SUMMITANALYTICAL	**STATISTICAL ANALYSIS PLAN** **CONFIDENTIAL**

STATISTICAL ANALYSIS PLAN

Protocol Title: A 16-Week, Single-Blind Randomized, Placebo- Controlled Food Study of the Safety and Tolerability of AXA1125 and AXA1957 in Subjects with Non-Alcoholic Fatty Liver Disease (NAFLD)

Protocol Number: AXA1125-003

Sponsor: Axcella Health Inc.
840 Memorial Drive, 3rd Floor
Cambridge, MA 02139 USA

Author: Summit Analytical, LLC
8354 Northfield Blvd.
Bldg. G Suite 3700
Denver, CO 80238

SAP Date: 2020-03-09

Status: FINAL 1.0

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

3. ABBREVIATIONS

ADaM	Analysis Data Model
AE	Adverse Event
ALT	Alanine Aminotransferase
ALP	Alkaline Phosphatase
ANOVA	Analysis of Variance
AST	Aspartate Aminotransferase
AUC	Area Under the Curve
BID	Twice a Day
BHB	Beta-Hydroxybutyrate
BMI	Body Mass Index
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence Interval
CK	Cytokeratin
CMH	Cochran-Mantel-Haenszel
CRF	Case Report Form
CRP	C-Reaction Protein
cT1	Corrected T1
CSR	Clinical Study Report
ECG	Electrocardiogram
EDC	Electronic Data Capture
ELF	The Enhanced Liver Fibrosis
eGFR	Estimated Glomerular Filtration Rate
EOS	End of Study
ET	Early Termination
GGT	Gamma-Glutamyl Transferase
GI	Gastrointestinal
GLP-1	Glucagon Like Peptide 1
HA	Hyaluronic Acid
HbA1c	Hemoglobin A1c
HOMA-IR	Homeostatic Model Assessment Insulin Resistance

IR	Insulin Resistance
MCP	Monocyte Chemoattractant Protein
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
NAFLD	Nonalcoholic Fatty Liver Disease
NEFA	Non-Esterified Fatty Acids
oGTT	Oral Glucose Tolerance Test
PIIINP	Amino-terminal Propeptide of Type III Procollagen
PBMC	Peripheral Blood Mononuclear Cells
PDFF	Protein Density Fat Fraction
PE	Physical Examination
PK	Pharmacokinetic
PP	Per-Protocol
PT	Preferred Term
QC	Quality Control
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SEM	Standard Error of the Mean
SOC	System Organ Class
T2D	Type 2 Diabetes
TIMP 1	Metalloproteinases 1
TZD	Thiazolidinediones
VAS	Visual Analog Scale

4. INTRODUCTION

4.1. Preface

This document presents a statistical analysis plan (SAP) for Axcella Health Inc. Protocol AXA1125-003 (A 16-Week, Single-Blind Randomized, Placebo-Controlled Food Study of the Safety and Tolerability of AXA1125 and AXA1957 in Subjects with Non-Alcoholic Fatty Liver Disease (NAFLD).

Reference materials for this statistical plan include the Protocol version 4.0 dated 01APR2019 and electronic case report forms (eCRF) version 4.0 dated 07MAR2019.

The SAP described hereafter is an *a priori* plan. The SAP will be finalized and approved prior to approved database lock. Statistical programming may occur as study data accumulate in order to have analysis programs ready at the time the study finishes. In such an event, arbitrary product administration group assignments must be randomly linked to subjects, effectively rendering any output of programs meaningless.

4.2. Purpose of Analyses

This study is intended to assess the safety and tolerability of proprietary amino acid food products, AXA1125 and AXA1957, in subjects with NAFLD. Results from the defined analyses will be included in the final clinical study report (CSR) for Study AXA1125-003, and may also be utilized for regulatory submissions, manuscripts, or other clinical development activities.

Post-hoc exploratory analyses not identified in this SAP may be performed to further examine the study data. These analyses will be clearly identified, where appropriate, in the final CSR, if included. Additional analyses may be performed on an exploratory/ad hoc basis to explore group level, or individual level comparisons of one or more biomarkers and clinical measures. Additional analyses not identified in this SAP may also be completed for publications, or regulatory or funding inquiries. These analyses, if performed, may not be reported in the CSR, but will be fully documented in the documents containing the additional analyses.

4.3. Summary of Statistical Analysis Changes to the Protocol

The analyses described in this analysis plan are consistent with the analyses described in the study protocol.

