STATISTICAL ANALYSIS PLAN

(1.0 Version)

Arachidonic acid supplementation improves cognitive impairment of schizophrenia: a pilot study

ETHICAL APPLICATION:

PROTOCOL NUMBER:

PREPARED BY:

INSTITUTION:

DATE:

ACKNOWLEDGEMENT AND SIGNATURE SHEET

Approved: **Date:**

Approved: **Date:**

Approved: **Date:**

TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Term
AA	Arachidonic Acid
BMI	Body Mass Index
CRF	Case Report Form
CSR	Clinical Study Report
FAS	Full Analysis Set
GC-MS	Gas Chromatography-Mass Spectrometry
LC-MS	Liquid Chromatography-Mass Spectrometry
PPS	Per-Protocol Set
SAP	Statistical Analysis Plan
SS	Safety Set
SZ	Schizophrenia

2. PURPOSE OF THE ANALYSES

This statistical analysis plan (SAP) provides a detailed description of the strategy and statistical methodology to be used for analysis of data from the protocol of arachidonic acid supplementation improves cognitive impairment of schizophrenia: a pilot study, **version 3.0, dated 25NOV2019.**

The purpose of the SAP is to describe the pre-specified statistical approaches to the analysis of study data prior to database lock. This analysis plan is meant to supplement the study protocol. If differences occur between analyses described in the SAP and the current protocol, those found in this SAP will assume primacy. Any deviations from this plan will be described in the Clinical Study Report (CSR).

3. PROTOCOL SUMMARY

The schizophrenia (SZ) cohort in this study was designed as a pilot study using a randomized, double-blind, placebo-controlled design. SZ patients could undergo a baseline (week 0) visit if they met all inclusion criteria and none of the exclusion criteria. Approximately 80 SZ patients were randomly assigned in a 1:1 ratio to either the intervention group or the placebo group: 40 to the arachidonic acid (AA) oral supplementation group (intervention group) receiving 338 mg/day, and 40 to the fatty acid-based placebo oral supplementation group (placebo group) without AA. The intervention period was 6 weeks. Cognitive assessments were conducted at baseline (week 0), midpoint (week 3), and endpoint (week 6); peripheral venous blood samples were collected at baseline (week 0) and endpoint (week 6); niacin skin tests were performed at baseline (week 0), midpoint (week 3), and endpoint (week 6).

The healthy cohort in this study plans to recruit 14 subjects who meet all inclusion criteria and none of the exclusion criteria for AA oral supplementation at a dose of 1400 mg. The supplementation period is 8 weeks, followed by a 4-week follow-up after supplementation ends. Peripheral venous blood samples will be collected once every 2 weeks starting from baseline (week 0), for a total of 7 time points; niacin skin tests will be conducted once every 2 weeks starting from baseline (week 0), for a total of 7 time points.

4. GENERAL ANALYSIS AND REPORTING CONVENTIONS

This section discusses general policies to be employed in the analysis and reporting of the data from the study. Departures from these general policies will be described, if applicable, in the appropriate sections of this SAP. When this situation occurs, the rules set forth in the specific section take precedence over the general policies.

For categorical variables, summary statistics will consist of the number and percentage of subjects in each category; percentages will be out of the number of subjects in the population being reported, unless otherwise noted. All percentages will be rounded to one decimal point. The number and percentage of subjects will always be presented in the form XX (XX.X%) where the percentage is in parentheses. To ensure completeness, all summaries for categorical and discrete variables will include all categories, even if none of the subjects had a response in a particular category. Denominators for each analysis will be based on the population of interest (e.g., safety population; subjects with non-missing data).

For continuous variables, summary statistics will consist of the number of subjects with data, mean, median, standard deviation (SD), minimum, and maximum values. The summary statistic n will be the number of subjects with non-missing values. All means and medians will be reported to one more significant digit than the values being analyzed. Standard errors (SE) and standard deviations will be reported to two more significant digits than the values being analyzed. The minimum and maximum will be reported to the same number of significant digits as the values being analyzed.

