STATISTICAL ANALYSIS PLAN
Protocol RM-493-015
An Open Label, 1-Year Trial, Including a Double-Blind Placebo-Controlled Withdrawal Period, of Setmelanotide (RM-493), a Melanocortin 4 Receptor (MC4R) Agonist, in Leptin Receptor (LEPR) Deficiency Obesity due to Bi-Allelic Loss-of-Function *LEPR* Genetic Mutation

Protocol Number: (Version Date)	RM-493-015 (10 APR 2018, Version 1.1, Amendment 3)
Name of Test Drug:	Setmelanotide (RM-493, Melanocortin-4 Receptor Agonist)
Phase:	3
Methodology:	Pivotal, 1 year open-label active treatment with a double-blind, placebo-controlled withdrawal period
Sponsor:	Rhythm Pharmaceuticals, Inc. 500 Boylston Street, 11th Floor Boston, MA 02116, USA
Sponsor Representative:	[redacted]
Document Date:	16 April 2019
Document Version:	Version 2.0

SIGNATURE PAGE

Protocol Title:	An Open Label, 1-Year Trial, Including a Double-Blind Placebo-Controlled Withdrawal Period, of Setmelanotide (RM-493), a Melanocortin 4 Receptor (MC4R) Agonist, in Leptin Receptor (LEPR) Deficiency Obesity due to Bi-Allelic Loss-of-Function *LEPR* Genetic Mutation
Sponsor:	Rhythm Pharmaceuticals, Inc. 500 Boylston Street, 11th Floor Boston, MA 02116, USA
Protocol Number:	RM-493-015
Document Date/Version:	10 APR 2018 Version 1.1 Amendment 3
Author: [redacted]	Signat [redacted] Date: 16

Sponsor Approval

By signing this document, I acknowledge that I have read the document and approve of the planned statistical analyses described herein. I agree that the planned statistical analyses are appropriate for this study, are in accordance with the study objectives, and are consistent with the statistical methodology described in the protocol, clinical development plan, and all applicable regulatory guidances and guidelines.

I have discussed any questions I have regarding the contents of this document with the biostatistical author.

All changes to the planned analyses will be described in the clinical study report (CSR).

Sponsor Signatory:

[redacted]cals, Inc.
500 Boylston Street, 11th Floor
Boston, MA 02116, USA

Telephone: [redacted]

Signature: [redacted]

Date: _____

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviation	Definition
[redacted]	
ACTH	Adrenocorticotropic hormone
AE	Adverse event
ALT	Alanine transaminase
AST	Aspartate transaminase
ATC	Anatomic Therapeutic Class
AUC	Area under the curve
BIA	Bio-electrical impedance analysis
BMI	Body mass index
BP	Blood pressure
BPM	Beats per minute
BTD	Breakthrough Therapy Designation
BQL	Below the limit of quantification
BUN	Blood urea nitrogen
CGIC	Caregiver Global Impression of Change
CGIS	Caregiver Global Impression of Severity
CI	Confidence Interval
CLIA	Clinical Laboratory Improvement Amendments
CM	Centimeters
[redacted]	
CO_2	Carbon dioxide
CPK	Creatine phosphokinase
CRF	Case report form
CS	Completer Set
CSR	Clinical study report
[redacted]	
DEXA	Dual-energy x-ray absorptiometry
DUS	Designated Use Set
ECG	Electrocardiogram
EU	European Union
FAS	Full Analysis Set
FDA	Food and Drug Administration
FSH	Follicle Stimulating Hormone
GGT	Gamma-glutamyltranspeptidase
HbA1c	Hemoglobin A1c
HOMA	Homeostatic model assessment
HR	Heart rate
[redacted]	
ICH	International Conference for Harmonization
INR	International normalized ratio
[redacted]	
KG	Kilograms
LDH	Lactate dehydrogenase
LepR	Leptin Receptor

LOCF	Last observation carried forward
LOF	Loss of function
MC4R	Melanocortin Receptor type 4
MedDRA	Medical Dictionary for Regulatory Activities
NCE	New Chemical Entity
NRS	Numeric rating scale
OGTT	Oral glucose tolerance test
ORO	Observer-Related Outcomes
[redacted]	
PD	Pharmacodynamic
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
[redacted]	
PI	Primary Investigator
[redacted]	
POC	Proof-of-concept
POMC	Pro-opiomelanocortin
PP	Per Protocol
PT	Prothrombin time
PTT	Partial thromboplastin time
RS1	Responder's set 1
RS2	Responder's set 2
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical Analysis Software
SD	Standard deviation
[redacted]	
SMB	Safety Monitoring Board
SOA	Schedule of assessments
SOC	System/Organ/Class
[redacted]	
USPI	US Package Insert
WHO	World Health Organization
[redacted]	

1. INTRODUCTION AND OBJECTIVES OF ANALYSIS

1.1. Introduction

This document presents the statistical analysis plan (SAP) for Rhythm Pharmaceuticals, Inc. (Rhythm), Protocol No. RM-493-015: An Open Label, 1-Year Trial, Including a Double-Blind Placebo-Controlled Withdrawal Period, of setmelanotide (RM-493), a Melanocortin 4 Receptor (MC4R) Agonist, in Leptin Receptor (LEPR) Deficiency Obesity due to Bi-Allelic Loss-of-Function *LEPR* Genetic Mutation. It contains the details and methods to be used to perform the proposed efficacy and safety analyses, including planned summary tables and by-patient listings.

Study RM-493-015 is a pivotal Phase 3 study designed to demonstrate statistically significant and clinically meaningful effects of setmelanotide on percent body weight change in patients with LEPR deficiency obesity due to rare bi-allele or loss-of-function mutations at the end of 1 year of treatment. Effects on hunger score suppression will also be evaluated over the course of treatment.

The extreme rarity of LEPR deficiency obesity has led to the small target number of LEPR deficiency patients to be enrolled and evaluated in this open-label clinical study of setmelanotide. We have found that patients are difficult to identify and recruit. In many ways, the set of patients under study will constitute a collection of detailed clinical case reports with a comprehensive baseline and past medical history assessment and complete clinical efficacy, safety and laboratory evaluations conducted for each patient. Patient numbers may be increased or decreased to satisfy regulatory requirements for registration, depending upon the number of affected patients identified.

It is planned that approximately ten treated pivotal patients will receive approximately one-year of setmelanotide at daily doses that will be individually titrated for each patient during their initial dose titration, with additional supplementary patients who will receive less than one-year of treatment at the time of filing; the primary analyses will only include the pivotal patients, and the primary analyses will later be repeated with the supplemental patients as supportive analyses. In addition, personalized therapeutic response assessments of weight reduction and hunger score changes after 12 weeks, and at approximately one year of dosing will also be assessed. Thus, many analyses will involve descriptive summary statistics derived from absolute and percentage responses for each patient over differing periods of time.

There will also be an 8-week double-blind, placebo-controlled withdrawal period consisting of 4 weeks of active therapy and 4 weeks of masked placebo substitution conducted in each patient serving as their own control. Assessments of each patient during this period will involve defined comparisons to weight and hunger scores just before and just after this sequence, which will be the only closed (blinded) design element of this study.

After this withdrawal period is completed, long-term efficacy and safety analyses will evaluate responses for total treatment periods extending for approximately 1 year. Overall, data will be tabulated and depicted to highlight individual patient responses as well as key group mean results to fully characterize key study data from this set of LEPR deficiency obesity patients.

The protocol includes elements to assess the expected robust treatment effects of setmelanotide despite the open-label design and small patient numbers. All these elements -- including prior patient histories documenting unsuccessful control of weight and hunger, other key medical and

natural history elements, potential responses (or the lack thereof, as in the Phase 2 study) during individualized dose titration from sub-therapeutic to personalized target efficacy doses, data from the double-blind, placebo-controlled withdrawal period, and the anticipated extremely robust weight and hunger responses during long term open-label treatment – are expected to help demonstrate and substantiate the effects of setmelanotide in this rare genetic form of obesity.

Access to efficacy data from this study will be restricted as defined in this SAP (further details in Section 2.4), in order to insure integrity of this Phase 3 study.

This SAP was based on Amendment 3 to the final protocol dated 10 April 2018. [Note that this amendment, which includes approval for enrollment of POMC deficiency obesity patients age 6 years and older, has not been approved at all clinical trial sites. Nonetheless, the sponsor Rhythm has included analyses pertaining to the evaluation of this younger age group in this SAP to prospectively identify these planned analyses, which will be conducted on younger age ranges of patients only if they are approved for enrollment in the trial.]

The reader of this SAP is encouraged to read the clinical protocol, and other relevant documents for details on the planned conduct of this study. The objectives listed below (primary, secondary and exploratory) are found in the protocol. Other than the schedule of study procedures, which is provided in Appendix 1, operational aspects related to collection and timing of planned clinical assessments are not repeated in this SAP unless relevant to the planned analyses.

1.2. Objectives of Statistical Analysis

This SAP is designed to outline the methods to be used in the analysis of study data to answer the study objectives. Populations for analysis, data handling rules, statistical methods, and formats for data presentation are provided. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results sections of the clinical study report (CSR) for this trial.

This SAP also will outline any differences in the currently planned analytical objectives relative to those planned in the study protocol.

2. STUDY DESIGN

2.1. Synopsis of Study Design

This is a multicenter, open-label, pivotal Phase 3 trial, with a double-blind placebo-controlled withdrawal period, evaluating the efficacy and safety of setmelanotide on percent body weight change at approximately the end of 1 year of treatment in patients at least 6 years old. The study includes a comprehensive screening period to evaluate potential participants with LEPR deficiency obesity. Given the rarity of such patients, a thorough understanding of the genetics and disease history for each patient dating to as early in childhood as possible is critical to understand the natural history of patients with this early-onset, extreme genetic obesity. The sponsor medical monitor will be in close contact with Site Primary Investigators (PIs) for the initial screening and assessment of potential candidate patients. Details to be jointly reviewed and verified by PIs with the medical monitor include body weight history (to be as comprehensive as feasible in these patients to help understand a complete weight and body mass index (BMI) trajectory prior to setmelanotide initiation), review of key disease features and other past medical history elements, and pertinent inclusion and exclusion criteria details for each participant in this pivotal trial. In addition, DNA sequence details for each LEPR deficiency obesity patient will be reviewed to ensure that appropriate bi-allelic mutations that define the specific *LEPR* gene variants have been identified for the LEPR deficiency patients enrolled in this clinical trial. These mutations will be subsequently verified using Clinical Laboratory Improvement Amendments (CLIA)-approved DNA sequencing of a genomic DNA sample collected from each patient.

The treatment phase of the study will begin with an initial period of dose titration in 0.5 mg increments and lasting between 2 and 12 weeks (ranging from as brief as a single dose step at either 1.0 mg for adults or 0.5 mg for pediatric (ages 6 – 11 years) and adolescent (ages 12 – 18 years) subjects, to as many as 6 dose steps to reach 3.0 mg QD starting from the 0.5 mg starting dose in pediatric and adolescent subjects). Thus, each subject will have his/her dose individualized depending on the number of dosing increments administered, where each individual patient's therapeutic dose will be established by upwards dose titration in 2 week intervals (refer to Appendix 11.8 of the protocol). Thereafter, patients will continue on active treatment at their optimal therapeutic dose for an additional 10 weeks, for a total combined dosing duration of 12 weeks at the individual patient's established therapeutic dose [i.e., the last 2 weeks during dose titration plus 10 weeks of open label treatment]. Patients who demonstrate at least 5 kg weight loss (or 5% weight loss change from baseline if baseline body weight is < 100 kg) at the end of this 12-week Open Label Treatment Period will continue onto the double-blind, placebo-controlled, withdrawal period lasting 8 weeks, including a 4-week period of placebo treatment. Patients who do not demonstrate at least 5 kg weight loss (or 5% weight loss change from baseline if baseline body weight is < 100 kg) will continue in the study but will be withdrawn from treatment. Following the withdrawal period, all patients will complete approximately 1 year of total therapeutic dosing. It is planned that the individual patient's therapeutic dose, established during the initial period of dose titration, will continue throughout the study. Please refer to the protocol for more detailed information.

It is anticipated that ~10 pivotal patients will be enrolled in this study (with approximately 1-year of treatment), due to the ultra-rare incidence of this population of patients. The pivotal patient set will be the first ~10 patients enrolled (in order of enrollment), and once determined,

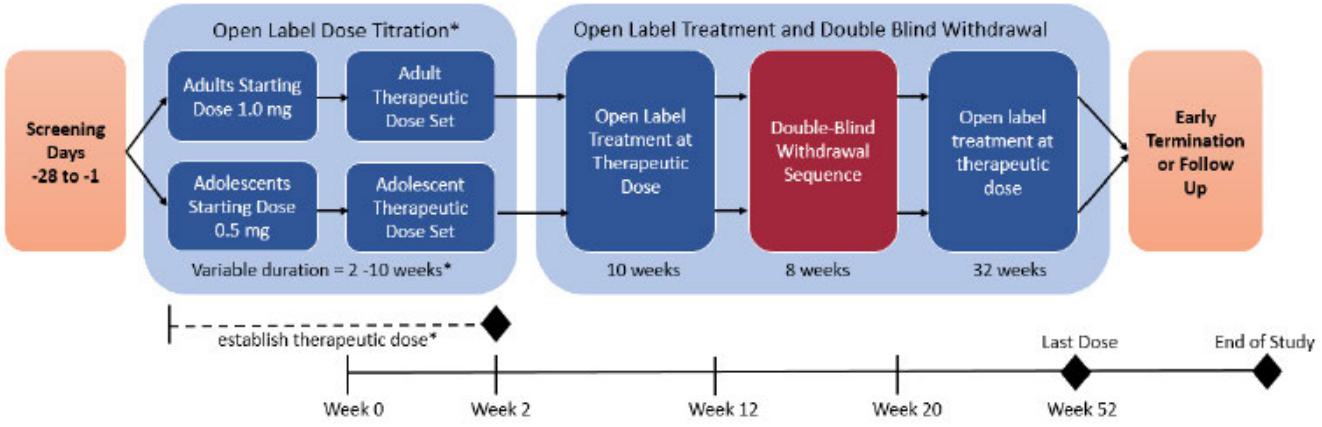
will be clearly documented. Enrollment will be kept open after the first ~10 pivotal patients begin, for the purpose of collecting additional supporting data.

Patients as young as 6 years of age will be enrolled at clinical sites where Amendment 3 is approved. Note that Amendment 3, which proposes enrollment of patients 6 years and older, remains under regulatory review with some European Union (EU) health authorities; this SAP has included elements that proactively address proposed specific assessments in this younger age cohort, assuming it will eventually be approved in most locations.

