a balanced number of patients on RAD011 (at dose level recommended by the DMC) and placebo. Patients aged 8 to <12 will not be included in the primary efficacy analysis.

10. CHANGE FROM PROTOCOL

Protocol:	Change:
Coagulation analysis is included in the Safety and Tolerability endpoints section for laboratory tests.	Removed the summary analysis for coagulation since it is only collected at Screening. Only listing will be provided.

11. REFERENCES

1. Fehnel S, Brown TM, Nelson L, Chen A, Roof E, Kim DD, et al. Development Of The Hyperphagia Questionnaire For Use In Prader-Willi Syndrome Clinical Trials. Race/ethnicity. 2015;6:100.0.
2. Berglund P, Heeringa SG. Multiple imputation of missing data using SAS: SAS Institute; 2014.
3. Wei R, Ogden CL, Parsons VL, Freedman DS, Hales CM. A method for calculating BMI z-scores and percentiles above the 95(th) percentile of the CDC growth charts. Ann Hum Biol. 2020; 47:514-21.
4. Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. Stat Med. 1992; 11:1305–19.
5. Cole TJ. The LMS method for constructing normalized growth standards. Eur J Clin Nutr. 1990/01/01. 1990; 44:45–60.
6. Posner K, Brent D, Lucas C, Gould M, Stanley B, Brown G, et al. Columbia-suicide severity rating scale (C-SSRS). New York, NY: Columbia University Medical Center. 2008;10.
7. Mary E. Nilsson, et al. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide, version 2.0, 2013.

12. APPENDICES

12.1 Appendix A Part 1: Schedule of Assessments – Screening and Tolerability Periods

Period	Screening	Tolerability		
Visit Number	1	2	3	4
Study Time Point	D -21 to 0 (W -3 to 0)	D 1 (W 1)	D 15 (W 3)	D 29 (W 5)
Visit Window (Days)	0	±3	±3	±3
Informed consent/assent (if appropriate)	X			
Review of inclusion and exclusion criteria	X			
Medical/surgical history including PWS genetic diagnosis, seizure history	X			
Demographics	X			
Review of prior/concomitant medications	X	X	X	X
Start collection/recording of AEs	X			
Review of AEs		X	X	X
Weight and height[a]	X	X	X	X
Record vital signs[b]	X	X	X	X
Physical exam[c]	X			
HQ-CT Questionnaire[d]	X	X pre-dose	X	X
Review of HQ-CT scores for eligibility[e]		X		
ABC Questionnaire[d]	X			
ESS-CHAD Questionnaire[d]		X pre-dose		
CGI-S Hyperphagia Questionnaire[d]		X pre-dose		
CaGI-S Hyperphagia Questionnaire	X			
Skin Picking Questionnaire[d]		X pre-dose		
C-SSRS[d,f]		X pre-dose		
Resting 12-lead ECG	X			
Serum pregnancy test[g]	X			
Urine drug screen[h]	X			
Urinalysis	X			
Chemistry panels[i]	X			
Hematology[j]	X			
HbA1c	X			

Period	Screening	Tolerability		
Visit Number	1	2	3	4
Study Time Point	D -21 to 0 (W -3 to 0)	D 1 (W 1)	D 15 (W 3)	D 29 (W 5)
Visit Window (Days)	0	±3	±3	±3
PT-INR	X			
Serologies for HBV, HCV, and HIV	X			
FSH for menopausal women only[k]	X			
Appetite hormones[l]	X			
Biomarkers[m]	X			
Review of eligibility[n]	X			
Placebo dispensing		X		
Placebo dosing[o]		X		
Drug accountability			X	X
Food Diary Reminder[p]				X
Food Diary Distribution				X

Abbreviations: ABC = Aberrant Behavior Checklist questionnaire; AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BMI = body mass index; BP = blood pressure; CBC = complete blood count; CGI-S = Clinical Global Impression of Change – Severity questionnaire; C-SSRS = Columbia-Suicide Severity Rating Scale; D = day; ECG = electrocardiogram; EDC = electronic data capture; ESS-CHAD = Epworth Daytime Sleepiness Scale; FSH = follicle-stimulating hormone; HBV = hepatitis B virus; HCV = hepatitis C virus; HIV = human immunodeficiency virus; HbA1c = glycosylated hemoglobin; HQ-CT = Hyperphagia Questionnaire for Clinical Trials; HR = heart rate; MCH = mean corpuscular hemoglobin; MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume ; PT-INR = prothrombin time-international normalized ratio; PWS = Prader-Willi Syndrome; RBC = red blood cells; RDW = red cell distribution width; RR = respiratory rate; THC = tetrahydrocannabinol; WBC = white blood cells; W = week; WOCBP = women of childbearing potential.