For tests of hypothesis of treatment group differences, the associated p-value will be reported. All p-values will be rounded to three decimal places; p-values that round to 0.000 will be presented as "<0.001". P-values are descriptive

In general, the baseline value will be considered the last non-missing measurement observed prior to the first dose of study treatment.

Data will be listed by treatment and subject. In general, listings will be sorted in the order that columns are displayed, starting with the first column on the left (treatment). Subject listings of data will be presented for all randomized subjects unless specified otherwise.

Unless otherwise specified, summaries will include the following treatment groups:

- AA intervention group
- Placebo group

R statistical software, version 4.0 or higher, will be used for all analyses.

4.1. Assessment Time Windows

Assessments should be performed within the windows stated in the protocol and will be analyzed by the visit/time point that they are entered into the case report form (CRF) under. For assessments other than the cognitive function, if an assessment is missing and an unscheduled or early termination assessment falls within the protocol-specified window, it will be assigned to that visit/time point for the purposes of summarization. If more than one unscheduled visit/ early termination visit falls into a window, the one closest to the scheduled time point will be used (with the earlier one used in case of a tie); again, scheduled visits will always take precedence, regardless of timing.

5. ANALYSIS SAMPLES

Full Analysis Set (FAS)

Following the intent-to-treat principle, the FAS comprises all subjects who are randomized. This set is utilized as a secondary population for efficacy evaluation.

Per-Protocol Set (PPS)

The PPS includes all participants who adhere to the treatment protocol and complete the planned follow-ups without significant deviations that could influence the treatment's efficacy. These deviations, which are identified during data auditing and may affect efficacy, can include scenarios such as:

- Hospital Discharge
- Adverse Event
- Death
- Lack of Efficacy
- Lost to Follow-up
- Physician Decision
- Pregnancy
- Protocol Deviation
- Study Terminated by Sponsor
- Technical Problem
- Withdrawal by Subject

The PPS is the primary population for efficacy analysis.

Safety Set (SS)

The SS encompasses all individuals who have received at least one session of AA intervention and have subsequent safety assessment data. This set is the key group for safety evaluations within this study.

6. STUDY SUBJECTS

6.1. Disposition of Subjects

The number and percentage of subjects screened, enrolled, randomized, received treatment, completed the study, and discontinued from the study, will be reported, along with the reason for discontinuation (counts only will be reported for screened and enrolled subjects). Percentages will be out of the number of subjects randomized. Additionally, subjects screened and enrolled will be reported overall; the remaining items will be reported by treatment group and overall. Subjects randomized will include all subjects for whom randomization is checked on the CRF; subjects receiving treatment will include all subjects that had any administered; subjects completing the study will be based on the recorded disposition.

Reasons for discontinuation include the following:

- Hospital Discharge
- Adverse Event
- Death
- Lack of Efficacy
- Lost to Follow-up
- Physician Decision
- Pregnancy
- Protocol Deviation
- Study Terminated by Sponsor
- Technical Problem
- Withdrawal by Subject
- Other

Additionally, the number and percent of subjects in each analysis population will be reported by treatment. Screen failures will also be included in a by-subject listing.

6.2. Demographic and Other Baseline Characteristics

Demographic and baseline characteristics will be collected during the Screening Visit. Descriptive statistics will be provided for all demographic and baseline characteristics based on the Safety Population. For categorical variables, the number and percentage of subjects in

each category will be presented. For continuous variables, summaries will include the number of subjects with data, mean, median, standard deviation, minimum, and maximum.

Variables to be summarized include:

- Age at screening (years) as recorded on the CRF
- Sex at birth
- Weight (kg)
- Height (cm)
- BMI (kg/m^2)

All demographic and other baseline characteristics will be provided in a listing.

