SIMPLIFY
Master
Statistical Analysis Plan
for the
Final and Interim Statistical Reports

PROTOCOL NUMBER: SIMPLIFY-IP-19

PROTOCOL TITLE: A Master Protocol to Test the Impact of Discontinuing Chronic Therapies in People with Cystic Fibrosis on Highly Effective CFTR Modulator Therapy (SIMPLIFY)

PRINCIPAL INVESTIGATORS:

Dave Nichols, MD
University of Washington/Seattle Children's

Nicole Mayer-Hamblett, PhD
University of Washington/Seattle Children's

Alex Gifford, MD
Dartmouth-Hitchcock Medical Center

FUNDING AGENCIES: Cystic Fibrosis Foundation, Inc. (CFF)

PREPARED BY:

CF Therapeutics Development Network Coordinating Center:

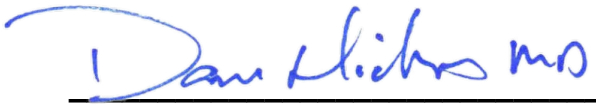
Nicole Mayer-Hamblett, PhD
Katherine Odem-Davis, PhD
Michelle Skalland, MS
Renee Russell, MS

RELEASE DATE: July 6, 2020

NCT04378153

Signature Page

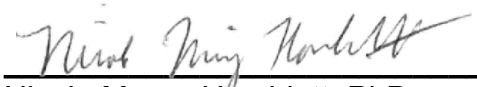
Analysis plan approved by:



06 July 2020

Dave Nichols, MD
Principal Investigator
University of Washington/Seattle Children's

Date



06 July 2020

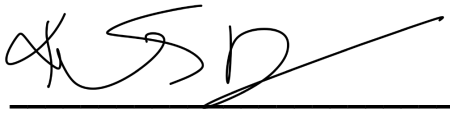
Nicole Mayer-Hamblett, PhD
Principal Investigator
University of Washington/Seattle Children's

Date

10 July 2020

Alex Gifford, MD
Principal Investigator
Dartmouth-Hitchcock Medical Center

Date



06 July 2020

Katherine Odem-Davis, PhD
Senior Biostatistician
Seattle Children's

Date

1. Overview

1.1 Study Rationale and Design

As transformative CFTR modulator drug therapies have become increasingly available to the CF population, many in the CF Community (patients, families and caregivers) are asking if any of their pre-existing therapies can be reduced or eliminated. Motivated by survey results from both the CF Community and Clinician-Investigator groups which indicate very high support for a randomized trial testing the withdrawal of chronic therapies after highly effective modulators, the SIMPLIFY master protocol was developed.

SIMPLIFY is a master protocol with two concurrent randomized trials. It is designed to evaluate the independent effects of discontinuing hypertonic saline (Study A) and dornase alfa (Study B) in people with CF age 12 and older taking elexacaftor/tezacaftor/ivacaftor (ETI) for at least 90 days prior to study screening. Individuals with CF ages 12-17 years with FEV_1% predicted 70% or greater and those 18 years and older with FEV_1% predicted 60% or greater may enroll. There is no upper limit for FEV_1% predicted. Study A and Study B are identical randomized, open label two-arm trials consisting of a 2-week screening period, randomization to either continue or discontinue hypertonic saline (Study A) or dornase alfa (Study B), followed by a 6-week study period. Only those that remain clinically stable and maintain adequate reported adherence to inhaled drug therapy between screening and Visit 1 will be eligible for randomization.

At study entry, participants currently being treated with only hypertonic saline or dornase alfa will be enrolled in Study A or Study B (as applicable) and will be randomized 1:1 to either continue or discontinue their current prescribed therapy. At study entry, participants who are currently being treated with both hypertonic saline and dornase alfa will remain on both therapies during the screening period and then be randomized to Study A (hypertonic saline) or Study B (dornase alfa) as well as randomized (1:1) to continue vs. discontinue the applicable therapy. The randomization to Study A or Study B among participants on both therapies is not optional and is essential to reduce indication bias and ensure comparable populations across studies. After completion of the first study, these participants may subsequently enroll in the alternate study if they meet eligibility criteria.

