Statistical Analysis Plan (SAP)

THE EFFECT OF VITAMIN E AND DOCOSAHEXAENOIC ACID ETHYL ESTER ON NON-ALCOHOLIC FATTY LIVER DISEASE (NAFLD) - A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP CLINICAL TRIAL (PUVENAFLD)

Protocol 2017-1088 Version 3.0, May 20, 2021

Internal Number: BIO-1812

Statistical Analysis Plan Version 4.0, August 10, 2022

ClinicalTrials.gov Identifier: NCT04198805

IND Sponsor: Dr Naga Chalasani, Indiana University School of Medicine

Funded by: DSM Nutritional Product LTD

Authored by:

Traci Blonquist, MS
Biostatistician, Biofortis Research
traci.blonquist@mxns.com

800A South Rohlwing Road, Addison, Illinois 60101
630-880-1720 (Tel)

Signature Page

IND SPONSOR-INVESTIGATOR	Naga Chalasani, MD
Signature	DocuSigned by: Naga Chalasani 876D01431117422...
Date	25-Aug-2022 | 7:10:39 PM BST

DSM	Szabi Péter MD, PhD
Signature	DocuSigned by: Szabolcs Péter 79A68FC430B84D0...
Date	25-Aug-2022 | 10:17:38 PM CEST

BIOFORTIS, INC. - BIOSTATISTICIAN	Traci Blonquist, MS
Signature	DocuSigned by: Traci Blonquist AD4020131968489...
Date	26-Aug-2022 | 6:53:01 AM PDT

Table of Contents

1.0 Abbreviations and Definitions

AE	adverse event
ALK-P	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ANOVA	analysis of variance
AST	aspartate transaminase
AUDIT	alcohol use disorders identification test
Bili	Bilirubin
CBC	complete blood count
cm	centimeter
ConMeds	concomitant medications
DHA	docosahexaenoic acid
EE	ethyl ester
EPA	eicosapentaenoic acid
FFQ	food frequency questionnaire
FIB-4	fibrosis score
HOMO-IR	homeostatic model of insulin resistance
ITT	intent-to-treat
kg	kilogram
MRI	magnetic resonance imaging
PBMC	peripheral blood mononuclear cells
PBO	Placebo
PDFF	proton density fat fraction
PP	per protocol
NAFLD	non-alcoholic fatty liver disease
SAE	serious adverse event
TEAE	treatment emergent adverse event

2.0 Introduction

This study is designed to evaluate the combination of vitamin E acetate and docosahexaenoic acid ethyl ester (DHA EE) against placebo to demonstrate efficacy and safety. Each active product is also included as an additional treatment arm to gain further insight into its contribution to changes in NAFLD parameters and to collect additional safety data. For the remainder of this document:

- PBO refers to the placebo treatment arm
- vitamin E refers to the vitamin E acetate alone treatment arm
- DHA EE + vitamin E refers to the combination treatment arm
- DHA EE refers to the DHA alone treatment arm

3.0 Study Objectives and Endpoints

3.1 Primary Objective

To determine the efficacy of the combination of vitamin E and omega-3 fatty acid (DHA EE + vitamin E) versus placebo (PBO) on reducing liver fat content after 6 months of intervention in adults with NAFLD. The efficacy of the treatment will be assessed by a decrease in hepatic fat fraction (%) relative to baseline measured by MRI-PDFF after 6 months of intervention.

3.2 Secondary Objective(s)

- To determine effect of vitamin E versus PBO on reducing liver fat after 6 months of intervention in adults with NAFLD.
- To determine effect of DHA EE versus PBO on reducing liver fat after 6 months of intervention in adults with NAFLD.
- To determine the change after 6 months of DHA EE and /or vitamin E intervention in anthropometric, metabolic, hepatologic, nutrient, and inflammatory parameters, and quality of life status of adults with NAFLD.

3.3 Safety Objective(s)

To examine any treatment-emergent adverse event (TEAEs) and changes in clinical safety laboratory parameters. Safety will be assessed with vital signs, serum hematology and biochemistry, and AE/SAE reporting.