5. STUDY OBJECTIVES AND ENDPOINTS

5.1. Study Objectives

This study is intended to assess the safety and tolerability of proprietary amino acid food products, AXA1125 and AXA1957, in adult subjects with NAFLD.

5.1.1. Study Assessments/Endpoints

Safety and Tolerability Endpoint

- Incidents of reported SAEs, AEs and product-emergent AEs;
- Other safety assessments such as safety laboratory measures, vital signs and ECG.

Structure and function Endpoints

- Liver biomarkers: ALT, AST, ALP and GGT;
    - A binary endpoint (ALT Response, Y/N) is defined as at least 17 U/L absolute reduction in the post-baseline ALT from the baseline ALT;
- Metabolic biomarkers: oGTT, MRI-PDFF, HOMA-IR, Body weight, BMI and Waist circumference; Fasting lactate, pyruvate, glycerol, BHB, NEFA, and Adipose tissue IR
    - The endpoint of oGTT is paired $AUC_{0\text{-}120min}$ of glucose and insulin after an overnight fast;
    - A binary endpoint (PDFF Response – 1, Y/N) is defined as at least 5% absolute reduction in the post-baseline PDFF from the baseline PDFF;
    - A binary endpoint (PDFF Response – 2, Y/N) is defined as at least 30% relative reduction in the post-baseline PDFF from the baseline PDFF;
- Inflammation biomarkers: CK18-M65, CK18-M30, CRP, MCP-1, YKL-40 and cT1;
    - A binary endpoint (cT1 Response-1, Y/N) is defined as the post-baseline cT1 decreases at least 40 ms from the baseline cT1;
    - A binary endpoint (cT1 Response-2, Y/N) is defined as the post-baseline cT1 decreases at least 80 ms from the baseline cT1.
- Fibrosis biomarkers: ProC3 and ELF score;
    - ELF total score and ELF subscores (PIIINP, HA, TIMP).

- Fibrosis-4 (Fib-4) index defined as

$$\text{Fib-4} = \text{Age (years)} \times \text{AST (U/L)} / [\text{PLT}(10^9/\text{L}) \times \text{ALT}^{1/2} \text{ (U/L)}]$$

- Other laboratory evaluations listed in the table below

Table 1. Clinical Laboratory Evaluations

Chemistry (fasted)	**Hematology (CBC)**	**Lipid panel (fasted)**	**Metabolites (fasted)**	**Metabolic Panel (fasted)**	**Biomarkers (fasted)**	**Biomarkers (fasted)**	**Urinalysis**
Albumin Alkaline phosphatase (ALP) ALT AST Blood urea nitrogen (BUN) Calcium Chloride Creatinine Gamma-glutamyl transferase (GGT) Glucose (serum) Insulin (serum) Phosphorus Potassium Sodium Bilirubin (total) Bilirubin (direct) Total CO2 (measured as bicarbonate) Total protein Uric acid eGFR	Hematocrit Hemoglobin Mean corpuscular hemoglobin (MCH) Mean corpuscular hemoglobin concentration (MCHC) Mean corpuscular volume (MCV) Platelet count Red blood cell count White blood cell (WBC) count WBC differential (% & absolute) Basophils Eosinophils Lymphocytes Monocytes Neutrophils PBMC isolation Prothrombin time/INR	Total cholesterol High-density lipoprotein cholesterol (HDL-C) Low-density lipoprotein (LDL-C) Triglycerides Non-HDL-C Apolipoprotein B (ApoB) Apolipoprotein CIII (Apo-CIII) Lipoprotein a [Lp(a)]	Lactate Glycerol Pyruvate	Insulin (plasma) Glucose (plasma) Beta-hydroxybutyrate Non-esterified fatty acids (NEFA) Free amino acids Hemoglobin A1c (HbA1c) Homeostasis model assessment of insulin resistance (HOMA-IR; calculated) Adipose tissue insulin resistance (Adipo-IR; calculated) OWL NASH Index	Adiponectin FGF-21 IL-1beta High sensitivity C-reactive protein (hsCRP) MCP-1 CK-18 (M30 and M65) YKL-40 Alpha-2 macroglobulin, Enhanced liver fibrosis (ELF) Score (TIMP-1, PIIINP & hyaluronic acid)	N-terminal fragment of type III collagen (ProC3) Internal epitope in the 7S domain of type IV collagen (ProC4) Released N-terminal pro-peptide of type VI collagen (ProC6) Internal epitope in the N-terminal pro-peptide of type 1 collagen (P1NP)	pH Specific gravity Protein Glucose Ketones Bilirubin Blood Nitrate Urobilinogen Leukocyte esterase

6. STUDY METHODS

6.1. General Study Design and Plan

As background for the statistical methods presented below, this section provides an overview of the study design and plan of study execution. The protocol is the definitive reference for all matters discussed in what follows.