In addition, for patients at sites where feasible, and who agree to participate, there are optional sub-studies to evaluate robust [redacted], and blood pressure/heart rate (BP/HR) by [redacted] data. It is anticipated that only a modest number of patients and sites will participate in each sub-study, and care should be taken not to overburden individual patients.

Thus, patients will progress sequentially through the study as depicted in Figure 1 below.

Figure 1 - Study Schema



*The last two weeks of the Open-label dose titration Phase in which the therapeutic dose for an individual patient is established will be considered the first two weeks of the Open Label Treatment. Subsequently patients will then receive an additional 10 weeks of active treatment in the Open Label Treatment for a total of 12 weeks of treatment at the therapeutic dose before transitioning into the Double-Blinded Withdrawal Phase.

2.2. Randomization Methodology

There will be no randomization in this trial.

2.3. Blinding During the 8-Week Double-Blind Withdrawal

The only double-blind study period in Study RM-493-015 is timed to occur immediately following the 12-week initial therapeutic test period after individualized up titration to a therapeutic dose. A more detailed description of the analysis of this important period follows in Section 4.6.3

This 8-week placebo-controlled withdrawal period is blinded to patients and site staff; thus, all site and investigator staff and study participants will be blinded and unaware of the sequence each participant is receiving. In contrast, sponsor personnel are aware of the sequence, with the 4-week active therapy preceding the 4-week placebo withdrawal period for all patients during this 8-week study interval, but will have no access to efficacy data from this period until the study is complete and the database locked. The rationale for this fixed sequencing with continuing active treatment preceding withdrawal to masked placebo in all subjects is to allow each participant to experience a total of 16 weeks of continuous active treatment at an individualized therapeutic dose established for each patient prior to the introduction of the placebo.

Breaking the blind of the 8-week double-blind placebo controlled phase for the clinical site staff will only be done in the event of a medical emergency where the identity of study drug is necessary to appropriately treat the patient. The request to break the blind will be discussed with the medical monitor and Rhythm, whenever possible. All details of when and how the blind was broken will be documented in the CSR.

2.4. Data Handling to Ensure Study Integrity

The open-label design of this pivotal clinical trial makes it important for the sponsor to be blinded to key efficacy data, to ensure bias is not introduced into this study by data reviews.

To this end, efficacy and safety data from this study will be carefully controlled both within the sponsor and outside:

- No combined or integrated patient data will be tabulated by the sponsor, shared with any investigator, investigator site staff, or outside personnel until the study is complete and the database for 1-year data is locked and finalized.
- Safety information will be shared with the Safety Monitoring Board (SMB), and any blinded safety data (i.e., from the double-blind, placebo-controlled withdrawal period) will not be unblinded until the study is complete and the database for 1-year data is locked and finalized. The exception can be if the outside member(s) of the SMB requires unblinded data to fully evaluate a safety concern, in which case the process, details and reason for unblinding will be carefully outlined in the CSR, and such data will only be available to the outside (independent) member of the SMB.
- The sponsor medical monitor and study clinical operations lead will have access to all safety and efficacy clinical data *only* during open-label dose titration phase. This close engagement by two key Rhythm study personnel in early periods of the RM-493-015

study for every patient is necessary to advise clinical site investigators or study staff on appropriate dose titration decisions. No other Rhythm personnel or any other individuals will have access to, or be allowed to obtain any access to this information in any form. Recall that the period covering the last 2 weeks of titration also is identical to the first 2 weeks of individualized therapeutic dosing, followed by 10 additional weeks to complete the entire 12-week dosing period at each participant's personalized therapeutic dose.

- Thereafter, all accruing study weight and hunger response data, during the 10 additional weeks of the initial 12-week dosing period, will no longer be available for the Rhythm medical monitor or study clinical operations lead, and will remain unavailable in any form for any other individuals (i.e., no Rhythm personnel will have access to the weight and hunger data). Specifically, from the time each study participant establishes their individualized therapeutic dose (with consensus between each Principal Investigator and the sponsor medical monitor), and at the beginning of 10 additional weeks at this therapeutic dose level, all clinical data files for primary and secondary efficacy endpoint parameters that are maintained by the data management vendor (MedSource) will no longer be accessible to Rhythm personnel or other blinded individuals.
- All weight and daily hunger data past the dose-titration phase will not be accessible to the lead or consulting biostatistician.
- Similarly, all accruing study weight and hunger response data will continue to remain unavailable to the Rhythm medical monitor or study clinical operations lead during the 8-week double-blind withdrawal period, and during the subsequent period of continued open label treatment until the full 1-year efficacy treatment period is completed at the end of the trial. As noted above, clinical data files for these primary and secondary efficacy endpoint parameters that are maintained by the data management vendor (MedSource) will not be accessible. These data will also remain unavailable in any form for any other individuals (i.e., no Rhythm personnel will have access to the weight and hunger data).
- In addition, no data will be unblinded from the 8-week double-blind placebo controlled withdrawal period for any reason (except for safety reasons outlined below) until the study is complete, all data has been cleaned, and the database is locked.
- The Rhythm study monitor and study operations lead will continue to monitor patient safety and laboratory data with the PI and during clinical site interactions in these subsequent phases of the study, but they will not inquire or communicate anything regarding ongoing efficacy, in particular weight and hunger responses with clinical site personnel. All site personnel will be instructed not to provide any insight, information or indications about any efficacy data during the course of the study except during the open-label titration phase.
- In the event an Investigator judges that a dose adjustment may be necessary, Investigators will be instructed to contact the Independent SMB member for consultation, who can then engage the Rhythm Medical Monitor once all efficacy and potential unblinding data, including patient identity, have been redacted. Any dose adjustments will be documented in the SMB minutes in an anonymized manner. It is anticipated that this type of clinical review request would be a rare event; if such cases

occur, the process, details and reason for a dose adjustment discussion will be carefully documented, and all data shared would be included in a note to file on such discussions.

- A clinical operational specialist(s) will be informed by the investigator at each site if a patient achieves the ≥ 5 kg weight loss threshold at the end of the 12-week initial treatment period and will be continuing on active treatment. This information will be used to arrange drug supply for patients who continue on treatment in the study (note that all patients will be encouraged to continue study visits, even if not on treatment, but those who do not achieve the ≥5 kg weight loss threshold will not receive any further active or placebo treatment.)

2.5. Planned Trial Conduct Assessment and Safety Monitoring

The open-label design of this pivotal clinical trial for all but 8 weeks of the double-blind, placebo-controlled withdrawal period and the very small number of investigative clinical sites involved in evaluating these extremely rare LEPR deficiency patients should enable the timely and thorough review of accruing clinical safety data. A SMB for the RM-493-015 study has been set up and details on SMB operation and analysis support will be documented in a separate SMB charter.

2.6. Study Procedures

The schedule of assessments (SOA), as outlined in the study protocol, is provided in Appendix 1. Table 1 illustrates the schedule for screening and dose titration, Table 2 illustrates the 10-week active treatment and 8-week double-blind placebo controlled withdrawal period, and Table 3 illustrates the additional ~32-week open label treatment period. All footnotes can be found following Table 3.

2.7. Efficacy, Pharmacokinetic, and Safety Variables

2.7.1. Efficacy Variables

The primary objective of this study is to demonstrate the efficacy of setmelanotide. The primary endpoint is the proportion of patients in the full analysis set (FAS) population who meet the ≥10% weight loss threshold (responders) after approximately 1 year of treatment (pivotal patient set), compared to the proportion from historical data (at most 5% responders in the null population).

If this analysis reaches statistical significance (1-sided test, alpha=0.05), then the proportion of patients in the FAS population who show ≥10% weight loss after approximately 1 year of treatment will be numerically compared (point estimate) to the 35% proportion of responders as defined in the "Success Criteria" (description and rationale are provided in Section 4.7.1).

Key Secondary Endpoints (with analysis to be performed at approximately 1 year) are as follows:

- The first key secondary endpoint is the mean percent change from baseline in body weight (kg) at the end of approximately 1 year of treatment in the Designated Use Set (DUS) population (see Sections 3.1 and 3.2 for a description and rationale for the primary population).

  If this analysis reaches statistical significance (P<0.05 compared to the null hypothesis of 0% weight loss; 1-sided test, alpha=0.05), then the mean percent change from

baseline at the end of approximately 1 year of treatment in the DUS population will be numerically compared (point estimate) to the ≥10% mean reduction from baseline in body weight "Success Criteria" for this study (description and rationale are provided in Section 4.7.1).

- The second key secondary endpoint is the mean percent change in weekly average hunger (using "most hunger over the last 24-hours" daily response) from baseline at the end of approximately 1 year of treatment. Hunger assessments will be determined on the weekly mean of daily responses, evaluated in patients aged 12 years and older in the DUS population.

  If this analysis reaches statistical significance (P<0.05 compared to the null hypothesis of 0% weight loss; 1-sided test, alpha=0.05), then the mean percent change in weekly average hunger at the end of approximately 1-year of treatment in the DUS population will be numerically compared (point estimate) to the 25% reduction from baseline "Success Criteria" (a description and rationale is provided in Section 4.7.1).

- The third key secondary endpoint is a categorical analysis of the proportion of patients in the FAS population who meet the ≥25% improvement in hunger threshold (responders) at the end of approximately 1 year of treatment, compared to the proportion from historical data (at most 5% responders in the null population).

  If this analysis reaches statistical significance, then the proportion of patients in the FAS population who show ≥25% improvement in hunger at the end of approximately 1 year of treatment will be numerically compared (point estimate) to the to the 35% proportion of responders as defined in the "Success Criteria" (a description and rationale is provided in Section 4.7.1).

Secondary endpoints (with analyses to be performed at approximately 1 year, except for the withdrawal analysis) are as follows:

- The proportion of patients who meet categorical thresholds of 5%, 15%, 20%, 25%, 30%, 35% and 40% weight loss from baseline.
- Global hunger as assessed by the Patient Global Impression of Severity (PGIS) and the Patient Global Impression of Change (PGIC) in patients ≥ 12 years of age.
- [redacted]
- Global hunger as assessed by the Caregiver Global Impression of Severity (CGIS) and the Caregiver Global Impression of Change (CGIC) in patients 6 years old up to < 12 years of age.
- During withdrawal from drug: reversal of weight gain and hunger reduction during the double- blind placebo-controlled withdrawal period.
- Percent change in body fat and body mass as measured by either dual-energy x-ray absorptiometry (DEXA) or bio-electrical impedance analysis (BIA), site-specific.
- Percent change in BMI.
- Change in glucose parameters: fasting glucose, glycated hemoglobin (HbA1c), oral

glucose tolerance test (OGTT) with concomitant insulin measurements to focus on parameters of insulin sensitivity.

- Change in waist circumference (cm).
- Safety and tolerability of setmelanotide (including BP and HR) via proportions of patients with AE's and shift from normal range status in lab parameters.

[redacted]

2.7.2. [redacted]

[redacted]

2.7.3. Safety Variables

Safety and tolerability of setmelanotide will be assessed by frequency and severity of AEs as well as changes in physical examinations, electrocardiograms (ECGs), vital signs (including resting BP and HR), clinical laboratory evaluations, monitoring for changes in neurocognition and pubertal development, including bone age progression in patients ≤16 years of age and injection site reactions.

3. SUBJECT POPULATIONS

3.1. Population Definitions

There are 4 important populations from the RM-493-015 pivotal clinical trial that Rhythm intends to evaluate for efficacy and safety. One key population of this trial, which is the population that will be indicated for use of setmelanotide, is the one that includes a pre-defined weight loss at 12 weeks of treatment. *The rationale for this population is outlined below*. Because this population will be the designated population indicated for setmelanotide treatment, it is referred to as the Designated Use Set (DUS) Population.

Other key populations, defined below, include a comprehensive safety analysis set (SAS), the full analysis set (FAS), and a completer's set (CS) who will have completed the full 1-year clinical trial period.

3.1.1. Key Populations:

Full Analysis Set (FAS): all subjects who received at least 1 dose of study medication and had a baseline assessment. This population includes patients who do and *do not* demonstrate at least 5 kg weight loss over the 12-week Open-Label Treatment Period and proceed into the double-blind, placebo controlled withdrawal period. It also includes all patients who do not provide a week 12 observation.

- The **FAS population**, (or modified intent-to-treat population) will provide information on the effect of setmelanotide on all patients who initiate treatment and have at least a baseline assessment, regardless of whether the patient meets the threshold for continuing into the longer-term study following individualized dose titration and the 12-week therapeutic response period on the personalized therapeutic dose determined for each patient. This population will include patients who do not complete the study for any reasons (due to AEs, stop treatment due to failure to achieve the necessary ≥5 kg weight loss from baseline after 12 weeks of therapeutic treatment, etc.). Every effort will be made to continue following all patients (on treatment or not) through the full year of this pivotal trial to obtain the (approximately) 1-year observations.

Designated Use Set (DUS): The DUS is intended to define the population of patients who demonstrate at least 5 kg weight loss (or at least 5% if baseline body weight, especially in pediatric subjects, is < 100 kg) over the 12-week Open-Label Treatment Period and subsequently proceed into the double-blind, placebo controlled withdrawal period, regardless of later disposition.

- Thus, the **DUS population** includes only patients who achieve a qualifying response necessary for continued treatment after the initial therapeutic trial at each patient's individualized therapeutic dose (i.e., ≥ 5 kg weight loss, or ≥ 5% weight loss if baseline weight is < 100 kg, after the open-label run-in). This population will provide the most accurate efficacy evaluation data according to the intended use of setmelanotide in LEPR deficiency obesity (see larger rationale below).

Safety Analysis Set (SAS): all subjects who received at least 1 dose of study medication**.**

- The **SAS Population** will be the trial population defined by those who receive at least one dose of setmelanotide, and will be used for all evaluations of safety parameters.

Ultimately, this population may be identical to the FAS population depending on final patient experience in this trial. The SAS is the primary population for the analysis of [redacted] and safety endpoints.

Completers Set (CS): all subjects in the DUS who both demonstrate at least 5 kg weight loss over the 12-week Open-Label Treatment Period and continue in the trial on active treatment to complete a full year (approximately) of treatment.

- The **CS population** includes participants who complete the entire treatment period for this trial and remain on treatment at the last (approximately 1-year) treatment visit. The rationale for this population is that scientifically it best represents the effect of setmelanotide on weight loss and hunger in the LEPR deficiency patient population who demonstrate at least a 5 kg weight loss (or 5% if baseline weight is < 100 kg) by 12 weeks and receive an approximately full year treatment course for evaluation of the primary weight endpoint. Hence, the CS population provides an estimate of what would be expected from one year of treatment in any patient who demonstrates at target therapeutic response for weight loss by 12 weeks and is able to continue long-term setmelanotide therapy for approximately one year of treatment.