4.0 Definition of Endpoint Measures

4.1 Primary Endpoint

The primary objective will be assessed by the hepatic fat fraction [%], as measured by MRI-PDFF, after 6 months of intervention of vitamin E + DHA combination compared to the placebo adjusted for baseline.

4.2 Secondary Endpoint(s)

- Hepatic fat fraction [%], as measured by MRI-PDFF, after 6 months of intervention of vitamin E compared to the placebo adjusted for baseline
- Hepatic fat fraction [%], as measured by MRI-PDFF, after 6 months of intervention of DHA compared to the placebo adjusted for baseline
- The difference after 6 months of intervention between specified treatment groups (vitamin E + DHA combination vs. placebo, vitamin E vs. placebo, DHA vs. placebo) for the following measures:
    - Anthropometric measures
        - Body weight (kg)
        - Waist circumference (cm)
        - Waist-to-hip ratio
        - Body mass index (BMI) (kg/m^2)
    - Insulin resistance
        - $HOMO - IR = \frac{fasting\ insulin\ (\mu U/mL) * fasting\ glucose\ (mmol/L)}{22.5}$
        - HbA1c (%)
        - Fasting glucose (mg/dL)
        - Fasting insulin (µIU/mL)
    - Liver biochemistries
        - alanine aminotransferase (ALT) (U/L)
        - aspartate aminotransferase (AST) (U/L)
        - alkaline phosphatase (ALK-P) (U/L)
        - total bilirubin (mg/dL)
    - Fibrosis-4 (FIB-4) score
    - Plasma vitamin E (α-tocopheryl) concentration (ng/mL)
    - Plasma DHA concentration (ug/mL)
    - Lipid profile
        - total cholesterol (mg/dL)
        - HDL-cholesterol (mg/dL)
        - LDL-cholesterol (mg/dL)
        - oxidized LDL (U/L)
        - triglycerides (mg/dL)
    - Short-Form Health Survey (SF-36) domains in which scoring will be conducted as specified by the scale developers[2]
        - Physical functioning (items 3-12)
        - Role limitations due to physical health (items 13-16)
        - Role limitations due to emotional problems (items 17-19)
        - Energy/fatigue (items 23, 27, 29, 31)
        - Emotional well-being (items 24, 25, 26, 28, 30)
        - Social functioning (items 20, 32)

        - Pain (items 21, 22)
        - General Health (items 1, 33, 34, 35, 36)
    - LC-PUFA-dietary intake from the food frequency questionnaire (FFQ)
        - Docosahexaenoic acid (DHA) (mmg/d)
        - Eicosapentaenoic acid (EPA) (mg/d)
    - Inflammatory markers
        - Cytokeratin 18 (CK-18) (U/L)
        - Cytokines, TNFα (pg/mL)
        - Cytokines, IL-6 (pg/mL)

4.3 Exploratory Endpoint(s)

- The difference after 6 months of intervention between specified treatment groups (vitamin E + DHA combination vs. placebo, vitamin E vs. placebo, DHA vs. placebo) on the hepatic fat fraction [%], adjusted for baseline, stratified by the following single nucleotide polymorphisms (SNPs):
    - PNPLA3 genotype (CC versus CG + GG)
    - HSD17b genotype (GA + GG versus AA)
    - Haptoglobin genotype (*2/*2 versus *1/*2 + *1/*1)
- The difference after 6 months of intervention between specified treatment groups (vitamin E + DHA combination vs. placebo, vitamin E vs. placebo, DHA vs. placebo) on the hepatic fat fraction [%}, adjusted for baseline, stratified by the following demographic and baseline variables:
    - Type II diabetes (yes/no)
    - Sex (male/female)
    - Baseline FIB-4 score (low risk [<1.3] versus indeterminate/high risk [≥1.3])
    - BMI (<30 kg/m^2 [underweight/healthy] versus ≥30 kg/m^2 [overweight/obese][7,8])
    - Age (cutoff defined by the median of the baseline distribution)
    - ALT level at baseline (cutoff defined by the median of the baseline distribution)
- Transcriptome analysis will be performed using the gene expression data of peripheral blood mononuclear cells in a subset of approximately 40 subjects that complete the intervention. Changes in gene expression from baseline to end of treatment will be examined.
- The difference after 6 months of intervention between specified treatment groups (vitamin E + DHA combination vs. placebo, vitamin E vs. placebo, DHA vs. placebo) on the LC-PUFA derived oxylipins

