Statistical Analysis Plan

A Phase 3, randomized, double-blind, placebo-controlled, multi-center trial to assess efficacy and safety of octreotide subcutaneous depot (CAM2029) in patients with acromegaly

Short Title:	A trial to assess efficacy and safety of octreotide subcutaneous depot in patients with acromegaly
Trial Code:	HS-18-633
EudraCT Number:	2019-001191-11
NCT Number:	NCT04076462
Document Date:	08-Feb-2023
Development Phase:	Phase 3
Investigational Medicinal Product:	Octreotide subcutaneous depot (CAM2029)
Indication:	Acromegaly
Sponsor:	Camurus AB

[redacted]

STATISTICAL ANALYSIS PLAN (SAP)

A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTI-CENTER TRIAL TO ASSESS EFFICACY AND SAFETY OF OCTREOTIDE SUBCUTANEOUS DEPOT (CAM2029) IN PATIENTS WITH ACROMEGALY

Protocol version: 5.0

Effective Date:	**Version:**
08 Feb 2023	**4.0**

TABLE OF CONTENTS

1. HISTORY OF CHANGES

All versions of the document have to be documented in the table of history of changes. The owner of the document has to act in accordance with the latest version of the document.

Status	Version	Effective Date (dd.mm.yyyy.)	Change description
[redacted]			

Status	Version	Effective Date (dd.mm.yyyy.)	Change description
[redacted]			

Status	Version	Effective Date (dd.mm.yyyy.)	Change description
[redacted]			

2. ABBREVIATIONS

Acronyms, Abbreviations, Initials	Description / Explanation
[redacted]	[redacted]
AcroQoL	*Acromegaly Quality of Life Questionnaire*
AE	*Adverse event*
AIS	*Acromegaly Index of Severity*
ANCOVA	*Analysis of covariance*
ATC	*Anatomical Therapeutic Chemical*
ATG	*Autogel*
BS	*Biostatistics*
CI	*Confidence interval*
Core Review Team	*Consists of Project Manager, Head of Drug Safety Unit, Clinical Data Manager, Biostatistician*
$C_{ss,trough}$	*Trough concentration during a dosage interval at steady state*
CT	*Computerized tomography*
CTCAE	*Common Terminology Criteria for Adverse Events*
CTR	*Clinical Trial Report*
ECG	*Electrocardiogram*
eCRF	*Electronic case report form*
EOT	*End of Trial*
EQ-5D-5L	*EuroQoL 5-dimension 5-level*
FAS	*Full analysis set*
GH	*Growth hormone*
HCV	*Hepatitis C virus*
ICH	*The International Council for Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use*
IGF-1	*Insulin-like growth factor-1*

Acronyms, Abbreviations, Initials	Description / Explanation
IMP	*Investigational medicinal product*
ITT	*Intention-to-treat*
LAR	*Long-acting release*
LLOQ	*Lower limit of quantification*
MCMC	*Monte Carlo Markov Chain*
MedDRA	*Medical Dictionary for Regulatory Activities*
MI	*Multiple imputation*
MNAR	*Missing not at random*
MRI	*Magnetic resonance imaging*
NA	*Not applicable*
PCR	*Polymerase chain reaction*
PD	*Pharmacodynamic*
PK	*Pharmacokinetic*
PP	*Per-protocol*
PT	*Preferred term*
QTcF	*QTc interval corrected by Fridericia's formula*
SAE	*Serious adverse event*
SAP	*Statistical Analysis Plan*
SD	*Standard deviation*
SIAQ	*Self-Injection Assessment Questionnaire*
SOC	*System organ class*
SOP	*Standard Operating Procedure*
T3	*Triiodothyronine*
T4	*Thyroxine*
TEAE	*Treatment-emergent adverse event*
TESAE	*Treatment-emergent serious adverse event*

Acronyms, Abbreviations, Initials	Description / Explanation
TSQM	*Treatment Satisfaction Questionnaire for Medication*
ULN	*Upper limit of normal*
VAS	*Visual analog scale*
WBC	*White blood cells*
WIN	*Work instruction*

3. General

3.1. Distribution list

Statistical Analysis Plan is [redacted] ~~confidential~~ document.

[redacted]

3.2. Scope

[redacted]

[redacted]

3.3. Purpose

The purpose of this SAP is to outline the planned analyses to be completed to support the completion of the Clinical Trial Report (CTR) for protocol HS-18-633. The planned analyses identified in this SAP will be included in regulatory submissions and may be included in future manuscripts. Also, exploratory analyses not necessarily identified in this SAP may be performed to support the clinical development program. Any post-hoc, or unplanned, analysis not identified in this SAP will be clearly identified in the respective CTR.

[redacted]

3.4. Preparation, Review and Changes to the SAP

[redacted]

3.5. Changes to the Protocol

If the clinical trial protocol undergoes a major amendment affecting statistical analysis procedures, a new version of the SAP is issued.

Final approval of changes in the protocol is done by Camurus AB.

4. PREFACE

This SAP describes the planned analysis and reporting for Camurus AB protocol HS-18-633. This Phase 3 trial is being performed to assess the safety and efficacy of CAM2029 for the treatment of acromegaly in patients with acromegaly who are biochemically controlled on treatment with long-acting somatostatin analogues (octreotide or lanreotide). The structure and content of this SAP provides sufficient detail to meet the requirements identified by the United States Food and Drug Administration and International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH): Guidance on Statistical Principles in Clinical Trials. The following documents were reviewed in preparation of this SAP:

- Clinical Trial Protocol HS-18-633 (Final Version 5.0, 31-01-2023)
- eCRFs for Protocol HS-18-633
- ICH E9: Guidance on Statistical Principles for Clinical Trials.
- ICH E9 (R1): Estimands and Sensitivity Analysis in Clinical Trials
- ICH E3: Structure and Content of Clinical Study Reports

5. TRIAL OBJECTIVES AND ENDPOINTS

5.1. Trial Objectives

5.1.1. Primary Objective

- The primary objective is to assess the superiority of CAM2029 compared to placebo in biochemical response for insulin-like growth factor-1 (IGF-1) in patients who are biochemically controlled on treatment with long-acting somatostatin analogues (octreotide or lanreotide) and who have prior evidence of active acromegaly disease.

5.1.2. Secondary Objectives

- To assess the superiority of CAM2029 compared to placebo in biochemical response for IGF-1 irrespective of dose of investigational medicinal product (IMP)
- To assess the superiority of CAM2029 compared to placebo in biochemical response for IGF-1 and GH
- To assess time to loss of response of CAM2029 compared to placebo
- To assess the efficacy of CAM2029 versus placebo for IGF-1 levels over time
- To assess the efficacy of CAM2029 versus placebo for GH levels over time
- To assess use of rescue medication
- To evaluate the effects of CAM2029 and placebo on tumor size
- To assess the effects of CAM2029 and placebo on clinical signs and symptoms of acromegaly
- To measure patients' satisfaction with CAM2029 and placebo
- To measure the effects of CAM2029 and placebo on quality of life
- To measure patients' satisfaction with self-administration of CAM2029 and placebo
- To assess supervised self- or partner-administration of CAM2029 and placebo
- To assess the plasma concentrations of octreotide after administration of CAM2029 in patients with acromegaly

5.1.3. Safety Objectives

- To assess the safety and tolerability of CAM2029 and placebo
- To assess the local tolerability of CAM2029 and placebo
- To assess the treatment effects of CAM2029 and placebo on prolactin
- To assess the immunogenicity of octreotide from CAM2029 in patients with acromegaly

5.2. Trial Endpoints (Target Variables)

5.2.1. Primary Endpoint

- Proportion of patients with mean IGF-1 levels ≤1x upper limit of normal (ULN) at Week 22 and Week 24 (average of the 2 measurements)

5.2.2. Key Secondary Endpoints

- Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24, including patients who have their dose decreased to 10 mg CAM2029 (or 0.5 mL placebo)
- Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 and mean GH cycle levels <2.5 µg/L at Week 24

5.2.3. Secondary Endpoints

- Time to IGF-1 levels >1xULN
- Time to IGF-1 levels ≥1.3xULN
- Change from baseline to Week 24 in IGF-1 levels
- Change from baseline to Week 24 in GH levels
- Proportion of patients with mean GH cycle levels <2.5 µg/L at Week 24
- Proportion of patients with mean GH cycle levels <1.0 µg/L at Week 24
- Proportion of patients who receive rescue medication with standard of care
- Change in tumor size from screening to Week 24
- Severity scores of clinical signs and symptoms of acromegaly over time
- Ring size over time
- Treatment Satisfaction Questionnaire for Medication (TSQM) scores over time using all 4 domains of TSQM (effectiveness, side effects, convenience and satisfaction)
- Patient satisfaction scale scores at Week 24
- Change from baseline in Acromegaly Quality of Life Questionnaire (AcroQoL) and EuroQoL 5-dimension 5-level (EQ-5D-5L) scores
- Change from PRE module to the POST module in Self-Injection Assessment Questionnaire (SIAQ) scores
- Proportion of patients/partners declared competent by a healthcare professional to administer CAM2029 or placebo out of those trying
- Octreotide plasma concentrations over time

5.2.4. Safety Endpoints

- Incidence of adverse events (AEs) and laboratory and electrocardiogram (ECG) abnormalities
- Changes in laboratory values, vital signs, ECG readings and gallbladder imaging
- Incidence of injection site reactions (erythema and swelling) and level of injection site pain
- Prolactin levels
- Qualification and quantification of anti-octreotide antibodies