This is a 16-week, single-blind, randomized, placebo-controlled food study of the safety and tolerability of AXA1125 and AXA1957 in adult subjects with NAFLD.

Subjects will sign an Informed Consent Form and be screened for eligibility per the inclusion and exclusion criteria below, up to 6 weeks before the start of the Administration Period.

Following completion of all Screening procedures, eligible subjects will be randomized in a 2:2:2:1 ratio to receive either AXA1125 24g BID, AXA1957 20.3g BID, AXA1957 13.5g BID or isocalorically matched placebo 24g BID. Randomization will occur via an interactive web response system (IWRS) after eligibility is confirmed and approximately 3-5 days prior to the Day 1 visit. Assigned study food product (AXA1125, AXA1957 or placebo) will be shipped to the clinical site upon randomization of each subject.

Once randomization has occurred, subjects will present to the study site on Day 1 (Baseline/Visit 2) for their baseline assessments per the schedule of events prior to the first administration of study product in Table 1. Study Day 1 is the first day of the 16-week Administration Period.

Subjects will return to the study site at Week 1 (Visit 3), Week 2 (Visit 4), Week 4 (Visit 5), Week 8 (Visit 6), Week 12 (Visit 7) and Week 16 (Visit 8) to receive their study food product and/or return any unused study food product, to provide blood samples for biomarker and other laboratory testing, undergo liver imaging, and to complete other study safety assessments per the Schedule of Events described in the protocol.

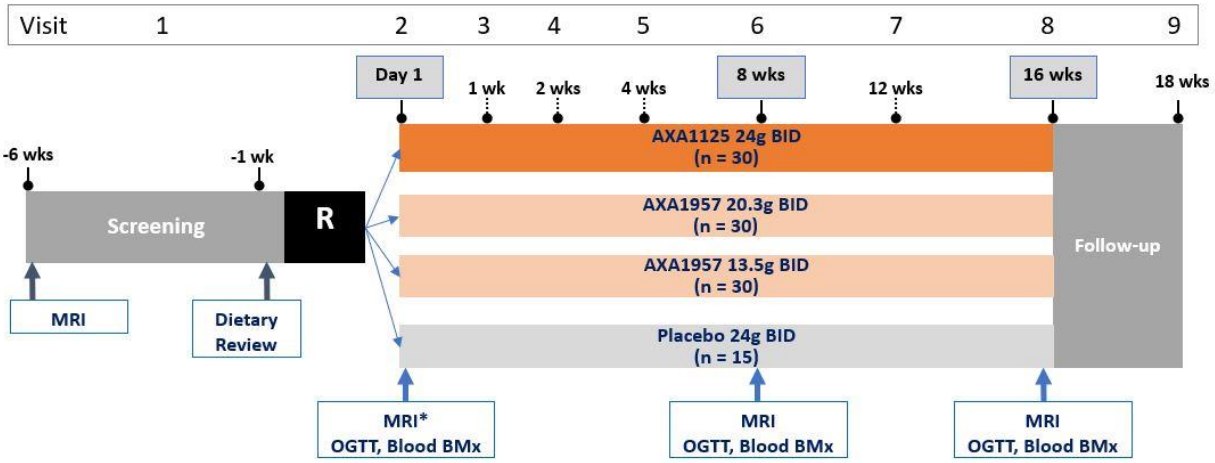
There will be a safety Follow-up Visit, approximately 2 weeks after the last visit in the Administration Period (i.e., after the Week 16 visit or at the time of early termination) which will be the End of Study (EOS) Visit (Visit 9).

There will be a total of 9 study visits, including the Screening and Follow up visits.

A diagram of the study design is provided in Figure 1 below and schedule of assessments is presented in Figure 1.

Study assessments will be performed per the Schedule of Assessments in Table 1. At each visit, AEs will be recorded and ongoing compliance to the study protocol will be assessed.

Figure 1. Study Design



*Screening MRI will be used as the Baseline MRI assessment if the Screening MRI occurs within approximately 14 days prior to Day 1.

MRI = magnetic resonance imaging; OGTT = oral glucose tolerance test; BMx = biomarkers

6.2. Inclusion – Exclusion Criteria and General Study Population

The study sample size will be approximately 105 subjects 18 years or older. The inclusion and exclusion criteria defined in the protocol apply to all subjects regardless of strata and are not repeated herein the SAP. Reference is made to the final protocol for the specific inclusion and exclusion criteria for study subjects.