4.4 Post hoc analysis

From the additional blood samples collected at Visit 1 and Visit 4, in the event of a positive treatment effect as defined by the primary endpoint, the following biomarkers will be analyzed:

- ELF (enhanced liver fibrosis)
- Pro-C3

5.0 Trial Design and Visit Structure

This is a multi-center, randomized, double-blind, placebo-controlled, parallel 4-arm intervention superiority trial in adults with NAFLD. Subjects are randomized to one of four treatment arms (vitamin E alone, DHA alone, vitamin E + DHA, or placebo) stratified by Type II Diabetes diagnosis.

The total study duration per subject is 6 months with 6 subject contact points including 2 screening visits, randomization/baseline visit, a one-month phone call, a 3-month visit, and a 6-month visit. The following schedule of assessments is from section 6.2 of the study protocol.

Table 1: Schedule of assessments

Visits	V0 Screening+	2nd Screening V1+	V2 Randomization	Phone call at 1 mos3	V3 3mos	V4 6mos
Day**	-28 to -14	-21 to -5		30±3	90 -7 to +3	180 -7 to +3
Informed Consent	X					
Medical History	X					
Physical Examination	X				X	
Liver Fibroscan	X					
MRI-PDFF for Liver fat		X				X
Clinical chemistry*	X	ALK-P, AST, ALT, Bili only			X	X
HbA1c and Insulin	X					X
Hep A (HAVIgM)#	X					
Hematology*	X				CBC	X
Coagulation: INR	X					
Urine pregnancy test	X		X			
Skinner questionnaire[2]	X					
AUDIT questionnaire^	X			X	X	X
Vital signs	X		X		X	X
Lipid panel* including oxidized LDL		X				X
Plasma/serum for inflammatory markers		X				X
Anthropometrics	X				X	X
Randomization			X			
DHA & α-tocopheryl determination and oxylipins[1]		X			X	X
Product dispense			X			

Compliance-Pill count				X	X	X
Adverse event reporting			X	X	X	X
ConMeds			X	X	X	X
FFQ	X		X	X	X	X
SF-36 questionnaire			X			X
Biosample collection ¶		X				X

* fasted blood draw required and ALT, AST, Alk-P, bilirubin is measured 2x

[+]values obtained at screening (except vital signs and FFQ) will be considered 'baseline' values

[#]HAVIgM to be performed unless negative results in past 3 years

[1]oxylipins measured only at V1 and V4.

[2]Skinner Lifetime Drinking Questionnaire and AUDIT given at V0 Screening

[3] Reminder phone calls at months 2,4 and 5 are advised to assist with retention and check on subject compliance, no diet fluctuations, ConMeds, and AEs.

** Additional time may be allowed between study visits with approval and subsequent protocol deviation.

^ Audit questionnaire - Self-report version is collected at V0, V3 and V4 and the interview is only conducted during the 1-month phone call

¶ Refer to lab manual for sample collection table - Exploratory analyses include gene expression at V1 and V4 in a sub-set of 40 subjects, and PNPLA3, HSD17b, haptoglobin genotyping at V1.

6.0 Definitions of Analysis Sets and Subgroups

A total of three populations will be defined including the intent-to-treat population (ITT), per-protocol (PP) population, and the safety population.

6.1 Intent-to-Treat Population

The intent-to-treat (ITT) population will include all randomized subjects and will be analyzed according to their randomized treatment assignment. In the event that randomized subjects never started treatment, a modified intent-to-treat (mITT) population will be used which is defined as the ITT that received at least one dose of study product and at least one post randomization measurement. The primary efficacy analysis will be conducted on the ITT/mITT population.

