STATISTICAL ANALYSIS PLAN
NCT04517396

FEnofibRate as a Metabolic INtervention for Coronavirus Disease 2019 (COVID-19): A randomized controlled trial (FERMIN trial)

April 6, 2022

This document details the statistical analysis plan in detail, including analyses that were pre-specified prior to data locking and unblinding. The document follows the original protocol analysis plan, but includes more methodologic detail and additional pre-specified analyses that are based on data not available at the time the protocol was written. These include: (1) A stratified analysis according to the specific fenofibrate/fenofibric acid preparation used, given recent data suggesting that there may be differential potency of fenofibrate as opposed to fenofibric acid *in vitro,* at least in some viral strains and/or culture systems;[1] (2) Stratified analysis suggested by the trial DSMB during the course of the study, such as according to country, baseline disease severity, and COVID-19 pandemic epochs in which specific strains predominated.

ENDPOINTS

Primary Endpoint

The primary endpoint is a composite global rank score that ranks patient outcomes according to the most severe outcome for each patient based on 5 factors: (1) time to death (ranked from shortest to longest, up to 30 days post-randomization), (2) the number of days supported by mechanical ventilation (invasive or non-invasive) or extracorporeal membrane oxygenation (until hospital discharge, up to 30 days post-randomization, ranked from longest to shortest); (3) The inspired concentration of oxygen/percent oxygen saturation (FiO2/SpO2) ratio area under the curve (until hospital discharge, up to 30 days post-randomization, ranked from highest to lowest); (4) For participants enrolled as outpatients who are subsequently hospitalized, the number of days out of the hospital during the 30 day-period following randomization (ranked from lowest to highest); (5) For participants enrolled as outpatients who don't get hospitalized during the 30-day observation period, the modified Borg dyspnea scale (mean value of assessments at ~5, and ~10 and ~15 days).

Secondary Endpoints (Pre-specified in the protocol)[1]

Secondary endpoints pre-specified in the protocol are:

(1) Number of days alive, out of the intensive care unit, free of mechanical ventilation (invasive and non-invasive), extracorporeal membrane oxygenation (ECMO) or maximal available respiratory support in the 30 days following randomization.
(2) A seven-category ordinal scale consisting of the following categories: 1, not hospitalized with resumption of normal activities; 2, not hospitalized, but unable to resume normal activities; 3, hospitalized, not requiring supplemental oxygen; 4, hospitalized, requiring supplemental oxygen; 5, hospitalized, requiring nasal high-flow oxygen therapy, noninvasive mechanical ventilation, or both; 6, hospitalized, requiring extracorporeal membrane oxygenation (ECMO), invasive mechanical ventilation, or both; and 7, death.
(3) A global rank score similar to the primary endpoint, but using a more comprehensive COVID-19 symptom scale instead of the dyspnea Borg scale (Appendix 1 of Protocol).

Exploratory endpoints will include:
(1) Time to all-cause death.
(2) Time to hospitalization among subjects initially randomized as outpatients
(3) Time to discharge among subjects initially randomized as inpatients
(4) Number of days alive and out of the hospital during the 30 days following randomization.
(5) A global rank score similar to the primary endpoint, but built only with factors 1-4

ANALYSIS

Intention to Treat Principle Unless otherwise specified, all analyses are conducted on an 'as-randomized' basis, irrespective of adherence to treatment. This analysis follows the intention to treat principle pre-specified in the protocol.

Descriptive Statistics (Intention to treat cohort)

Analyses will be for the overall cohort, stratified by arm and stratified by in/out patient status at randomization. The outcomes and baseline characteristics of the subjects will be described using proportions for categorical variables and means/medians as well as standard deviations (SD) and interquartile range (IQR) for continuous variables.

Primary Analysis: This analysis will use a two-sided van Elteren test, the stratified version of the Wilcoxon rank sum test. For each stratum, estimation, and 95% confidence intervals of differences between arms will use both the Hodges-Lehman median difference and the Mann-Whitney Parameter, the probability that a subject from the treatment arm has a score greater or equal to that of a subject for the control.[2] In addition to the primary outcome, the Wilcoxon test will be used to test hypotheses involving the continuous secondary and exploratory outcomes.

Secondary Analysis:

Time-to-event outcomes. Kaplan-Meier curves and cumulative event plots will be used to graphically explore differences between arms. The Cox model will be used to assess differences by arm in the risk of all-cause death and the risk of hospitalization for subjects who were outpatient at the time of randomization. Subjects who withdrew from the study will be considered censored. The proportional hazards assumption will be assessed, and violations addressed using a time by intervention term or a parametric model. Time to discharge from hospital will be modelled using a competing risk model with death as the competing risk.

Regression models

Linear regression analysis to estimate the treatment effect for each continuous outcome of interest will adjust for covariates including age, sex, inpatient vs. outpatient status at enrollment, FiO_2/SpO_2 at the time of enrollment, ethnicity, body mass index, altitude above sea level and history of diabetes at baseline. Although not specified in the initial protocol, at the recommendation of the DSMB, we will add Country as a fixed effect with the goal of adjusting for a broad indicator of treatment practices, timing of variants, and timing of surges. We will also perform sensitivity analyses clustered study site.