For participants randomly assigned to continue their therapy during a given study, this therapy should be taken at least once daily according to each participant's pre-existing, clinically prescribed regimen (e.g. daily, twice daily). The concentration of hypertonic saline will also be according to clinical prescription (e.g. 7% sodium chloride or 3.5% sodium chloride). Hypertonic saline concentration must be at least 3%.

Clinical outcomes (FEV_1, antibiotic use, pulmonary exacerbations, and patient reported outcomes), safety (adverse events) and participants' perception of how stopping HS or dornase alfa (or both) would impact their daily life will be evaluated at all sites during each study. Additional measurements will be conducted at selected study sites with the capabilities to conduct these procedures:

- Multiple Breath Washout to evaluate changes in lung clearance index (LCI)
- Mucociliary Clearance (MCC) scans using inhaled radio-labeled particles and imaging techniques to evaluate changes in mucociliary clearance

The primary objectives of the protocol are:

- *Study A:* To determine whether discontinuing hypertonic saline is non-inferior to continuing hypertonic saline among participants on chronic ETI, as measured by the 6-week absolute change in FEV_1 % predicted

- *Study B:* To determine whether discontinuing dornase alfa is non-inferior to continuing dornase alfa among participants on chronic ETI, as measured by the 6-week absolute change in FEV_1 % predicted

The secondary objectives of each study are to evaluate:

- The safety of discontinuing vs. continuing hypertonic saline (Study A) or dornase alfa (Study B)

- The effect of discontinuing vs. continuing hypertonic saline (Study A) or dornase alfa (Study B) on lung clearance index (LCI)

- The effect of discontinuing vs. continuing hypertonic saline (Study A) or dornase alfa (Study B) on other clinical outcomes (e.g., antibiotic events, pulmonary exacerbations, and patient reported outcomes)

1.2 Interim Data Monitoring Committee Reviews

Safety oversight for this trial will be conducted by the Cystic Fibrosis Foundation (CFF) Data Safety Monitoring Board (DSMB; Chair, Lynne M. Quittell, MD). A subcommittee, the Data Monitoring Committee (DMC), will serve on the review board for this trial. A DMC will consist of at least 2 physicians experienced in treating CF and a biostatistician experienced in clinical trial monitoring with an option for adding ad-hoc expertise. The DMC is responsible for safeguarding the interests of study participants, assessing the safety and efficacy of study procedures, and for monitoring the overall conduct of the study.

Interim safety reports will be provided for each study on a semi-annual basis starting after the first participant is randomized. These reports will include a summary of screening, enrollment metrics, baseline characteristics, participant withdrawals, protocol violations, and AEs and SAEs tabulated by treatment group. The proportion of patients with significant pulmonary function declines will be summarized as well for these reviews. An unblinded, open review with the DMC and the Sponsor-Investigators of the Screening and Enrollment Report will take place. The safety data summarized by intervention arm will be presented in the closed section of the DMC meeting as detailed in an Interim Report SAP.

In addition, for each study, the scheduled interim review following enrollment and Week 6 visit completion of 25% and 50% of planned sample size allowing for 12% loss to follow-up will include a formal evaluation of excess harm of treatment withdrawal. After interim analysis, if DMC approves, a separate cohort (lower lung function cohort) of approximately 120 subjects ≥ 18 years old with FEV_1 40 to < 60 % predicted will be enrolled into Study A.

2. Report Generation

2.1 Data Flow

TDNCC utilizes Medidata Solutions, Inc. (Medidata) Rave® for their EDC studies. The Medidata Rave EDC system is designed to be US Code of Federal Regulations (CFR) 21 Part 11 compliant, with a robust audit trail system and electronic signature capabilities. Study personnel at each site will enter data from a subject's visit onto electronic CRF screens via a web browser.