6.2 Per Protocol (PP) Population

The per protocol (PP) will include subjects who completed the study according to the protocol. Exclusion from the PP will occur if subject fails to meet one or more of the inclusion criteria at enrollment (section 4.2.1 of the protocol) or meets one or more of the exclusion criteria (section 4.2.2 of the protocol). Data will be excluded if the subject has other significant protocol deviations or significant non-compliance with study product. Compliance (%) will be determined by capsule counts calculated as: (the number of capsules provided – the number of capsules returned) divided by the number of expected capsules taken multiplied by 100. The expected number of capsules will be defined as the duration of treatment (end of study medication- start of study medication +1) multiplied by three. Subjects must take a minimum of 80% of expected capsules to be considered compliant. The per protocol (PP) population will be determined at a blinded data review meeting.

6.3 Safety Population

The safety population will be comprised of the intent-to-treat population and will include all randomized subjects. Subjects will be analyzed according to the product actually received during the study. All safety analyses will be performed on the safety population.

6.4 Subgroups

Subgroup analyses will be performed for the exploratory endpoint pertaining to the transcriptomic analysis of the PBMC gene expression. In order to obtain 40 subjects (~10 subjects / treatment arm) in the PBMC gene expression analysis set, samples from the first 55 subjects enrolled and randomized will be obtained and analyzed.

6.5 Unblinded compliance population

Blood levels for DHA and Vitamin E (α-tocopheryl) will be used to confirm the accuracy of capsule counts used in the determination of the PP population. The unblinded compliance population will be defined after database lock. For each subject, the change from baseline to each follow-up time point (3 and 6 months) will be calculated following the cutoffs in Table 2. For subjects randomized to placebo, a change in DHA >26µg/mL or in Vitamin E (α-tocopheryl) >18µmol/L at 6-month will be considered non-compliant. For subjects randomized to DHA EE, subjects with an increase in α-tocopheryl blood level (>18µmol/L) and/or a small change in DHA level (<10 µg/mL) will be considered non-compliant. For subjects randomized to vitamin E, a significant increase in DHA blood level (>26µg/mL) and/or a small change in Vitamin E (<7µmol/L) will be considered non-compliant. Finally, for subjects randomized to DHA EE + vitamin E, if neither the Vitamin E level (<7µmol/L) nor the DHA (<10 µg/mL) increases sufficiently, the subjects will be considered non-compliant.

Table 2. Cutoff levels to determine blood level compliance

Observed change in	Placebo: exclude if more than	Treated: exclude if less than
DHA (µg/ml)	26	10
Vitamin E (µmol/l)	18	7

The disagreement in the classification of subjects from the PP and compliance population will be described qualitatively (e.g. not included in PP due to non-compliance reason, blood level suggest control non-compliance, etc.). McNemar's test will be used to compare inclusion in the PP and the compliance population. If significantly different, driven by product compliance reason, analyses will also be conducted in the blood level compliance population.

7.0 Determination of Sample Size and Randomization

As indicated in section 9.3.7 of the study protocol, with 60 subjects per arm for the combination (DHA EE + vitamin E) and PBO groups, there is 80% power to detect a 3.64% absolute reduction in liver fat assuming a standard deviation of 7% and a two-sided 0.05 significance level with a two-sample t-test (G*power 3.1.9.2). To estimate the secondary endpoints of the PBO compared to the components (DHA EE, vitamin E) with a given precision of ±3% of liver fat, a total of 30 subjects per arm are needed for the

DHA EE and the vitamin E groups. Accounting for an expected 10% attrition, the resulting required sample size was 200 subjects.

Subjects will be randomly assigned to one of four treatment arms following an unequal allocation of 65/65/35/35 for combination (DHA EE + vitamin E), PBO, DHA EE, and vitamin E, respectively. A computer generated randomization scheme (i.e. interactive web based response system [IWRS]) will be used to assignment stratified by Type II diabetes.

8.0 Statistical Methods

Unadjusted descriptive statistics will be provided per measured time point. Continuous variables will be summarized with mean, standard deviation, median, and range. Categorical variables will be summarized with counts and percentages. All tests of significance are two-sided and considered significant at the 0.05 level. All confidence intervals presented will be 95% and two-sided. SAS version 9.4 or higher will be used for statistical analyses unless otherwise stated.