3.2. Rationale for the DUS Population as Population of Interest for Percent Change in Weight and Hunger over 1-year (Key Secondary Endpoints)

The **DUS** population will be the primary population for the analysis of key secondary efficacy parameters that evaluate mean change in both body weight and hunger from baseline. This population will also be applicable for the evaluation of these parameters in the assessment of success of this study (see Section 4.7.1).

The FAS population will be used for the primary efficacy analysis, as well as the analysis of all categorical efficacy endpoints for body weight and hunger score assessments (i.e., responder analyses).

All key parameters will also be evaluated for the FAS and CS populations.

[redacted]

Rhythm proposes that the **DUS population** represents the most accurate and scientifically compelling population to evaluate for the proposed indication and to characterize the quantitative magnitude of full efficacy data in the final US Package Insert (USPI). According to the clinical protocol, this population best defines the target population for labeling according to the intended use of setmelanotide. Therefore, this population will address the key scientific and

clinical question: "What is the expected effect of setmelanotide on body weight reduction and hunger suppression in the designated population when used as indicated in the label?" Hence, this cohort will be a vital analysis population for the SAP.

In contrast, the FAS population is not as suitable as an efficacy population for mean weight loss, for evaluations of mean change from baseline, for many important reasons:

- Most importantly, the FAS population may include patients who do not meet the 12-week rule for continuing treatment based on a ≥ 5 kg weight loss response OR who drop out early in the study. As a result, the FAS population would include a mixture of patients, both with and without actual treatment experience, because the FAS population includes all patients who enter the study through approximately one year (including subjects who drop out for any reason, or who do not pass the stipulated weight reduction response rule, or who do not proceed with active treatment past the 12-week Open Label Treatment Period). The resulting distribution of percent change in body weight is likely to be bi-modal, and provides significant loss of power for this endpoint with N=10 patients (the feasible limit for recruitment that is the current target of this protocol in such an ultra-rare indication). Table 6 and Table 7 of Appendix 5 provide estimated 1-sided confidence limits for a DUS and a simulated bi-modal FAS population, respectively. These tables show the large increase in the confidence interval (CI) half-width when a bi-modal distribution is simulated (as may occur in the FAS population).
- In addition, the use of FAS population for assessing mean efficacy changes from baseline may inadvertently bias against the intended effect of setmelanotide treatment. Based on prior case descriptions and limited natural history data for rare LEPR deficiency patients, LEPR deficiency patients do not necessarily remain stable but often actively gain weight without therapy. For example, ongoing weight gains of 5 to 10 kg per year are commonplace in fully characterized cases (described fully in the Rhythm Breakthrough Therapy Designation (BTD) application, submission 0041, November 2, 2015). Because the main weight endpoint after approximately 1 year of treatment in this pivotal trial is not placebo-compared, the inclusion in the FAS population of prospectively defined yet untreated patients (because they did not achieve this initial response threshold) would add subjects with highly divergent weight trajectories in the final efficacy evaluations. As a consequence, the analysis of this FAS population could be confounded and inappropriately underestimate the effects of setmelanotide treatment in the intended treatment population. The DUS population of patients who are intended to continue for long term treatment may also include subjects who discontinue for AEs or other reasons, but at least this population would not include subjects who were never intended to use setmelanotide long-term in the first place based on the failure to achieve 5 kg or more weight loss during an initial therapeutic dose test.
- The natural history of this disorder, with ongoing weight gains of +5 to +10 kg per year (with both a mean and median yearly weight gain of +7 kg; based on analysis of weight curves since birth for existing patients in our trials) in the absence of setmelanotide treatment, also clearly supports that the FAS population likely will be bimodal if any patients are discontinued early due to lack of early weight loss. This will likely result in non-normal distributions and the potential for some confusion of the treatment effect in the label.

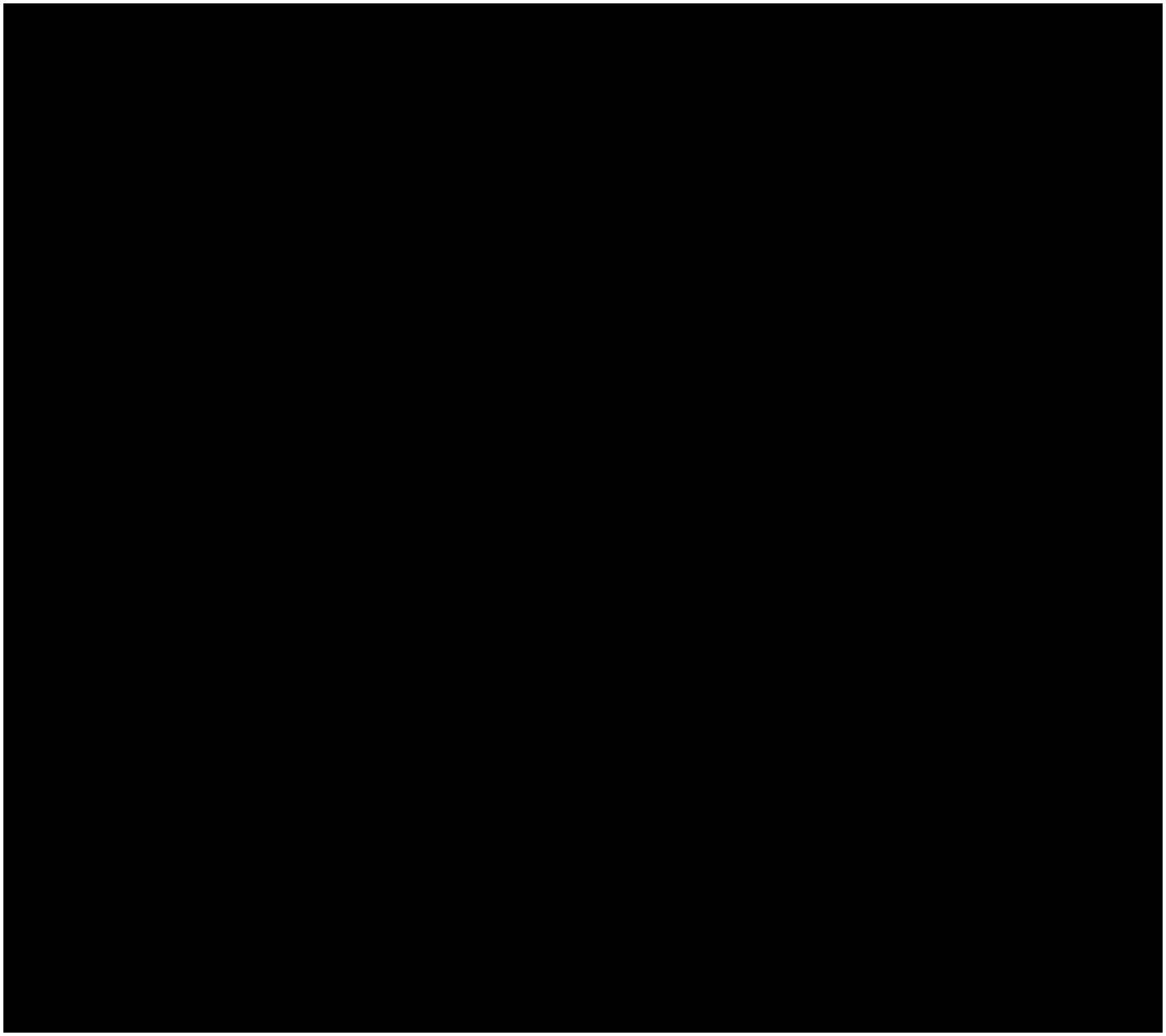
- While Rhythm is proposing that several key analyses for percent reduction in weight (first key secondary endpoint) and hunger (key secondary endpoint) will be assessed in the DUS population, for each of these endpoints we will also provide a sensitivity analysis to be performed in the FAS population, using a non-parametric rank analysis as described below, which should mitigate the issues raised by the occurrence of a bimodal distribution. We propose that this analysis be accompanied with a tabulation of the percent of patients who do not meet the ≥5 kg threshold (i.e., those who do not achieve the threshold for long-term dosing after the first 12-weeks of therapeutic dosing), to provide useful information needed in the label.

Thus, similar mean change from baseline analyses, albeit via non-parametric rank tests (specifically, Wilcoxon 1-sample signed rank tests comparing the mean weight percent change from baseline against 0%), will be conducted for the FAS population, as for the primary DUS population, along with disposition tables of subjects who did not meet the criteria for inclusion in either the DUS or the CS populations and the reasons why (e.g., unable to pass the ≥ 5 kg (5% in subjects with baseline body weight < 100 kg) stopping rule to be included in DUS, other lack of efficacy, withdrawal due to AEs, etc.). These factors will likely be the most informative data gathered from the FAS and CS populations, as it will provide information on patients who either do not respond to treatment or have issues with tolerating the study drug.

The following exploratory subject population will also be evaluated and used for presentation and analysis of the data:

Exploratory Population:

Per Protocol Set (PP): all subjects in the FAS without any major protocol violations (see Section 3.3).



3.4. Protocol Violations

At the discretion of the sponsor, major protocol violations as determined by a review of the data prior to unblinding of the study results and the conduct of statistical analyses may result in the removal of a subject's data from the PP Population. In this open-label study with only limited double-blind treatment data, decisions on protocol violators may be subject to bias, and hence only major deviations will be considered. The sponsor, or designee, will be responsible for producing the final protocol violation file (formatted as a Microsoft Excel file), in collaboration with Cytel and the data monitoring group as applicable; this file will include a description of the protocol violation, and clearly identify whether a violation warrants exclusion from the PP Population. This file will be finalized prior to hard database lock.

Major protocol violations may include, among others:

1. Violation of key inclusion/exclusion criteria, as reviewed by study clinician.

2. Violation of legal documents including, but may not be limited to informed consent and patient privacy.

All protocol violations will be presented in the data listings.

4. STATISTICAL METHODS

This SAP may not describe all analyses that will be performed for this study. It is expected that additional exploratory analyses may be requested by the Sponsor and performed to supplement the results outlined in this SAP, because the purpose will be to explore all possible data from the limited number of subjects. [redacted]

4.1. Sample Size Justification

The primary endpoint is proportion of patients who demonstrate at least 10% weight loss at ~1 year from baseline. The primary research hypothesis is that this proportion is at least 5%. The null hypothesis is that this proportion is at most 5%. Rationale for the 5% historical control value for this comparison is as follows.

The available data suggests that 0% of LEPR deficiency patients would demonstrate at least 10% weight loss in a single year, even with extraordinary measures. The exception might be after bariatric surgery, as some patients have had very short-term improvements, though over time their overwhelming hunger leads to weight regain. In the absence of such measures in our clinical trial (exclusions from the trial), and the general contraindication of bariatric surgery in this population, it is defensible to consider that no (0%) patients would meet these criteria and the statistical comparison for this endpoint should be a comparison to a 0% proportion. Despite this strong rationale based on the natural history of these patients, since the data are limited and the number of patients in the trial is small, it is probably more appropriate to go beyond this very conservative estimate of 0% responders, to do the historical control comparisons assuming that a small proportion of the population (i.e. 5%, or a null of 0.05) *might* show a 10% weight loss in any given calendar year; hence it is assumed a categorical comparison to a 5% responder estimate in untreated historical patients.

Assuming LEPR deficiency patients will be a total sample size of ~10 patients in the one year study, all power statements will assume n=10 patients in the study. The actual numbers for this ultra-rare disease may be larger or smaller.

It is expected that treatment with RM-493 for 1 year is associated with a TRUE underlying probability of at least 10% weight loss at 1 year of at least 50%. That assumption yields at least 94% power to yield a statistically significant (alpha=0.051-sided) difference from the null hypothesis 5% value for N=10 patients. If the TRUE probability of at least 10% weight loss at 1 year is 40%, then power is ~83%.

4.2. General Statistical Methods and Data Handling

4.2.1. General Methods

All output will be incorporated into Microsoft Word files, sorted and labeled according to the International Conference on Harmonisation (ICH) recommendations, and formatted to the appropriate page size(s).

Tabulations will be produced for appropriate demographic, baseline, efficacy, pharmacokinetic and safety parameters. For categorical variables, summary tabulations of the number and percentage within each category (with a category for missing data) of the parameter will be presented. For continuous variables, the mean, median, standard deviation, minimum and

maximum values will be presented. Formal statistical hypothesis testing will be performed on the primary and secondary efficacy endpoints with all tests conducted at the 1-sided, 0.05 level of significance, for a comparison of change from baseline. Note that no comparator population is included in this study.

However, the endpoint pertaining to the 8-week placebo withdrawal period will be analyzed and presented differently, as each patient will serve as their own control and data will be examined relative to both the preceding and the ensuing active treatment periods. Summary statistics will be presented, as well as confidence intervals on selected parameters, as described in the sections below.

4.2.2. Definition of Baselines, Study Periods, and Reporting

The last value (of weight, last week of mean hunger scores, etc.) obtained prior to the first dose of any active treatment in the open-label dose titration phase will be considered the baseline value for statistical analysis for the primary analyses for one-year data. Note that if the weekly average hunger score prior to the first administration of study drug is missing, the daily hunger score collected on Day 1 (first day of active dose) will be used as "Baseline."

For the specific analysis of the withdrawal period for the key primary weight analysis, the baseline will be the weight loss (kg) over the first 4-weeks of the double-blind, placebo-controlled withdrawal period (i.e., the weight change (loss) that that occurs over the last 4 weeks of active treatment), which will be compared to the weight change (gain) during the 4-weeks of placebo-controlled withdrawal.

For the specific analysis of the withdrawal period for key secondary supporting hunger analysis, baseline hunger ("most hunger over past 24-hours") will be considered the last evaluable week of hunger data obtained just prior to the beginning of the off-treatment 4-week placebo-controlled interval for each patient.

For other interval analyses (for example, the exploratory analysis of weight loss over the initial 12-week therapeutic dose treatment), the last value prior to the beginning of that interval will be considered the baseline.

Visit 13 will be considered the end of ~1 year of treatment. Therefore, the approximate one year of treatment includes the final 2 weeks of the dose titration period (Visit 2a, 2b, etc.), the 10 weeks of the open label treatment period (Visit 3 – Visit 5), the 8 weeks of the placebo-controlled withdrawal period (Visit 6 – Visit 7) and the final ~32 weeks of open label treatment (Visit 8 – Visit 13). The actual total time on treatment will be ~48 weeks, due to the 4 weeks of placebo withdrawal. The pre-specified visit window for Visit 13 will be 10-14 months of treatment.