Study subjects will not be identified by name in the study database or on any data capture screens but will be identified by a unique subject identification number. Only study personnel at the individual sites will be able to link the study ID to the subject's name. TDNCC also utilizes the Medidata Rave eCOA/ePRO system, a regulatory compliant system which allows subjects in a study using Medidata Rave EDC to complete and submit forms and data for patient-reported outcomes electronically on a mobile device to the Medidata Rave EDC System. Study personnel at each site will register subjects using their unique subject identification number which generates an activation code unique to that subject. Study site personnel provide the subjects with their activation code. The subject downloads the Medidata Rave eCOA/ePRO app to their mobile device and uses their unique activation code to create their ePRO login and password.
The Biostatistics and Clinical Data Management group of the TDNCC will be responsible for data processing, in accordance with procedural documentation. Database lock will occur once quality assurance procedures have been completed. All procedures for the handling and analysis of data will be conducted using good computing practices for the handling and analysis of data for clinical trials.

2.2 Randomization

At the randomization visit (Week 0), those subjects who are eligible and taking only either hypertonic saline or dornase alfa will be enrolled in the appropriate study (A or B). If both studies remain open to enrollment, those taking both and enrolling for the first time will be randomized to study (A or B) via stratified block randomization, with blocks of size 4. If only one study is open for enrollment, eligible subjects taking both therapies may be enrolled into only the open study. Those taking both and who are eligible and enrolling for the second time may enroll in the study that they were not previously in if that study remains open for enrollment. Within each study (A or B), subjects will be assigned 1:1 to continue or discontinue the applicable therapy by stratified randomization in blocks of size 2. Treatment assignment lists for the cohorts specified by stratification groups (each combination of levels across strata) will be created in SAS 9.4 and uploaded to RTSM in Medidata Rave. Stratifying factors include Week 0 FEV_1 % predicted (≥90%, <90%), treatment combination at Week -2 (single or concurrent use of hypertonic saline and/or dornase alfa), prior study participation (yes/no), and age at Week 0 (≥18 vs < 18). Subjects enrolled in the lower lung function cohort will be similarly randomized to continue or discontinue hypertonic saline and stratified by current dornase alfa use.

2.3 Report Generation

The final statistical reports will describe and justify any deviations from the original statistical plan described herein. Analyses will be performed using SAS 9.4 software and most current version of R. No adjustments for multiple comparisons will be made. All programs used to produce this report will be documented, tested, and archived and all tables, figures and listings will be validated before considered final.

2.4 Definition of the Analysis Populations

Enrollment and screening summaries will be generated using all screened participants. All participant disposition, secondary, exploratory, and safety summaries will be performed using an intent to treat (ITT) population, defined as all participants randomized at Visit 1 (Day 0). The primary analyses in both Study A and Study B will be performed using a per-protocol analysis (PPA) population, as defined below. Sensitivity analyses, repeating the primary analyses, will be done on the ITT population. Secondary analyses will likewise be run on the PPA population and repeated on the ITT population.

PPA is defined by the following criteria:
1. Daily diary completion ("Compliance") from Week 0 (Visit 1) to Week 6 (Visit 3)
    a. >=70% non-missing data
    b. >=70% non-missing in last 2 weeks

2. Daily diary responses from Week 0 (Visit 1) to Week 6 (Visit 3) aligned with randomized treatment ("Adherence")
   a. to assigned treatment regimen (HS or Dnase) among non-missing days overall (>=70%)
   b. to assigned treatment regimen (HS or Dnase) among non-missing days in 2 weeks (>=70%)
3. No initiation of new acute oral, inhaled, or IV antibiotics for respiratory symptoms (rate from ETI trial was ~5%) from Week 0 (Visit 1) to Week 6 (Visit 3)
4. Non-missing FEV at Week 6 (Visit 3)
5. Given the correct randomization instructions
6. Minimum 70% use of ETI among non-missing days in last 2 weeks
7. Eligible at Week -2 (Screening Visit) and at Week 0 (Visit 1, Randomization)

Data from subject disposition visits will be allocated to the nearest subsequently scheduled visit. Missing outcome data in the ITT population for the final primary analysis and for key secondary analyses (CRISS, CFQ-R, and LCI) will be imputed using the least favorable treatment mean in arms discontinuing treatment and using the most favorable treatment mean in arms continuing treatment. Complete case results, including participants based on availability of non-missing values, will also be reported.