8.1 Demographic and Baseline Variables

Subject demographics of randomized subjects will include age, sex, diabetes strata, ethnicity, and race collected at the first screening visit overall and by treatment assignment. Summary demographics will be provided by treatment group for the study populations defined in section 6.0.

The following describes the time point used as the baseline measures for the outcome variables:

- Baseline hepatic fat fraction[%] as measured by MRI-PDFF is collected at the second screening visit (Visit 1)
- Baseline anthropometric measures are collected at the first screening visit (Visit 0)
- Baseline HbA1c, fasted insulin and fasting glucose are collected at the first screening visit (Visit 0)
- Baseline liver enzymes (AST, ALT, ALK-P) and bilirubin are collected at the first and second screening visits with the allowance for a 3rd unscheduled pre-treatment visit. The measurement closest to the date of randomization will serve as the baseline measure.
- Baseline Fibrosis-4 score is collected at the first screening visit (Visit 0)
- Baseline plasma vitamin E and DHA concentrations are collected at the second screening visit (Visit 1)
- The baseline lipid panel profile is collected at the second screening visit (Visit 1)
- The baseline responses to the short-form health survey (SF-36) is collected at the randomization visit (Visit 2)
- The baseline food frequency questionnaire (FFQ) for the LC-PUFA is collected at the randomization visit (Visit 2)
- Baseline inflammatory markers are collected at the second screening visit (Visit 1)
- Baseline AUDIT questionnaire total score collected at the first screening visit (Visit 0)

8.2 Primary Endpoint Analysis

The 6-month hepatic fat fraction [%] will be evaluated with a linear mixed analysis of covariance (ANCOVA) model. Baseline hepatic fat fraction [%], intervention group, and baseline diabetes status will be included as fixed effects. Study site will be included as a random effect. The residuals of the model will be analyzed visually (QQ-plots, residual plots) to verify model assumptions (normality, constant variance, homogeneity) are met. If necessary, a variance stabilizing transformation (i.e. log) may be considered. The impact of the random effect will be evaluated with the likelihood ratio test and may be removed if not significantly adding to the model. The model derived pairwise comparison between the combination (DHA EE + vitamin E) and PBO will be estimated using least squares means and tested for significance with a t-test. Since the primary outcome is the pre-specified comparison of DHA EE + vitamin E compared to PBO at 6-month, no adjustment for multiple comparisons is planned.

Note that the estimate of the treatment effect is computationally identical to an analysis that utilizes the change score (6-month vs. baseline) as the response variable adjusted for baseline, intervention group, and baseline diabetes status as fixed effects and site as a random effect[3]. However, the 6-month response variable was selected to allow for easier variance stabilizing transformations if necessary.

8.3 Secondary Endpoint Analyses

- The linear mixed model described in section 8.2 for the 6-month hepatic fact fraction [%] will also be used to estimate the difference between DHA EE and PBO, and vitamin E and PBO. The respective model derived pairwise comparisons will be estimated along with a 95% confidence interval. A Dunnett correction for multiple comparisons will be used.
- The between group differences in the anthropometric measures, liver enzymes, FIB-4 score, plasma Vitamin E (α-tocopherol), plasma DHA, and lipid panel will be estimated at 3 and 6 months with a linear mixed ANCOVA model. The respective follow-up measure, at each of 3 and 6 months, will be the response vector. Fixed effects will include the respective baseline anthropometric measure, intervention group, diabetes stratification, time point, and intervention group by time point interaction. Clinic site will be the random effect. Model derived pairwise comparisons of DHA EE + vitamin E vs. PBO, vitamin E vs. PBO, and DHA EE vs. PBO at 3 and 6 months will be estimated along with 95% confidence intervals. The Holm method will be used to adjust for the 6 planned comparisons (i.e. the LSMESTIMATE statement in SAS). Note, the purpose of the model is to estimate the effect at 3 and 6 months, adjusted for baseline, and not evaluate the trend which would require a joint test of intervention group and the intervention group by time interaction[6].
- At 6-month, the differences of each intervention group to PBO in insulin resistance measures, SF-36 domains, and inflammatory markers will be estimated with a linear mixed ANCOVA model. The respective baseline measure, intervention group, and baseline diabetes status will be included as fixed effects. Study site will be included as a random effect. Pairwise comparisons of DHA EE + vitamin E vs. PBO, vitamin E vs. PBO, and DHA EE vs. PBO will be estimated along with 95% confidence intervals. A Dunnett correction for multiple comparisons will be used.