[a] Weight and height must be documented at every in-person visit as the EDC system will automatically calculated BMI (or BMI z-score) using both variables (Protocol Section 10.1.3).

[b] Vital signs include BP, HR, RR, oxygen saturation, and temperature (Protocol Section 10.1.2).

[c] A full physical exam should be performed at Screening Visit 1 (Protocol Section 10.1.4). Other physical exams should be abbreviated, symptom -directed.

[d] If a caregiver is required to assist the patient in completing the questionnaire, the same caregiver should be available throughout the study.

[e] The EDC system will automatically calculate mean HQ-CT score from Visit 1 and Visit 2. Patients with a mean HQ-CT score ≥13 will be eligible to continue in the Tolerability Period. Further details are described in Protocol Section 6.4.

[f] If the patient is the mental equivalent of 6 or younger then the children's version of the C-SSRS should be used.

[g] For WOCBP only.
[h] Drugs of abuse include THC. May be repeated at the Investigator's discretion.
[i] Chemistry panels include renal and liver function panels. Renal function panel includes sodium, potassium, chloride, calcium, bicarbonate, glucose, phosphorus, blood urea nitrogen, creatinine, and estimated glomerular filtration rate. Should be obtained after a 10-hour fast if possible. Liver function panel includes ALT, AST, albumin, alkaline phosphatase, direct bilirubin, total bilirubin, indirect bilirubin, and total protein (Protocol Table 9)
[j] Hematology panel includes hematocrit, hemoglobin, MCV, MCH, MHCH, RDW, RBC, WBC, platelets, and differential (Protocol Table 9).
[k] FSH levels must be ≥30 mIU/mL with absence of menstruation for ≥1 year to be considered menopausal status (Protocol Section 10.1.11).
[l] Includes insulin, ghrelin, and leptin (Protocol Table 9). Ghrelin samples will only be collected for Phase 3 patients at sites where proper collection is feasible.
[m] Will only be performed in patients enrolled in the Phase 3 part of the study and be optional. Biomarker assessments will not be conducted in the Phase 2 part of the study.
[n] Review of eligibility at Visit 1 includes all inclusion and exclusion criteria with the exception of HQ-CT. Eligibility related to HQ-CT scores will be assessed at Visit 2 and Visit 5 (Protocol Section 6.4).
[o] Dosing should be completed at the site when patients are present for a study visit during Visit 2 and Visit 5 (Protocol Section 7.1)
[p] Patients will be instructed to record the time of their evening and morning doses of study medication on the night prior to and the day of their next clinic visit with the type of meal administered. Patients should also be contacted approximately 1 week prior to their next visit to remind them to complete their food diary prior to their next visit (Protocol Section 10.1.14).

12.2 Appendix A Part 2: Schedule of Assessments - Dose Escalation, Maintenance, Taper, and Follow-Up Periods

	DOUBLE-BLIND TREATMENT PHASE[a]										TAPER PERIOD[b] (2 weeks)	FOLLOW-UP PERIOD[b] (2 weeks)	
Period	DOSE ESCALATION PERIOD[a] (3 Weeks)			MAINTENANCE PERIOD (24 Weeks)									
Visit Number	5[c,d]	6[e]	7[e,f]	8[c]	9[e]	10[c]	11[e,f]	12[c]	13[e,f]	14[c,g,aa,bb]		15[c]	16[e]
Study Time Point	D 43 (W 7)	D 50 (W 8) (Remote)	D 57 (W 9) (Remote)	D 64 (W 10)	D 99 (W 15) (Remote)	D 127 (W 19)	D 155 (W 23) (Remote)	D 183 (W 27)	D 204 (W 30) (Remote)	D 232 / ET/EOS (W 34)	D 232 (W 34)	D 246[c] (W 36[c])	D 260[e]/ EOS[e] (W 38) (Remote)
Visit Window (D)	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3
Review of concomitant medications	X	X	X	X	X	X	X	X	X	X		X	X
Review of AEs	X	X	X	X	X	X	X	X	X	X		X	X
Weight and height[h]	X			X		X		X		X			
Record vital signs[i]	X			X		X		X		X			
Physical exam[j]	X			X		X		X		X		X[k]	
HQ-CT Questionnaire[l,m]	X[pre-dose]			X		X		X		X			
Review of HQ-CT scores and I/E [n,o]	X												
C-SSRS[l,m,p]	X[pre-dose]			X		X		X		X			
Randomization[o]	X[q]												
ABC Questionnaire[l,m]	X			X		X		X		X			
ESS-CHAD Questionnaire[l,m]	X					X				X			
CGI-S Hyperphagia Questionnaire[l,m]	X			X		X		X		X			
CaGI-S Hyperphagia Questionnaire[l,m]	X			X		X		X		X			
CGI-C Hyperphagia Questionnaire[l,m]										X			
CaGI-C Irritability Questionnaire[l,m]										X			
CaGI-C Questionnaire[l,m]										X			