6.3. Randomization and Blinding

Following completion of all Screening procedures, eligible subjects will be randomized in a 2:2:2:1 ratio to receive either AXA1125 24g BID, AXA1957 20.3g, BID AXA1957 13.5g BID or placebo 24g BID using a stratified randomization in blocks of size 7, with type 2 diabetes as the stratification factor. Randomization will occur via an interactive web response system (IWRS) after eligibility is confirmed and approximately 3-5 days prior to the Day 1 visit (note: Day 1 is also the baseline visit). Assigned study food product (AXA1125, AXA1957 or placebo) will be shipped to the clinical site upon randomization of each subject.

Subjects who are randomized but discontinue before administration of any study food product may be replaced per the pre-specified treatment allocation schedule.

6.4. Study Blinding

This is a single-blind, placebo-controlled study. Subjects will be blinded as to whether they receive AXA1125, AXA1957 or placebo food products. Site personnel dispensing the product to subjects for administration and sponsor will be unblinded to subject assignment.

6.5. Sample Size

It is anticipated to enroll and administer study product to approximately 105 subjects in a 2:2:2:1 stratified randomization schema (AXA1125 24g BID, AXA1957 20.3g BID, AXA1957 13.5g BID and placebo 24g BID, respectively) using blocks of size 7. Approximately 90 subjects will receive active study product and approximately 15 subjects will receive placebo. This study is exploratory in nature, and the sample size is based on clinical judgement that this number of subjects may be sufficient to provide a characterization of the product tolerability and safety.

7. GENERAL CONSIDERATIONS

7.1. Analysis Populations

7.1.1. Safety Analysis Population (SAFETY)

The "Safety Analysis Population" will be comprised of any subject who receives at least one administration of the study food product.

The Safety Analysis Population will serve as the primary analysis population for all planned analyses. In Safety Analysis Population, subjects will be analyzed according to the actual product administration received on the Day 1 Visit.

7.1.2. Per-Protocol Analysis Population (PP)

The "Per-Protocol Analysis Population" will be comprised of all randomized subjects who meet the following:

- Have at least one post-baseline MRI image measurement.
- Between $\geq$80% and $\leq$120% compliance rate of the assigned study food product administration
- No major protocol deviations.

In PP population, subjects will be analyzed according to the product administration that subjects were randomized to.

7.2. Covariates and Subgroups

7.2.1. Planned Covariates

Analyses, especially structure and function endpoint analyses, may be performed with the adjustment of covariates such as stratification factor (T2D status).

7.2.2. Planned Subgroups

Subgroup analyses (e.g., stratification factor, T2D status) may be performed for selected endpoints, especially in structure and function analyses.

7.3. Management of Analysis Data

7.3.1. Unscheduled Visits

Unscheduled visits will be counted toward the closest scheduled visit where a given measure is supposed to be taken according to Table 1. If there is already a measure for a subject at the

scheduled visit, then the unscheduled visit result will not be counted for the summary table. All results will be provided in listings as collected.

Early termination visit, not considered as unscheduled visits, will be mapped to Week 8 or Week 16 visits, with a +/- one-week window. If there is already a measure on those visits, then the early termination measure will be summarized separately.

7.3.2. Repeated Measures

For measures where only one sample is expected for selected visit, any repeated measures on the same day will be treated as repeats using methods below.

- For continuous measure, repeated measures for the same day will be averaged. If a measure is below the limit of quantification, then the limit of quantification is used for averaging purposes.

- For measures that are categorized as “Abnormal”, “Normal”, “High”, “Low” etc, the last measure will be used.

If a repeated measure is taken on a separate day, then it will be treated as unscheduled visit. For measures where multiple measures are taken on a given day, only measures closest to the target timepoints (e.g. 120 min post dose), would be included.

7.3.3. Missing Data

In general, there will be no substitutions made to accommodate missing data points aside from the few situations described below. For example, to compare percentage change from baseline of PDFF at Week 8 between placebo and 1125, only subjects in either placebo or 1125 group with non-missing PDFF data at both baseline and Week 8 (closely mapped early termination visit) will be included for this analysis.

Early termination measures which are able to be mapped to either Week 8 or Week 16 based on visit window will be included

All data recorded on the case report form (CRF) will be included in data listings that will accompany the clinical study report.