3. Overview of Planned Analyses

3.1. Screening Report

3.1.1. Outline of Screening and Enrollment

The overall flow from screening to enrollment is illustrated by a CONSORT diagram. The number of participants screened and eligible are summarized by site. The status of second study screening and enrollment among participants initially on both hypertonic saline and dornase alfa that completed their first study is summarized by first study and by intervention arm.

3.1.2. Screen Failures, Run-in Loss to Follow-up, and Reasons Not Randomized

Screen failure reasons are summarized for Week -2, the initial screening visit. Run-in periods initiated are tracked over time among study participants who are eligible at Week -2. Then, follow-up from Week -2 to Week 0 (Visit 1) is tabulated among participants eligible at initial screening. Participants not completing the Week 0 visit within the allowed window (21 days after the Week -2 visit) are categorized as an incomplete run-in. Finally, reasons for ineligibility at randomization or decision to not randomize are given. Ineligibility reasons are summarized for Week 0 (Visit 1) and by current therapy and prior enrollment status.

3.1.3. Enrollment, Demographics, and Follow-up Overview

Total participants randomized in each study and their eligibility before randomization (Study A, Study B, or both) is summarized overall and broken down by therapy regimen/prior enrollment. Demographic and Week 0 characteristics for unique participants enrolled are summarized among all randomized participants by study. All measures were recorded at Visit 1 unless specified otherwise.

3.2. Study Reports

3.2.1. Summary of Randomization and Study Visit Completion

A CONSORT diagram for the corresponding study (A or B) delineates counts of participants from randomization to per-protocol analysis population inclusion. The cumulative monthly enrollment of

participants randomized into the study is graphically summarized. Participants randomized, withdrawn, and completing the study are tabulated by intervention arm and site.

Completion of each study visit and clinic spirometry at each visit are summarized by intervention arm and overall. Participants are considered to have completed the visit if there is a CRF page with a study date corresponding to that visit.

An overview of analysis populations summarizes the number of participants, by intervention arm and in total, excluded from the per-protocol analysis population for one or more reasons. The reasons for exclusion are also summarized.

3.2.2. Demographics and Characteristics at Week 0

Intervention arms are described and compared with respect to Week 0 demographic and clinical characteristics including age, sex, CFTR genotype, race, height, weight, and all randomization strata. For all summarizations, Week 0 clinical characteristics are defined as measurements obtained at Visit 1 unless specified otherwise.

3.2.3. Summary of Withdrawals, Treatment Assignment and Other Therapy Adherence

The number of participants who withdrew early from the study is tabulated by intervention arm. The reasons for withdrawal and time to withdrawal are also summarized.

The daily therapy ePRO questionnaire completeness, intervention assignment and ETI adherence, and dornase alfa (or hypertonic saline, if Study B) and airway clearance use are summarized by intervention arm for participants who have completed study or withdrawn.

The difference between intervention arms in participants meeting adherence criteria is estimated using logistic regression methods and reported with confidence intervals accounting for randomization strata.

3.2.4. Adverse Events

All reported SAEs and AEs are coded using MedDRA and grouped by system organ class (SOC). The number of (S)AEs is summarized by each intervention arm as follows: (i) The proportion of participants with at least one (S)AE, (ii) The average number of (S)AEs per participant, and (iii) The rate of (S)AEs per participant week of follow-up. Histograms showing the frequency of the number of (S)AEs in each intervention arm are included. The incidence and rate of (S)AEs in each intervention arm is summarized by SOC and preferred term, relationship to arm, and severity. Poisson regression modeling is used to derive rate ratios and 95% confidence intervals. Rates by arm are compared using a two-sided 0.05 level test for Poisson count data. Proportions of participants with at least one (S)AE are also compared between arms with estimated differences and corresponding 95% confidence intervals by the Newcombe-Wilson method without continuity correction; p-values are from a two-sided 0.05 level Fisher's Exact test.