Reporting of measurements in summary tables will be performed with all individual therapeutic doses combined due to small sample sizes. The primary focus for analysis will be on the impact at the end one year of treatment (i.e., at the end of all treatment intervals (titration, 12-week open label, 8-week double blind and 32-week open label) to complete the approximate one-year treatment period. Additional analyses will focus on the specific intervals separately (screening, titration, initial 12-open label treatment period, the 8-week double blind period, and the 32-week open label treatment period).

Supplemental exploratory analyses will be examined for data collected during the screening period, as well as the dose titration period to better understand baseline variability of LEPR deficiency subjects in a clinical trial, and to understand the therapeutic dose response during titration. Specifically, hunger and weight variability/stability or loss/gain, as well as patient medical history will be examined from before study entry (patient history from birth, if available), during the 4 weeks of screening and the 2-12 weeks of titration.

4.2.3. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software (Version 9.4 or higher), unless otherwise noted. Medical History and AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 20.0. Concomitant medications will be coded using World Health Organization (WHO) Drug version March 2016.

4.2.4. Methods of Pooling Data between Different Protocols for Supplemental Exploratory Analyses

It is expected that patients who continue to tolerate setmelanotide after approximately 1 year of treatment, and who show clinically important weight loss, will continue on long-term extensions (to be fully outlined in a future extension protocol, as noted in Section 5.5 of the protocol), on a year by year basis. Additional analyses will summarize the long-term experience as appropriate.

It is also expected that some LEPR deficiency obesity patients will be participating in a separate, essentially identical proof-of-concept protocol (Study RM-493-011) with a long-term extension phase, and possibly in other studies. After assessing study conduct of subjects in the RM-493-011 study and upon determining essentially identical treatment procedures to this current pivotal study, these sentinel patients from Study RM-493-011 will also be identified and included in a pooled supplemental analysis. Clinical data from these 3 patients in the RM-493-011 study will be summarized and pooled (based on treatment duration) with data from the RM-493-015 study patients in a supplemental analysis.

In addition, if any other LEPR deficiency patients are treated in any other Rhythm study, their data will be summarized and may be included as well in this supplemental analysis.

Appropriate statistical methods will be employed in order to take into account study differences in the pooled analysis. The following pooled analyses will be done:

1. Percent change from baseline in body weight (kg) and hunger score at the end of approximately 1 year of treatment: Although the placebo withdrawal period does not occur in the same design between the two studies, the data collected during this period will be included in the pooled analysis. It is expected that weight loss will taper toward the end of approximately 1 year on treatment, and both weight fluctuations from dose changes or 4 weeks on placebo will be negligible in the examination of body weight at this time point. Only morning hunger will be examined in the pooled analysis as global hunger (PGIS/CGIS, PGIC/CGIC) [redacted] Both hunger and body weight will be analyzed as described in Section 4.6.2. The pooled analyses will be performed in two ways. RM-493-011 does not have a nominal Visit 13; for both analyses, approximately 1 year in this study will be derived by relative day (study day in relation to first administration of study drug). However, there will be two approaches in pooling with the RM-493-015 study data: 1) using the nominal Visit 13 (or, for diary

daily hunger scores, the closest week to the nominal Visit 13) and 2) by relative day, where 1 year (Week 52) will be derived programmatically based on the date of first dose, and the closest nominal visit to this Week 52 derivation based on actual time spend on drug will be considered the 1 year timepoint.

2. Body weight (kg) and morning hunger score at the end of the titration period: As length of time spent during the titration period will differ between participants, rate of weight loss per week and its 95% CI will be calculated for this phase of the study. If more appropriate, this can be summarized in a figure rather than performing inferential statistical analysis. Change from baseline in hunger score will be analyzed as described in Section 4.6.2, with the inclusion of an adjustment for number of weeks in titration.
3. Descriptive statistics for all participants: Mean, standard deviation, median, Q1, Q3, minimum and maximum for weight, BMI and daily hunger will be examined for all patients across both studies. Participant plots will also be output for these 3 endpoints to examine trends across the course of approximately 1 year (and beyond) for both studies.

4.2.5. Adjustments for Covariates

For continuous endpoints, the only covariate considered will be the baseline measurement of the endpoint of interest. Additionally, as a sensitivity analysis, age and BMI, provided the sample size is adequate, may be considered as covariates in a sensitivity analysis for continuous endpoints. If the sample size does not support this, these characteristics may be examined in a subgroup analysis.

4.2.6. Multiple Comparisons/Multiplicity

Multiplicity for the primary analysis is not of concern as there is a single primary efficacy analysis as outlined in Section 4.6.1.

This study will consist of ~10 pivotal patients, and in many ways, the set of patients under study will constitute a collection of detailed clinical case reports with a comprehensive baseline and past medical history assessment and complete clinical efficacy, safety and laboratory evaluations conducted for each patient. Based on the ultra-rare prevalence of this disease, and the small number of patients to be enrolled in this study (representing a large proportion of the known, published patients with this disorder), the ability to use extremely rigorous statistical approaches to address multiplicity for many endpoints is limited. Therefore, for publication, nominal-p-values will be used to interpret *each endpoint separately* in this small study where sample size is limited by feasibility.

Rhythm acknowledges that this approach may increase the probability of potential Type 1 error for the *set* of endpoints being analyzed. This may be a regulatory concern, so we have pre-specified a step-down procedure to control Type-1 error, if needed for this purpose.

For the endpoints below, for regulatory evaluations, we control for multiplicity, *by stepping down in the following order*, if statistical significance is achieved in the primary endpoint:

1. The first key secondary endpoint is change from baseline in body weight at approximately 1-year in the DUS population.
2. The second key secondary endpoint is change from baseline in weekly “most hunger in the last 24 hours” hunger scores over approximately 1-year of treatment in the DUS

population.

3. The third key secondary endpoint is the categorical percent of responders analysis of hunger (using the 25% improvement in hunger threshold) in the FAS population at approximately 1-year.

Another key set of analyses is from the double-blind, placebo controlled withdrawal period, which serves as an internal control for the weight loss/hunger improvements in this study. Rhythm believes the analyses outlined for this withdrawal period are underpowered, and hence will be more supportive and descriptive, and therefore will not be adjusted for multiplicity.

No interim analysis is currently planned for this protocol, so no adjustment for multiplicity is needed for any interim look.

4.2.7. Subpopulations

The following subgroups will be examined, provided the sample size is adequate in each group:

- Age (6 – 12, 12 – 18 and 18+ years of age)
- Sex (male, female)
- Race (white, non-white)
- Site (US, Europe)
- Met stopping rule (< 5 kg weight loss, ≥ 5 kg weight loss; or 5% for subjects with baseline weight < 100 kg) – differences with respect to demographics and baseline characteristics
- BMI (< 50; ≥ 50 at baseline)

Subgroup analyses will be explored for the combined pivotal and supplemental cohort, these are described in Section 3.3.

4.2.8. Withdrawals, Dropouts, Loss to Follow-up

Subjects who withdraw from the study will not be replaced. However, due to the small sample size, every possible effort will be made to follow any subjects who may withdraw and discontinue treatment for the remainder of the study in order to examine both safety and efficacy measures. Similarly, for any patient who meets the stopping rule and withdraws from active therapy, every possible effort will be made to follow these subjects through the approximate1-year timepoint. These follow-up assessments will include all study visits and procedures, including weight and hunger assessments.

4.2.9. Missing, Unused, and Spurious Data

As visit windows are very flexible, we do not anticipate a significant amount of missing data. However, missing data for the primary and key secondary endpoints will be handled in the following ways:

- The primary method for handling missing primary/key secondary endpoint data at approximately 1 year will first examine the reason for missingness. If unrelated to treatment (e.g.: patient moved), the endpoint will either be extrapolated using a linear model (y=a+b*study week) based on existing data points or imputed using the

longitudinal mixed model for analysis. If the reason for missingness is directly related to treatment (lack of efficacy or an AE), weight change at approximately 1 year will be conservatively imputed as 0 kg. Likewise, hunger change at approximately 1 year will be imputed as 0. If less than 3 months of data are available for the supplemental patients at the time of first analysis of the combined cohorts, these patients will not be imputed and will be left out of the analysis until more data is available.

  - Example 1: A patient drops out of the trial after 6 months after experiencing an AE deemed related to treatment; their 1 year body weight change from baseline will be imputed as 0 kg and hunger change will be imputed as 0, as the reason for missingness is related to treatment (treatment-related AE).
  - Example 2: A patient drops out of the trial after 9 months due to moving out of the area; their 1 year body weight and hunger changes from baseline will be imputed using a linear model to extrapolate based on data up through their 9 months in the trial as the reason for missingness is not related to treatment (patient moved).
  - Example 3: A supplemental patient has 2 months of data available at the time of the first analysis of the combined cohorts (pivotal and supplemental); their 1 year body weight and hunger changes from baseline will *not* be imputed at that time. Data will be displayed in summary tables and listings, but this patient will not be included in a formal analysis of the primary and secondary endpoints.

- As sensitivity analyses, the following methods will also be examined, regardless of the reason for missingness of data:
  - Weight loss/hunger trends will be extrapolated to determine the weight at 1 year; a linear model will be fit to the individual patient's observed data.
  - For categorical endpoints, all patients not ongoing on test treatment and missing their data at 1 year will be considered 'failures,' for example: for the primary analysis, it will be assumed those missing 1 year data did not achieve at least a 10% weight loss from baseline. Ongoing patients will have weight/hunger imputed as described above.
  - Last observation available will be carried forward and considered as the weight/hunger at approximately 1 year.

For non-primary or key secondary endpoint analyses utilizing a longitudinal mixed model, if a visit it missed, an assumption is made that data will be missing at random, consistent with longitudinal analyses. The first key secondary analysis of the DUS population employs a longitudinal mixed model, and all available data will be included. However, because assuming all missing data is missing at random is too liberal, missing 1 year outcome data will be first imputed as described in the first bullet above (the reason for missingness will first be discerned). As a sensitivity analysis, the model will also be run with observed data; no manual or model-based imputation will be performed.

AE dates that are missing or incomplete are derived as follows:

- If the start date of an AE is partially or completely missing, the date will be compared as far as possible with the date of the start of administration of study drug. The AE will be

assumed to be treatment-emergent if it cannot be definitively shown that the AE did not occur or worsen during the treatment-emergent period (worst case approach).

The following general rules will be used:

- If the start date is complete, an AE will only be excluded as being treatment-emergent if the start date is before the date of study drug administration or if the stop date is before study drug administration.
- If the start day is missing but the start month and year are complete, an AE will only be excluded as being treatment-emergent if the start month/year is before the month/year of study drug administration or if the stop date is before study drug administration.
- If the start day and month are missing but the start year is complete, an AE will only be excluded as being treatment-emergent if start year is before the year of study drug administration or if the stop date is before study drug administration.
- If the start date is completely missing, an AE will be considered treatment-emergent unless the stop date is before study drug administration.

4.2.10. Visit Windows

As this patient population is very rare, visit windows will not be employed in the analysis, i.e., data will not be excluded from the analysis based on whether the data fell outside of visit windows. However, the goal will be for visits to occur within +/- 3 days during the Dose Titration phase, and within +/- 4 weeks during the 32-week Open Label Treatment Period. For the one year timepoint, the 1-year visit data will be included in the primary and key secondary endpoints if it occurs between 10 and 14 months after the start of the study. Throughout this SAP, 'approximately 1 year' refers to this 1-year timepoint data collection, which can occur during months 10 – 14.

4.3. Data Verification Procedures during Screening and the Titration Period

It is important to verify data handling for proper handling and compliance. Therefore, an early look of the screening and initial clinical data (only during the initial titration period) in raw form (including individual patient listings but not summarized in tables) will be taken based on enrollment and planning. Data from the screening period will be examined when 2-3 participants have completed this stage of the study as discussed in Section 4.2.2 to examine missing data percentages, completeness of past weight and growth data, including pediatric weight curves, past medical history, associated disorders, and pre-existing concomitant hormonal therapies and other medications. Additionally, vital signs, medical history, hunger and weight data will be examined once 2-3 patients finish their initial titration. Special attention will be devoted to the completeness of dosing titration and therapeutic level decision making at the end of this titration period based on early weight and hunger change data and BP or HR data recorded during titration clinic visits. Combined or exploratory results will not be allowed to be viewed by clinical site staff or other individuals who have clinical influence in reporting the study results.

No further efficacy data reviews will be conducted from the time each patient is decided to have attained a therapeutic dose and starts the 10 weeks of personalized therapeutic level dosing followed by the double-blind, placebo-controlled withdrawal period.

4.4. Subject Disposition

A tabulation of subject disposition will be tabulated, including the number screened, the number dosed with setmelanotide, the number in each subject population for analysis, the number that withdrew prior to completing the study, and reasons for withdrawal.

A by-subject listing of study completion information, including the reason for premature study withdrawal, if applicable, will be presented.

4.5. Demographic and Baseline Characteristics

Baseline, demographic and medical history information will be summarized for the DUS, CS and FAS populations using descriptive statistics. No formal statistical comparisons will be performed.

Demographic and baseline data will be provided in data listings.

4.6. Efficacy Evaluations

Efficacy analyses will be conducted primarily in the FAS and DUS populations as specified below; analyses in the CS population will also be performed for key endpoints (if it differs from the DUS), but analyses in the PP Population will be performed only if the need is indicated by the number of protocol violations; specifically, only if the number of subjects in the PP population differs from the FAS by more than 15% on the primary endpoint, then the primary endpoint and key secondary endpoints will be analyzed in the PP population as well. All efficacy analyses will be performed in both pivotal cohort, supplemental cohort, and combined pivotal and supplemental cohorts (refer to Section 3.3 for more information on the supplemental cohort). Note that the primary and key secondary endpoints will be conducted in the FAS, DUS, and CS populations.

For each binary endpoint, a tabular summary will be provided; this will include number and proportion of subjects who meet the categorical criteria and binomial test results (p-values and two-sided 90% CIs).

For each continuous endpoint, a tabular summary will be provided; this will include summary statistics (N, mean, SD, median, minimum and maximum) and model results (p-values and two-sided 90% CIs).

Efficacy endpoints being examined as percent change from approximately 1 year will primarily be defined using the nominal Visit 13. For diary data (daily hunger scores), the closest week to the nominal Visit 13 will be determined and used as the approximate 1 year timepoint. As a sensitivity analysis, all primary and key secondary endpoints will also be analyzed by relative day (study day in relation to first administration of study drug). Specifically, Week 52 will be derived programmatically based on the date of first dose, and the closest nominal visit to this Week 52 derivation based on actual time spend on drug will be considered the 1 year timepoint. Refer to Section 4.2.10 regarding visit windows.