7.3.3.1. Handling of Missing Date Values

<u>Partial or Missing Dates</u>

The following conventions will be used to impute missing portions of dates for AEs and concomitant medications/therapies, if warranted. Note that the imputed values outlined here may not always provide the most conservative date. In those circumstances, the imputed value may be replaced by a date that will lead to a more conservative analysis.

Using the following imputation rules, if the start and stop dates are still both missing for an AE, then the AE is considered product-emergent. Similarly, a reported medication under such circumstances would be concomitant.

Using the following imputation rules, if an AE stop date is before the first administration of study product, then the AE will not be considered product emergent. Similarly, if a reported medication has an imputed stop date prior to the screen date, then the medication will be considered prior medication, as opposed to concomitant.

A. Start Dates
   1) If the year is unknown, then the date will not be imputed and will be assigned a missing value.
   2) If the month is unknown, then:
      i) If the year matches the first administration date year, then impute the month and day of the first administration date.
      ii) Otherwise, assign 'January.'
   3) If the day is unknown, then:
      i) If the month and year match the first administration date month and year, then impute the day of the first administration date.
      ii) Otherwise, assign the first day of the month.

B. Stop Dates
   1) If the year is unknown, then the date will not be imputed and will be assigned a missing value.
   2) If the month is unknown, then assign 'December.'
   3) If the day is unknown, then assign the last day of the month.

7.3.3.2. Missing Baseline Data

Every effort will be made to ensure that accurate baseline information of the subject is collected. In the event that a subject's baseline information is missing, the subject will be included in the safety analysis but excluded from any shift tables or change from baseline analyses directly pertaining to the missing parameter.

7.3.3.3. Missing Severity and Relationship to Study Product for AE

If the severity or relationship to study product is missing for AEs, then that AE will be marked as missing and included in the category of missing for the severity and/or relationship in the summary table. The relationship and/or severity will be presented as missing in listings.

7.3.4. Handling of Early Termination Visit Information

In the event that a subject is terminated early from this study, if the termination visit takes place instead of a scheduled visit and during the corresponding scheduled visit window, then the results collected will be counted toward that visit. Otherwise, the early termination visit results will be summarized separately after all other visits. All early termination (ET) visit results and information including the date of the ET and total weeks on study product at the time of ET will be presented as collected in listings.

7.3.5. Pooling of Investigational Sites

The data from all study centers will be pooled together for all planned analyses. Study site will not be included as a covariate in the statistical modeling. Summary statistics may be performed for an individual site if the number of subjects in each group enrolled from that site is deemed to be clinical meaningful.

Sensitivity analyses may be performed by excluding some site(s) with potential issues or patterns.

7.3.6. Coding Conventions for Events and Medications

All adverse events, and medical history will be mapped to the Medical Dictionary for Regulatory Activities (MedDRA Version 21.0) system for reporting (preferred term and body system).

Prior and Concomitant medications will be coded using WHO-DD (Drug Dictionary) (C3 Format – 01Mar2018).

7.3.7. Analysis Software

Data manipulation, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will be performed primarily using SAS (release 9.4 or higher) for Windows. If the use of other software is warranted, the final clinical study report will detail what software was used and for what purposes.

7.3.8. Study Data

Study data identified in the schedule for time and events (Table 1) are collected, and source verified, on the electronic data capture tool: Fusion eClinical Suite.

A separate analysis plan will be prepared for data collected by different vendor OWL.

Study data will not be completely CDISC compliant for the purpose of current planned analyses, but ADaM-like datasets will be programmed and QC'ed to support final TFLs.

7.4. Planned Study Analyses

7.4.1. General Statistical Consideration

All analyses will be performed at the significant level of two-sided 0.05, which will be considered exploratory in nature, with the purpose of generating hypotheses and guiding a future program development path.

Descriptive summaries of variables will be provided where appropriate. For continuous variables, the number of non-missing values (n), mean, standard deviation, median, minimum, and maximum will be tabulated by product administration group. For categorical variables, the counts and proportions of each value will be tabulated by product administration group and visit. Expansion of descriptive table categories within each product administration group may occur if such elaborations are thought to be useful.

All study data, except for OWL data, will be presented in data listings. Study related data not subject to analyses according to this plan may not appear in any tables or figures.

7.4.2. Interim Analyses and Data Monitoring

Axcella plans to perform a non-binding interim analysis when all the subjects reach their Week 8 visit and all the assessments' results from that visit are available or the data snapshot of October 15, 2019, whichever comes first. Axcella does not plan to make any modifications of this study based on this non-binding interim look, but rather to monitor the safety of study subjects and potentially make decisions on both AXA1125 and AXA1957 programs' further development path.