- The differences between intervention groups (DHA EE + vitamin E vs. PBO, vitamin E vs. PBO, DHA EE vs. PBO) in dietary intake of LC-PUFA, specifically DHA and EPA, will be assessed over time with a linear mixed model following an ANOVA approach. Fixed effects will include the intervention group, categorical time point (baseline, 1 month, 3 months, 6 months), and intervention group by time point interaction. A random intercept and a random time effect for subject nested within site will be included. For the random time effect, the covariance structure (e.g compound symmetry, unstructured) will be selected such that the AICC (corrected Akaike information criterion) is reduced. If a significant interaction effect is indicated, an interaction plot will be produced and the relevant pairwise comparisons between intervention groups and PBO at each time point (i.e. DHA EE + vitamin E vs. PBO, vitamin E vs. PBO, DHA EE vs. PBO) will be estimated along with a 95% confidence interval (CI) at each modeled time point. The Benjamini-Hochberg (false discovery rate) adjustment will be used for the 12 planned comparisons. Additionally, the model derived change from baseline to each follow-up time point will be estimated along with a 95% confidence interval.

For each model, the residuals of the model will be analyzed visually (QQ-plots, residual plots) to verify model assumptions (normality, constant variance, and homogeneity) are met. If necessary, a variance stabilizing transformation may be considered. The impact of the random effect(s) will be evaluated with the likelihood ratio test and may be removed if not significantly adding to the model.

8.4 Exploratory Endpoint Analyses

The interaction between intervention groups and each of PNPLA3, HSD17b, and haptoglobin SNPs, as well as with the specified categorical demographic and baseline variables (section 4.3), on the 6-month hepatic fat fraction will be evaluated with the model specified in the primary endpoint analysis (section 8.2) where the interaction between intervention group and respective variable (i.e. SNP, categorical variable) will also be included. A separate model for each will be constructed for each exploratory interaction. Pairwise comparisons of DHA EE + vitamin E vs. PBO, vitamin E vs. PBO, and DHA EE vs. PBO will be estimated for each genotype / category along with 95% confidence intervals and presented in forest plots.

Transcriptomics (Tx) analysis of the PBMC gene expression will follow a discovery approach. The analyses will be performed internally at DSM. In brief, the gene expression workflow will be performed as defined by the Partek Software (https://www.partek.com/). The large data set will be first visually explored using different techniques like heat maps, clustering and principal component analysis (predefined by the Tx software application). This will also trigger the thresholds to be used in further processing. ANOVA pairwise group comparisons (placebo vs the different treatments) will be applied with the goal to identify genes that are significantly differently expressed. The contrast method of Fisher's least significant difference (LSD) will be used (Gene expression workflow, Partek Software) and q-Values (FDR) will be calculated to correct for multiple testing. The q-values will be set after inspection of the whole Tx data set, but will be between q = 0.05 – 0.15, if possible. The filtering will use the change fold of +1.2 fold up- and -1.2 fold down regulation versus the placebo group with the lowest p-values

possible. The generated gene lists (combination vs. placebo, vitamin E vs. placebo, DHA vs. placebo) will be further analyzed in pathway exploration programs (MetaCore and IPA).

The oxylipin levels will be analyzed similarly to the hepatic fat fraction analysis with a linear mixed analysis of covariance. Additionally, the correlation between the change in oxylipin levels from baseline to 6-month and the change in hepatic fat fraction will be evaluated. Specifically, the Spearman correlation coefficient will be estimated along with a 95% confidence interval for each treatment group.