All efficacy endpoints will also be displayed in by-patient listings.

Weight and daily hunger data will be graphed in both overall and by-patient plots; by-patient plots will include historical data prior to the start of study. Cumulative proportion of responder plots will also be shown for a range of thresholds/cutoffs that encompass all patients' values.

4.6.1. Primary Efficacy Evaluation

The primary endpoint is the proportion of patients in the FAS population who achieve at least a 10% weight loss from baseline at approximately 1 year threshold compared to the null hypothesis that 5% of patients will achieve this threshold. This will be analyzed via the exact binomial test which will test whether the percentage of patients who reach at least 10% weight loss is greater than 5%. Two-sided, 90% CIs will be calculated using the exact Clopper-Pearson method.

Once the primary endpoint analysis is complete, and if it reaches statistical significance, the proportion of patients in the FAS population who show ≥10% weight loss at approximately 1 year will be numerically compared (point estimate) to the 35% proportion of responders as defined in the "Success Criteria;" no formal statistical analysis will be performed.

Rationale for 1-sided alpha=0.05: This study will consist of ~10 pivotal patients, and in many ways, the set of patients under study will constitute a collection of detailed clinical case reports with a comprehensive baseline and past medical history assessment and complete clinical efficacy, safety and laboratory evaluations conducted for each patient. Based on the ultra-rare prevalence of this disease, and the small number of patients to be enrolled in this study (representing a large proportion of the known, published patients with this disorder), the power for extremely rigorous statistical approaches is limited. Based on the marked weight loss previously demonstrated in this population, Rhythm has chosen 1-sided, alpha=0.05 as the scientific approach, which we identify as the primary analysis. However, as this may not meet regulatory standards for approval, a 1-sided, alpha=0.025 comparison may also be required.

By-patient data listings and plots of weight across the course of the study will also be presented.

4.6.2. Key Secondary Efficacy Endpoints

Key secondary efficacy endpoints will be analyzed in the DUS population (percent change from baseline in body weight, "most hunger in the past 24-hours"), and in the FAS population (categorical analysis for a threshold of ≥25% improvement in hunger scores). Each of these will be separately numerically compared (point estimate) to the "Success Criteria" for this these endpoints (a description and rationale is provided in Section 4.7.1).

All key secondary efficacy endpoints will assess the effect the effect of setmelanotide over the course of approximately 1-year.

These analyses will be supported by the pre-specified similar analyses in all populations (DUS, FAS, and CS populations), and for hunger, by the additional hunger assessments (morning hunger, average hunger).

Specifically, daily hunger will be assessed on individual patients ages 12 and above, with support from their parent/caregiver, using a hunger questionnaire [redacted]

[redacted]

Prior to analysis, daily hunger scores for each of the 3 hunger assessments will be averaged separately by week. For a week of hunger scores to be considered evaluable, scores need to be recorded and available for analysis on at least 1 of 7 days to provide sufficient data to determine mean values. In order to look at impact of missing hunger data, a sensitivity analysis will be done to examine if there are any differences in conclusions drawn by data available; primary and key secondary analysis for daily hunger scores will be repeated for patients with at least 3 of 7 days of data present, 4 of 7 days of data present, 5 of 7 days of data present, 6 of 7 days of data present, and 7 of 7 days of data present.

Percent change from baseline in body weight (kg) at the end of approximately 1 year of treatment in the DUS population will be assessed as the first secondary efficacy evaluation. If statistical significance is achieved for this endpoint, the mean percent change from baseline at the end of approximately 1-year of treatment in the DUS population will be numerically compared (point estimate) to the ≥10% mean change from baseline in body weight "Success Criteria" for this study (a description and rationale is provided in Section 4.7.1).

The protocol specifies that the weight measurement at any given visit will be measured at the clinic in triplicate, and averaged of the 3 measurements taken as the visit weight.

A linear mixed model repeated measures analysis of variance with a fixed term for time and baseline weight and a random effect for subjects will be used to assess the primary efficacy endpoint at approximately 1-yeart. All weight measurements obtained during the study will be included in the model and missing data will be handled as described in Section 4.2.9. An unstructured covariance matrix will be used to model the expected different variances among the participants. In the event the mixed model does not converge with an unstructured covariance matrix; a compound-symmetric then Toeplitz covariance matrix will be employed instead. Additionally, a paired t-test will be derived from the model and compared to no change (0% mean percent weight change from baseline) and will use Satterthwaite's degrees of freedom estimates (1-sided, compared to alpha of 0.05 for the primary success of this endpoint; a comparison to an alpha of 0.025 will also be provided).

The assumption of normality will be visually assessed via a visual diagnostic of closeness to normality, not as a conditional test associated with efficacy endpoint analyses. Graphical assessments of residuals from the model fit may be examined. If a substantial departure from normality is observed, a transformation such as log (post/pre) or rank or other non-parametric test (such as 1-sample Wilcoxon signed rank test versus a 0% change from baseline) may be used to analyze the data as a sensitivity analysis for continuous endpoints; however, the analysis on the original scale of observation will be reported.

Once this analysis is complete, and if it is statistically significant, the mean percent change from baseline at the end of 1-year of treatment in the DUS population will be numerically compared (point estimate) to the ≥10% mean change from baseline in body weight; no formal statistical analysis will be performed.

Results from the linear mixed model repeated measures analysis of variance will be presented in a table.

Rhythm is proposing that this first secondary endpoint also be performed in the FAS population, using a non-parametric 1-sample Wilcoxon signed rank approach (comparing the mean weight percent change from baseline against 0%), which should mitigate the issues raised by the occurrence of a bimodal distribution. We propose that this analysis be applied as a sensitivity analysis, along with a tabulation of the percent of patients who do not meet the ≥5 kg threshold, to provide useful information needed in the label.

The second key secondary endpoint is the mean percent change from baseline in weekly average hunger (using "most hunger over the last 24 hours" daily response) in the DUS population (for patients 12 years of age and older). This endpoint will be measured as described above and will be analyzed in a similar manner as the key secondary endpoint. Specifically, it will be analyzed using a linear mixed model for repeated measures analysis of covariance with weekly average hunger percent change from baseline as the outcome and fixed terms for time and baseline hunger and a random effect for subject. "Hunger in the morning" and "average hunger over a 24-hour period" will be analyzed in the same manner as additional secondary endpoints.

Once the second key secondary analysis is complete, and if it is statistically significant, the mean percent change in hunger from baseline at the end of approximately 1-year of treatment in the DUS population will be numerically compared (point estimate) to a 25% mean change from baseline in hunger; no formal statistical analysis will be performed.

As a sensitivity analysis, the second key secondary endpoint will also be analyzed for the FAS population, using a non-parametric 1-sample Wilcoxon signed rank approach (specifically, Wilcoxon 1-sample signed rank tests comparing the mean weight percent change from baseline against 0%), which should mitigate the issues raised by the occurrence of a bimodal distribution.

The third key secondary endpoint is the proportion of patients in the FAS population who achieve at least a 25% hunger improvement from baseline threshold compared to the null hypothesis that 5% of patients will achieve this threshold at the end of approximately 1-year of treatment. This will be analyzed via the exact binomial test which will test whether the percentage of patients who reach at least 25% hunger improvement is greater than 5%.

Once the third key secondary analysis is complete, and if it reaches statistical significance, the proportion of patients in the FAS population who show ≥25% hunger improvement will be numerically compared (point estimate) to the 35% proportion of responders at the end of approximately 1-year of treatment as defined in the "Success Criteria;" no formal statistical analysis will be performed.

4.6.3. Secondary Efficacy Endpoints

Secondary efficacy endpoints will be analyzed first for the DUS population, to be supported by the pre-specified similar analyses in the FAS and CS populations; other populations will be analyzed as indicated.

Secondary objectives are to assess the effect of setmelanotide over the course of approximately 1 year on the following:

- The proportion of patients who meet categorical thresholds of 5%, 15%, 20%, 25%, 30%, 35% and 40% weight loss from baseline.
- Global hunger as assessed by the Patient Global Impression of Severity (PGIS) and the Patient Global Impression of Change (PGIC) in patients ≥ 12 years of age.
- Hunger and/or hyperphagia assessments in patients 6 years old up to < 12 years of age (to be enrolled when Amendment 6, the pediatric age amendment, is implemented).
- Global hunger as assessed by the Caregiver Global Impression of Severity (CGIS) and the Caregiver Global Impression of Change (CGIC) in patients 6 years old up to < 12 years of age.
- Percent change in body fat and body mass as measured by either DEXA or BIA (site- or sub-study specific).
- Percent change in BMI.
- Change in glucose parameters: fasting glucose, HbA1c, OGTT with focus on parameters of insulin sensitivity.
- Change in waist circumference (cm).
- Safety and tolerability of setmelanotide (including BP and HR).

Secondary objectives are to assess the effect of setmelanotide during the double-blind, placebo-controlled withdrawal period as follows:

- During withdrawal from drug: reversal of weight and hunger reduction during the double-blind placebo-controlled withdrawal period.

Safety and Tolerability

Frequencies of AEs and injection site reactions, as well as percentage of abnormal ECGs, vital signs, laboratory evaluations, glucose parameters and [redacted] in patients ≤16 years of age will be examined in order to assess safety and tolerability of the treatment. Additionally, shifts between each study visit where the above is collected and the baseline measurement will be tabulated. Please see Section 4.9 for more detail regarding summary of safety measurements.

[redacted]

[redacted]

Data plots and displays will be aligned by week according to the timing of when the therapeutic dose for each individual participant was achieved; thus, each patient graphic display should include a total of 12 weeks of open-label therapeutic dosing following a variable number of titration steps for each patient. Thereafter, the 8-week double-blind placebo or active periods will be depicted, followed by the remaining open-label treatment period on active therapeutic level dosing.

Patient Global Impression of Severity (PGIS) and Change (PGIC)

Additionally, two global hunger questions will also be administered to patients ≥ 12 years of age to assess static and current states of hunger, PGIS and PGIC, at screening (question 1 only), Visits 6 and 7 during the placebo withdrawal period, Visit 8, 9 and 13 during the second open label treatment period and the final visit. These categorical measurements identify level of hunger at the time the question is asked (question 1) as well as in comparison with prior to study start (question 2). PGIS (question 1) data will be transformed into a numeric index ranging from 0-3, with higher scores being more severe. For the raw categorical data, shift tables between each study visit where PGIS is collected and the baseline measurement will be tabulated. As there is no baseline comparator for PGIC and the question itself asks to compare hunger prior to study start, this data will be displayed in a table of frequencies for each visit and no formal statistical inference will be done. Numeric data for PGIS will be analyzed in the same manner as the key secondary endpoint, described in Section 4.6.2, in order to examine change from baseline at approximately 1 year.

Two global hunger questions, the Caregiver Global Impression of Severity (CGIS) and the Caregiver Global Impression of Change (CGIC), will also be administered to the parent or caregiver of patients 6-11 years of age to assess their perceptions of the child's hunger or feed-seeking behavior (both current status and change from baseline). These data will be examined in the same manner described above for the PGIS and PGIC.

Body Fat, Body Mass, BMI and Waist Circumference

Body fat and body mass will be measured by either DEXA or BIA, based on what is available at each study site, at screening, Visit 6 and Visit 13. BMI will be calculated for each individual based on their weight at each time point and the last height measurement available. Analyses for body fat, body mass, BMI and waist circumference will employ similar methods as the key secondary efficacy analysis as described in Section 4.6.2, with change from baseline at approximately 1 year including both absolute kg change from baseline and percent change from baseline.

Glucose Parameters

Glucose parameters include fasting glucose, HbA1c and an OGTT with focus on parameters of insulin sensitivity derived from insulin assay measurements. Glucose and HbA1c are collected at screening, Visits 2a, 4, 6, 9, 11, 13 and the final visit. An OGTT with concurrent insulin levels at all time points will be done at screening, Visit 6 and Visit 13. The main analysis will be on the area-under-the-curve (AUC) of insulin and glucose, compared to baseline. In addition, the glucose and insulin data will be presented in tables illustrating shifts between each study

visit starting from the baseline measurement, and tabulations of abnormal measurements will be shown. These data will also be displayed in a figure (glucose and HbA1c plots over time versus study week, plus OGTT plots with glucose and insulin levels over time during the OGTT). Additional analyses such as HOMA (homeostatic model assessment) measures of insulin sensitivity may be included as additional descriptive statistics.

Analysis of Reversal of Weight and Hunger Reduction during the Double-Blind Placebo-Controlled Withdrawal Period

This study includes a single, double-blind, placebo-controlled period lasting 8 weeks in duration as described in Section 2.1 and shown in Figure 1. The sequence of this double-blind 8-week period is masked to patients and clinic site personnel, but follows the same sequence in all patients: continued active therapy for 4 weeks (weeks 13-16), followed by the 4-week placebo withdrawal (weeks 17-20), then re-treatment at the therapeutic dose, which for this analysis includes just the first four weeks after restarting open-label setmelanotide (weeks 21-24). Retreatment at the therapeutic dose continues long term for the remainder of the study to ~ 1 year. The focused analysis of this on-off-on treatment period is to be comprehensive and assess weight and hunger changes preceding, during, and following the double-blind withdrawal period, and therefore weight and hunger changes will also be assessed for this analysis for the first 4 weeks of open-label re-treatment following the placebo withdrawal in every patient.

The key endpoint for weight during this period will be the absolute difference between change in weight (kg) over the last 4-weeks of active treatment (i.e., nominally the weight loss during weeks 13-16) and the change in weight during the 4-week placebo-controlled withdrawal period (i.e., nominally weeks 17-20). This will be examined using summary statistics (N, mean, median, min-max) for the absolute change in weight during this period. If the withdrawal period is <4 weeks, the data will be normalized to a 4-week period.

Supplemental analyses for weight will include summary statistics for the weight change in each of the three periods described above: the last 4-weeks of active treatment before placebo withdrawal, the 4-week placebo-controlled withdrawal period, and the first four weeks of active re-treatment.

The key endpoint for hunger during this period will be the absolute difference (in score) from the last active treatment measurement *prior* to the withdrawal phase and the last hunger assessment measurement at end of the withdrawal phase. Specifically, this will be the absolute change in mean weekly "most hunger in the last 24 hours" hunger score for the last evaluable week prior to the withdrawal of active treatment (nominally the data from the full week 16), compared to the mean weekly hunger score for the last week of the 4-week withdrawal period (nominally the data from the full week 20). Similar supporting analyses will be performed for the other hunger assessments (e.g., daily morning hunger score). This will be examined using summary statistics (N, mean, median, min-max) for the absolute change in weight during this period. If the withdrawal period is <4 weeks, the last hunger value off of treatment in this period will be considered the end value for the withdrawal period (e.g., if the withdrawal period last 3 weeks, the last weight value on placebo withdrawal will be used in this analysis).