In such event, all planned analyses will be performed primarily on Week 8 data. The same planned analysis may be performed on limited Week 16 data, as well as some potential trend analyses on selected structure and function endpoints.

7.4.3. Stopping Rules

No stopping rule is planned for this study.

7.4.4. Final Analysis and Publication of Study Results

The final analysis will be completed after all subjects have completed the study, and database is locked.

7.5. Multiple Testing Procedures

All analyses will be performed at the significant level of two-sided 0.05, which will be considered exploratory in nature, with the purpose of generating hypotheses and guiding future program development path. There will be no multiplicity adjustment in this study.

7.6. Subject Disposition

A summary of subject disposition will include the number and percentage of subjects for the following categories by randomized product administration groups: subjects randomized, subjects in the Safety analysis population, Per-Protocol population, subjects completed study through the week 18 follow up Visit, subjects terminated early from administration period and from study, respectively. For subjects that terminate early, the exact date of termination and days/weeks on the assigned study product prior to termination will be provided in listings. All percentages will be based on the number of subjects that received the product. Reasons for early termination will also be summarized and percentages are calculated using the total number of early terminated subjects as denominators. All randomized subjects will be included in this summary. Number and percentage of screen fails will be summarized in each category of screening failure reasons.

Above information will also be presented in data listings.

7.7. Protocol Deviations

All subject-level protocol deviations will be presented in a document listing for all randomized and product-administered subjects.

A summary table of all major subject-level protocol deviations will be also presented.

7.8. Demographics and Baseline Characteristics

The demographic data and baseline characteristics will be summarized for the Safety Analysis Population and Per-protocol Population. Individual subject demographics and baseline characteristics will be provided in listings. The demographic and baseline characteristics of all screened subjects (including screen failure) may be summarized, if available.

The demographics consist of age (year), gender, childbearing potential, race, and ethnicity. The baseline characteristics consist of baseline height (cm), baseline body weight (kg), baseline body mass index (BMI) (kg/m^2), and baseline waist circumference (cm). Body mass index is calculated as (body weight in kilograms) / (height in meters)^2.

A subject's age in years is calculated using the date of the informed consent and date of birth. Age, baseline height, baseline body weight, baseline waist circumference, and baseline BMI will be summarized using descriptive statistics, including the provision of ranges, T2D, NAFLD score at baseline. The number and percentage of subjects by gender, race and ethnicity will also be reported. Percentages will be based on the total number of subjects in the safety analysis population presentation.

Same summary statistics tables will be repeated in the Per-Protocol Population.

All demographic and baseline information will be listed by product administration group received and by subject.

7.9. Medical History

Medical history will be coded using the MedDRA Version 21.0. The number and percentage of subjects with any medical history will be summarized overall and for each system organ class (SOC) and preferred term (PT). Percentages will be calculated based on number of subjects in the Safety analysis set.

Subject medical history data including specific details will be presented in a listing.

7.10. Concomitant Medication

Prior Medications are medications started prior to first product administration. Concomitant medications and therapies are those taken at or after the first product administration. If start date is missing, medication/therapies are treated as concomitant to be conservative.

All prior and concomitant medications will be summarized by ATC and preferred term, as well as listed.

7.11. Concomitant Procedures

All concomitant procedures will be listed but not summarized.

7.12. Product Administration Compliance and Exposure

There are two sources of data collected regarding study product administration, carton count will be used as the primary source of compliance calculation, and eDiary data will be listed only.

Continuous descriptive summaries will be presented for the Safety Analysis Population by each of product administration groups and overall combined. Product administration compliance will be calculated between Day 1 and the Week 16 visit. Product administration compliance will be summarized by the compliance rate < 80%, >=80% and <=120%, and > 120%.

Compliance is calculated using the carton accountability CRF and is defined as:

(Total number of sticks dispensed – Total number of sticks returned) /2/#Stick packs per administration / (Date of current visit – Date of first dose + 1), where Total number of sticks returned is defined as number of full cartons x 42 + sticks returned from partial cartons.

stick packs per administration is 4, 3, 2, 4 for AXA1125 24g BID, AXA1125 20.3g BID, AXA1125 13.5g BID and Placebo 24g BID respectively.

Total number of days subjects are exposed to study product based on the patient diary will also be summarized as a continuous variable.

8. SAFETY ANALYSES

All Safety analyses will be conducted using the Safety Analysis Set.

8.1. Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), version 21.0.