6.3. Prior and Concomitant Medications

All medications taken during the screening period through the end of the study or early termination will be recorded in the concomitant medications log in the CRF, with the exception of rescue analgesic medications, which are collected on the rescue medications page. All medications will be coded using the World Health Organization Drug Dictionary (WHO-DD). Prior medications will be summarized in a table by treatment group using PPS individuals summarizing it by intervention group and placebo group. A subject will be counted only once for each medication

6.4. Medical History

Medical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA). Conditions will be listed, including the verbatim investigator description of the relevant medical condition, the coded terms, start date, end date, and whether or not the condition is ongoing.

7. STUDY OPERATIONS

7.1. Protocol Deviations

A protocol deviation is any noncompliance with the clinical trial protocol requirements. The noncompliance may be either on the part of the participant, the investigator, or the study site staff. Protocol deviations will be recorded on the source documents and with an explanation for the deviation. All protocol deviations will also be recorded as specified in the monitoring plan.

Protocol deviations will be identified by site staff, through medical reviews, and by clinical research associates during site monitoring. Deviations will be classified as minor or major prior to the database lock. Major protocol deviations are defined as those that result in harm to the study patients or significantly affect the scientific value of the reported results of the study. Other deviations will be considered minor.

All protocol deviations will be summarized in a table and well as presented in a listing, including their assigned severity (major/minor).

7.2. Randomization

The study employs a validated system to generate a randomization list for participants, automatically assigning participant numbers to randomization numbers, which are linked to various intervention plans. Utilized the Minirand package in R for dynamic randomization based on minimized differences, the randomization probabilities are adjusted dynamically based on predetermined covariates such as age, gender, and BMI, ensuring balance in baseline characteristics across treatment groups. Participants are ultimately randomly assigned in a 1:1 ratio to either the 338mg AA intervention group or the placebo group. The randomization numbers are kept confidential from patients and researchers and are not disclosed to any on-site staff. A by-subject listing of randomized treatment group and randomization number will be presented.

7.3. Measures of Treatment Compliance

AA and the placebo are both centrally distributed by the hospital, accompanied by a medication information intake form. This form is completed by the responsible physician or medical staff, meticulously documenting the time and dosage of each administration of medication, as well as whether the subject has taken the medication correctly according to the study protocol. These records serve as the primary basis for assessing subject compliance and are summarized and analyzed at the conclusion of the study.

Patient exposure to the study duration is calculated as follows:

Study duration (days) = (Last follow-up time - Informed consent date) + 1.

8. ENDPOINT EVALUATION

8.1. Overview of Efficacy Analysis Methods

8.1.1. Multicenter Studies

This is a single-center study and no special considerations are required for pooling of data.

8.1.2. Timing of Analyses

All final, planned analyses will be performed after the last participant has completed all study assessments, all relevant study data have been processed and integrated into the analysis database, and the database has been locked.

Any post-hoc, exploratory analyses completed to support planned study analysis, which were not identified in this SAP, will be documented and reported in appendices to the CSR. Any results from these unplanned analyses (post-hoc) will also be clearly identified in the text of the CSR.

8.1.3. Multiple Comparisons/Multiplicity

Unless otherwise indicated, all statistical tests will be carried out at a two-sided significance level of 5%. Given that all efficacy assessments are considered as exploratory only and are intended to inform future trials, no multiplicity adjustment shall be applied to these exploratory analyses.

8.2. Primary Efficacy Endpoint

8.2.1. Primary Efficacy Endpoint and Statistical Analysis

All efficacy endpoints are exploratory. However, the primary endpoint of interest is cognitive function assessed using the Cambridge Neuropsychological Test Automatic Battery (CANTAB) for AA subjects compared with placebo subjects.

To assess the changes in cognitive function from baseline to endpoint within the intervention group, the study will utilize either a paired t-test or a paired non-parametric test. The decision to use a paired t-test or a paired Wilcoxon signed-rank test will be guided by the outcome of normality tests on the data distribution. If the data are normally distributed, a paired t-test will

be conducted; otherwise, a paired Wilcoxon signed-rank test will be employed to analyze the data.