We do not have good estimates for the magnitude or variability of effects of a withdrawal period, such as included in this protocol, as such procedures have not been included in previous protocols. The only guiding information was an open label withdrawal for the first patient in the Phase 2 study, where the one patient gained 4.8 kg over the first ~2-3 weeks of the withdrawal

period, and her hunger scores increased from 0-1 to 7 points. The magnitude of effects on weight are limited by the length of the withdrawal period (i.e., 4 weeks) and therefore, the change in weight is likely to be small in additional patients. For hunger, the magnitude of effect might be larger, as shown by changes in hunger over short periods of time (i.e., weeks) for the patients in Study RM-493-011. However, as shown in Appendix 5, the variability of hunger scores is quite large. As a result, any statistical test is likely to be underpowered relative to the overall variability of weight and hunger change in this study. Therefore, the key weight and hunger data will be the summary statistics as outlined above; a paired t-test will also be employed to determine if weight and hunger differences (between the two periods described above) are different from zero. All comparisons using the t-test will be within-patient comparisons.

Weight and weekly hunger scores (actual values and change for each individual will be tabulated, presented in listings and displayed in a figure across time for this 12-week period. A plot of weights (by week) and hunger scores (by mean weekly scores) will also be presented with each of the three 4-week periods labelled. Changes pre- and post-withdrawal will be examined to evaluate possible hysteresis of weight and hunger changes during these transitions off and on therapy in each patient. Additional exploratory analyses may be examined, including calculation of the rate of weight gain and hunger reversal during the 4 weeks on placebo to illustrate how quickly weight is gained and hunger reappears once treatment is halted. In addition, changes in PGIC and PGIS will be summarized.

Other Secondary Analyses

Evaluations by interval will also be performed for key weight and hunger parameters. Changes from baseline during screening, dose titration, open-label treatment periods (initial 12-week; later 32 week) and the withdrawal phase will all be looked at individually. The model described in Section 4.6.2 will be implemented, for each phase of the study separately, so the following will be likely examined: change in weight between baseline and end of dose titration, baseline and first open-label treatment period, baseline and double-blind withdrawal period and finally, baseline and second open-label treatment period.

By-patient data listings of other secondary endpoints will also be presented.

[redacted]

4.7. Rationale for Key Elements of the SAP

4.7.1. Definitions of Success for Clinical Benefit in the Assessment of Efficacy for the Primary and Secondary Endpoints

This section defines what results would need to be seen in this study for it to be regarded as a success. The first key step in the definition of success is to reach statistical significance on the primary and key secondary endpoints. However, statistical significance is necessary but not sufficient, and it is important to consider whether the data are clinically meaningful in the context of this indication. This section provides a guide for a threshold for clinical meaningfulness for this indication, for the primary and key secondary endpoints.

Due to several contributing factors -- a) the small size of this study, b) the potential for wide confidence limits (especially in some populations of interest, such as the FAS, which may be a bi-modal population), c) the lack of experience in this ultra-rare orphan population, and d) limited data on the natural history of this disease (to determine good baselines in historical data) -- a rigorous statistical approach to clinically meaningful bounds is not proposed. Specifically, for the thresholds proposed below, once statistical significance is achieved, a numerical comparison of the results (point estimate) of the key endpoints will define overall success in this study.

Most importantly, the SAP provides trial success criteria for the primary endpoint, which will determine the success of this study. All data and rationales provided for key secondary endpoints are to provide context only.

Context for the Assessment and Definition of Trial Success

For an open-label pivotal clinical trial in a limited number of rare patients, Rhythm understands and agrees with the FDA perspective that designating a prospective definition of trial success is

critical. Rhythm considers three components to be crucial to understand in order to provide context for setting this prospective success target.

First, it is clearly important to consider the FDA guidance for general obesity products that has guided development of current obesity medications is also informative. Key points from the FDA guidance and relevant general obesity clinical trial data are briefly provided in the next section to highlight this perspective.

Second, to help gauge what is observed in clinical trials that led to FDA approvals, it is also important to consider the magnitude of weight changes demonstrated in pivotal clinical trials that led to the approval of current general obesity products. Relevant general obesity clinical trial data are also briefly described to highlight this perspective, including how relatively small in magnitude such differences between active therapy compared to placebo have been, and how difficult it is for general obese patients to maintain any long-term weight reduction from baseline over time. The modest long-term efficacy of current obesity therapies has been one of their major limitations.

Finally, it is useful to emphasize the relentless progression and refractoriness of the early-onset, extreme obesity that develops in LEPR deficiency obesity (often +5% to +10% body weight gain during each year of life, estimates based on limited weight curves from infancy in LEPR patients; see figure below). These aspects were described comprehensively in the BTD application and are described only briefly here. Given the therapeutic challenges in the general obesity population, it is not at all surprising that diet plus exercise is almost uniformly futile in LEPR deficiency or other monogenic obesity disorders characterized by stronger major gene effects than seen in general polygenic human obesity.

FDA Draft Guidance on Developing Products for Weight Management: Mean Differences and Categorical Responses for Weight Change from Baseline

The FDA Draft Guidance from 2007 cites a 5% efficacy benchmark as the target minimum percent decrease for weight loss agents seeking approval for broad use in general common obesity, noting that "in general, a product can be considered effective for weight management if after 1 year of treatment the following occurs: The difference in mean weight loss between the active-product and placebo-treated groups is at least 5 percentage points and the difference is statistically significant." In addition to setting this target value for mean difference in weight loss, categorical weight responses are also cited as important to consider for approving weight loss therapies. The following goals for categorical weight reduction response are noted, namely that "the proportion of subjects who lose greater than or equal to 5 percent of baseline body weight in the active-product group is at least 35 percent, is approximately double the proportion in the placebo-treated group, and the difference between groups is statistically significant."

Rhythm considers these two interrelated weight reduction targets -- the first based on a 5% mean reduction vs. placebo after approximately 1 year of treatment and the second based on categorical 5% weight reduction responses from baseline among the treated population, again over approximately 1 year – *as the appropriate paradigms (in approach, if not in magnitude) to set targets* for determining success of setmelanotide in treating LEPR deficiency obesity in the pivotal RM-493-015 clinical trial.

Clinical Data Over Long-Term Dosing / Reference Points from General Obesity Clinical Trials for Setting Success Criteria:

The magnitude of mean weight reduction achieved in general obesity clinical trials is also informative for setting an appropriate success target for setmelanotide in LEPR deficiency. Five New Chemical Entities (NCEs) or combination products are currently approved by FDA for chronic weight management. These agents and their weight reduction impact vs. placebo after approximately 1 year of chronic dosing consist of orlistat (2.9 – 3.4 kg; 2.9 – 3.4%), lorcaserin (3.6 kg, 3.6%), phentermine / topiramate fixed-combination (6.6 - 8.6 kg; 6.6 – 8.6%), naltrexone / buproprion (4.8%) and liraglutide injection (5.8 kg). These medicines were approved by FDA between 1999 through 2014 (Apovian, 2015). Overall, the weight change demonstrated by these therapies above diet and lifestyle alone after approximately 1 year on active treatment ranges from 2.9% to 8.6% weight reductions. [The specific kg values and/or % ranges for these agents are included adjacent to the individual products noted above; note that in nearly all trial populations the mean baseline weights for the general obese adult populations studied were approximately 100 kg, leading to similar kg and percent weight reductions (Apovian, 2015)]. These clinical data provide a framework of the magnitude of mean weight loss changes achieved by current obesity medication approved for use in general human obesity. As such, they also inform Rhythm's proposal to define a success target for weight loss with setmelanotide over a similar time course for chronic treatment of approximately one year in LEPR deficiency obesity.

LEPR Deficiency: Major Gene Effects Conferring Early-Onset Extreme Obesity

Overall, the size of the LEPR deficiency obesity patient population to be enrolled and evaluated in this study is very small (N ~ 10 pivotal patients). LEPR deficiency is one of several monogenic forms of obesity impacting the leptin – melanocortin pathway that causes extreme, early-onset obesity, often beginning in infancy and progressing in severity throughout childhood and into adolescence and adulthood without any abatement; detailed clinical and genetic features of LEPR deficiency obesity have been previously submitted to FDA as part of the prior BTD application (submission Sequence 0071, March 27, 2017).

The pivotal study represents a small, open-label trial employing a "personalized medicine" type of dosing and treatment approach, with every patient is part of the DUS population needing to qualify for long-term treatment by demonstrating ≥ 5 kg weight reduction from baseline (or ≥ 5% weight reduction if baseline weight was < 100 kg) during an initial individualized therapeutic dosing and test period. The expected clinical results for these rare genetic obesity patients was considered by the FDA and the Sponsor to represent a set of informative case reports grouped in this nature because large clinically significant treatment effect sizes are expected for both weight reduction and hunger score improvements.

Success for the Primary and Key Secondary Endpoint: Interrelated Weight Reduction Goals Equal to Twice the FDA Guidance Targets

- **10% OBSERVED Percent Mean Weight Loss from Baseline in the DUS Population**
- **35% of Patients Achieve a 10% Weight Reduction from Baseline in the FAS Population**

Rhythm's general approach to defining success was to use the paradigm from the FDA guidance (defining a percent mean weight loss target and/or a percent of patients achieving a clinically important threshold) but including a greater magnitude of effect of setmelanotide in this rare disease population.

Rationale for ≥10% mean weight loss from baseline in the DUS population.

Rhythm's approach to "success" in this trial parallels the FDA guidance, but sets a higher threshold (≥10%) for all of the criteria.

Rhythm proposes that a ≥10% OBSERVED mean weight reduction (point estimate) from baseline to approximately 1-year of treatment and statistically significant difference from zero (one-sided alpha=0.05) in the DUS population should serve as one of the primary criteria for achieving the overall study success of the RM-493-015 pivotal trial.

To substantiate that the 10% change from baseline is a robust threshold (for both the mean and categorical approaches) in the absence of a parallel placebo control arm, Rhythm will collect detailed past growth / development curves and baseline weight histories for every LEPR deficiency patient. Accordingly, a success goal for weight loss set at ≥ 10% from baseline represents a valid and robust goal for the success of these pivotal clinical trial.

In some ways, Rhythm believes this is a conservative approach. Based on past medical / natural history data for these rare LEPR deficiency patients demonstrating long-standing refractoriness to diet, exercise, other non-specific weight loss therapies, the expected result after approximately 1-year untreated (and presumably also after approximately 1-year of placebo) would be a substantial weight *gain*. While natural history data for these patients is very limited, and our study may provide more natural history data from the patients being enrolled (not yet available), the data from the first three LEPR patients (Figure 2) shows 5-10% weight gain for almost every year of life (unless very active efforts were made, in which case weight sometimes could be stabilized for short periods). Hence, a comparison to baseline and requiring a substantial weight loss is a conservative comparison, and we hypothesize that a 10% mean weight loss might represent 15-20% relative weight loss for essentially all of this group of patients.

Ten percent mean weight loss from baseline is also designated as a target success threshold because this magnitude of mean weight loss (either as a mean change vs. placebo) represents a clinically meaningful threshold, and a level of response not previously attained among therapeutic weight loss clinical trials of either oral or injectable approved weight loss medicines in a general obesity population (Apovian, 2015). Thus, the achievement of a ≥ 10% weight reduction from baseline in LEPR deficiency obesity would be expected to be both statistically significant and clinically significant, especially in this severe, early-onset obesity patient population suffering from unremitting hunger. In fact, the relentless progression of extreme weight gain amounting to 5 – 10 kg per year beginning early in childhood in the majority of genetically identified cases makes this goal highly impactful in such patients. There are very few, if any, cases where conventional diet and exercise management and/or non-specific weight loss therapies can provide lasting weight reduction. Thus, the impact of 10% weight loss can be significant on co-morbidities and quality of life in general common obesity and it is likely to be as or possibly more impactful in this early-onset, extreme form of genetic obesity.

Thus, the efficacy threshold Rhythm has designated for setmelanotide in LEPR deficiency obesity is approximately > 2 times higher than the efficacy goal cited by FDA for general obesity (with caveats). Rhythm proposes that this ≥10% response from baseline weight threshold is rigorous and appropriate for the genetically defined and extreme LEPR deficiency obesity population, especially in the absence of any alternative therapies or interventions for this rare genetic form of extreme obesity.

Rationale for at least 35% of patient attain a ≥10% mean weight loss from baseline in the FAS population.

The paragraphs above help to establish the "≥10% weight loss" vs. baseline as a clinically meaningful threshold for this indication, and this rationale also applies to a categorical analysis. For this analysis, it is more appropriate to consider the FAS population to understand the full efficacy of setmelanotide in this categorical approach.

Therefore, analogous with FDA guidance but with greater percent weight loss as a threshold, and understanding the desire for "intention to treat" analyses, we also propose that an analogous categorical response criterion could also be set for success of RM-493-015, namely, that at least 35% of patients from the FAS population demonstrate ≥10% weight reduction from baseline after approximately 1-year on treatment.

Here, the complex issue is how to compare vs. historical controls. The available data suggests that 0% of LEPR deficiency patients would demonstrate at least 10% weight loss in a single year, even with extraordinary measures. The exception might be after bariatric surgery, as some patients have had very short term improvements, though over time their overwhelming hunger leads to weight regain. In the absence of such measures in our clinical trial (exclusions from the trial), and the general contraindication of bariatric surgery in this population, it is defensible to consider that no (0%) patients would meet these criteria and the statistical comparison for this endpoint should be a comparison to a 0% proportion. Despite this strong rationale based on the natural history of these patients, since the data are limited and the number of patients in the trial is small, it is probably more appropriate to go beyond this very conservative estimate of 0% responders, to do our historical control comparisons assuming that a small proportion of the population (i.e. 5%, or a null of 0.05) *might* show a 10% weight loss in any given calendar year;

hence we propose a categorical comparison to a 5% responder estimate in untreated historical patients.

Statistical tables to support the categorical weight analysis (Table 4 and Table 5 of Appendix 5) show the p-values to reject the null hypothesis for a given observed number of successes, and the power for this endpoint, respectively.