	DOUBLE-BLIND TREATMENT PHASE[a]										TAPER PERIOD[b] (2 weeks)	FOLLOW-UP PERIOD[b] (2 weeks)	
Period	DOSE ESCALATION PERIOD[a] (3 Weeks)			MAINTENANCE PERIOD (24 Weeks)									
Visit Number	5[c,d]	6[e]	7[e,f]	8[c]	9[e]	10[c]	11[e,f]	12[c]	13[e,f]	14[c,g,aa,bb]		15[c]	16[e]
Study Time Point	D 43 (W 7)	D 50 (W 8) (Remote)	D 57 (W 9) (Remote)	D 64 (W 10)	D 99 (W 15) (Remote)	D 127 (W 19)	D 155 (W 23) (Remote)	D 183 (W 27)	D 204 (W 30) (Remote)	D 232 / ET/EOS (W 34)	D 232 (W 34)	D 246[c] (W 36[c])	D 260[e]/ EOS[e] (W 38) (Remote)
Visit Window (D)	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3
Skin Picking Questionnaire[l,m]	X					X				X			
Resting 12-lead ECG	X			X		X				X			
Urinalysis	X			X		X		X		X			
Urine dipstick pregnancy test[r]	X			X		X		X		X			
Urine drug screen[s]	X												
Chemistry panels[t]	X			X		X		X		X			
Hematology[u]	X			X		X		X		X			
HbA1c				X				X		X			
Appetite hormones[v]	X			X				X		X			
Biomarkers[w]										X			
Pharmacokinetic	X			X[x]				X		X			
Food Diary Collection	X			X				X		X			
Food Diary Reminder[f]			X				X		X				
Food Diary Distribution			X[cc]				X[cc]		X[cc]				
Drug accountability	X			X		X		X		X		X	
IP dispensing	X			X	X	X	X	X	X		X		
Dosing[y]	X												
Seizure Patient Impression										X			
DEXA scan[z]	X									X			
Tapering of IP over 14 days[b]											X[b,g]		

Abbreviations: ABC = Aberrant Behavior Checklist questionnaire; AE = adverse event; ALT = alanine aminotransferase; AST =

aspartate aminotransferase;BMI = body mass index; BP = blood pressure; CaGI-C = Caregiver Global Impression of Change; CBC = complete blood count; CGI-C= Clinical Global Impression of Change; CGI-S= Clinical Global Impression of Change – Severity questionnaire; C-SSRS = Columbia-Suicide Severity Rating Scale; D = day; DEXA = dual-energy x-ray absorptiometry; DMC = Data Monitoring Committee; EDC = electronic data capture; ECG = electrocardiogram; ESS-CHAD = Epworth Daytime Sleepiness Scale; EOS = End of Study; HbA1c = glycosylated hemoglobin; HQ-CT = Hyperphagia Questionnaire for ClinicalTrials; HR = heart rate; IP = investigational product (RAD011 or placebo); MCH = mean corpuscular hemoglobin; I/E = Inclusion/Exclusion Criteria; MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume; PK= Pharmacokinetic; PWS = Prader-Willi Syndrome; RBC = redblood cells; RDW = red cell distribution width; RR = respiratory rate; THC = tetrahydrocannabinol; W = week; WBC = white blood cells.