The number and percent of participants changing their assigned therapy due to an adverse event is summarized by intervention arm. The number and percent of participants changing their assigned therapy when directed by a physician because of an adverse event is also summarized.

3.2.5. Hospitalizations and Pulmonary Function Decline

The number of hospitalizations is summarized within each intervention arm as follows: (i) The proportion of participants with at least one hospitalization, (ii) The average number of hospitalizations per participant, (iii) The rate of hospitalizations per participant week of follow-up, and (iv) The number of days hospitalized per participant. Poisson regression modeling is used to

derive rate ratios and 95% confidence intervals. Rates by arm are compared using a two-sided 0.05 level test for Poisson count data. Proportions of participants with at least one hospitalization are also compared between arms with estimated differences and corresponding 95% confidence intervals by the Newcombe-Wilson method without continuity correction; p-values are from a two-sided 0.05 level Fisher's Exact test. The difference between intervention arms in participants hospitalized is estimated using logistic regression methods and reported with confidence intervals accounting for randomization strata.

The proportions of participants with a significant decline in FEV_1 % predicted from Week 0 are also summarized by intervention arm.

3.2.6. Analyses of Primary Endpoint

The primary endpoint is the difference between arms in the change in FEV_1 % predicted from Week 0 (Visit 1) to Week 6 (Visit 3). The primary analysis for non-inferiority is conducted on the per-protocol analysis (PPA) population. An ANOVA model is used to adjust for dichotomous randomization strata: Week 0 FEV_1 % predicted, treatment combination at screening, prior study enrollment, and Week 0 age. The estimated effect of discontinuation and corresponding 95% confidence interval are reported, and the p-value is evaluated for a one-sided alpha-level 0.025 test of non-inferiority with a margin of -3% absolute change in FEV_1 % predicted. An unadjusted estimate is also provided.

The primary analysis is repeated in the ITT population. Missing outcome data in the ITT population were imputed using the least favorable treatment mean in arms discontinuing treatment and the most favorable treatment mean in arms continuing treatment. In this case, the least (most) favorable treatment mean is defined as the mean of the arm with the greater negative (positive) change from Week 0 to Week 6. Complete case results, including participants based on availability of non-missing values, are also reported.

3.2.7. Spirometry Results

Absolute and relative changes in spirometry measures from Week -2 (screening) to Week 0 (Visit 1) and from Week 0 (Visit 1) to each post-randomization visit are summarized. Mean differences between intervention arms at each visit and corresponding 95% confidence intervals are presented. Changes from Week 0 (Visit 1) to Week 6 (Visit 3) are assessed using a two-sample, two-sided t-test. Additionally, the absolute change in FEV_1 % predicted from Screening to Week 0 (Visit 1) is compared between intervention arms using an ANOVA model adjusting for randomization strata. A forest plot qualitatively comparing treatment effects by subgroup will be shown for the following characteristics: Week 0 FEV_1 % predicted, treatment combination at screening, prior study enrollment, Week 0 age, sex at birth, pseudomonas aeruginosa positive culture in past year, genotype, concurrent chronic airway clearance therapy and randomization strata.

3.2.8. Exacerbation and Concomitant Medication Parameters

The number of protocol-defined and physician identified pulmonary exacerbations (PEx) experienced by participants from Week 0 (Visit 1) to Week 6 (Visit 3) are summarized by intervention arm as follows: (i) The total number of PEx, (ii) The rate of PEx per participant week of follow-up, (iii) The average number of PEx per participant, (iv) The proportion of participants with at least one PEx, and (v) the total requiring antibiotics (acute IV, oral, or inhaled) or hospitalization. Poisson regression modeling is used to derive rate ratios and 95% confidence intervals. Rates by arm are compared using a two-sided 0.05 level test for Poisson count data. Proportions of participants with at least pulmonary exacerbation are also compared between arms with estimated differences and corresponding 95% confidence intervals by the Newcombe-

Wilson method without continuity correction; p-values are from a two-sided 0.05 level Fisher's Exact test.