Success for Secondary Efficacy Endpoints – Hunger Score Change from Baseline

Besides extreme weight gain, pronounced and unrelenting hunger / hyperphagic feeding behaviors represent the most debilitating symptoms that plague patients with early-onset extreme obesity caused by LEPR deficiency. Hunger and increased appetite symptoms can be even more challenging when patients and parents are not able to convince pediatricians and other health professionals of their impact, especially as hunger is a normal daily experience in nearly everyone that has converted to a pathological extreme, almost addictive, symptom in LEPR deficiency patients. Therefore, adult and pediatric age range efficacy assessments relying on self-reported hunger scores and/or parent/caregiver observer-related outcome (ORO) assessments represent important secondary endpoints for the RM-493-015 pivotal trial (and exploratory endpoints for patients age 6 through age 11 years given challenges in ORO assessments of this symptom). [redacted]

Nonetheless, in recognition of their importance for daily symptoms and health impact in LEPR deficiency patients, Rhythm proposes that mean reduction in "most hunger in the last 24 hours" hunger scores from baseline ≥ 25% over approximately 1-year would define a prospective percent reduction that would represent a clinically meaningful response on incessant and debilitating hunger symptoms in LEPR deficiency obesity patients.

While not a criterion for success, robust supporting data from the initial 12 weeks of therapeutic treatment, as well as a robust magnitude of change upon double-blind, placebo-controlled withdrawal would provide additional corroboration.

For this endpoint, it is especially important to emphasize that success is a numerical comparison of the mean changes in hunger assessments. At present, we have almost no variability data to support a power analysis for this endpoint, and only limited data to estimate a magnitude. As such, this endpoint might be under powered. Using all current data, an updated estimate of the 1-sided confidence limits for hunger percent change from baseline is included in Table 6 of Appendix 5. This table shows that the half-width of the confidence interval for hunger change from baseline might exceed 70 percentage points, and Table 9 (power table, Appendix 5) supports that for this study, a true underlying mean percent change from baseline of -32.5% achieves 80% power (alpha=0.05) to yield an OBSERVED difference that achieves statistical significance (-38% for an alpha=0.025).

[redacted]

[redacted]

[redacted]

4.8. [redacted]

[redacted]

4.9. Safety Analyses

Safety analyses will be conducted using the SAF Population and data will be collected as described in Appendix 1.

4.9.1. Adverse Events

AEs will be coded using MedDRA and displayed in tables and listings using System/Organ/Class (SOC) and Preferred Term.

Analyses of AEs will be performed for those events that are considered treatment emergent, where treatment emergent is defined as any AE with onset after the administration of study medication through the end of the study, specifically, 30 days after the last dose of setmelanotide, or any event that was present at baseline but worsened in intensity or was subsequently considered drug-related by the investigator through the end of the study.

AEs are summarized by subject incidence rates, therefore, in any tabulation, a subject contributes only once to the count for a given AE (SOC or preferred term).

The number and percentage of subjects with any treatment-emergent AE, with any treatment-emergent AEs assessed by the Investigator as related to treatment (definite, probable, or possible relationship), and with any SAE will be summarized by treatment group and overall. In these tabulations, each subject will contribute only once (i.e., the most related occurrence or the most

intense occurrence) to each of the incidence rates in the descriptive analysis, regardless of the number of episodes.

No formal hypothesis-testing analysis of AE incidence rates will be performed.

All AEs occurring during the study will be listed in subject data listings.

By-subject listings also will be provided for the following: subject deaths; SAEs; and AEs leading to withdrawal.

An additional listing will include onset and resolution dates, verbatim term, preferred term, treatment, severity, relationship to treatment, action taken, and outcome will be provided.

Summary tables and figures (by-patient) illustrating frequency of AEs of special interest will also be examined.

4.9.2. Laboratory Data

Clinical laboratory values will be expressed using SI units.

Safety laboratory data will include CBC with platelet count and standard indices, a chemistry panel (includes sodium, potassium, chloride, CO_2, albumin, total protein, glucose, BUN, creatinine, uric acid, AST, ALT, GGT, CPK, alkaline phosphatase, total bilirubin, direct bilirubin, LDH, calcium, phosphorus), urinalysis if positive findings on dipsticks warrant further examination, a coagulation profile (prothrombin time [PT] or international normalized ratio [INR], and partial thromboplastin time [PTT]), fasting samples where feasible, [redacted] and HbA1c. These will be collected at screening, visits 2a, 4, 6, 9, 11, 13 and the final visit.

The actual value and change from baseline (as defined by the last measurement obtained prior to the first dose of the active treatment during the open-label dose titration phase) to each on study evaluation will be summarized for each clinical laboratory parameter, including hematology and clinical chemistry. In the event of repeat values, the last non-missing value per study day/time will be used. Screening values will also be included in these tables and figures.

A shift from baseline table will be presented (N and %) for the following categories, where each patient is counted in one cell only. Missing data will not be imputed. There will be no inferential analysis.

- Hematology
- Chemistry

All laboratory data will be provided in both data listings, as well as figures, which will illustrate all measurements taken in both the pre-dose (screening) period and treatment period.

A subset listing will be presented for all clinically significant and abnormal laboratory values.

4.9.3. Vital Signs and Physical Examinations

Vital sign data will include BP, HR, body temperature, respiration rate and will be collected at all visits.

The actual value and change from baseline (as defined by the last measurement obtained prior to the first dose of the active treatment during the open-label dose titration phase) to each on study evaluation will be summarized for vital signs.

Repeat measures or more frequent monitoring can be implemented for significant increases in BP or HR.

By-subject listings of vital sign measurements will be presented in data listings.

Physical examination results at each time point will be summarized; shifts from baseline in physical examination findings to each on study visit will also be presented. All physical examination findings will be presented in a data listing.

If greater than 4 patients participate in the 24-hour [redacted] sub-study, the data will be summarized; otherwise, individual data will be presented.

4.9.4. Electrocardiogram

ECG data will be collected at screening, visits 2a, 6, 9, 13, and the final visit.

Results will be summarized descriptively, including the number and percent of subjects with normal, abnormal and clinically significant abnormal results at baseline and each study visit. In particular, analysis of QTc intervals will be computed both for mean changes from baseline, as well as categorical analyses.

All ECG data for each subject will be provided in data listings.

4.9.5. Concomitant Medications

A review of concomitant medications will be conducted during the screening period and at every study visit. Any medications taken by study patients will be recorded in source documents and on the appropriate CRF.

Concomitant medications will be coded using the WHO Drug dictionary. Results will be tabulated by Anatomic Therapeutic Class (ATC) and preferred term.

The use of concomitant medications will be included in by-subject data listing.

6. REFERENCES

Apovian CM, Aronne LJ, Bessesen DH, McDonnell ME, Murad MH, Pagotto U, Ryan D, Still C. Pharmacological Management of Obesity: An Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metab. 2015, 100(2): 342-362.

Kuehnen P, Clement K, Wiegand S, Blankenstein O, Gottesdiener K, Martini LL, Mai K, Blume-Peytavi U, Gruters A, Krude H. Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. NEJM 2016; 375: 240-246.

7. CLINICAL STUDY REPORT APPENDICES

Appendix 1: Schedule of Assessments

Table 1 - Screening and Dose Titration[#]

Study Period	**Screening**	**Open Label Dose Titration[6]**
Procedure Visit Number (V)	**V1**	**V2a[6]**
Start of Dose Titration Week (Dose Titration Study Day ± 3 days)	**-4 to 0 (-28 to -1)**	**1 (1)**
Informed consent/Assent	X	
Inclusion/Exclusion review	X	X[5]
Medical history review	X	X[5]
[redacted]		
Pregnancy test	X	X[4, 5]
Physical examination[1]	X	
Height[1]	X	
Comprehensive skin exam[2]	X	
Fitzpatrick scale	X	
Open label placebo practice	X	
Dose Titration Decision[24]		X[5]
[redacted]		
Weight/waist circumference[11]	X	X[5]
Archive sample for storage[12]	X	
Study treatment administration[13]		X
Injection site inspection[14]		X
Vital signs[15]	X	X[5,16]
ECG (12-lead)[17]	X	X[5, 17]
Safety laboratory tests[18]	X	X[5]
OGTT[28]	X	
[redacted]		
Daily Hunger Questionnaire[20]	X	X[5]
Global Hunger Questions[33]	X	
[redacted]		
Body Composition[23]	X	
[redacted]		
Anti-RM-493 antibody samples	X	X[5, 26, 31]
Adverse Event assessment[25]	X	X
Concomitant meds review	X	X
Telephone contact		X

Optional Sub-Studies

Table 2 - 10 Week Active Treatment and 8 Week Double-Blind Placebo Controlled Withdrawal

Study Period	**Open Label**			**Double Blind**	
	Active Treatment[7]			**Placebo-Controlled Withdrawal[7]**	
Procedure Visit Number (V) Start of Week (Study Day)	**V3 3 (15)**	**V4* 5 (29)**	**V5 9 (57)**	**V6 13 (85)**	**V7 17 (113)**
Pregnancy test	X[4,5]	X[4,5]	X[4,5]	X[4,5]	X[4,5]
Physical examination[1]				X	
Height[1]				X	
Comprehensive skin exam[2]					
[redacted]					
Weight/waist circumference[11]	X[5]	X[5]	X[5]	X[5]	X[5]
Archive sample for storage[12]				X[5]	
Therapeutic Dose Established	X				
Study treatment administration[13]	X	X	X	X	X
Injection site inspection[14]	X	X	X	X	X
Vital signs[15]	X[5]	X[5]	X[5]	X[5]	X[5]
ECG (12-lead)[17]				X[5]	
Safety laboratory tests[18]		X[5]		X[5]	
OGTT[28]				X[5]	
[redacted]					
Daily Hunger Questionnaire[20]	X[5]	X[5]	X[5]	X[5]	X[5]
Global Hunger Questions[33]				X[5]	X[5]
[redacted]					
Body Composition[23]				X	
[redacted]					
Anti-RM-493 antibody samples		X[5, 26]		X[5, 26]	
Adverse Event assessment[25]	X	X	X	X	X
Concomitant meds review	X	X	X	X	X
Telephone contact	X	X	X	X	X

Optional Sub-Studies

[redacted]

Table 3 - Additional 32 Week Open Label Treatment

Study Period	**Open Label**						**Early Termination / Final Visit[9]**
	Active Treatment[8]						
Procedure Visit Number (V) Start of Week (Study Day)	**V8 21 (141)**	**V9 27 (183)**	**V10* 33 (225)**	**V11 39 (267)**	**V12* 45 (309)**	**V13 53 (365)**	
Pregnancy test	X[4, 5]	X[4,5]	X[4,5]	X[4, 5]	X[4,5]	X[4, 5]	X
Physical examination[1]		X				X	X
Height[1]		X		X		X	
Comprehensive skin exam[2]				X		X	X
[redacted]							
Weight/waist circumference[11]	X[5]	X[5]	X[5]	X[5]	X[5]	X[5]	X
Archive sample for storage[12]						X[5]	
Study treatment administration[13]	X	X	X	X	X	X	
Injection site inspection[14]	X	X	X	X	X	X	
Vital signs[15]	X[5]	X[5]	X[5]	X[5]	X[5]	X[5]	X
ECG (12-lead)[17]		X[5]				X[5]	
Safety laboratory tests[18]		X[5]		X[5]		X[5]	X
OGTT[28]						X[5]	
[redacted]							
Daily Hunger Questionnaire[20]	X[5]	X[5]	X[5]	X[5]	X[5]	X[5]	X
Global Hunger Questions[33]	X[5]	X[5]				X[5]	X
[redacted]							
Body Composition[23]						X	
[redacted]							
Anti-RM-493 antibody samples		X[5, 26]		X[5, 26]		X[5, 26]	X[5, 26]
Adverse Event assessment[25]	X	X	X	X	X	X	X
Concomitant meds review	X	X	X	X	X	X	X
Telephone contact	X	X	X	X	X	X	X

Optional Sub-Studies

[redacted]

Schedule of Assessments Footnotes

1 A complete physical examination will be conducted at screening and at the end of study. [redacted] Height will be measured during the Screening Period only for those patients ≥18 years of age. Height will be measured according to the SOA for those patients <18 years of age.

2 A comprehensive skin evaluation will be performed by a dermatologist. Any concerning lesions identified during the screening period will be biopsied and results known to be benign prior to first dose of setmelanotide. If the pre-treatment biopsy results are of concern, the patient will be excluded from the study.

3 Optional sub-studies for Investigative Sites that are able to carefully perform these assessments and have patients willing and able to participate. In the event a site can perform multiple sub-studies, care needs to be taken in order to not overburden with multiple additional assessments/visits.

4 Urine pregnancy test may be performed in order to expedite availability of results prior to dosing on Dose Titration Day 1. All other pregnancy tests will be serum; dosing may continue with results pending.

5 Prior to study drug administration.

6 The Dose Titration phase will be a variable schedule lasting a minimum of 2 weeks and a maximum of 10 weeks, in which patients will return to the clinic approximately every 2 weeks in order to establish the individual patient's therapeutic dose according to Appendix 11.8. Given the variable number of dose titration steps in the Dose Titration Phase, each Dose Titration Visit Number (V) will remain V2 with an alphabetized suffix added to each titration visit (i.e.; first dose titration at start of Week 1 = V2a, second dose titration at start of Week 3 = V2b, etc.). This will allow for the Visits to be appropriately tracked. Additionally, each dose titration visit will have the same pre and post dose assessments as outlined in the SOA (with the exception of Anti-RM-493 antibodies and [redacted]).

7 Once the patient's individual therapeutic dose is established the patient will enter the Open Label Active Treatment phase for 10 additional weeks, for a combined total of 12 weeks of dosing at the therapeutic dose. During this time, the study calendar will be reset, starting when the therapeutic dose was initiated (i.e. the last 2 weeks of dose titration when the therapeutic dose was established). Therefore, the Open Label Active Treatment phase starts at the beginning of Week 3 (V3). Patients losing 5 kg of weight at the end of the Open Label Treatment phase will enter the Double-Blind Placebo Controlled Withdrawal phase lasting 8 weeks.

8 Upon completion of the Double Blind Placebo Controlled Withdrawal phase, patients will resume Open Label Active Treatment for an additional ~32 weeks.

9 Early Termination: For those patients who withdraw consent or are withdrawn and not willing to complete the remaining study visits, the early termination visit assessments should be performed, when possible. Additionally, patients who withdraw and are not willing to return for the remaining clinic visits can be contacted via phone, if amenable, to collect self-reported patient data (i.e.: weight, hunger, AEs, etc.). Final Visit: For patients who complete the study but do not wish to enroll into the future long term extension study (as noted in Section 5.5), patients will be required to return for a Final Visit ~30 days after the last dose of setmelanotide, for a final follow-up safety assessment. Any ongoing AEs reported at this visit should be monitored as outlined in Section 7.4. For patients who enroll into the long term extension study, this visit is not required.