Product-emergent AEs are defined as any AE onset at the time of or after the first administration of the study food product.

If the onset of an AE is on Day 1, then the time of first administration, collected from the study product administration CRF, will be compared to the time of onset to determine if the AE is Product-emergent.

AEs with missing dates or otherwise unclear temporal relations to the start of study product, will be listed, but not included in the summary tables.

The number and percent of subjects with any product-emergent AEs will be summarized by system organ class (SOC) and preferred term (PT) by product administration group. At each level of tabulation (ex. at the preferred term level) subjects will be counted only once if they had more than one such event reported during the AE collection period.

Missing data are treated as described in section 8.3.2.

The following summary tables and subject level listings will be presented:

- Overall summary (Number of AEs, Product-emergent AEs, related product emergent AEs, SAEs, Product-emergent AEs leading to study product discontinuations, interruption or product administration amount reduction).
- Summary table of product-emergent AEs by SOC and PT
- Summary table of SAEs (both product-related and non-related) by SOC and PT and causalities to the administered product
- Summary table of product-emergent AEs by highest relationship level to study product by SOC and PT (Definitely related and Possibly related will be pooled as Related).
- Summary table of product-emergent AEs by severity by SOC and PT
- Summary table of product-emergent AEs leading to study product discontinuation, interruption or product administration amount reduction by SOC and PT
- Summary table of the average onset day and average duration for diarrhea, nausea and abdominal pain.
- Summary table of diarrhea, nausea and abdominal pain by relationship and by severity separately.

- Listing of product-emergent AEs for diarrhea, nausea and abdominal pain.
- Listing of SAE for diarrhea, nausea and abdominal pain.

All AEs will be listed.

8.2. Deaths, Serious Adverse Events and Other Significant Adverse Events

8.2.1. Deaths

All deaths, regardless of causality, will be provided in a listing.

8.2.2. Serious Adverse Events (SAEs)

SAEs will be summarized by SOC and PT and the causalities to the administered product. A listing of Serious Adverse Events (SAEs) (both product-related and non-related) will be provided.

8.2.3. Adverse Events Leading to Discontinuation, Interruption or Reduction of Study Product

A listing of all AEs leading to study discontinuation, interruption or reduction of study product will be presented.

8.2.4. Adverse Events of Diarrhea, Nausea and Abdominal Pain

The number of days between baseline and the AE start dates are summarized in addition to average duration for those three specific AEs will be summarized. Only product-emergent, and PTs diarrhea, nausea and abdominal pain are included. For the average duration summary, only AEs with an end date will be included.

8.3. Pregnancy Examination

For subjects with childbearing potential, serum pregnancy test results and screening FSH results will be listed by product administration group and subject.

8.4. Vital Signs

In a listing, all vital signs, including body weight (kg), BMI (kg/m^2), waist circumference (cm), systolic blood pressure (mmHg), diastolic blood pressure (mmHg), temperature (℃), heart rate (bpm) and respiration rate (breaths/minute) will be presented for the Safety Analysis Population.

8.5. 12-Lead ECG

Shift from baseline tables will also be presented presenting the three possible responses Normal, Abnormal Not Clinically Significant, and Abnormal Clinically Significant.

All 12-lead ECG data by patient will be presented in a listing. Unscheduled visit results will be presented per Section 8.3.1.

8.6. Physical Examinations

Since abnormal physical exams will be summarized as AEs, no separate summary tables will be presented. All physical exam results will be listed for the Safety Analysis Population.

8.7. Clinical Laboratory Evaluation

Normal ranges are designated by the Central Laboratory. Laboratory shift tables for the worst post-baseline from each status at baseline may be provided for some laboratory parameters.

In circumstances where multiple results are collected for the same visit of the same subject, the last result within the visit window will be used.

Unscheduled visit laboratory results will be handled per Section 8.3.1.

All laboratory data will be listed per subject by product administration group and visit.

9. STRUCTURE AND FUNCTION BIOMARKER ANALYSES

All structure and function biomarker analyses will be performed in the Safety Analysis Population and may be repeated in the Per-protocol Population as sensitivity analysis.

All analyses may be stratified by baseline T2D status and may be repeated in each of T2D subgroups.

Some biomarkers (i.e., liver, metabolic, inflammation and Fibrosis) below, not limited to will be summarized for each product administration group by time point for the observed non-missing value as well as for the change (absolute and percent) from baseline value. The same analyses may be repeated for other biomarkers collected but not specified below. Shift tables for the worst post-baseline contrasted by status at baseline may be presented for some biomarkers, such as ALT, AST and Bilirubin.