6. TRIAL METHODS

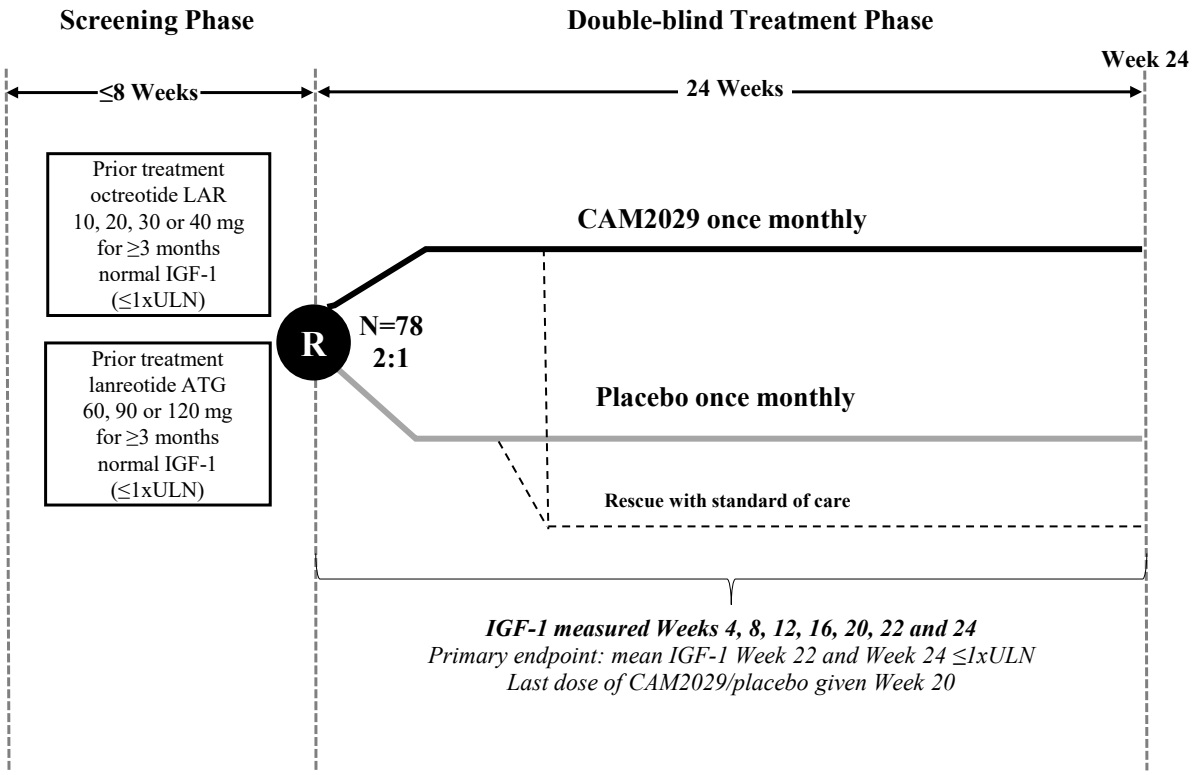
6.1. Overall Trial Design and Plan

This is a Phase 3, randomized, double-blind, placebo-controlled, multi-center trial that assesses the efficacy and safety of CAM2029 versus placebo in patients with acromegaly. To be eligible for the trial, patients need to be biochemically controlled (with IGF 1 levels ≤1xULN and mean GH cycle levels <2.5 µg/L at screening), be on treatment with long-acting somatostatin analogues (octreotide or lanreotide) for at least 3 months before screening and have prior evidence of active acromegaly disease. Eligible patients will be randomized in a 2:1 ratio to treatment with either CAM2029 or placebo in a 24-week Double-blind Treatment Phase. The starting dose of CAM2029 is 20 mg once monthly, regardless of the patient's prior treatment and dose. The dose may be down titrated from 20 mg to 10 mg based on safety and tolerability. Patients who discontinue treatment will continue to be followed in the trial until Week 24.

During the 24-week Double-blind Treatment Phase, IGF-1 will be measured on Day 1 (before the first dose of CAM2029/placebo) and at Weeks 4, 8, 12, 16, 20, 22 and 24. The primary endpoint (IGF-1 levels ≤1xULN) will be evaluated as an average value of the measurements at Week 22 and Week 24. GH will be measured as random samples or cycles during the trial. Safety will be evaluated continuously and plasma samples for assessment of octreotide concentration will be taken. Patient-reported treatment satisfaction will be assessed using a general questionnaire and will be compared between CAM2029 and placebo and with previous treatment. Health-related quality of life parameters will also be assessed.

At Week 24 (or 4 weeks after the last dose in case of premature withdrawal from the trial), the patients will come to the trial site for an End of Trial (EOT) visit. Patients who complete the trial will be offered to continue with treatment with CAM2029 as roll-over patients in a long-term safety trial (HS-19-647). Alternatively, patients will be switched back to their previous treatment for acromegaly.

An overview of the trial design is shown in Figure 1.



ATG: autogel; IGF-1: insulin-like growth factor-1; LAR: long-acting release; R: randomization; ULN: upper limit of normal

Figure 1: HS-18-633 trial design

6.2. Selection of Trial Population

Main Inclusion Criteria

1) Diagnosis of acromegaly by historical evidence of (persistent or recurrent) acromegaly as documented by:

   a) Pituitary adenoma on magnetic resonance imaging (MRI) or pathology report. Computerized tomography (CT) is accepted if MRI could not be performed

   b) A lack of suppression of GH nadir to <1 µg/L after an oral glucose tolerance test with 75 g of glucose (not applicable for patients with diabetes)

      or

      a mean GH concentration of a 5-point profile within a 2-hour time period of >2.5 µg/L

      or

      IGF-1 levels >1xULN (for patients who have undergone pituitary surgery, the measurement should be performed at least 3 months after the surgery)

2) Treatment with a stable dose of octreotide LAR (10 mg, 20 mg, 30 mg or 40 mg) or lanreotide ATG (60 mg, 90 mg or 120 mg) for at least 3 months as monotherapy prior to screening
3) IGF-1 levels ≤1xULN (adjusted for age and sex) (mean value of the first measurement at screening and the second measurement at 2 weeks before Day 1)
4) Adequate liver, pancreatic, renal and bone marrow functions
5) Normal ECG

Main Exclusion Criteria

1) GH ≥2.5 µg/L at screening (cycle)
2) Have received medical treatment for acromegaly with pasireotide (within 6 months prior to screening), pegvisomant (within 3 months prior to screening), dopamine agonists (within 3 months prior to screening) or other investigational agents (within 30 days or 5 half-lives prior to screening [whichever is longer])
3) Patients who usually take octreotide LAR or lanreotide ATG less frequently than every 4 weeks (e.g. every 6 weeks or 8 weeks)
4) Patients with compression of the optic chiasm causing any visual field defect for whom surgical intervention is indicated
5) Patients who have undergone major surgery/surgical therapy for any cause within 1 month prior to screening
6) Patients who have undergone pituitary surgery within 6 months prior to screening
7) Patients who have received prior pituitary irradiation
8) Patients with poorly controlled diabetes mellitus (hemoglobin A1c >8.0%)

6.3. Method of Treatment Assignment and Randomization

Adult (≥18 years of age) male and female patients will be enrolled and randomized into one of two treatment arms (CAM2029 or placebo) in a 2:1 ratio, using an interactive randomization system (interactive web or voice response system). Randomization will be stratified based on previous treatment (octreotide LAR versus lanreotide ATG).

6.4. Treatment Masking (Blinding)

This is a double-blind trial. The randomized treatment assignment to CAM2029 or placebo and all individual IGF-1 and GH results will remain concealed to patients, investigators and the trial team until the trial is completed and a decision is made to unblind.

7. SEQUENCE OF PLANNED ANALYSIS

7.1. Interim Analysis

No interim analyses are planned for this trial.

7.2. Data Monitoring Committee Meetings

Not applicable.

8. SAMPLE SIZE DETERMINATION

For the sample size calculation, assumptions on absolute levels of biochemical response, defined as proportion of patients with mean IGF-1 levels at Week 22 and Week 24 ≤1xULN by treatment arm are made. It should be noted that the treatment difference in biochemical response is based on the assumption of a very low response in the placebo treatment arm. It should also be noted that the actual endpoint calculated as the average of 2 measurements (Week 22 and Week 24) is not common in the literature. It is, thus, assumed that the proportion with biochemical response with this endpoint is very similar to response based on 1 measurement.

The proportion of responders at Week 22 and Week 24 in the CAM2029 treatment arm is assumed to be 0.80. For the group of patients receiving placebo it is assumed that the proportion of responders is 0.40 at Week 22 and Week 24. It is also assumed that the missing data frequency will be low and have a limited impact on the sample size assumptions.

With 78 patients randomized in a 2:1 ratio, i.e., 52 patients in the CAM2029 treatment arm and 26 patients in the placebo treatment arm, the power is 90% to show a difference between CAM2029 and placebo, if the percentages of responders, as defined by the primary endpoint, are 80% for CAM2029 and 40% for placebo. The power calculation is based on the Chi-squared test (with a continuity correction).

The primary analysis of the primary endpoint will be performed by the common proportion difference test using Mantel-Haenszel stratum weights, stratified by prior treatment (octreotide LAR or lanreotide ATG). Due to the stratification the efficiency of the analysis may increase (Kernan et al. 1999), and the power of the applied model is assumed to be higher than 90%.

9. ANALYSIS POPULATIONS

9.1. Definition of Analysis Populations

The following analysis populations are planned:

Intention-to-treat Analysis Set

The intention-to-treat (ITT) analysis set comprises all patients who have been randomized to a treatment arm. Analyses based on this population will group patients according to the treatment

they were randomized to receive, regardless of actual treatment received. All efficacy analyses will be based on the ITT analysis set.

Full Analysis Set

The full analysis set (FAS) comprises all randomized patients in the ITT analysis set who were administered at least one dose of the randomized IMP.

Modified Intention-to-treat Analysis Set

The modified intention-to-treat (mITT) analysis set comprises all patients from the ITT analysis set who do not have COVID-19-related missing IGF-1 measurements at Week 22 and Week 24 (i.e. they either have at least one available measurement, or missingness is not due to COVID-19).

Per Protocol Analysis Set

The per protocol (PP) analysis set is defined as all patients in the ITT analysis set with no major protocol deviations that would have an impact on the efficacy assessment. Protocol deviations will be reviewed before data are unblinded.

Safety Analysis Set

The safety analysis set comprises all patients who were administered at least one dose of IMP. Analyses based on this population will group patients according to the actual treatment the patients received. The safety analysis set will be used for analysis of safety data.

9.2. Use of Analysis Populations in Different Analyses

All efficacy analyses for randomized patients will be based on the ITT analysis set, with treatment arm assignment “as randomized”. Efficacy analysis of the primary endpoint will be repeated on the mITT analysis set, PP analysis set and FAS as supportive analyses if the analysis set differs from the ITT analysis set.

Safety analyses will be based on the safety analysis set, with treatment arm assignment ”as treated”. If the actual treatment is not consistent for a patient during the treatment phase, the patient will belong to the treatment arm (CAM2029 or placebo) from which the patient received more of the treatment. If the patient receives the same number of CAM2029 and placebo doses, he/she will belong to the CAM2029 treatment arm.

Demography and background characteristics summaries will be prepared for all analysis sets that are different from each other.

10. GENERAL ISSUES FOR STATISTICAL ANALYSES

For all efficacy endpoints, descriptive measures and 95% confidence interval (CI) as specified in the separate sections for the difference between the 2 treatment arms (CAM2029 and placebo) will be presented, unless otherwise specified.

The risk differences and 95% CI for the primary endpoint and key secondary endpoints will be given together in a separate table and forest plot.

10.1. Definition of Visit Windows

No additional visit windows for analysis variables will be defined, they will be used as defined in the protocol. The EOT visit should take place at Week 24 or 4 weeks after the last dose in case of premature withdrawal from the trial. If a patient did not attend the Week 24/EOT visit, the last post-baseline visit will be considered as the EOT visit. If a patient did not have any post-baseline visit, the EOT visit will be considered as missing. If a patient attended the Week 24/EOT visit at any other time than at Week 24, all data from the Week 24/EOT visit will be re-allocated to the scheduled visit closest in time to when withdrawal was done.

10.2. Handling of Missing Data

Procedural attempts with the aim to avoid missing data and keep the frequency of missing data to a minimum will be implemented in the trial (see protocol Section 6.11.2). Missing data may be due to withdrawal of the patient from the trial (possible reasons: patient withdrew consent, AE, Investigator's decision, need for administration of a non-permitted concomitant medication, protocol deviation, pregnancy, lost to follow-up, inability to fulfil trial requirements and procedures, death, other) or because the assessment is not done or data is not available (reasons will be specified on the protocol deviation page of the eCRF).

Missing data will be imputed for:

- The primary endpoint (proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24)
- The key secondary endpoints of
    - Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24, including patients who have their dose reduced to 10 mg CAM2029 (or 0.5 mL placebo)
    - Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 and mean GH cycle levels <2.5 µg/L at Week 24
- Other efficacy endpoints as described in corresponding sections
- Missing AE start and end dates

These procedures are parameter-specific, and they are described in detail in the corresponding sections of the SAP.