Also summarized is the frequency of signs and symptoms for protocol-defined PEx for each intervention arm and overall. The difference between intervention arms in participants experiencing a protocol-defined PEx is estimated using logistic regression methods and reported with confidence intervals accounting for randomization strata.

The difference between intervention arms in participants initiating acute antibiotics is estimated using logistic regression methods and reported with confidence intervals accounting for randomization strata.

3.2.9. Summary of CRISS and CFQ-R scores

Absolute changes in CRISS and CFQ-R, respiratory domain, from screening to Week 0 (Visit 1) and from Week 0 (Visit 1) to each post-randomization visit are summarized. Mean differences between intervention arms at each visit and corresponding 95% confidence intervals are presented. Changes from Week 0 (Visit 1) to Week 6 (Visit 3) are assessed using a two-sample, two-sided t-test. Additionally, the changes in CRISS and CFQ-R, respiratory domain, between intervention arms are compared between intervention arms using an ANOVA model adjusting for randomization strata.

3.2.10. Summary of Anthropometric Measures

Absolute changes in weight (kg), weight percentile, BMI (kg/m^2), and BMI percentile from screening to Week 0 (Visit 1) and from Week 0 (Visit 1) to each post-randomization visit are summarized. Mean differences between intervention arms at each visit and corresponding 95% confidence intervals are presented. Changes from Week 0 (Visit 1) to Week 6 (Visit 3) are assessed using a two-sample, two-sided t-test.

3.2.11. Summary of Lung Clearance Index (LCI)

Demographics and baseline characteristics of the subset participants with multiple breath washout procedure completed are summarized as described in Section 3.1.3. Absolute and relative change in Lung Clearance Index (LCI) from screening to Week 0 (Visit 1) and from Week 0 (Visit 1) to each post-randomization visit are summarized. Mean differences between intervention arms at each visit and corresponding 95% confidence intervals are presented. Changes from Week 0 (Visit 1) to Week 6 (Visit 3) are assessed using a two-sample, two-sided t-test. Additionally, the changes in LCI between intervention arms are compared between intervention arms using an ANOVA model adjusting for randomization strata.

3.2.12. Listings

Listings will include protocol violations and deviations. A listing of pregnancies will also be included if any are reported during follow-up.

3.2.13. Lower Lung Function (LLF) Cohort

The lower lung function cohort monitoring and outcomes will be summarized in a separate report from Study A. Among those enrolled in the lower lung function cohort, descriptive safety summaries will be provided for the differences between treatment arms in the change in FEV1 % predicted, adverse event rates, and proportion of participants non-adherent to assigned therapy after randomization. Non-adherence will include definitions based on three distinct outcomes: a change action relative to assigned treatment following an adverse event, <70% adherence overall post-randomization, and <70% adherence in the last 2 weeks prior to Week 6. Treatment arms will be formally compared within this cohort to determine if assignment to STOP Taking

hypertonic saline results in clinically meaningfully inferior outcomes as compared to the KEEP Taking hypertonic saline. The primary analysis population will be ITT, with PP analyses reported if indicated for sensitivity. For all analyses with a model adjusted for randomization strata, the only stratification variable included is treatment combination at screening.

The following exhibits will be excluded from the LLF reports:
Screening Tables 1.2 and 3.1. Screening Figure 2.1. Study Figure 3.1. Also, Study Table 6.1 and Figure 6.1 will be excluded from interim reports.

The following are other modifications to the LLF report exhibits:

Screening exhibits remove components referencing Study B or use of dornase alfa only, which are not applicable to the LLF cohort. Screening tables present overall summaries only (i.e., do not include columns further broken down by treatment regimen and prior enrollment).

Demographic summaries remove the age category of >=12 to <18 years and modify the FEV_1 % predicted categories (<40, >=40 to <50, >=50 to <60, >60). Study Table 4.11 also uses modified FEV_1 % predicted categories (<50, >=50).