

Document title:

Statistical Analysis Plan

Study Number/Cyprus National Bioethics Committee Approval Number:

EEBK/ΕΠ/2019/50

Review Bioethics Committee approving the Program:

Biomedical and Clinical Research Committee A

Study Title:

Open-Label, Single-Arm Pilot Study of the Effect of Cholecalciferol on Mucin-1 Levels in Individuals With Autosomal Dominant Tubulo-Interstitial Kidney Disease Due to MUC1 Mutations (ADTKD-MUC1)

Acronym:

CyCHOLEMUC1

Study Type:

Interventional

Date generated/Date submitted:

14 May 2019

Date of final approval:

01 July 2019

Date of last change:

24 July 2026

Sponsors:

University of Cyprus

Broad Institute of MIT and Harvard

Purpose:

This document presents the sample size calculation and statistical methods specified in the approved clinical study protocol, provided as a separate attachment. It also provides a detailed description of the statistical analyses subsequently performed on the collected study data to evaluate the primary endpoint, namely the reduction of MUC1 and MUC1-fs levels in urinary extracellular vesicles following cholecalciferol administration.

Statistical Analysis Plan

1. Statistical analyses specified in the approved study protocol

1.1 Sample size calculation and justification

“Based on plasma mucin-1 determinations at the Broad Institute, the mean plasma mucin-1 level was 14.1 ± 7.3 for 22 individuals. If both genders are considered together, to obtain 80% power to identify a difference in plasma mucin-1 levels of 50%, 35 participants would be required. The same calculations are applied for mucin-1 levels in urinary EVs.”

1.2 Statistical analysis

“Statistical analysis will include a paired T-test, comparing baseline mucin-1 levels in urinary EVs (average of values from visits on days -10 to 0) to values obtained at days +3, +7 as the primary endpoints. Further testing will be performed to compare mucin-1 levels in urinary EVs obtained at 7 and 10 days after administration of cholecalciferol. Continuous variables will be summarized using descriptive statistics including the number of observations, mean, standard deviation, minimum, median, and maximum values. Categorical values will be summarized using number of observations and percentages. All statistical tests will be performed at the two-sided, 5% significance level, unless otherwise specified.”

Protocol context. The approved study design specified an open-label, single-arm pilot study targeting 45 participants with ADTKD-*MUC1*. The sample-size justification above estimated that 35 participants would be required. The planned enrollment target and the calculated minimum sample size are separate protocol quantities.

2. Statistical analyses performed and intended for publication

Methodology: Three 5-mL aliquots of centrifuged urine from different timepoints were pooled and processed using a modified protocol based on a commercially available kit for urinary extracellular vesicle (uEV) isolation. Following protein lysis, total protein concentrations were determined using a standard bicinchoninic acid (BCA) assay, and protein inputs were normalized prior to tryptic digestion. Target peptides were selectively enriched using specific antibodies recognizing two wild-type MUC1 and two MUC1-fs peptide sequences. A validated immunoaffinity enrichment coupled with multiple reaction monitoring mass spectrometry (iMRM) assay was employed for the sensitive, specific, and reproducible quantification of these uEV-derived peptides, enabling longitudinal assessment of changes in protein abundance following VDRA administration.

Status of analyses: The methods in this section describe the distinct and detailed analyses subsequently performed to characterize uEV iMRM assay performance, longitudinal biomarkers and clinical laboratory measurements, cohort comparisons, and response heterogeneity. Methods were not specified in the statistical analysis paragraph of the approved CyCHOLEMUC1 protocol but are listed in this document as they were eventually performed for publication purposes.

2.1 Assay performance in the VDRA trial

Technical reproducibility of the iMRM assay was assessed using paired peptide measurements through multiple approaches. Spearman correlation coefficients were calculated between technical replicates,

defined as same-type (MUC1 or MUC1-fs) uEV-derived peptides at each time point considered per cohort. Coefficients of variation (CVs) were computed for each peptide at each time point as the standard deviation divided by the mean of the raw peptide intensities, expressed as a percentage, and compared across time points, cohorts, and members of each replicate pair to assess measurement consistency. Concordance between paired peptides was further evaluated using receiver operating characteristic (ROC) curves, with each peptide of a pair used in turn to predict the rank of its partner (forward and reverse comparisons), summarized by the area under the curve (AUC). Differences in iMRM peptide quantification between cohorts and between peptides representing the same protein form were evaluated using nested t-tests.