10 [redacted]

[redacted]

11 Weight is to be measured at the clinic using the same scale after patients have emptied their bladder and while fasting. Patients are to wear light clothing or underwear, no shoes, and will be weighed at approximately the same time of day. Weight measurements are to be done in triplicate; waist circumference will be single measures.

12 Extra retain samples will consist of 2 serum and 2 plasma (K2EDTA) vacutainer tubes.

13 Study drug is administered by patients/caretakers beginning the morning of Day 1 and for the duration of dosing. Patients/caretakers will draw up and self-administer/administer the drug once on a daily basis in the morning. On days with clinic visits, the patients/caretakers will administer the drug in the clinic in the presence of the clinical staff to assure proper technique. Patients/caretakers will return all used vials to the clinic when they visit (the number recorded) and both clinic administered study drug, as well as outpatient study drug administration will be recorded in a study diary.

14 Injection site evaluations and scoring (by the clinical staff) will include identification and measurement of areas of erythema, edema and induration, as well as the presence of localized pain, tenderness and itching. Additional evaluation data can be collected at any visit where there are injection site reactions even if not a timepoint for formal assessment.

15 All BP and HR measurements are to be obtained in the sitting position following at least 5 minutes of rest. All measurements will be taken in triplicate, approximately 2 minutes apart. When possible, BP should be taken in the non-dominant arm throughout the study, using the same methodology (automated or manual) according to Appendix 11.7. Body temperature (°C) and respiration rate (breaths/minute) will be obtained in the sitting position following at least 5 minutes of rest.

16 During Day 1 and for any dose titration, vitals will be collected prior to dosing and then approximately hourly post-dose for up to 8 hours.

17 A single 12-lead ECG will be performed in the supine position following a period of at least 10 minutes of rest. On days in which dose titration occurs, measures will be obtained prior to dosing and approximately 8 hours post-dose

18 Safety laboratories will include: CBC with platelet count and standard indices, chemistry panel (includes sodium, potassium, chloride, CO_2, albumin, total protein, glucose, BUN, creatinine, uric acid, AST, ALT, GGT, CPK, alkaline phosphatase, total bilirubin, direct bilirubin, LDH, calcium, phosphorus), urinalysis with microscopic analysis if positive findings on dipsticks warrant further examination. Safety laboratories shall also include a coagulation profile (prothrombin time [PT] or international normalized ratio [INR], and partial thromboplastin time [PTT], also referred to as activated partial thromboplastin time [aPTT]. Fasting samples (8 hour minimum) are required at all time points where feasible. [redacted] panel and HbA1c will also be included.

19 [redacted]

20 Daily hunger questionnaire scores will be assessed by asking the patient to score their hunger [redacted] Daily hunger questionnaire scores will be recorded on a daily basis, prior to the patient's morning meal.

21 [redacted]

[redacted]

22 [redacted]

23 Body composition may be performed using an appropriate method available at sites (e.g. BIA, DXA, etc.) Refer to Section 6.3.3 regarding appropriate methodology for assessing this patient population.

24 Once all the pre-dose assessments have been performed, the decision to dose titrate according to Appendix 11.8 will be made. If the patient's therapeutic dose has been established, the patient will transition into the 10-week Open Label Active Treatment Phase, receive their therapeutic dose, and complete the V3 post-dose assessments as defined in the SOA. If the patient's therapeutic dose has not been established according to Appendix 11.8, the patient will be administered study drug, complete the dose titration post-dose assessments as defined in the V2 SOA, and return to the clinic in ~2 weeks for the next sequential Visit 2 (i.e.; V2b, V2c, etc.).

25 Adverse events will be recorded from the time a patient provides informed consent. AEs reported after dosing on Day 1 will be considered treatment-emergent AEs.

26 Any patients with positive anti-drug antibodies will be followed ~every 3 months until titers resolve or return to baseline.

27 Telephone contact by site on a monthly basis, or more frequently, if needed.

28 Following collection of pre-meal (time 0) blood samples, patients will be given a standard oral glucose tolerance test. The following blood samples will be obtained during each OGTT: Blood glucose and insulin at approximately 30, 60 90 and 120 minutes after meal start.

29 ~8 hours post-dose on Day 1 only.

30 ~8 hours post-dose.

31 To be collected on the first two, two-week dose titration visits (V2a and V2b [or V3, if the patient's therapeutic dose has been established]).

32 A blood sample will be obtained at Screening for genotyping for mechanisms considered to be possibly related to the safety or efficacy response to the study medication (e.g., other obesity related genes).

33 [redacted]

34 [redacted]

be measured only if BP increases are noted).

Once a patient's therapeutic dose has been established according to Appendix 11.8, no further dose titrations will occur, and patients will transition directly into the 10 week Open Label Active Treatment phase. Therefore, the Dose Titration phase will be a minimum of 2 weeks and a maximum of 10 weeks.

* For patients that reside a considerable distance from the clinic, these visits are optional clinic visits, and may be performed by the patient's local physician or home health care professionals.

+ If the visit is performed by a home health care professional, assessment may be performed at the next schedule in clinic visit.

§ For patients previously enrolled, the assessment should be performed at the patient's next scheduled visit.

Appendix 2: Supplemental Statistical Analyses supporting the primary and key endpoints.

Table 4 - Calculating p-values for the categorical (binary endpoint) proportion of patients with (i) weight loss from baseline of ≥10% at 1-year, or (ii) proportion of patients with ≥25% improvement in hunger score at 1-year

In each case comparing to null hypothesis value that assumes 5% (0.05) of *untreated* patients with weight loss from baseline of ≥10% (≥25% improvement in hunger score) at approximately 1-year (for N=10 patients total)

For total sample size N=10 (i.e., the FAS population), the table below shows p-values to reject the null hypothesis in favor of the alternative that proportion exceeds 0.05 for a given observed number of successes using a 1-sided exact binomial test:

No. of patients who achieve weight loss ≥ 10% OR 25% improvement in hunger score	P-value
1	0.401
2	0.086
3	0.012
4	0.001
5	0.00006
6	0.000003
7	<0.000001
8	<0.000001
9	<0.000001
10	<0.000001

Table 5 - Power for 1-sample unconditional exact binomial test of % responders versus null hypothesis value 5% (weight and hunger) for alpha=0.05 and 0.025, 1-sided

For total sample size N=10, the table below shows the power for variety of TRUE underlying proportions (0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9) for test versus 0.05.

TRUE underlying proportion	Table entries are power to reject indicated null hypothesis and conclude TRUE proportion is higher	
	null hypothesis rate=0.05	
	alpha=0.05 1-sided*	alpha=0.025 1-sided*
0.2	0.322	0.322
0.3	0.617	0.617
0.4	0.833	0.833
0.5	0.945	0.945
0.6	0.988	0.988
0.7	0.998	0.998
0.8	>0.999	>0.999
0.9	>0.999	>0.999

*Note that alpha = 0.05 and 0.025 yield the same power because of the discrete nature of the response.

Table 6 - Calculating 1-sided upper confidence limits for mean weight percent change from baseline

Using <u>all</u> prior data of all possible patients (2 POMC, 3 LepR, 5 Bardet-Biedl) from Studies RM-493-011 and -014 (i.e., using N=10 patient's data for variability estimates).

The purpose of this analysis is to use all data available, including patients with other upstream MC4 pathway disorders, to estimate variability and to provide the most realistic evaluation of the half-width of the confidence interval for percent change in weight from baseline at approximately 1-year. Note that this data includes these patients, all of whom show important weight loss, and assumes all will complete approximately 1 year of treatment. Please compare with Table 7, which assumes a more realistic FAS-type of population, where a small proportion of patients will not be responders or will not complete approximately 1-year of treatment.

Prior assumptions to calculate SD: Total sample size N=10, with various follow-up times. Using a linear mixed effects model with week as a fixed effect and patient as a random effect, we can extrapolate to 44 weeks of weight loss (exclude 8 weeks of placebo controlled phase). The estimated percent change is -21.53% from baseline and the conservative estimate of standard deviation is 8.79%. 1-sided confidence limit uses all data available. SD of 8.79 implies an observed range of 4x8.79=35.2 for the 10 weight change values. The actual observed range is equal to 34 so this seems to be a reasonable approximation.

Sample size	T-statistic for 95% CI	T-statistic for 97.5% CI	95% 1-sided upper confidence limit using SD =8.79	97.5% 1-sided upper confidence limit using SD = 8.79
5	2.13	2.78	18.72	24.44
6	2.02	2.57	17.76	22.59
7	1.94	2.45	17.05	21.54
8	1.89	2.36	16.61	20.74
9	1.86	2.31	16.35	20.30
10	1.83	2.26	16.09	19.87

Table 7 - Calculating 1-sided upper confidence limits for mean weight percent change from baseline

Using prior data of several patients (2 POMC, 3 LepR, 5 Bardet-Biedl) from Studies RM-493-011 and -014 (i.e., using N=7 patient's data for variability estimates), then including fabricated data from N=3 participants who gain at least 5% of weight from baseline. This represents what is likely to be seen for the FAS population.

Like Table 6, the purpose of this analysis is to use all data available, including patients with other upstream MC4 pathway disorders, to estimate variability and to provide the most realistic evaluation of the half-width of the confidence interval for percent change in weight from baseline at approximately 1-year. However, unlike Table 6, this analysis models a more realistic FAS-type population that includes N=7 treated patients who respond and complete approximately 1-year of treatment, along with N=3 patients who demonstrate at least 5% weight gain from baseline at approximately 1-year (either because they are non-responders at 12 weeks and are discontinued from treatment, or because of early drop-out for other reasons). This results in a bi-modal distribution, and markedly larger estimates of the half-width of the CI.

Prior assumptions to calculate SD: Total sample size N=10, with various follow-up times. Using a linear mixed effects model with week as a fixed effect and patient as a random effect, we can extrapolate to 44 weeks of weight loss (exclude 8 weeks of placebo controlled phase). The estimated percent change is -12.55% from baseline and the conservative estimate of standard deviation is 16.46%. 1-sided confidence limit uses 7 true patient's data, 3 patients fabricated data. SD of 16.46 implies an observed range of 4x16.46=65.8 for the 10 weight change values. The actual observed range is equal to 42.9 so there is no need for a conservative estimate. Note that a pooled analysis of 3 non-responders and 7 responders is likely a bi-modal distribution; hence, analysis via the usual normal distribution may not be appropriate. Hence, the precision indicated in the table below may be overly conservative since use of a rank transformation or an analysis of proportion of responders may be more appropriate and more precise.

Sample size	T-statistic for 95% CI	T-statistic for 97.5% CI	95% 1-sided upper confidence limit using SD = 16.46	97.5% 1-sided upper confidence limit using SD = 16.46
5	2.13	2.78	35.06	45.76
6	2.02	2.57	33.25	42.30
7	1.94	2.45	31.93	40.33
8	1.89	2.36	31.11	38.85
9	1.86	2.31	30.62	38.02
10	1.83	2.26	30.12	37.20

*Note that the 3 fabricated had assumed % change from baseline of 5.88%, 7.41% and 10%.

Table 8 - Calculating 1-sided upper confidence limits for mean hunger percent change from baseline

Using prior data of some patients (2 POMC, 3 LepR, 3 Bardet-Biedl) from Studies RM-493-011 and -014 (i.e., using N=8 patient's data used for variability estimates).

The purpose of this analysis is to use all data available, including patients with other upstream MC4 pathway disorders, to estimate variability and to provide the most realistic evaluation of the half-width of the confidence interval for percent change in hunger at approximately 1-year. Note that this data includes patients all of whom show important improvements in hunger, and assumes all will complete approximately 1 year of treatment.

Prior assumptions to calculate SD: Total sample size N=8, with various follow-up times. Using a linear mixed effects model with week as a fixed effect and patient as a random effect, we can extrapolate to 44 weeks of hunger (exclude 8 weeks of placebo controlled phase). The estimated percent change is -70.52% from baseline and the conservative estimate of standard deviation is 38.11%. 1-sided confidence limit uses all data available. SD of 38.11 implies an observed range of 4x38.11=152.4 for the 8 hunger change values. The actual observed range is equal to 50 so there is no need for a conservative estimate.

Sample size	T-statistic for 95% CI	T-statistic for 97.5% CI	95% 1-sided upper confidence limit using SD = 38.11	97.5% 1-sided upper confidence limit using SD = 38.11
5	2.13	2.78	81.17	105.95
6	2.02	2.57	76.98	97.94
7	1.94	2.45	73.93	93.37
8	1.89	2.36	72.03	89.94
9	1.86	2.31	70.88	88.03
10	1.83	2.26	69.74	86.13

Table 9 - Power for 1-sample mean percent change from baseline versus null hypothesis value 0% (weight/hunger percent change from baseline) for alpha=0.05 and 0.025, 1-sided

For total sample size N=10, the table below shows the power for variety of TRUE underlying mean percent change from baseline of -2% through -12% for test versus 0% (i.e., focused on weight percent change from baseline; similar information about hunger percent change from baseline is included in the footnote).

TRUE underlying mean percent change from baseline	Table entries are power to reject indicated null hypothesis			
	null hypothesis percent mean percent change from baseline in weight=0, assuming SD=8.79		null hypothesis mean percent change from baseline in hunger=0, assuming SD=38.11	
	alpha=0.05 1-sided	alpha=0.025 1-sided	alpha=0.05 1-sided	alpha=0.025 1-sided
-2	0.164	0.094	0.068	0.035
-3	0.259	0.161	0.079	0.041
-4	0.377	0.251	0.091	0.048
-5	0.507	0.363	0.104	0.056
-6	0.636	0.487	0.118	0.065
-7	0.751	0.612	0.134	0.075
-8	0.843	0.727	0.152	0.086
-9	0.909	0.821	0.170	0.099
-10	0.952	0.892	0.190	0.112
-11	0.977	0.939	0.212	0.127
-12	0.990	0.969	0.235	0.143

Note about power for hunger scores: *For hunger*, alpha=0.05, the minimal OBSERVED difference to achieve statistical significance is -21.5 mean percent change from baseline. The TRUE underlying mean percent change from baseline of -32.5% achieves 80% power to yield an OBSERVED difference that achieves statistical significance. -38.5% achieves 90% power. *For hunger*, alpha=0.025, the minimal OBSERVED difference to achieve statistical significance 50% power is -26.5 mean percent change from baseline, -38% TRUE underlying difference achieves 80% power, and -44% achieves 90% power.