8.5 Post hoc analysis

If a positive intervention effect is indicated based on the primary outcome, the 6-month biomarker levels (i.e. ELF and Pro-C3) will be evaluated with a linear mixed analysis of covariance (ANCOVA) model. Baseline biomarker level, intervention group, and baseline diabetes status will be included as fixed effects. Study site will be included as a random effect. Pairwise comparisons of the combination (DHA EE + vitamin E) vs PBO, vitamin E vs. PBO, and DHA EE vs. PBO will be estimated along with 95% confidence intervals. A Dunnet correction for multiple comparisons will be used.

8.6 Safety Analysis

Adverse events (AEs) will be collected and classified according to the Common Terminology for Adverse Events (CTCAE) system[4,5]. Note that each CTCAE v5.0 term is a MedDRA LLT (Lowest Level Term). All adverse events (AEs) will be summarized by intervention group according to the safety population. The number of subjects experiencing a given event will be summarized overall (i.e. any event of any grade) and by the maximum recorded grade. A separate listing will be provided of the subset of events that were indicated as serious adverse events (SAEs).

8.7 Responder Analysis

A responder will be defined as a relative decrease in hepatic fat fraction (MRI-PDFF) by 30% or higher on the end of treatment MRI as compared to the baseline MRI. The proportion of responders will be estimated, along with a 95% confidence interval, for each intervention group.

8.8 Handling of Missing Data

Missing data will not be imputed. Only observed data will be included in the statistical analysis. For a given endpoint, if data is collected during at least 2 post randomization time points, statistical models that account for unbalanced data will be used. Specifically, missing data will be handled by maximum likelihood estimation of mixed models which assumes the data is missing at random.

9.0 Deviations from Statistical Plan and Other Issues

During the analysis and reporting process, any deviations from the statistical analysis designed for this protocol will be described and justified in the statistical report.

10.0 Data management plan

Data handling and cleaning prior to statistical analysis is outlined in the data management plan Version 2.2 (2017-1088 Data Management Plan Version 2.2_12Oct2021).

11.0 References

1. Muniyappa R, Madan R. Assessing Insulin Sensitivity and Resistance in Humans. [Updated 2018 Jul 10]. In: Feingold KR, Anawalt B, Boyce A, et al., editors. Endotext [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK278954/
2. https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form/scoring.html
3. Laird, N. Further comparative analyses of pre-test post-test research designs. The American Statistician 1983, 37: 329-330; O'Connell, N. Methods for Analysis of Pre-Post Data in Clinical Research: A Comparison of Five Common Methods. Journal of Biometrics and Biostatistics, 2017. 8(1):1-8.
4. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf
5. https://livertox.nih.gov/Severity.html
6. Fitzmaurice, Garrett M., et al. *Applied Longitudinal Analysis*. Wiley, 2011.
7. https://www.cdc.gov/healthyweight/assessing/index.html
8. https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi

12.0 Listing of tables

The following serves as a general guide to the provided tables and figures for the primary and secondary endpoints. Slight variations are possible depending on the data at the time of analysis such as: changing table numbers, re-ordering of tables, or deletion of tables if appropriate.

Table 1.1 Subject disposition:

a. Number of subjects by visit and treatment arm
b. Number of subjects by treatment arm and group:
    i. Screened
    ii. Randomized
    iii. Discontinued
    iv. Completed
c. Premature discontinuations by reason and treatment arm

Table 1.2 Protocol violations by treatment arm and population

Table 1.3 Compliance by treatment arm and population

Table 1.4 Medical History listing

Table 1.5 Prior and concomitant medications listing

12.1 Example tables for primary and secondary endpoints

For all example tables, the descriptive statistics and the model derived estimates may be split into two tables. Any log transformed variables will be back transformed to the original scale.

The following table will be produced for the 6-month outcomes analyzed with the ANCOVA approach. Note that if the table includes multiple outcome variables, a column will be added to identify the specific outcome variable (e.g. insulin resistance measures includes HOMO-IR, fasting glucose, fasting insulin, and HbA1c). Note, a supplemental table will be created to provide descriptive statistics for the absolute and percent change from baseline.