[a] In the Phase 2 part of the study, all patients will be initiated on IP (RAD011 or placebo) 10 mg/kg/day (0.1 mL/kg/day) for a period of 1 week at Visit 5 (Week 7). Patients randomized to low dose or low volume IP (10 mg/kg/day) will continue at a dose of 10 mg/kg/day (0.1 mL/kg/day) until the end of Week 9. Patients randomized to mid dose or mid volume IP (20 mg/kg/day) will have their dose increased to 20 mg/kg/day (0.2 mL/kg/day) at Visit 6 (Week 8) and will continue at that dose until the end of Week 9. Patients randomized to high dose or high volume IP (40 mg/kg/day) will have their dose increased to 20 mg/kg/day (0.2 mL/kg/day) at Visit 6 (Week 8) and to 40 mg/kg/day at Visit 7 (Week 9). Details of the dose escalation schema are provided in Protocol Table 4.

In the Phase 3 part of the study, all patients will be initiated on IP (RAD011 or placebo) 10 mg/kg/day IP (0.1 mL/kg/day) for a period of 1 week at Visit 5 (Week 7). Dose escalation will occur similarly to the Phase 2 part of the study with the 1 or 2 doses recommended by the DMC following their review of safety and tolerability data from the Phase 2 part of the study. Details of the dose escalation schema are provided in Protocol Table 4.

No dose decreases will be allowed during the Dose Escalation Period (Protocol Section 5.7.1).

[b] The Taper and Follow Up Period visits (Visits 14 to 16) are only for patients who do not proceed to enroll in the active long-term extension study SCOUT-016 and who are tapering off the IP. Patients will return to the site for assessment of concomitant medications, AEs, and to return and reconcile IP on Visit 15. Visit 16 may be done remotely (Protocol Section 5.8.6).

[c] To be performed in person at the research site.

[d] All assessments must be performed prior to dosing, unless specified otherwise.

[e] Assessments may be conducted remotely via a phone call, video call, or other remote means of communication. Dispensing of IP may be done at the research site or remotely, using direct-to-patient shipping methods in accordance with all local laws and regulations. Sites may elect to conduct the entire visit in person.

[f] At Week 9 (Visit 7), Week 23 (Visit 11), and Week 30 (Visit 13), patients will be instructed to record the time of their evening and morning doses of study medication on the night prior to and the day of their next clinic visit with the type of meal administered. Patients should also be contacted approximately 1 week prior to their next visit to remind them to complete their food diary prior to their next visit (Protocol Section 10.1.14).

[g] At visit 14, patients may be offered participation in the long-term extension study SCOUT-016. Patients eligible and electing to

participate in the long-term extension study SCOUT-016 should continue treatment with IP as per the SCOUT-016 protocol. Patients who are not eligible or decide not to participate in the long-term extension study SCOUT-016 should proceed to the Taper and Follow-Up Periods (Protocol Section 5.8.5.1).

[h] Weight and height must be documented at every in-person visit (Protocol Section 10.1.3).

[i] Vital signs include BP, HR, RR, oxygen saturation, and temperature (Protocol Section 10.1.2).

[j] A full physical exam should be performed at Week 7 (Visit 5). Other physical exams should be abbreviated, symptom directed (Protocol Section 10.1.4).

[k] To be performed at Visit 15 (Week 36) only.

[l] If a caregiver is required to assist the patient in completing the questionnaire, the same caregiver should be available throughout the study.

[m] When possible, study questionnaires should be completed prior collection of ECG and blood samples.

[n] Patients meeting all inclusion and exclusion criteria should proceed to randomization.

[o] The EDC system will automatically calculate mean HQ-CT score from Visit 3, Visit 4 and Visit 5 inclusively. Patients with a mean HQ-CT score ≥13 will be eligible for randomization. In addition, the EDC system will also automatically calculate the change in HQ-CT score from the Tolerability Period to the Screening Period. Patients must have a decrease in HQ-CT score ≤7 to be eligible for randomization. Further details are defined in Protocol Section 6.4.

[p] If the patient is the mental equivalent of 6 or younger then the children's version of the C-SSRS should be used.

[q] Randomization cannot be performed prior to evaluation and review of HQ-CT questionnaire scores for eligibility. Prior to assessing randomization eligibility, patients must complete the Visit 5 HQ-CT questionnaire (Protocol Section 6.4.2).

[r] For WOCBP only. If positive, a quantitative serum pregnancy test must be sent to the central laboratory for confirmation.

[s] Drugs of abuse include THC. May be repeated at the Investigator's discretion.