8.3. ENDPOINT EVALUATION

8.3.1. Fatty Acid Detection and Statistical Analysis

Fatty acids detection, as one of the secondary endpoints in this study, aims to assess the changes in fatty acid levels and their relationship with treatment outcomes.

Blood samples will be collected from subjects at baseline and scheduled follow-up time points, and stored at -80°C in an ultra-low temperature freezer to maintain the stability of the fatty acids. A standardized chemical extraction procedure will be used to extract fatty acids from the blood samples, ensuring the purity and recovery rate during the extraction process to obtain reliable analytical results. Quantitative analysis of fatty acids will be conducted using Gas Chromatography-Mass Spectrometry (GC-MS) to determine the types and concentrations of fatty acids in the samples.

R will be employed as the primary software to evaluate and visualize the relationship between changes in fatty acid levels and different treatment groups. Descriptive statistical analysis will be performed, including calculations of the mean, median, maximum, and minimum values of fatty acid levels. Additionally, based on the distribution characteristics of the data, t-tests or non-parametric tests will be used to compare fatty acid level differences between different time points and treatment groups; if the data do not conform to a normal distribution, the non-parametric Wilcoxon signed-rank test will be applied. Furthermore, Pearson correlation analysis will be conducted between fatty acid data and cognitive level data to explore their relationship with treatment effects.

8.3.2. Mitochondrial Lipidome Detection and Statistical Analysis

Mitochondrial lipidome detection, as one of the secondary endpoints in this study, aims to assess the changes in mitochondrial lipidome and their relationship with treatment outcomes.

Blood samples will be collected from subjects at baseline and scheduled follow-up time points, and mitochondrion will be extracted from the white blood cells of these samples, which will be store at -80°C in an ultra-low temperature freezer to maintain the stability of the lipids. A standardized chemical extraction procedure will be used to extract lipids from the mitochondria, ensuring the purity and recovery rate during the extraction process to obtain reliable analytical

results. Quantitative analysis of mitochondrial lipidome will be conducted using Liquid Chromatography-Mass Spectrometry (LC-MS) to determine the types and concentrations of lipids.

R will be employed as the primary software to evaluate and visualize the relationship between changes in fatty acid levels and different treatment groups. Descriptive statistical analysis will be performed, including calculations of the mean, median, maximum, and minimum values of fatty acid levels. Additionally, based on the distribution characteristics of the data, t-tests or non-parametric tests will be used to compare fatty acid level differences between different time points and treatment groups; if the data do not conform to a normal distribution, the non-parametric Wilcoxon signed-rank test will be applied. Moreover, Weighted Gene Co-expression Network Analysis (WGCNA) combined with Pearson correlation analysis will be conducted between lipid data and cognitive level data to explore their relationship with treatment effects.

8.3.3. Transcriptome Detection and Statistical Analysis

Transcriptome detection, as one of the secondary endpoints in this study, aims to explore the changes in the transcriptome and their potential association with treatment outcomes.

Blood samples will be collected from subjects at baseline and scheduled follow-up time points, with a particular focus on isolating total RNA from white blood cells in these samples. These samples will be stored at -80 ℃ in an ultra-low temperature freezer to maintain the integrity and stability of the RNA. A standardized RNA extraction procedure will be employed, using column purification techniques or similar methods to extract total RNA from white blood cells, ensuring the purity and recovery rate during the extraction process to obtain reliable analytical results. The extracted RNA will undergo quality checks, including concentration and purity assessments, to ensure the quality of subsequent transcriptome sequencing.