For patients who withdraw from the trial before Week 24, efficacy data for the primary and key secondary efficacy endpoints collected at the EOT/Week 24 visit will be re-allocated to the visit following the last visit the patient performed.

10.3. Definition of Treatment Arms

The two treatment arms will be denoted as follows in the tables:

- ”CAM2029” – CAM2029 (octreotide subcutaneous depot) once monthly
- ”Placebo” – matching placebo once monthly

10.4. Definition of Baseline

Baseline values for IGF-1 will be calculated as the mean value of pre-treatment measurements (i.e., the measurement at Week -2 and the pre-dose measurement on Day 1). If one of the IGF-1 values at Week -2 or Day 1 is missing, the other value will be used as baseline value. For GH, the pre-dose value on Day 1 will be used as baseline. If this is missing, the screening value will be used. Baseline values of ECG measurements will be calculated (as for all other time points) as the mean of triplicate ECG. For all other parameters, the baseline value will be the closest measurement in time to the date of first dose of IMP and preceding the first dose of IMP.

10.5. Multi-center Trials

This is a multi-center trial. However, no by center analyses/displays are planned.

10.6. Multiple Comparisons and Multiplicity

Three evaluations of biochemical control are planned. To control the overall significance level of 5%, a stepwise closed testing procedure is planned. The following test order, from 1 to 3, will be used:

1) Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 (average of the 2 measurements; primary endpoint)
2) Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24, including patients who have their dose reduced to 10 mg CAM2029 (or 0.5 mL placebo)
3) Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 and mean GH cycle levels <2.5 µg/L at Week 24

In this testing procedure, a comparison will be eligible for superiority testing only if all previous comparisons, if any, have established superiority at the one-sided significance level of $p<0.025$.

Since this is a hierarchical closed testing procedure, no adjustment of the type I error rates of hypothesis tests is needed.

10.7. Planned Subgroup Analyses

Patients will be stratified by prior treatment (octreotide LAR or lanreotide ATG). For all endpoints included in the test order (Section 10.6), exploratory subgroup analyses (summary statistics) will be performed by sex, race, age and prior treatment with octreotide LAR or lanreotide ATG.

For the subgroup of patients having IGF-1 levels $\leq$1xULN at baseline (average of the 2 measurements at Week -2 and Day 1), the proportion of patients having IGF-1 levels $\leq$1xULN at Week 22 and Week 24 (average of the 2 measurements) will be analyzed in a similar way as the primary endpoint without imputation, if the subgroup differs from ITT analysis set (excluding the sensitivity and supportive analyses).

For the subgroup of patients having GH levels <1.0 µg/L at baseline, the proportion of patients having GH levels <1.0 µg/L at Week 24 will be analyzed in a similar way as the primary endpoint without imputation, if the subgroup differs from ITT analysis set (excluding the sensitivity analyses).

10.8. Analysis Software

All analyses will be performed using SAS® Software version 9.4 or above.

11. TRIAL PATIENTS

11.1. Disposition of Patients

All patients who provide informed consent will be accounted for in this trial. The number of screening failures will be presented. The number and percentage of patients in each analysis set will be presented by treatment arm (percentage of the number of patients randomized) and overall. The number and percentage of patients completing treatment and trial will be presented by treatment arm and overall. The number and percentage of patients who discontinued treatment or withdrew from the trial will be presented by the coded term for main reason of discontinuation or withdrawal and overall, by treatment arm and overall. Patients who withdrew from the trial due to COVID-19 will be listed separately.

The number and percentage of patients who continued to trial HS-19-647 will be summarized.

All patients excluded from one or more analysis datasets will be listed.

Randomization errors will be listed.

11.2. Protocol Deviations

Prior to database lock, the Medical Monitoring Team will review the protocol deviation list and define/document which deviations are considered minor/major protocol deviations in accordance with the Protocol Deviation Handling Plan. Handling of protocol deviations due to COVID-19 will also be addressed prior to database lock.

All protocol deviations will be listed and grouped by category. Protocol deviations due to COVID-19 will be listed separately.

11.3. Inclusion and Exclusion Criteria

Patients who meet all the inclusion criteria and none of the exclusion criteria are eligible to participate in the clinical trial.

Deviations from inclusion/exclusion criteria will be listed.

12. DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

12.1. Demographics

For variables assumed to be continuous, like age at baseline, weight, height and body mass index, descriptive statistics will be prepared by treatment arm, stratum and overall (including n, N, arithmetic mean, standard deviation (SD), minimum, median and maximum, where n denotes the number of patients with available data in a certain category and N denotes the number of patients in the respective analysis set and corresponding treatment arm).

For categorical variables like sex, race and ethnicity, frequency tables (n/N and percentage) will be prepared by treatment arm, stratum and overall.

All demographic data will be listed.

12.2. Prior and Concomitant Medications and Treatments

The number and percentage of patients who took prior and concomitant medications will be summarized by Anatomical Therapeutic Chemical (ATC) level using the WHO Drug Dictionary March 2022 release. ATC level 2, ATC level 4 and Preferred Name, corresponding to preferred base code XXXXXX01001, will be presented in decreasing order of frequency by treatment arm. The number and percentage of patients who took prior and concomitant non-pharmacological treatments will also be presented.

Prior medications/treatments are defined as those for which the end date is prior to the date of the first IMP administration. Concomitant medications/treatments are defined as those with start date on or after the date of the first IMP administration and those with start date prior to the first IMP use but with end date on or after the date of the first IMP administration.

If medication/treatment dates are incomplete and it is not clear whether the medication/therapy was concomitant, it will be assumed to be concomitant.

For the calculation of time intervals using partial dates with year and month, a partial start date will be updated to the earliest possible full date. Similarly, a partial end date will be updated to the latest possible full date. In case of partial dates with year only, time intervals will not be calculated.

All prior and concomitant medications and non-pharmacological treatments will be listed.

Disease-specific Prior Medications and Treatments

A complete medication/treatment history for acromegaly will be collected, including the date and time of the last dose of octreotide LAR or lanreotide ATG prior to the first dose of IMP. The number and percentage of patients who took prior medications and non-pharmacological treatments for acromegaly will be presented by treatment arm, stratum and overall.

Time since onset of the first treatment with octreotide LAR or lanreotide ATG medication to the start of treatment (Day 1) (in years) will be presented using descriptive statistics (including n, arithmetic

mean, SD, minimum, median and maximum) and categorical (<10 years, 10-20 years, ≥20 years) by treatment arm, stratum and overall.

For the last dose of prior medication for acromegaly (octreotide LAR or lanreotide ATG) taken before start of IMP treatment, the number and percentage of patients on each dose level will be presented by treatment arm and overall. The number and percentages of patients who received low doses (10 mg octreotide LAR or 60 mg lanreotide ATG) or mid/high doses (20-40 mg octreotide LAR or 90-120 mg lanreotide ATG) as their last dose will also be presented.

12.3. Baseline and Screening Conditions

12.3.1. Baseline Medical History

Medical history (recorded at screening) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 25.0.

All medical history events will be listed by system organ class (SOC) and preferred term (PT). The number and percentage of patients with medical history events will be presented by SOC PT, treatment arm, stratum and overall.

Hepatobiliary diseases ongoing at screening will be tabulated separately by SOC, PT, treatment arm and stratum.

12.3.2. Baseline Physical Examination

Physical examination data at screening, including all major body systems (head, eyes, ears, nose, throat, general appearance, skin, neck, including thyroid, lungs, heart, abdomen, back, lymph nodes, extremities and nervous system), together with evaluations will be listed only.

12.3.3. Baseline Disease Characteristics and Screening History

Acromegaly history data (method of diagnosis, time since diagnosis, IGF-1 values at diagnosis and after surgery, tumor size at diagnosis, method of pituitary tumor confirmation, pituitary surgery [yes/no], and prior GH values) will be listed and tabulated.

12.4. Drug Exposure and Measurement of Treatment Compliance

Total exposure and compliance will be calculated per patient and summarized by treatment arm, stratum and overall.

Duration of exposure in weeks will be calculated as:

[redacted]

The duration of exposure in weeks will be summarized by treatment arm and stratum using descriptive statistics (n, arithmetic mean, SD, minimum, median and maximum) for the safety analysis set. In addition, the number and percentage of patients with a duration of exposure of ≥4 weeks, ≥8 weeks, ≥12 weeks, ≥16 weeks, ≥20 weeks and ≥24 weeks will be presented.

The number of injections of each dose (20 mg and 10 mg) and the total number of injections received will also be summarized.

The number and percentages of patients with dose adjustments (reductions, interruption or permanent discontinuation of IMP) and the number and percentages of patients with dose increase back to previous dose after dose adjustments, along with the reasons will be summarized by treatment arm, stratum and overall.

Treatment compliance (percentage) will be calculated by the following formula:

total administered dose (mg)/total prescribed dose (mg) x 100

Date, time, dose and other parameters of IMP administration will be listed.

13. EFFICACY ANALYSIS

13.1. Primary Efficacy Variable Analysis

The primary efficacy endpoint is the proportion of patients with biochemical response, defined as the proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 (average of the 2 measurements). Classification will be based on the numeric value of IGF-1. The primary analysis of the primary endpoint will be performed by the common risk difference test using Mantel-Haenszel stratum weights, stratified by prior treatment (octreotide LAR or lanreotide ATG). The ULN derived from the patient's sex and age at screening will be used for the assessment of the individual patient's response.

The analysis will be based on the ITT analysis set.

Definition of the Primary Estimand

Given the nature of the expected intercurrent events, a composite variable strategy has been chosen. The items that define the primary estimand are:

a) ***Population of interest:*** patients with acromegaly as defined by the inclusion/exclusion criteria in the trial

b) ***Variable (or endpoint) of interest:*** biochemical response defined as the mean of the IGF1 values measured at Week 22 and Week 24 being ≤1xULN (adjusted for age and sex). A patient will be considered as non-responder if the mean of the IGF-1 values measured at Week 22 and Week 24 is >1xULN (adjusted for age and sex).

c) ***Specification of how intercurrent events are reflected in the scientific question of interest:***

- Discontinuation: Patients who discontinue treatment with CAM2029/placebo prior to Week 22 (regardless of IGF-1 values) will be regarded as non-responders
- Rescue medication: Patients who receive rescue medication will be regarded as non-responders (as it results in treatment discontinuation)
- Dose decrease: Patients who have had their dose decreased (due to safety or tolerability reasons) will be regarded as non-responders

d) ***Population level summary for the variable:*** difference in the proportion of patients with biochemical response between the CAM2029 and placebo arms, stratified by prior treatment.

The *primary estimand* is thus the difference in the proportion of patients who maintained the initial CAM2029/placebo dose and had biochemical response (average of the Week 22 and Week 24 values being ≤1xULN) between CAM2029 and placebo for patients with acromegaly**.** This estimand assumes that occurrence of intercurrent events other than those affecting dose received are considered irrelevant in defining the treatment effect of interest.

Statistical Analysis Methods

The frequency of intercurrent events and the number of missing data will be summarized separately.