		PBO	DHA EE + vitamin E	DHA EE	vitamin E
Baseline	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
6-Month	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
ANCOVA	6-month Estimate (95 % CI)[+]				
	Estimated difference to PBO[++]				
	(95% CI)				
	p-value*				
	adjusted p-value**				

* p-value for the pairwise comparisons to the PBO ** Dunnett correction

+ Estimated 6-month value adjusting for baseline

++ Computationally equivalent to the estimated difference from placebo in the change from baseline

Supplemental table for 6-month outcomes analyzed with the ANCOVA approach

		PBO	DHA EE + vitamin E	DHA EE	vitamin E
Absolute change from Baseline	N				
	Mean (sd)				
	Median (min,max)				
	Q1,Q3				
Percentage change from Baseline	N				
	Mean (sd)				
	Median (min,max)				
	Q1,Q3				

The following table will serve as a guide for the repeated measures ANCOVA outcomes collected at baseline, 3 months, and 6 months. Note that the outcome includes multiple variables, a column will be added to identify the specific outcome variable (e.g. anthropometric measures including body weight, waist circumference, waist-to-hip-ratio, and body mass index). Note, a supplemental table will be created to provide descriptive statistics for the absolute and percent change from baseline.

		PBO	DHA EE + vitamin E	DHA EE	vitamin E
Baseline	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
3-Month	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
6-Month	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
r-ANCOVA+	3-month estimate (95 % CI)				
	3-month difference++				
	95% CI				
	p-value*				
	adjusted p-value**				
	6-month estimate (95 % CI)				
	6-month difference++				
	95% CI				
	p-value*				
	adjusted p-value**				

+ Mixed model for repeated measures ANCOVA

++ Model derived estimated difference from the PBO; computationally equivalent to the estimated difference from placebo in the change from baseline

* p-value for the pairwise comparisons to the PBO

** Holm correction for multiple comparisons between arms at multiple time points

Supplement table for outcomes analyzed with repeated measures ANCOVA collected at baseline, 3-month, and 6-month

			PBO	DHA EE + vitamin E	DHA EE	vitamin E
3-Month	Absolute change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
	Percentage change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
6-Month	Absolute change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
	Percentage change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				

The following table will serve as a guide for the repeated measures ANOVA outcomes (i.e. dietary intake of DHA and EPA) collected at baseline, 1-month, 3-month, and 6-month. As specified in the SAP, the pairwise comparisons will only be conducted if there is a significant interaction effect. Additionally, an interaction plot will be produced. Note, a supplemental table will be created to provide descriptive statistics for the absolute and percent change from baseline.

		PBO	DHA EE + vitamin E	DHA EE	vitamin E
Baseline	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
1-Month	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
3-Month	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
6-Month	N				
	mean (sd)				
	median (min, max)				
	Q1, Q3				
r-ANOVA^{+}	Time-treatment interaction p-value				
	Baseline estimate (95 % CI)				
	Baseline difference^{++}				
	95% CI				
	p-value*				
	adjusted p-value**				
	1-month estimate (95 % CI)				
	1-month change from baseline (95% CI)				
	1-month difference^{++}				
	95% CI				
	p-value*				
	adjusted p-value**				
	3-month estimate (95 % CI)				
	3-month change from baseline (95% CI)				

	3-month difference++				
	95% CI				
	p-value*				
	adjusted p-value**				
	6-month estimate (95 % CI)				
	6-month change from baseline (95% CI)				
	6-month difference++				
	95% CI				
	p-value*				
	adjusted p-value**				

+ Mixed model for repeated measures ANCOVA

++ Model derived estimated difference from the PBO

* p-value for the pairwise comparisons to the PBO

** FDR correction for multiple comparisons between arms at multiple time points

Supplement table for outcomes analyzed with repeated measures ANCOVA collected at baseline, 1-month, 3-month, and 6-month

			PBO	DHA EE + vitamin E	DHA EE	vitamin E
1-Month	Absolute change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
	Percentage change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
3-Month	Absolute change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
	Percentage change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
6-Month	Absolute change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				
	Percentage change from Baseline	N				
		mean (sd)				
		median (min, max)				
		Q1, Q3				