Transcriptome data analysis will begin with raw data obtained from high-throughput sequencing platforms, processed through a series of standardized bioinformatics procedures. Initially, sequencing data will undergo quality control to remove low-quality reads and adapter sequences. Subsequently, high-quality reads will be aligned to the reference genome to determine the expression levels of each gene. Quantitative analysis will assess differences in gene expression and identify genes with significant expression changes at different time points and between treatment groups. Additionally, multivariate statistical methods, such as clustering analysis and principal component analysis, will be used to reveal correlations between samples and gene expression patterns. To further explore the biological significance of gene expression

changes, functional enrichment analyses, such as Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses, will be conducted to identify potential biological processes and signaling pathways.

8.3.4. Niacin Skin Flushing Test and Statistical Analysis

Niacin skin flushing test, as one of the secondary endpoints in this study, aims to explore in the degree of niacin skin reaction and their potential association with treatment outcomes.

The niacin skin flushing test will be conducted on subjects at baseline and at scheduled follow-up time points. A niacin skin flushing test device will be utilized to evaluate the local skin response of subjects to niacin. During the test, the subjects' arms will be stimulated with niacin, and continuous images of the skin post-stimulation will be captured. These images will be transformed into key feature values through image recognition technology.

R will be used as the primary software to assess and visualize the changes in niacin skin flushing degree feature values and their relationship with different treatment groups. Descriptive statistical analysis will be conducted, including calculations of the mean, median, maximum, and minimum values of the niacin reaction degree feature values. Additionally, based on the distribution characteristics of the data, t-tests or non-parametric tests will be used to compare differences in niacin reaction degree feature values between different time points and treatment groups; if the data do not conform to a normal distribution, the non-parametric Wilcoxon signed-rank test will be applied. Furthermore, Pearson correlation analysis will be performed between the niacin skin flushing degree feature values and cognitive level data to explore their relationship with treatment outcomes.

8.4. Examination of Subgroups

There are no preplanned analyses of subgroups; exploratory analyses of subgroups may be performed, but will be post-hoc.

9. SAFETY EVALUATION

9.1. Overview of Safety Analysis Methods

Safety assessments will include adverse events (AEs), vital signs, clinical laboratory assessments, and physical examination.

9.2. Adverse Events

An AE is defined as any untoward medical occurrence in a subject or clinical investigation subject administered a pharmaceutical product that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of a product, whether or not related to the product.

Pre-existing diseases or conditions will not be considered AEs unless there is an increase in the frequency or severity, or a change in the quality, of the disease or condition. (Worsening of a pre-existing condition is considered an AE.) Events that occur in subjects treated with placebo are also considered AEs.

All AEs will be coded using Medical Dictionary for Regulatory Activities (MedDRA). The number of cases, occurrences, and incidence rates of adverse events for both groups will be calculated based on the coded system and standard terms. A tabular description will be provided for all adverse events during the treatment period, including the types, severity, frequency, and relationship to the study drug. Special notation will be made for cases that discontinue the study due to adverse events and for those experiencing severe or serious adverse events.

9.3. Vital Signs

Vital signs including blood pressure, heart rate, respiratory rate and temperature will be collected at the Screening Visit, Baseline Visit (at check-in) and the Follow-up Visit. Values collected at Baseline Visit will serve as the baseline; if missing, the Screening Visit or last unscheduled visit prior to baseline will be used. Additional vital signs will be collected as part of the LAST assessments (see below), but will not be summarized.

Descriptive summaries (mean, SD, median, minimum, and maximum) of observed (absolute) values and changes from baseline values will be presented for vital sign values for each treatment group at each time point.

9.4. Clinical Laboratory Evaluation

Clinical laboratory values will be collected at the Screening Visit and the Follow-up Visit. Values collected at Screening will serve as the baseline. Descriptive summaries (mean, SD, median, minimum, and maximum) of observed (absolute) values and changes from baseline values will be presented for clinical laboratory values for each treatment group at each time point. The number of subjects with clinical laboratory values categorized as below, within, or above normal ranges, will be tabulated showing change from baseline to follow-up for each clinical laboratory analyte by treatment group. Only laboratory values that are outside the normal range will be flagged in the data listings and presented with the corresponding normal ranges.

10. APPENDIX

None.