If one of the IGF-1 values at Week 22 or Week 24 is missing, the other value will be used to define a responder/non-responder for the analysis. The variable (or endpoint) of interest will be missing only if no IGF-1 value can be obtained from both the Week 22 and Week 24 samples. If this is the case, the following imputations will be made:

- If a patient had missing IGF-1 values at both Week 22 and Week 24 either due to being randomized and not treated or due to intercurrent events other than those affecting dose received, then these will be imputed under MAR, i.e., missing relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening) will be imputed by MI using a monotone regression model from the treatment group the patient was randomized to. A two-stage MI will be performed for missing Pre-Dose IGF-1 values at Day 1, Week 4, 8, 12, 16 and 20 and Post-Dose IGF-1 values at Week 22 and 24 (i.e. excluding the Post-Dose profile at Week 20):
    1. A Monte Carlo Markov Chain (MCMC) method is applied to obtain monotone missing data patterns (100 imputations)
    2. Monotone regression is applied to each imputed dataset with monotone missing data patterns obtained in step 1, resulting in 100 complete datasets

  The following variables will be included in the imputation model: treatment, stratum, IGF-1 values from all scheduled visits available.
- For the continuous endpoint, a total of 100 complete datasets will be created by imputation. The responder/non-responder status for each patient will be determined from the imputed continuous response, i.e., the average of the values at Week 22 and Week 24 being ≤1xULN or >1xULN. Results from the common risk difference test using Mantel-Haenszel stratum weights, stratified by previous treatment, will be combined using Rubin's rule (Rubin 1987).
- All other patients with missing IGF-1 values at both Week 22 and Week 24 where it was either known or assumed that treatment was affected will be regarded as a non-responder.

The following statistical hypothesis with a one-sided alternative will be tested at level 0.025 to address the primary efficacy objective:

H_0: $\Theta \leq 0$ versus H_1: $\Theta > 0$

where Θ is the difference in the proportion of patients with biochemical response at Week 22 and Week 24 (average of the 2 measurements) between CAM2029 and placebo. The weighted common difference in proportions across strata (two-sided 95% CI) will be estimated using the commonriskdiff (CL=MH) option of proc freq.

The primary analysis of the primary endpoint will be performed by the common proportion difference test using Mantel-Haenszel stratum weights, stratified by prior treatment (octreotide LAR or lanreotide ATG). The trial null hypothesis, H_0, will be rejected if the associated one-sided p-value is <0.025.

The odds ratio and its 95% CI (Mantel-Haenszel) will also be given. If there are empty cells, a Logit-type common odds ratio will be calculated instead of Mantel-Haenszel-type common odds ratio, across strata, stratified by prior treatment (octreotide LAR or lanreotide ATG).

In addition, the proportion of patients with IGF-1 ≤1xULN (adjusted for age and sex) will be provided for all assessments including baseline (average of the measurements at Week -2 and Day 1) and end of the trial (average of the measurements at Week 22 and Week 24) by treatment arm. The proportion of patients with IGF-1 levels at the end of the trial (average of the measurements at Week 22 and Week 24) below or equal to the IGF-1 level at baseline will also be presented.

13.2. Sensitivity Analyses for the Primary Efficacy Variable

To assess the impact on the results in the primary analysis for patients with missing data for the primary endpoint that were set to be non-responders, the following sensitivity analyses will be performed:

Sensitivity Analysis I. The assumption that patients with missing data on the primary endpoint are non-responders will be assessed using a retrieved dropout analysis to impute missing data in the treatment arm using relevant baseline information and data from the retrieved dropout patients in the corresponding arm (within the same treatment arm and stratum). This imputation method is assumed to be the most accurate, however, to have a reliable assumption for the imputation, a sufficient number of retrieved dropout patients is needed within the corresponding arm (same treatment arm and stratum, Wang and Hu 2022).

- Therefore, if the number of retrieved dropouts is higher than 3 per stratum in both treatment and placebo arm, a sensitivity analysis will be performed by MI of relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening), assuming informative missingness (data missing not at random) and creating retrieved dropout-based imputation patterns.

  A two-stage MI will be performed:

  1. A Monte Carlo Markov Chain (MCMC) method is applied to obtain monotone missing data patterns (100 imputations)
  2. Monotone regression is applied to each imputed dataset with monotone missing data patterns obtained in step 1, resulting in 100 complete datasets

  The following variables will be included in the imputation model: treatment, stratum, IGF-1 values from all scheduled visits available.

- For the continuous endpoint, a total of 100 complete datasets will be created by imputation. The responder/non-responder status for each patient will be determined from the imputed continuous response, i.e., the average of the values at Week 22 and Week 24 being ≤1xULN or >1xULN. Results from the common risk difference test using Mantel-Haenszel stratum

weights, stratified by previous treatment, will be combined using Rubin's rule (Rubin 1987). SAS codes are presented in Appendix A.

Sensitivity Analysis II. A control-based "wash-out" analysis will also be used to impute missing relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening), in the treatment arm using relevant baseline information and data from the placebo arm. This is a conservative imputation method. (Ratitch and O'Kelly 2011)

- A sensitivity analysis will be performed by MI of IGF-1 values, assuming informative missingness (data missing not at random) and creating control-based imputation patterns. That is, an imputation model for the missing observations in both treatment arms will be constructed based on the observed data in the placebo arm.

  A two-stage MI will be performed:

  1. A Monte Carlo Markov Chain (MCMC) method is applied to obtain monotone missing data patterns (100 imputations)
  2. Monotone regression is applied to each imputed dataset with monotone missing data patterns obtained in step 1, resulting in 100 complete datasets

  The following variables will be included in the imputation model: treatment, stratum, IGF-1 values from all scheduled visits available.

- For the continuous endpoint, a total of 100 complete datasets will be created by imputation. The responder/non-responder status for each patient will be determined from the imputed continuous response, i.e., the average of the values at Week 22 and Week 24 being ≤1xULN or >1xULN. Results from the common risk difference test using Mantel-Haenszel stratum weights, stratified by previous treatment, will be combined using Rubin's rule (Rubin 1987). SAS codes are presented in Appendix A.

Sensitivity Analysis III. A sensitivity analysis will also be performed for patients with missing data only due to reasons other than AEs or lack of efficacy. For these patients, relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening) will be imputed by MI, assuming non-informative missingness (data missing at random). For these patients, response will be evaluated based on the imputed Week 22 and Week 24/EOT values. Patients who discontinue treatment due to AEs or lack of efficacy but have no missing Week 22 and Week 24/EOT data will be regarded as non-responders.

Sensitivity Analysis IV. The tipping point method

For the tipping point analysis, a graphical method will be applied (Liublinska & Rubin, 2014) to evaluate the potential impact of missing data. On the two axes, all the possible numbers of responders in the two treatment arms for patients with missing data will be presented. The common risk difference across strata (95% CI) will be estimated (as specified for the primary analysis) for each pair of outcomes. These estimates will be classified into 3 different categories represented by 3

different symbols on the graph: favoring the active treatment arm (significantly), favoring the placebo arm (significantly), or no significance. The magnitude of these areas covered by the three types of results may suggest which of the three conclusions is more likely to occur, and whether missing data may change the conclusion about the statistical significance of the estimated effect.

Sensitivity Analysis V. The assumption that occurrence of any other intercurrent event is considered irrelevant in defining the treatment effect of interest will be assessed by the following sensitivity analyses: A sensitivity analysis will be performed for patients with missing data due to COVID-19. For these patients, relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening) will be imputed by MI, assuming non -informative missingness (data missing at random). Patients with missing data due to other reasons than COVID-19 will be regarded as non-responders.

13.3. Supportive Analyses for the Primary Efficacy Variable

The responder rates by treatment arm at Week 22 and Week 24 (average of the 2 measurements) and the difference in responder rates between the 2 treatment arms will also be estimated together with their respective 2-sided 95% CI, based on:

- Supportive Analysis I: Patients who received their Week 20 dose of IMP
- Supportive Analysis II: The FAS
- Supportive Analysis III: The mITT analysis set
- Supportive Analysis IV: The PP analysis set
- Supportive Analysis V: The ITT analysis set using ULN for IGF-1 values based on actual age category at the time of each assessment and which are not carried forward from screening. This analysis will reflect potential changes due to age increase.
- Supportive Analysis VI: Estimation of the influence of different demographic and baseline factors on the primary endpoint: Supportive analysis of biochemical response will be performed by logistic regression, with treatment arm, prior treatment (octreotide LAR or lanreotide ATG), sex and race as fix factors, and time since diagnosis, age and baseline IGF-1 value as covariates.

13.4. Secondary Efficacy Variable Analysis

13.4.1. Proportion of Patients with Biochemical Response Based on IGF-1 Levels, Irrespective of IMP Dose

Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 (average of the 2 measurements), including patients who have their dose decreased to 10 mg CAM2029 (or 0.5 mL placebo), will be analyzed in the same way as the primary efficacy endpoint (including sensitivity

analyses).
The analysis will be based on the ITT analysis set.

Definition of the Secondary Estimand

Given the nature of the expected intercurrent events, a composite variable strategy has been chosen. The items that define the secondary estimand are:

a) ***Population of interest:*** patients with acromegaly as defined by the inclusion/exclusion criteria in the trial

b) ***Variable (or endpoint) of interest:*** biochemical response defined as the mean of the IGF-1 values measured at Week 22 and Week 24 being ≤1xULN (adjusted for age and sex). A patient will be considered as non-responder if the mean of the IGF-1 values measured at Week 22 and Week 24 is >1xULN (adjusted for age and sex).

c) ***Specification of how intercurrent events are reflected in the scientific question of interest:***

- Discontinuation: Patients who discontinue treatment with CAM2029/placebo prior to Week 22 will be regarded as non-responders
- Rescue medication: Patients who receive rescue medication will be regarded as non-responders
- Dose decrease: Patients who have had their dose decreased (due to safety or tolerability reasons) will be evaluated based on the mean of the IGF-1 values measured at Week 22 and Week 24 being ≤1xULN (adjusted for age and sex)

d) ***Population level summary for the variable:*** difference in the proportion of patients with biochemical response between CAM2029 and placebo arms, stratified by prior treatment.

The *secondary estimand* is thus the difference in the proportion of patients that irrespective of IMP dose had biochemical response (average of the Week 22 and Week 24 being ≤1xULN) between CAM2029 and placebo for patients with acromegaly**.** This estimand assumes that occurrence of intercurrent events other than those affecting dose received are considered irrelevant in defining the treatment effect of interest.

13.4.2. Proportion of Patients with Biochemical Response based on IGF-1 Levels ≤1xULN and GH Levels <2.5 µg/L

Proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 (average of the 2 measurements) and mean GH cycle levels <2.5 µg/L at Week 24, will be analyzed in a similar way as the primary efficacy endpoint (including sensitivity analyses). For MI based sensitivity analyses, imputations will be performed for relative IGF-1 and GH values independently but using the relevant baseline information of both IGF-1 and GH, as the two variables are correlated.
The analysis will be based on the ITT analysis set.

Definition of the Secondary Estimand

Given the nature of the expected intercurrent events, a composite variable strategy has been chosen. The items that define the secondary estimand are:

a) ***Population of interest:*** patients with acromegaly as defined by the inclusion/exclusion criteria in the trial

b) ***Variable (or endpoint) of interest:*** biochemical response defined as the mean IGF-1 levels ≤1xULN at Week 22 and Week 24 and mean GH cycle levels <2.5 µg/L at Week 24. A patient will be considered as non-responder if the mean of the IGF-1 values measured at Week 22 and Week 24 is >1xULN (adjusted for age and sex), and/or the mean GH cycle levels measured at Week 24 is ≥2.5 µg/L.

c) ***Specification of how intercurrent events are reflected in the scientific question of interest:***

- Discontinuation: Patients who discontinue treatment with CAM2029/placebo prior to Week 22 (regardless of IGF-1 or GH levels) will be regarded as non-responders
- Rescue medication: Patients who receive rescue medication will be regarded as non-responders (as it results in treatment discontinuation)
- Dose decrease: Patients who have had their dose decreased (due to safety or tolerability reasons) will be regarded as non-responders

d) ***Population level summary for the variable:*** difference in the proportion of patients with biochemical response between CAM2029 and placebo arms, stratified by prior treatment.