The distribution fit of raw peptide intensities was assessed using quantile-quantile (QQ) plots. Observed intensities were plotted against quantiles expected under a normal distribution and, in a separate panel, against those expected under a log-normal distribution. Closer adherence to the reference diagonal under the log-normal model supported the use of log2-transformed intensities in downstream analyses. Analyses were performed in R, with the tidyverse, pROC, and corrplot packages (analysis environment version: 2024.09.1+394).

To account for variation in urine concentration, raw peptide intensities were normalized by urinary specific gravity and urinary creatinine. For each sample, peptide values were divided by the matched urine specific gravity and creatinine measurements to generate normalized count datasets. Inter-peptide concordance at each time point was then evaluated using Spearman correlations among the four peptides and heatmaps generated with GraphPad Prism v11.0.0 (Dotmatics, USA).

2.2 Statistical analysis

Statistical analyses were performed to assess longitudinal changes in clinical laboratory parameters and iMRM peptide levels during the clinical trial, as well as differences between patients and controls. For patients, measurements obtained at baseline and at two subsequent trial time points were analyzed as repeated measures. iMRM peptide levels and clinical laboratory values were compared across time points using repeated-measures one-way or two-way ANOVA, as appropriate. Normality was assessed using the Shapiro-Wilk and Kolmogorov-Smirnov tests, and variables deviating from normality when all time points were considered were further log2-transformed before parametric statistical testing.

Changes from baseline were also evaluated by calculating Δ values for each post-baseline time point (Δ = time point value – baseline). These Δ values were used to assess the magnitude of change over time and to compare responses between patients and controls, where applicable. For analyses involving both group status and time point, two-way ANOVA was used. When significant main effects or interactions were detected, post hoc pairwise comparisons were performed using Tukey's multiple-comparison test or Bonferroni correction, as appropriate.

The same analytical approach was applied within each cluster to assess cluster-specific temporal changes in absolute peptide levels, Δ peptide levels, and patient-control differences. Analyses and graphical visualizations were performed using GraphPad Prism v11.0.0 (Dotmatics, USA). Cluster-specific temporal patterns were visualized using ggplot2 in R.

2.3 Linear mixed-effects modeling

Longitudinal changes in peptide intensities and clinical variables were modeled using linear mixed-effects models (LMMs) in R with the lme4 and lmerTest packages, including time point as a fixed effect and patient as a random intercept. Measurements obtained from patients at baseline and at two subsequent clinical

trial time points were treated as repeated measures. Pairwise comparisons between each post-baseline time point and baseline were estimated using the emmeans package, with adjustment for multiple comparisons using the Bonferroni method.

Changes from baseline were also calculated for each post-baseline time point (Δ = post-baseline value – baseline value), and these Δ values were used to assess the magnitude of longitudinal change. For cluster-level analyses, a cluster $\times$ time point interaction was included to estimate cluster-specific β coefficients and pairwise contrasts.

2.4 Patient clustering

Longitudinal clustering was performed in R using the kml3d package to identify patient subgroups with distinct temporal uEV iMRM peptide profiles. Clustering was based on log2-transformed iMRM measurements of four peptides, measured at three time points per patient. The optimal number of clusters was assessed for $k = 2$ to 6 using the Calinski-Harabasz index, which evaluates cluster separation by comparing between-cluster and within-cluster variance.

The final three-dimensional clustering model included patient/sample identifiers, time points, and uEV peptide measurements generated by the iMRM assay, and permitted one missing value per time series. The resulting cluster assignments were used to characterize cluster-specific longitudinal peptide patterns and to support downstream statistical comparisons.

End of document