[t] Chemistry panels include renal and liver function panels. Renal function panel includes sodium, potassium, chloride, calcium, bicarbonate, glucose, phosphorus, blood urea nitrogen, creatinine, and glomerular filtration rate. Should be obtained after a 10-hour fast if possible. Liver function panel includes ALT, AST, albumin, alkaline phosphatase, direct bilirubin, total bilirubin, indirect bilirubin, and total protein (Protocol Table 9).

[u] Hematology panel includes hematocrit, hemoglobin, MCV, MCH, MHCH, RDW, RBC, WBC, platelets, and differential (Protocol Table 9).

[v] Includes insulin, ghrelin, and leptin (Protocol Table 9). Ghrelin samples will only be collected for Phase 3 patients at sites where proper collection is feasible.

[w] Will only be performed in patients enrolled in the Phase 3 part of the study and be optional. Biomarker assessments will not be conducted in the Phase 2 part of the study

[x] Must be drawn 6 ± 2 hours after the previous IP dose. Patients may take their IP dose at home prior to presenting to the site for their visit in order to comply with the PK timing requirements.

[y] Dosing should be completed at the site during Visit 2 and Visit 5. Dosing is not required to be done at the site when patients have remote or other visits (Protocol Section 7.1).
[z] Can be performed before or after dosing. The DEXA scan can be performed within ± 7 days of the visit time point (Protocol Section 9.3.6). The DEXA scan will be performed at a certain number of US sites only.
[aa] All assessments will be collected regardless of if the patient enters the taper period or completes the study and enter the long-term extension study.
[bb] The two-week taper period does not have any visits associated with it.
[cc] The food diary can be provided at the previous in-person visit.

12.3 Appendix B: Missing or Incomplete Dates Imputation Rules

Imputation of Partial or Completely Missing Start Dates for AEs and Concomitant Medications

- If only the month and year are specified and the month and year of first dose is not the same as the month and year of the start date, then use the 1st of the month.
- If only the month and year are specified and the month and year of first dose is the same as the month and year of the start, then use the date of first dose.
- If only the year is specified, and the year of first dose is not the same as the year of the start date, then use January 1 of the year of the start date.
- If only the year is specified, and the year of first dose is the same as the year of the start date, then use the date of first dose.
- If the start date is completely unknown, and the stop date is unknown or not prior to the date of first dose, then use the date of first dose.

Imputation of Partial or Completely Missing Stop Dates for AEs and Concomitant Medications

- If only the month and year are specified, then use the last day of the month.
- If only the year is specified, then use December 31 of that year.
- If the stop date is completely unknown, do not impute the stop date.
    - If an AE was resolved and the resolution date is completely unknown, then do not impute the resolution date.

Imputation for Missing Study Medication Start and Stop Dates

No imputation should be performed for missing study medication start dates. This field on the eCRF should not be partial or missing.

For partial or missing date of last dose of study medication, the following imputation rules will be applied for the purpose of calculating overall exposure:

If the day is missing (but month and year available), impute the last dose date as the minimum of the last day of the month or the date of last contact reported on the trial termination eCRF; if day and month are both missing (only year available), impute the last dose date as the minimum of the last day of the year or the date of last contact on the trial termination eCRF.

If a patient died and has a partial or missing last administration date, the date is to set to the date of death. If there is a partial date of last dose and the month/year are prior to the month and year of the date of death, follow partial date imputation rules.

If the last dose date is completely missing and no information could be obtained from data cleaning exercises, the last dose date should be imputed as the date of last contact according to the study termination eCRF. A review of the data for patients with completely missing last dose dates should be performed to ensure that the imputation does not result in an unrealistic value for duration of exposure.

Imputed date of last dose dates should only be used for calculation of the duration of exposure.

The date as recorded on the eCRF should be presented in patient data listings (no imputed dates should be included in patient data listings).