The *secondary estimand* is thus the difference in the proportion of patients that maintained the initial CAM2029/placebo dose and had biochemical response (average of the Week 22 and Week 24 values being ≤1xULN and mean GH cycle levels <2.5 µg/L at Week 24) between CAM2029 and placebo for patients with acromegaly**.** This estimand assumes that occurrence of intercurrent events other than those affecting dose received are considered irrelevant in defining the treatment effect of interest.

13.4.3. Time to Loss of Response

Two different definitions for time to loss of response will be applied:

- The earliest time when the IGF-1 levels of 2 consecutive measurements are >1xULN
- The earliest time when the IGF-1 levels of 2 consecutive measurements are ≥1.3xULN

The time point of the first of the 2 consecutive measurements will be used in the analysis.

These two endpoints corresponding to the two definitions will be analyzed separately.

Time to loss of response in weeks will be presented for each treatment arm separately.

Time to loss of response will be calculated as:

((Date of loss of response – Date of baseline visit) + 1)/7

For patients who are categorized as non-responder at their first and second consecutive measurement, time to loss of response will be Day 1.

Patients who are still receiving CAM2029/placebo and are still responders at Week 24 will be censored at Week 24.

For patients who discontinue treatment or withdraw from the trial, IGF-1 values from the visit following their last injection will be used in this analysis but all subsequent IGF-1 values (if any) will be set to missing. Patients who discontinue the treatment or withdraw from the trial without having loss of response confirmed will be censored at their last available IGF-1 measurement.

Descriptive statistics (n, N, arithmetic mean, 95% CI of mean, SD, minimum, median, maximum) for time to loss of response will be presented by treatment arm and stratum.

A comparison between the two treatment arms will be performed using a stratified log-rank test (according to treatment and stratum). Kaplan-Meier method will be used to summarize the estimated time to loss of response of 10-90 percentages of patients. In addition, a Cox proportional hazards model will be used for time to loss of response with treatment as factor and baseline IGF-1 as covariate, stratified by stratum. Hazard ratios will be presented from the model along with the associated 2-sided 95% CIs and p-values.

For the distribution of times to loss of response, a Kaplan-Meier graph will be prepared by treatment arm and stratum. The median time to loss of response and its 95% CI will be estimated if possible. All time to loss of response data will be listed.
The analysis will be based on the ITT analysis set.

13.4.4. Change from Baseline to Week 24 in IGF-1 and GH Levels

Data from retrieved dropouts (i.e., following intercurrent events) will be set to missing when actual levels of IGF-1 and GH are summarized and analyzed.

IGF-1 values from Week 22 and/or Week 24 and GH values from Week 24 for patients who withdraw from the trial after having received their Week 20 injection will be used in all analyses of this endpoint. Data for absolute IGF-1 levels, relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening) and GH levels over time will be presented by treatment arm and stratum.

Changes from baseline at Week 24 in relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening) and GH levels will be evaluated by a control-based imputation performed within a pattern-mixture model framework, assuming a missing not at random (MNAR) mechanism. A two-stage multiple imputation will be performed. First, monotone missing patterns will be created by the MCMC method, and 100 imputed datasets will be created. This first step may be skipped if monotone missingness already holds for each patient.

Starting from each dataset created in the first step, 100 complete datasets will be created in the second stage, creating control-based full patient profiles under MNAR assumption.

In both stages SAS proc mi will be used for imputation.

Changes from baseline at Week 24 in relative IGF-1 values (IGF-1/ULN adjusted for age and sex at screening) and GH levels will then be estimated for each complete dataset using an analysis of covariance (ANCOVA) model accounting for the fixed effect of treatment, prior treatment (octreotide LAR or lanreotide ATG) and baseline relative IGF-1 / GH levels as a covariate and results (estimates and their standard errors) will be combined using SAS proc mianalyze. Combined estimates (and 95% CIs and p-values) will be presented. All assessments will be listed by patient and time point.

SAS codes are presented in Appendix A.

Missing data patterns will be presented in a descriptive table.

In addition, graphs showing IGF-1/ULN and GH levels over time and change from baseline will be presented by patient and as mean (± SD) concentration by treatment arm and stratum. A scatter plot of IGF-1/ULN and GH levels will be also presented by treatment arm and stratum.

All IGF-1 and GH data along with change from baseline will be listed.

The analysis will be based on the ITT analysis set.

13.4.5. Proportions of Patients with Biochemical Response based on GH Levels <2.5 µg/L and <1.0 µg/L

The proportions of patients with mean GH cycle levels <2.5 µg/L and 1.0 µg/L at Week 24 will be analyzed in a similar way as the primary efficacy endpoint of proportion of patients with mean IGF-1 levels ≤1xULN at Week 22 and Week 24 (average of the 2 measurements) (excluding the sensitivity analyses). For this endpoint, intercurrent events will be handled in the same way as for the primary endpoint.

13.4.6. Proportion of Patients who Received Rescue Medication with Standard of Care

Number and proportion (with Clopper Pearson 95% CI) of patients who received rescue medication with standard of care due to loss of efficacy will be presented for each treatment arm and stratum.

13.4.7. Change in Tumor Size

Tumor size assessed by pituitary MRI (or CT) will be summarized at screening (or historical MRI/CT scanning performed within 3 months before screening) and Week 24/EOT by treatment arm and stratum, using descriptive statistics (n, arithmetic mean, 95% CI of mean, SD, minimum, median, maximum.

Relative change (%) from baseline in tumor size will be also summarized by treatment arm and stratum.

Relative change (%) from baseline in tumor size will be evaluated by a control-based imputation performed within a pattern-mixture model framework, assuming an MNAR mechanism. A two-stage multiple imputation will be performed. First monotone missing patterns will be created by the MCMC method, and 100 imputed datasets will be created. This first step may be skipped if monotone missingness already holds for each patient.

Starting from each dataset created in the first step, 100 complete datasets will be created in the second stage, creating control-based full patient profiles under MNAR assumption.

In both stages SAS proc mi will be used for imputation.

Relative changes (%) from baseline at Week 24 in tumor size will then be estimated for each complete dataset using an ANCOVA model accounting for the fixed effect of treatment, prior treatment (octreotide LAR or lanreotide ATG) and baseline tumor size as a covariate and results (estimates and their standard errors) will be combined using SAS proc mianalyze. Combined estimates (and 95% CIs and p-values) will be presented. All tumor size assessment will be listed by patient and time point.

The analysis will be based on the ITT analysis set.

13.4.8. Clinical Signs and Symptoms of Acromegaly over Time

Frequencies of acromegaly signs and symptoms (headache, sweating, fatigue, joint pain, paresthesia, and soft tissue swelling, will be summarized by treatment arm and stratum at each time point using descriptive statistics.

Data from retrieved dropouts (i.e., following intercurrent events) will be set to missing.

Shift tables using the categories 'None'/'Mild'/'Moderate'/'Severe'/'Missing' will be produced to compare baseline to each post-baseline value by treatment arm and stratum.

To shed further light on acromegaly signs and symptoms the Acromegaly Index of Severity (AIS) will be derived as described by Fleseriu et al. (2021). The AIS Overall Score will be calculated as:

- the sum of the scores all six symptoms, where none = 0, mild = 1, moderate = 2, and severe = 3 for each symptom. The AIS Overall Score for all symptoms ranges from 0 (no symptoms) to 18 (severe symptoms)
- the sum of the scores for the 5 symptoms of headache, sweating, fatigue, joint pain and soft tissue swelling, where none = 0, mild = 1, moderate = 2, and severe = 3 for each symptom. The AIS Overall five Score ranges from 0 (no symptoms) to 15 (severe symptoms)

The two AIS scores will be presented by treatment arm and stratum at each time point using descriptive statistics. Change from baseline will also be presented. The AIS scores and change from baseline in AIS scores over time by visit will be displayed in mean profiles. Absolute change from baseline in AIS will be evaluated by a control-based imputation performed within a pattern-mixture model framework, assuming an MNAR mechanism. A two-stage multiple imputation will be

performed. First monotone missing patterns will be created by the MCMC method, and 100 imputed datasets will be created. This first step may be skipped if monotone missingness already holds for each patient.

Starting from each dataset created in the first step, 100 complete datasets will be created in the second stage, creating control-based full patient profiles under MNAR assumption.

In both stages SAS proc MI will be used for imputation.

Absolute changes from baseline to Week 24 in AIS scores will then be estimated for each complete dataset using an ANCOVA model accounting for the fixed effect of treatment, prior treatment (octreotide LAR or lanreotide ATG) and baseline AIS as a covariate and results (estimates and their standard errors) will be combined using SAS proc mianalyze. Combined estimated changes from baseline with 95% CIs and the treatment difference in change from baseline (and 95% CIs and p-values) will be presented.

The proportion of patients who maintained or improved their overall AIS score at Week 24 (improvement defined as a reduction of at least 1 point in the AIS score) will be summarized and this variable will be analyzed by logistic regression, with treatment arm, prior treatment (octreotide LAR or lanreotide ATG) and baseline AIS as fix factors. Based on this model the adjusted odds for maintained or improved their overall AIS score at Week 24 after treatment with CAM2029 or Placebo, and the corresponding odds ratio between CAM2029 and Placebo will be estimated with a 95% CI.

Change from baseline in ring size will be summarized over time. Results will be presented by treatment arm and stratum at each assessment time point.

Change from baseline in ring size will be evaluated by a control-based imputation performed within a pattern-mixture model framework, assuming an MNAR mechanism. A two-stage multiple imputation will be performed. First monotone missing patterns will be created by the MCMC method, and 100 imputed datasets will be created. This first step may be skipped if monotone missingness already holds for each patient.

Starting from each dataset created in the first step, 100 complete datasets will be created in the second stage, creating control-based full patient profiles under MNAR assumption.

In both stages SAS proc mi will be used for imputation.

Change from baseline in ring size will then be estimated for each complete dataset using an ANCOVA model accounting for the fixed effect of treatment, prior treatment (octreotide LAR or lanreotide ATG) and baseline ring size as a covariate and results (estimates and their standard errors) will be combined using SAS proc mianalyze. Combined estimates (and 95% CIs and p-values) will be presented.

All data will be listed by patient and time point.
The analysis will be based on the ITT analysis set.

13.4.9. Treatment Satisfaction Scores over Time (TSQM)

The TSQM Version 1.4 will be used to compare the patients' satisfaction with CAM2029 as compared to their previous treatment (octreotide LAR or lanreotide ATG) and assess satisfaction with CAM2029/placebo during the trial.

The questionnaire has 4 domains (effectiveness, side effects, convenience and global satisfaction), each comprising 3 to 4 items. Responses are evaluated on a 5 to 7-point scale for each item.

All domain scores range from 0 to 100. Higher scores indicate a better outcome from the patient's perspective.

The calculation of domain scores based on raw item scores will be done according to the Treatment Satisfaction Questionnaire for Medication (TSQM) User Manual version 2.3.

