Statistical Analysis Plan Cover Page:

Official Title:	A Study to Evaluate the Effects of fixed dose Flavonoid Isoquercetin on thrombo-inflammatory biomarkers in subjects with stable Sickle Cell Disease
NCT number:	NCT04514510
Document Type:	Statistical Analysis Plan (SAP)
Document Date:	07/20/2022

Statistical Analysis Plan for 20-H-0137 Isoquercetin Study
7/20/22

This document has been prepared prior to the unblinding of treatment assignments.

Summary of Protocol Statements Regarding the Statistical Analysis

Primary Endpoint(s): The primary outcome will be the change in the plasma soluble P-selectin comparing the baseline to IQ response after 28 days in patients with SCD. The proposed statistical method is an analysis of covariance model with follow-up plasma soluble P-selectin P-selectin measurement (i.e. post-treatment measurement) as the dependent variable with baseline measurement and treatment assignment as the covariates. Significance in this pilot phase II study will be evaluated using a two-sided test with alpha level of 0.05.

Missing data considerations: If randomized subjects have missing data in either arm, then the primary analysis will employ a multiple imputation procedure that will be developed without knowledge of the treatment assignments (beyond assessing the percentage of missing outcomes in each arm). In these circumstances the analysis utilizing only available information would be secondary.

Secondary endpoints: Baseline and end of study measures of secondary end points will be similarly conducted using ANCOVA or Wilcoxon's rank sum test, if assumptions of normality are not warranted.

Secondary analyses under consideration for evaluating treatment effect of isoquercetin:

- Compare baseline and end of study plasma protein disulfide isomerase activity
- Quantify baseline and end of study number of plasma tissue factor positive extracellular vesicles
- Compare baseline and end of study tissue factor procoagulant activity
- Compare baseline and end of study inflammatory cytokines
- Compare baseline and end of study plasma D-Dimer and TAT
- Compare baseline and end of study antioxidant effects of Isoquercetin
- Compare baseline and end of study platelet aggregation (Aggregometer)
- Compare baseline and end of study clot formation (TEG)
- Compare baseline and end of study contemporary biomarkers of vascular function and atherothrombosis (EndoPAT and NIRS)
- Compare baseline and end of study plasma Quercetin level

Details of the Statistical Analysis

Primary Analysis:
For the primary analysis the outcome will be assessed via the R program's lm command of the form

lm(eos ~ baseline+group,data=isoquercetin)

where eos denotes the end of study isoquercetin level, baseline denotes the baseline isoquercetin value, and group is a factor with two levels: isoquercetin and placebo. The assumptions of normally distributed error terms appears reasonable based on empirical distributions of isoquercetin levels.

The protocol indicated all randomized participants would be analyzed. Although 46 participants were randomized, one person was randomized but clinical changes in health status changed their eligibility and precluded them from starting the study. Consequently she was unable to begin study procedures and no intervention (placebo or isoquercetin) was given. Because no treatment was administered, and no baseline measurements were obtained we will omit this individual from our analysis dataset. This approach avoids bias (Fergusson D, et al. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. BMJ. 2002 Sep 21;325(7365):652-4) and improves the precision of estimates when compared with the alternative of multiple imputation for someone untreated and without baseline or follow-up data. Consequently, our analysis data set contains 45 individuals. A different individual was randomized, treated, but was lost to follow-up, and has missing follow-up data. Because this individual was treated, a multiple imputation procedure will be used to impute their data as indicated in the protocol.

Imputation will be done using the **mice** package of the R programming language with m=20 imputations (with 30 iterations) using predictive mean matching based on age, gender, and baseline isoquercetin value.

The primary analysis will employ Rubin's method for pooling results across imputations (Rubin, DB, 1987, Multiple Imputation for Nonresponse in Surveys, John Wiley & Sons, New York, pp. 76-77).

Secondary Analyses:

For the secondary analyses listed above, the principle analyses will assess the treatment effects for these outcomes. The determination of whether to use a regression model based on normality or a non-parametric Wilcoxon test of the baseline to end of study change for the principle analyses will be determined by whether the p-value for a Kolmogorov-Smirnov test of normality is less than 0.01 (in which case the Wilcoxon test will be the principle analysis method).

If the Wilcoxon test is used for these analyses then ANCOVA using transformed data to achieve error distributions with closer adherence to an assumption of a normal distribution may be used as a secondary sensitivity analysis.