12.4 Appendix C: List of Preferred Terms for AESI

PT for Hepatic dysfunction AESI

PT
Acquired factor V deficiency
Alanine aminotransferase abnormal
Alanine aminotransferase increased
Ammonia abnormal
Ammonia increased
Ascites
Aspartate aminotransferase abnormal
Aspartate aminotransferase increased
AST/ALT ratio abnormal
Bacterascites
Bile output abnormal
Bile output decreased
Biliary ascites
Bilirubin conjugated abnormal
Bilirubin conjugated increased
Bilirubin urine present
Biopsy liver abnormal
Blood bilirubin abnormal
Blood bilirubin increased
Blood bilirubin unconjugated increased
Bromosulphthalein test abnormal
Child-Pugh-Turcotte score abnormal
Child-Pugh-Turcotte score increased
Computerised tomogram liver abnormal
Congestive hepatopathy
Foetor hepaticus
Galactose elimination capacity test abnormal
Galactose elimination capacity test decreased
Gamma-glutamyltransferase abnormal
Gamma-glutamyltransferase increased
Guanase increased
Hepaplastin abnormal
Hepaplastin decreased
Hepatic artery flow decreased
Hepatic enzyme abnormal
Hepatic enzyme decreased
Hepatic enzyme increased
Hepatic function abnormal
Hepatic hydrothorax

PT
Hepatic hypertrophy
Hepatic hypoperfusion
Hepatic mass
Hepatic pain
Hepatic sequestration
Hepatic vascular resistance increased
Hepatic venous pressure gradient abnormal
Hepatic venous pressure gradient increased
Hepatobiliary scan abnormal
Hepatomegaly
Hepatosplenomegaly
Hyperammonaemia
Hyperbilirubinaemia
Hypercholia
Hypertransaminasaemia
Kayser-Fleischer ring
Liver function test abnormal
Liver function test decreased
Liver function test increased
Liver induration
Liver palpable
Liver scan abnormal
Liver tenderness
Drug-Induced Liver Injury
Magnetic resonance imaging hepatobiliary abnormal
Magnetic resonance proton density fat fraction measurement
Mitochondrial aspartate aminotransferase increased
Molar ratio of total branched-chain amino acid to tyrosine
Oedema due to hepatic disease
Perihepatic discomfort
Retrograde portal vein flow
Total bile acids increased
Transaminases abnormal
Transaminases increased
Ultrasound liver abnormal
Urine bilirubin increased
White nipple sign
X-ray hepatobiliary abnormal
5'nucleotidase increased
AST to platelet ratio index increased
Blood alkaline phosphatase abnormal
Blood alkaline phosphatase increased

PT
Blood cholinesterase abnormal
Blood cholinesterase decreased
Cytokeratin 18 increased
Deficiency of bile secretion
Glutamate dehydrogenase increased
Glycocholic acid increased
Haemorrhagic ascites
Hepatic fibrosis marker abnormal
Hepatic fibrosis marker increased
Hepatic lymphocytic infiltration
Hepatitis A immunity confirmed
Hepatitis B immunity confirmed
Hepatitis E immunity confirmed
Hypoalbuminaemia
Leucine aminopeptidase increased
Liver iron concentration abnormal
Liver iron concentration increased
Liver opacity
Model for end stage liver disease score abnormal
Model for end stage liver disease score increased
Osteopontin increased
Periportal oedema
Peritoneal fluid protein abnormal
Peritoneal fluid protein decreased
Peritoneal fluid protein increased
Pneumobilia
Portal vein flow decreased
Portal vein pressure increased
Retinol binding protein decreased
Urobilinogen urine decreased
Urobilinogen urine increased

PT for Suicide ideation and behavior AESI

PT
Assisted Suicide
Columbia suicide severity rating scale abnormal
Completed suicide
Depression suicidal
Intentional overdose
Intentional self-injury
Poisoning deliberate

PT
Self-injurious ideation
Suicidal behaviour
Suicidal ideation
Suicide attempt
Suicide threat
Suspected suicide
Suspected suicide attempt

12.5 APPENDIX D: Marked Lab abnormalities

12.5.1 Hematology

PARAMETER	*AGE RANGE*	*UNIT (conventional)*	*ABNORMALITY CRITERIA (conventional)*	*UNIT (standard)*	*ABNORMALITY CRITERIA (standard)*
Hemoglobin	All	g/dL	≤10.0 >110% of ULN	g/L	≤100 >110% of ULN
WBC/Leukocytes	All	10^9/L	≤3.0	g/L	≤3.0
Neutrophils Absolute	>1m	10^9/L	<1.5	g/L	<1.5
RBC/ Erythrocytes	All	10^{12}/L	<90% of LLN	g/L	<90% of LLN

Note: Criteria for abnormality are defined either from Common Terminology Criteria for Adverse Events v5.0 Grade 1 if it exits or applies a >=110% or a <=90% rule.