SCALE SCORING ALGORITHM

TSQM Scale scores range from 0 to 100 and no computed score should be lower or higher than these limits.

CONVENIENCE

([Sum of Item 9 to Item 11) – 3] divided by 18) × 100

If one item is missing: ([(Sum of Item9? To Item11?)) – 2] divided by (12) × 100

EFFECTIVENESS

([(Item 1 + Item 2 + Item 3) – 3] divided by 18) × 100

If one item is missing: ([(Sum of Item 1? + Item 2? + Item 3?)) – 2] divided by (12) × 100

GLOBAL SATISFACTION

([Sum of Item 12 to Item 14) – 3] divided by (14) × 100

If Item 12 or 13 is missing

([(Sum(the two completed items)) – 2] divided by 10) × 100

If Item 14 is missing

([(Sum(Item 12 and Item 13)) – 2] divided by 8) × 10

SIDE EFFECTS

If Item 4 "As a result of taking this medication, do you experience any side effects at all?" is answered "No", then the Side Effects domain score will be equal to 100

If item 4 is answered yes: ([Sum of Item 5 to Item 8) – 4] divided by 16) × 100

If one item is missing: ([(Sum of Item 5? To Item 8?)) – 3] divided by 12) × 100

Question marks (?) in the calculation are used to highlight that one of the listed items is missing, but it is not defined which one.

The number and percentage of patients completing the questionnaires and the number and percentage of missing or incomplete assessments will be summarized by treatment arm and stratum for each assessment time point.

Data from retrieved dropouts (i.e., following intercurrent events) will be set to missing.

Descriptive statistics (n, arithmetic mean, 95% CI of mean, SD, minimum, median, maximum) will be used to summarize the domain scores by treatment arm and stratum at each assessment time point. Additionally, change from baseline in domain scores will be summarized by treatment arm and stratum at each assessment time point. 95% CIs for the mean change will be provided by treatment arm and stratum at each assessment time point. Individual item answer categories will also be summarized in frequency tables for each domain.

Change from baseline in domain scores will be evaluated by a control-based imputation within a pattern-mixture model framework, assuming an MNAR mechanism. A two-stage multiple imputation will be performed on the raw item scores. First monotone missing patterns will be created by the MCMC method, and 100 imputed datasets will be created. This first step may be skipped if monotone missingness already holds for each patient.

Starting from each dataset created in the first step, 100 complete datasets will be created in the second stage, creating control-based full patient profiles under MNAR assumption.

In both stages SAS proc mi will be used for imputation.

Domain scores will be calculated for each complete dataset.

Change from baseline in domain scores will then be estimated for each complete dataset using an ANCOVA model accounting for the fixed effect of treatment, prior treatment (octreotide LAR or lanreotide ATG) and baseline domain scores as a covariate and results (estimates and their standard errors) will be combined using SAS proc mianalyze. Combined estimates of LS Means for change from baseline with 95% CIs and p-values by treatment arm and mean difference between treatment arms (and 95% CIs and p-values) will be presented for all domain scores. A forest plot for LS Means for change from baseline by treatment arm and a forest plot for mean difference between treatment arms will be presented by domain scores by descending p-values.

SAS codes are presented in Appendix A.

Missing data patterns will be presented in a descriptive table.

The analysis will be based on the ITT analysis set.

13.4.10. Patient Satisfaction Scale

At Week 24/EOT, the patients will complete a satisfaction questionnaire, in which they will rate their overall experience with CAM2029/placebo compared to their previous treatment with octreotide LAR or lanreotide ATG. The patients will evaluate the treatment on a 5-unit scale from "much worse" to "much better".

Data from retrieved dropouts (i.e., following intercurrent events) will be set to missing.

Descriptive statistics (n, arithmetic mean, 95% CI of mean, SD, minimum, median, maximum) will be used to summarize patient satisfaction scale scores at Week 24/EOT by treatment arm and stratum.

Number and percentage of patients who provided each answer (from "much worse" to "much better") will be also summarized by treatment arm and stratum.

The analysis will be based on the ITT analysis set.

13.4.11. Change from Baseline in Health-related Quality of Life Scores

Health-related quality of life will be assessed by the AcroQoL questionnaire and utilities will be assessed using the EQ-5D-5L. The AcroQoL is a disease-specific scale for assessment of quality of life in patients with acromegaly.

Acromegaly Quality of Life Questionnaire (AcroQoL)

The AcroQoL is uni-dimensional and contains 22 items divided into 2 scales; physical (8 items) and psychological (14 items). The psychological scale is further divided into 2 sub-scales: physical appearance and impact of disease on personal relationships of the patient (7 items each). Each of the 22 items of the AcroQoL is answered using a 1 to 5 Likert scale measuring either the frequency of occurrence (always, most of the time, sometimes, rarely, or never) or the degree of agreement with the items (completely agree, moderately agree, neither agree nor disagree, moderately disagree, completely disagree).

Higher scores indicate better quality of life.

Calculation of disease-specific AcroQOL total score and sub-scale scores based on raw item scores will be done according to Badia et al. 2004.

SCALE SCORING ALGORITHM

TOTAL SCORE

([(Sum of Item 1 to Item 22) – 22] divided by (110-22)) x 100

PHYSICAL SCALE SCORE

([(Sum of the Physical Scale Items (1, 3, 9, 13, 14, 15, 19, 22) – 8] divided by (40-8)) x 100

PHYCHOLOGICAL SCALE SCORE

([(Sum of the Psychological Scale Items (2, 4, 5, 6, 7, 8, 10, 11, 12, 16, 17, 18, 20, 21) − 14] divided by (70-14)) x 100

PHYCHOLOGICAL/APPEARANCE SUBSCALE SCORE

([(Sum of the Psychological/appearance Subscale Items (2, 4, 7, 11, 12, 16, 17) – 7] divided by (35-7)) x 100

PHYCHOLOGICAL/PERSONAL RELATIONS SUBSCALE SCORE

([(Sum of the Psychological/personal relations Subscale Items (5, 6, 8, 10, 18, 20, 21) – 7] divided by (35-7)) x 100

If one or more item is missing:

([(Sum of the valid Items) – number of valid Items] divided by [(number of valid Items x 5)-number of valid Items]) × 100

If more than 25% invalid or missing answers were obtained (more than 5 in total, more than 2 for the physical dimension, more than 3 for the psychological dimension, or more than 1 for each psychological sub-dimension), the score will not be calculated.

Data from retrieved dropouts (i.e., following intercurrent events) will be set to missing.

Descriptive statistics (n, arithmetic mean, 95% CI of mean, SD, minimum, median, maximum) will be used to summarize change from baseline in AcroQoL scores by treatment arm and stratum.

Change from baseline in AcroQoL scores will be evaluated by a control-based imputation within a pattern-mixture model framework, assuming an MNAR mechanism. A two-stage multiple imputation will be performed on the raw item scores. First monotone missing patterns will be created by the MCMC method, and 100 imputed datasets will be created. This first step may be skipped if monotone missingness already holds for each patient.

Starting from each dataset created in the first step, 100 complete datasets will be created in the second stage, creating control-based full patient profiles under MNAR assumption.

In both stages SAS proc mi will be used for imputation.

Domain scores will be calculated for each complete dataset.

Change from baseline in AcroQoL scores will then be estimated for each complete dataset using an ANCOVA model accounting for the fixed effect of treatment, prior treatment (octreotide LAR or lanreotide ATG) and baseline AcroQoL scores as a covariate and results (estimates and their standard errors) will be combined using SAS proc mianalyze. Combined estimates of LS Means for change from baseline with 95% CIs and p-values by treatment arm and mean difference between treatment arms (and 95% CIs and p-values) will be presented by domain scores. A forest plot for LS Means for

change from baseline by treatment arm and a forest plot for mean difference between treatment arms will be presented by domain scores by descending p-values.

The analysis will be based on the ITT analysis set.

EQ-5D-5L

The EQ-5D-5L is a measure of health status from which utilities can be calculated for use in economic modeling. It has 2 components, the EQ-5D-5L descriptive system and the EQ-5D-5L visual analog scale (VAS).

The EQ-5D-5L descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems and unable to do activity or extreme problems. A utility index can be computed from the EQ-5D-5L descriptive system with utility scores ranging from <0 (worst imaginable health state) through 0 (dead) to 1 (best imaginable health state), with negatively valued health states representing health states valued as being worse than being dead. Validated country-specific utility scores will be used (EuroQol 2019, Rencz et al. 2020). Where country-specific utility scores are not specified, the following will be used:

IT – ES

HU – HU

GR – HU

RS – HU

RU – HU

TR – HU

PL – HU

The EQ- 5D-5L VAS records the patient's self-rated health state on a 100-point vertical VAS (0=worst imaginable health state; 100=best imaginable health state).

Computation of index scores will be performed by EQ-5D-5L Crosswalk Index Value Calculator.

Data from retrieved dropouts (i.e., following intercurrent events) will be set to missing.

Descriptive statistics (n, arithmetic mean, 95% CI of mean, SD, minimum, median, maximum) will be used to summarize change from baseline in EQ-5D-5L scores by treatment arm and stratum.

The analysis will be based on the ITT analysis set.

13.4.12. Self-injection Scores over Time (SIAQ)

The SIAQ is an instrument designed to measure overall patient experience with subcutaneous self-administration. It comprises a PRE self-injection module containing 7 items and a POST self-injection

module containing 21 items and will be completed by patients who choose to self-administer CAM2029 (or placebo).

The SIAQ PRE module will be completed by the patient before the first self-administration of CAM2029 (or placebo) and contains the following domains: Feelings about injections (item 1-3), Self-confidence (item 4-6) and Satisfaction with current way of taking medication (item 7).

The SIAQ POST module will be completed 20 to 40 minutes after the first self-administration and after the self-administration at Week 20. The POST module contains the PRE domains: Feelings about injections (item 1-3), Self-confidence (item 5-7) and Satisfaction with current way of taking medication (item 18), plus the following domains: Self-image (item 4), Injection site reactions (item 8a-9e), Ease of use (item 10-14) and Satisfaction with self-injection (item 15-21).

Responses are evaluated on a 5- to 6-point scale for each item. Scores range from 0 to 10. Higher scores indicate a better patient experience.

Calculation of domain scores based on raw item scores will be done according to Keininger and Coteur 2011.

Item scores will be transformed to obtain a score ranging from 0 (worst experience) to 10 (best experience) for each item. The domain score will be calculated as the mean of the item scores included in the domain.

Domain scores will be calculated only if at least half of the domain items were completed.

Data from retrieved dropouts (i.e., following intercurrent events) will be set to missing.

Descriptive statistics (n, arithmetic mean, SD, minimum, median, maximum, 95% CI) will be used to summarize the domain scores for the first self-administration and the administration at Week 20. Change in domain scores from the PRE module (completed before the first self-administration) to the POST module (completed after the first and later self-administrations) will be summarized for feelings about injections, self-confidence and satisfaction with self-injection domains.

All analyses of SIAQ data will be based on data from patients who were treated with IMP, had self-administered IMP at least once and had at least one evaluable assessment of the SIAQ instrument (PRE and POST module).
The analysis will be based on the ITT analysis set.

13.5. Other Endpoints and Analyses – Feasibility of Self-administration of IMP

For patients opting for self- or partner-administration, the proportion of patients/partners declared competent by the healthcare professional out of those trying will be presented by treatment arm and stratum. The number and proportion of patients/partners requiring 1, 2 or 3 attempts until success or failure after maximum of 3 attempts will be provided.