Sensitivity Analyses

1. **Per-Protocol and As-Treated Analysis.** The extent of non-adherence to protocol will be described by arm. Cumulative event plots will be used to compare non-adherence to protocol (as a binary event) by arm. If there is a substantial difference in non-adherence

across treatment arms, we will adjust for non-adherence in the secondary analyses of the primary endpoint

2. The FiO_2/SpO_2 ratio will be normalized for altitude above sea level (atmospheric pressure) using standard equations.

3. **Incomplete Data** While missing data is an inevitable problem in longitudinal studies, we will make every possible effort to ensure final assessments for all participants, including those opting to discontinue study participation. Possible mechanisms for missingness-missing completely at random (MCAR), missing at random (MAR), nonignorable or not missing at random (NMAR)-will be evaluated prior to implementing methodology intended to minimize bias from missing data.[3] We anticipate that <5% of randomized subjects will have missing data in the components required to compute the study outcomes. Imputation is not commonly used for rank-based statistical methods such as the van Eleteren or Wilcoxon Rank-sum test. If the missingness rate exceeds 5%, we will carry out a sensitivity analysis of the primary outcome using linear regression with multiple impuatation.[4] The primary analysis will be a complete case analysis.

Subgroup analyses

Exploratory subgroup analyses will follow the plans described above with emphasis on estimation and confidence intervals for the purpose of hypothesis generation in future studies. Subgroups include

1. Sex
2. Age (categorized by < or ≥ the median value in the study population),
3. Race
4. Diagnosis of diabetes at randomization,
5. Body mass index (categorized by obese or non-obese)
6. Inpatient vs. outpatient status at the time of enrollment,
7. FiO_2/SpO_2 at the time of enrollment (categorized by < or ≥ the median value in the study population)
8. Duration of symptoms prior to randomization (<7d vs. ≥7 days) and
9. Fasting triglyceride levels (when available, categorized by < or ≥ the standard median value).
10. Country
11. Baseline disease severity (low, medium, high based on WHO criteria) *
12. Enrollment epoch/period (will divide into 3 epochs, early, mid, late)
13. Fenofibrate formulation (fenofibrate nanoparticle formulation (used in the USA, Lebanon and Greece), fenofibrate micronized formulation (used in Peru and Mexico) or fenofibric acid (used in Colombia). We will also do a stratified analysis of all fenofibrate formulations vs. fenofibric acid. Although all preparations utilized are approximately dose-equivalent, these analyses are pre-specified due to potential pharmacokinetic differences between the formulations, as well as recent data suggesting that there may be differential potency of fenofibrate as opposed to fenofibric acid *in vitro,* at least in some viral strains and/or culture systems.[1] We note, however, that potential *in vitro* differences between fenofibrate and fenofibric acid are likely much less relevant *in vivo*, since fenofibric acid is the active metabolite of fenofibrate.

Country, Baseline Disease Severity, and Enrollment epoch/period were recommended as subgroups by our DSMB and prior to conclusion of enrollment. Their recommendation acknowledged variation in standard of care among different countries as well as subgroups that

were pre-defined prior to initiation of the study. Fenofibrate formulation was also added as a subgroup analysis due to the use of 3 different formulations depending on availability within a given country. We will also visually explore trends of epoch by country, plotting study month at enrollment versus summary statistics of the primary outcome by treatment and country.

Safety Analysis

All subjects in the study at will be included in the safety analysis. The frequencies of adverse event (AE) by type, body system, severity and relationship to study drug will be summarized Serious adverse events (SAE), if any, will be described in detail. AE incidence will be summarized along with the corresponding exact binomial 95% two-sided confidence intervals.

References

1. Davies SP, Mycroft-West CJ, Pagani I, Hill HJ, Chen YH, Karlsson R, Bagdonaite I, Guimond SE, Stamataki Z, De Lima MA, Turnbull JE, Yang Z, Vicenzi E, Skidmore MA, Khanim FL and Richardson A. The Hyperlipidaemic Drug Fenofibrate Significantly Reduces Infection by SARS-CoV-2 in Cell Culture Models. *Front Pharmacol*. 2021;12:660490.
2. Fay MP and Malinovsky Y. Confidence intervals of the Mann-Whitney parameter that are compatible with the Wilcoxon-Mann-Whitney test. *Statistics in medicine*. 2018;37:3991-4006.
3. Little RJ. Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*. 1995;90:1112-1121.
4. Li P, Stuart EA and Allison DB. Multiple Imputation: A Flexible Tool for Handling Missing Data. *JAMA*. 2015;314:1966-7.

SIGNATURES

Julio A. Chirinos, MD, PhD
Principal Investigator

Jordana Cohen, MD, MSCE
Director, Data Coordinating Center

Mary Putt , PhD, ScD
Senior Study Biostatistician