- Liver biomarkers: ALT, AST, ALP, GGT and beta-hydrooxybutyrate (BHB);
- Number of subjects who have ALT reduction of at least 17 U/L from baseline;
- Metabolic biomarkers: oGTT, MIR-PDFF, HOMA-IR, Body weight, BMI and Waist circumference;
- Inflammation biomarkers: CK18-M30, CK18-M65, CRP, MCP-1 and cT1;
- Fibrosis biomarkers: ProC3 and ELF score, as well as ELF subscores (TIMP 1, HA, PIIINP);

For categorical variables, the Cochran-Mantel-Haenszel test (CMH) with the adjustment of the stratification factor (T2D status at baseline) will be applied. The following SAS code will be used.

```
proc freq;
      tables t2d*dose*Response / cmh;
   run;
```

In such analyses, if the number of counts in at least one cell is below 5, which may be deemed too small for the assumption of normality, the exact CMH test will be applied to that analysis using the following SAS procedure.

```
proc multtest;
         class ordlevel;
         test ca(y / perm=5);
         strata t2d;
         contrast 'ordtreat' 1 2 3 4;
         run;
```

For continuous variables, one-way or two-way ANOVA with the adjustment of the stratification factor (T2D status at baseline) will be applied.

```
Proc ANOVA;
     Class t2d dose;
     Model response=t2d*dose;
Run;
```

If the normality assumption doesn't hold, nonparametric ANOVA will be also applied to the comparison.

```
Proc npar1way;
     By t2d;
     Class dose;
     Var response;
Run;
```

The point estimation (LSM) with 95% CI and corresponding p-value from the comparison will be reported.

Pair-wise comparisons include:

- AXA1957 20.3g BID vs. Placebo
- AXA1957 13.5g BID vs. Placebo
- AXA1125 24.0g BID vs. Placebo
- AXA1957 20.3g BID vs. AXA1957 13.5g BID
- AXA1125 24.0g BID vs. AXA1957 13.5g BID
- AXA1125 24.0g BID vs. AXA 1957 20.3g BID

If the observed difference between AXA1957 20.3g BID and AXA1957 13.5g BID is considered not clinical meaningful, additional comparisons which two AXA1957 groups are combined may be performed.

The spaghetti plot of all of each individual subject, as well as boxplots and line plots of mean (+/- SEM) by study product group at each time point may be presented. The spaghetti plot may be reproduced by each study product group as well as T2D status at baseline.

Additional collected biomarkers will be summarized and analyzed if applicable.

9.1. oGTT Analysis

oGTT including fasting insulin, fasting glucose, AUC glucose, AUC insulin, peak concentration comparison (within the same day) for insulin and glucose. Analyses will be

stratified by T2D status at baseline, or performed as subgroup analyses for subjects with/without T2D at baseline.

Fasting insulin, fasting glucose area Under the Curve ($AUC_{0\text{-}120min}$) for both glucose and insulin values derived using trapezoidal rule and peak concentration comparison will be summarized descriptively. If a sample has time over 130 minutes post-baseline, it will not be included in the AUC calculation.

Peak concentration is the highest concentration within the same day as dosing.

9.2. PDFF and cT1 Analysis

The following endpoints will be summarized for visit.

- PDFF reduction by at least 5 points from baseline;
- PDFF reduction by at least 30% from the baseline;
- cT1 decrease of at least 40 ms from the baseline.
- cT1 decrease of at least 80 ms from the baseline.

Analyses will be stratified by T2D status at baseline, and repeated for each of T2D subgroups.

9.3. Hunger and Satiety Visual Analogue Scale (VAS)

Starting on Day 1, and then one day each week of the Administration Period subjects will complete the VAS. On the days the VAS is completed subjects will fill-out the VAS worksheets in conjunction with each study food product administration for that day (e.g., morning and evening) at two timepoints:

1) immediately prior to the study food product administration; and

2) immediately prior to eating the meal where the study food product is administered.

All VAS results will be listed by product administration group and visit using the safety analysis population.

All VAS results will be listed only.

9.4. Summary of HOMA-Insulin Resistance (IR) Score

HOMA-IR score will be assessed at baseline, weeks 2, 4, 8, 12 and 16. The score and change from baseline will be summarized.

Subgroup analysis will be performed by baseline T2D status.

10. REFERENCES

ICH E9 Expert Working Group. Statistical Principles for Clinical Trials: ICH Harmonized Tripartite Guideline, September 1998