The number and proportion of injections administered by patient/partner/personnel will be presented by visit, treatment arm and stratum. The number and proportion of patients/partners who

had problems/had no problem while injecting and number and proportion of patients/partners who decided to withdraw from self/partner administration will be presented by administrator (self/partner/personnel), visit, treatment arm and stratum.

For patients/partners who were declared competent after at most 3 attempts, any unsuccessful administration will be listed.

14. SAFETY AND TOLERABILITY ANALYSIS

All safety analysis will be performed for patients in the safety analysis set.

14.1. Adverse Events

Adverse events (AEs) will be coded using the MedDRA central coding dictionary, version 25.0 or higher.

Only treatment-emergent adverse events (TEAEs) will be presented in tables. A TEAE is defined as an AE, which occurs during or after the first administration of the IMP (CAM2029 or comparator) and to the EOT visit or the next dose of any acromegaly treatment (ATC code H01CB, including CAM2029 in other trials), whichever comes first. An AE reported at any time after last dose of IMP will be considered treatment emergent if assessed as related to IMP. All AEs will be listed by patient. Non-treatment-emergent AEs will be listed separately.

Unless otherwise stated, AE analyses will be patient-based, each patient will be counted only once in frequency tables, and the denominator for percentages will be the number of patients in the safety analysis set, in the respective treatment arm (i.e., “as treated” assignment of treatment arms).

Incomplete start and end dates will not be imputed. Relative day and duration will be calculated only if dates are complete.

If an AE cannot be categorized as treatment-emergent or not treatment-emergent based on the partial date, the AE will be considered to be treatment-emergent.

The AE will be considered as ”ongoing” if both the end date and the outcome of an AE are missing.

14.1.1. All Adverse Events

Summary tables of all TEAEs including severity of TEAEs, IMP-related TEAEs, treatment-emergent serious adverse events (TESAEs), TEAEs leading to treatment discontinuation, TEAEs leading to withdrawal from trial, TEAEs resulting in death, TEAEs leading to dose reduction, injection site TEAEs, TEAEs of special interest and TEAEs related to COVID-19 will be presented by treatment arm and stratum. Number of patients (n) within the category, percentage of patients (%) within the category and the total number of events (E) within the category will be presented as well as the number of TEAEs (E) per patient-treatment-year by treatment arm and stratum (rate). In the overall summary table presenting data by number and percentage of patients, the weighted common difference in proportions across strata (two-sided 95% CI) will be estimated using the commonriskdiff (CL=MH) option of proc freq if there are at least 5 patients in total with any event within the category.

TEAEs will be summarized by treatment arm, stratum, SOC and PT (according to MedDRA) displaying number of patients in the treatment arm, number and percentage of patients with the TEAE and number of TEAEs.

IMP-related TEAEs and TEAEs of Grade 3 or higher will be summarized by treatment arm, stratum, SOC and PT. Each patient will be counted only once within each PT per trial period. If a patient

experiences more than one TEAE within a PT during the same recording period, only the TEAE with the strongest relationship or worst severity, as appropriate, will be included in the summaries by relationship and severity. In the table presenting related TEAEs by SOC and PT, the treatment difference will be given with 95% CI if there are at least 5 patients in total reporting the individual PT.

In addition, a table will be prepared by SOC and PT for the most frequent TEAEs (frequency of at least 5% in any treatment arm).

14.1.2. Adverse Events Leading to Discontinuation, Withdrawal or Dose Reduction

Separate listings will be provided for TEAEs leading to discontinuation of the IMP and TEAEs leading to withdrawal from the trial, displaying details of the event(s) captured in the eCRF.

TEAEs leading to discontinuation of the IMP or withdrawal from the trial will be summarized by treatment arm, stratum, SOC and PT. Similarly, TEAEs leading to dose reduction will be tabulated by treatment arm, stratum SOC and PT. In addition, the time to the first TEAE leading to treatment discontinuation, withdrawal from trial or dose reduction will be tabulated by treatment arm and stratum.

14.1.3. Serious Adverse Events

A listing of all SAEs will be provided, displaying details of the event(s) captured in the eCRF. SAEs collected during the pre-treatment period will be flagged.

Treatment-emergent SAEs (TESAEs) will be summarized by treatment arm, stratum, SOC and PT (according to MedDRA) displaying the number of patients in the treatment arm, number and percentage of patients with the TESAE and the total number of TESAEs.

IMP-related TESAEs and TESAEs of Grade 3 or higher will be tabulated separately by treatment arm, stratum, SOC and PT.

14.1.4. Deaths

If any patient dies during the trial, relevant information will be provided in a data listing, and appropriate SAE narratives.

14.1.5. Adverse Events in Connection with COVID-19

Adverse events within the Standardized MedDRA Query (SMQ) for COVID-19 will be listed separately.

14.1.6. Other AE Assessments

Time to the First TESAE

To quantify the uncertainty in the treatment comparison for SAEs, a Cox proportional hazards model will be used for time in days to the first TESAE. The median time to the first TESAE and its 95% CI will be estimated, if possible. Patients with no TESAEs will be censored at the last assessment of the AE status.

Treatment will be a fixed factor and stratification will be employed with the same strata as used in the primary efficacy analysis. Hazard ratios, a measure of the relative risk at a particular time point, will be presented from the model along with the associated 2-sided CIs with a confidence level of 95%.

If possible, the number of TESAEs will be analyzed in a negative binomial model. The 2 treatment arms will be compared in terms of mean number of TESAEs per patient-year within a generalized linear model framework (with a logarithmic link function, which is the default for the negative binomial distribution), accounting for the fixed effect of treatment arm, prior treatment (octreotide LAR or lanreotide ATG), and the logarithm of the follow-up time. The rates and rate ratio will be presented from the estimated model. The follow-up time is defined in the same way as for TEAEs in Section 14.1.

This is a descriptive approach; no specific safety hypothesis will be prespecified and p-values will generally not be presented for safety data.

Injection Site TEAEs

Injection site TEAEs and IMP-related injection site TEAEs will be tabulated by treatment arm, stratum and PT. Injection site TEAEs of Grade 3 or higher will be summarized separately by treatment arm, stratum, and PT. The treatment difference in injection site TEAEs will be given with 95% CI if there are at least 5 patients in total reporting the individual PT.

Injection site TEAEs will also be tabulated according to onset and duration (categorized as ≤1 day/ 2 to 6 days/ 1 week to 1 month/ >1 month to <2 months/ ≥2 months) by treatment arm, stratum and PT. Further, if at least 5 patients were down-titrated to the 10 mg dose (i.e., received a dose volume of 0.5 mL), injection site TEAEs will be tabulated by injection volume. Injection site TEAEs will also be tabulated by injection number, treatment arm, stratum and PT.

TEAEs of Gallbladder Abnormalities

TEAEs of gallbladder abnormalities will be defined before database lock and unblinding of the trial. TEAEs of gallbladder abnormalities will be tabulated by PT, treatment arm and stratum, and the treatment difference will be given with 95% CI if there are at least 5 patients in total with the individual PT. Furthermore, IMP-related TEAEs of gallbladder abnormalities and TEAEs of gallbladder abnormalities of Grade 3 or higher will be summarized separately by treatment arm, stratum and PT.

TEAEs of Special Interest

TEAEs of special interest will be identified before database lock and unblinding of the trial. TEAEs of special interest will be tabulated by SOC, PT, treatment arm and stratum, and the treatment difference will be given with 95% CI. Furthermore, IMP-related TEAEs of special interest and TEAEs of special interest of Grade 3 or higher will be summarized separately by treatment arm, stratum, SOC and PT. The time to the first TEAE of special interest will also be analyzed and presented in the same way as for TESAEs.

14.2. Pregnancies

Pregnancy data will be presented in a listing. No special analysis will be performed on the pregnancy data. Patients are to be withdrawn from the trial if they become pregnant.

14.3. Clinical Laboratory Evaluations

For summary statistics, numerical laboratory measurement values below the lower detection limit or above the upper detection limit (indicated in the database as <xx, ≤xx, >xx, ≥xx) will be replaced by the detection limit (i.e., xx) reported by the laboratory. In the listings, the values will be displayed as originally recorded.

Measurements resulting in negative numbers will be set to missing.

Parameters will be grouped as:

- **Hematology and Coagulation Panel:** hematocrit, hemoglobin, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, mean corpuscular volume, platelets, erythrocytes, erythrocyte distribution width, leukocytes, basophils, eosinophils, lymphocytes, monocytes, neutrophils as relative and absolute numbers, prothrombin time and international normalized ratio, activated partial thromboplastin time
- **Biochemistry Panel:** albumin, alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, amylase, bicarbonate, calcium, chloride, creatinine, creatine kinase, gamma glutamyl transferase, lipase, fasting plasma glucose, phosphate, magnesium, potassium, sodium, total bilirubin (direct bilirubin and indirect bilirubin only in patients with total bilirubin >2xULN), cholesterol, low density lipoprotein, high density lipoprotein, protein, triglycerides, urea, urate, hemoglobin A1c, vitamin B12
- **Urinalysis Panel:** macroscopic panel (bilirubin, blood, glucose, ketones, WBCs, pH, protein, specific gravity)
- **Thyroid Panel and Prolactin**: free triiodothyronine (T3), free thyroxine (T4), thyrotropin, and prolactin
- **Serology Panel:** hepatitis B surface antigen, hepatitis C virus (HCV) antibodies (polymerase chain reaction [PCR] required for HCV antibody-positive patients), human immunodeficiency virus antibodies)

Glomerular filtration rate will also be estimated using the MDRD (Modification of Diet in Renal Disease) formula.

Grading of laboratory values will be assigned programmatically as per National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE). The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account.

CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher.

For all laboratory tests results will be categorized as low clinically significant/low not clinically significant/normal/high not clinically significant/high clinically significant based on laboratory normal ranges.

The following summaries will be generated separately for **hematology** and **biochemistry** tests:

- Listing of all laboratory data with values flagged to show the corresponding CTCAE grades, if applicable, and the classifications relative to the laboratory normal ranges.
- For numerical laboratory measurements and the changes from baseline, descriptive statistics (n, arithmetic mean, SD, minimum, median and maximum) will be presented, by treatment arm and stratum at each assessment time point.

For laboratory tests, shift tables using CTCAE grading will be used for all laboratory parameters for which CTCAE grading is available. In addition, shift tables based on the categories low clinically significant/low not clinically significant/normal/high not clinically significant/high clinically significant will be produced for all parameters to compare baseline to the end of trial and to the worst high/low post-baseline value up to Week 24 visit or the next dose of any acromegaly treatment (ATC code H01CB, including CAM2029 in other trials), whichever comes first. In addition, for liver enzymes (alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, gamma glutamyl transferase and total bilirubin), shift tables will be produced both based on CTCAE grading and on the categories <2 ULN, ≥2 to <3 ULN, ≥3 to <5 ULN and ≥5 ULN to compare baseline to the end of trial and to the worst post-baseline value up to the EOT visit or the next dose of any acromegaly treatment (ATC code H01CB, including CAM2029 in other trials), whichever comes first.

For numerical laboratory measurements of prolactin and thyroid hormones, descriptive statistics (n, arithmetic mean, SD, minimum, median and maximum) will be presented by treatment arm and stratum at each assessment time point. Shift tables of prolactin and thyroid hormones will be produced to compare baseline to the end of trial and to the worst post-baseline value up to the EOT visit or the next dose of any acromegaly treatment (ATC code H01CB, including CAM2029 in other trials), whichever comes first. The treatment effects of CAM2029 vs placebo on prolactin levels at EOT/Week 24 will be assessed using an ANCOVA model accounting for the fixed effect of treatment, prior treatment (octreotide LAR or lanreotide ATG) and baseline prolactin as a covariate. Results will be presented including estimated levels at EOT/Week 24 and the estimated treatment difference with a 95% CI.

Urinalysis and serology data will be listed only.

All laboratory values will be presented in SI units.

Hematology and biochemistry laboratory values over time will be presented as boxplots (median, interquartile range [IQR] Q1+1.5IQR, Q3+1.5IQR, outliers), by treatment arm.

14.4. Vital Signs

Data on vital signs consist of body temperature, blood pressure (systolic and diastolic, mmHg), heart rate (beats/min), and respiratory rate (breaths/min) and will be tabulated and listed. Abnormal values will be flagged and categorized as 'abnormal not clinically significant' or 'abnormal clinically significant'.

Descriptive statistics (n, arithmetic mean, SD, median, minimum, and maximum) of actual values and changes from baseline will be presented by treatment arm and stratum at each assessment time point.

Shift tables using the categories low /normal/high will be produced to compare baseline to the end of trial and to the worst post-baseline value by treatment arm.

For the overall evaluation of vital signs, shift tables using the categories abnormal, clinically significant/abnormal, not clinically significant/normal will be produced to compare baseline to the EOT and to the worst post-baseline value by treatment arm.

14.5. Physical Examination

Physical examination data together with evaluations will be listed by patient and time point. Abnormal findings will be flagged (abnormal not clinically significant/abnormal clinically significant).

14.6. ECGs

Triplicate 12-lead ECGs, including heart rate (beats/min), QT interval (ms), QTc interval corrected by Fridericia's formula (QTcF; ms), PR interval (ms), RR interval (ms) and QRS interval (ms), will be obtained for each patient during the trial. All measured values and means of the triplicates will be listed, including categories for QTcF values (as defined below).

Mean value at a given visit will be calculated as a mean of all evaluable measurements at the given time point.

Actual time point of the triplicate measurement will be defined as the time point of the last evaluable value.

Descriptive statistics (n, arithmetic mean, SD, median, minimum, and maximum) of triplicates will be presented by treatment arm and stratum at each assessment time point.

Categorical analysis of QT/QTc interval data based on the number of patients meeting or exceeding predefined limits in terms of absolute QT/QTc intervals or changes from baseline will be presented by treatment arm and stratum at each assessment time point.

Categorical analysis of QTcF data will be presented as follows:

- Absolute QTcF interval:
    - ≤450 ms

- >450 ms to ≤480 ms
- >480 ms to ≤500 ms
- >500 ms

- Increase from baseline in QTcF interval:
  - ≤30 ms
  - >30 ms to ≤60 ms
  - >60 ms

14.7. Gallbladder Imaging

Number and proportion of patients with gallstones and gallbladder sludge will be summarized by treatment arm and stratum.

Shift tables of gallstones and gallbladder sludge using the categories absent/present will be prepared to compare baseline to Week 24/EOT by treatment arm.

Gallbladder data will also be listed.

14.8. Local Tolerability

Patients will self-assess the injection site pain using an 11-point numerical rating scale (ranging from 0 to 10, where 0 represents no pain and 10 represents worst possible pain) immediately and 1 hour after the injections. The Investigator will assess the degree of erythema and swelling at the injection site at the same time points. Any event that occurs or persists beyond 1 hour after the injection will be collected as an AE.

Local tolerability results (erythema, swelling and pain scores) will be listed.

Frequency counts (number and percentages of patients) will be provided for erythema and swelling by treatment arm and stratum, and severity at each assessment time point. Pain assessments and change from immediately to 1 h after treatment will be summarized (n, N, arithmetic mean, SD, minimum, 25 percentile, median, 75 percentile and maximum) by treatment arm and stratum at each assessment time point.

15. OTHER PLANNED ANALYSES

15.1. Pharmacokinetic Analysis

The plasma concentrations of octreotide will be summarized using descriptive statistics (n, n [below the lower limit of quantification (LLOQ)], arithmetic mean, SD, coefficient of variation, geometric mean, geometric CV%, minimum, median and maximum) over time by dose, if applicable. All concentrations below the LLOQ or missing data will be reported as such in the concentration data

listings. Concentrations below the LLOQ will be treated as zero in summary statistics of arithmetic mean and as LLOQ/2 in the calculations of geometric mean.

Individual and summarized octreotide concentration-time profiles will be displayed graphically on linear and logarithmic scale. The FAS will be used for these presentations.

A population PK model will be developed using data generated from the current trial together with data from other clinical trials. This analysis will be detailed in a separate document, and a separate report will be completed.

15.2. Exposure-response Analyses

The relationship between octreotide exposure and QTcF change from baseline will be explored using correlation plots presenting QTcF intervals (ms) versus octreotide concentrations and QTcF change from baseline (ms) versus octreotide concentrations.

Individual and summarized IGF-1 concentration (absolute values and IGF-1/ULN on the same plot) versus octreotide concentration profiles will be displayed graphically.

Individual and summarized GH concentration versus octreotide concentration profiles will be displayed graphically.

PK, pharmacodynamic (PD; IGF-1 and GH) and safety (ECG) data generated from this trial may be used together with PK, PD and safety data from other clinical trials for population PK and PK/PD assessments.

Exposure-response analyses will be detailed in a separate document, and a separate report will be completed.

15.3. Immunogenicity Assessment

The number and proportion of patients with positive results for anti-drug antibody assessment, and the number and proportion of patients with neutralizing anti-drug antibodies (if any) will be summarized by treatment arm and stratum at each assessment time point.

Shift tables will be prepared to compare the presence of anti-drug antibodies at Week 24/EOT to the presence at screening. If applicable, shift tables will also be prepared to compare the presence of neutralizing anti-drug antibodies at Week 24/EOT to the presence at screening.

For the duration of previous octreotide treatment, descriptive statistics (n, arithmetic mean, SD, minimum, median and maximum) will be prepared by the presence of anti-drug antibodies and by the presence of neutralizing anti-drug antibodies (if applicable) at screening.

The plasma octreotide concentrations at Week 24/EOT will be summarized by the presence of anti-drug antibodies and neutralizing anti-drug antibodies (if applicable) at Week 24/EOT by treatment arm and stratum. If antibodies are present, the mean plasma octreotide concentration vs time curve will be plotted for patients with antibodies and patients without antibodies.

Number and proportion of patients with mean IGF-1 levels <1xULN at Week 22 and Week 24 (average of the 2 measurements) will be summarized by the presence of anti-drug antibody and by the presence of neutralizing anti-drug antibodies (if applicable) at Week 24/EOT by treatment arm and stratum.

Number and proportion of patients with related TEAEs and SAEs will be summarized by the presence of anti-drug antibody and neutralizing anti-drug antibodies (if applicable) at Week 24/EOT by treatment arm and stratum.

The incidence of anti-drug-specific antibodies may be correlated with other safety or efficacy parameters collected across the trial, as applicable.

All data will be listed by patient and time point.

16. DEVIATIONS FROM THE ANALYSIS PLANNED IN THE PROTOCOL

Not applicable.

17. APPENDICES

17.1. Appendix A / SAS code for multiple imputation

[redacted]

18. LIST OF APPLICABLE QUALITY DOCUMENTS

18.1. List of Relevant SOPs and work instructions (WINs)

Document Code	Title	Effective Date
[redacted]		

19. REFERENCES

No.	Designation/ Code/Appendix	Title
1	NA	ICH Harmonised tripartite Guideline: Guideline for Good Clinical Practice
2	NA	ICH Harmonised tripartite Guideline: STRUCTURE AND CONTENT OF CLINICAL STUDY REPORTS E3
3	NA	ICH Harmonised tripartite Guideline: STATISTICAL PRINCIPLES FOR CLINICAL TRIALS E9
4	NA	FDA-CDER: GUIDELINE FOR THE FORMAT AND CONTENT OF THE CLINICAL AND STATISTICAL SECTIONS OF AN APPLICATION
5	NA	US Federal Register. (1998) International Conference on Harmonization; Guidance on Statistical Principles for Clinical Trials. Department of Health and Human Services: Food and Drug Administration [Docket No. 97D-0174]. Federal Register Volume 63, Number 179, pages 49583-49598. September 16, 1998.
6	NA	ASA. (1999) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, August 7, 1999 http://www.amstat.org/committees/ethics/index.html
7	NA	RSS. (1993) The Royal Statistical Society: Code of Conduct, August 1993. http://www.rss.org.uk/main.asp?page=1875
9.	NA	CPMP. (2000) Points to Consider on Switching between Superiority and Non-inferiority, July, 2000. www.tga.gov.au/docs/pdf/euguide/ewp/048299en.pdf
10	NA	ICH Harmonised Guideline Estimands and Sensitivity Analysis in Clinical Trials E9(R1), Step 2 Version dated 16 June 2017
11	NA	Treatment Satisfaction Questionnaire for Medication (TSQM) User Manual version 2.3, 02 March 2021
12	NA	Badia X, Webb SM, Prieto L, Lara N. Acromegaly Quality of Life Questionnaire (AcroQoL). Health Qual Life Outcomes. 2004;2:13.
13	NA	Clopper CJ, Pearson ES. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. Biometrika.1934;26(4):404-13

No.	Designation/ Code/Appendix	Title
14	NA	EuroQol Research Foundation. "EQ-5D-5L user guide." *Basic information on how to use the EQ-5D-3L instrument. Version 3.0* (2019)
15	NA	Fleseriu M, Molitch M, Dreval A, Biermasz NR, Gordon MB, Crosby RD, Ludlam WH, Haviv A, Gilgun-Sherki Y, Mathias SD. Disease and treatment-related burden in patients with acromegaly who are biochemically controlled on injectable somatostatin receptor ligands. Front. Endocrinol. 2021; 12: 627711.
16	NA	Keininger D, Coteur G. Assessment of self-injection experience in patients with rheumatoid arthritis: psychometric validation of the Self-Injection Assessment Questionnaire (SIAQ). Health Qual Life Outcomes. 2011;9:2.
17	NA	Kernan, WN, Viscoli, CM, Makuch, RW, Brass, LM, Horwitz, RI. Stratified randomization for clinical trials. Journal of clinical epidemiology. 1999; 52(2):19-26.
18	NA	Liublinska, V, Rubin, DB. Sensitivity analysis for a partially missing binary outcome in a two-arm randomized clinical trial. Statistics in medicine. 2014; 33(24):4170-4185.
19	NA	Ratitch, B, O'Kelly, M. Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures PharmaSUG2011. Paper SP04
20	NA	Rencz, F, Brodszky, V, Gulácsi, L, Golicki, D, Ruzsa, G, Pickard, AS, Law, EH, Péntek, M. Parallel valuation of EQ-5D-3L and EQ-5D-5L by time trade-off in Hungary. Value in Health. 2020; 23(9):1235-1245.
21	NA	Rubin D.B. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons; 1987
22	NA	Wang, S, Hu, H. Impute the missing data using retrieved dropouts. BMC medical research methodology. 2022 Dec;22(1):1-5
Comments:		

NA: not applicable