Statistical Analysis Plan

A Phase III Randomized Crossover Study to Validate the Clinical Safety, Performance, and Benefit of the Trillium Full Face Mask in Patients with Obstructive Sleep Apnea

ClinicalTrials.gov identifier

NCT07444749

Document version

1.0

Document date

22-Apr-2026

Source document identifier

SRC-301688-Trillium Pivotal

Public disclosure copy with privacy redactions

Administrative cover; scientific content and document version/date are retained.

Statistical Analysis Plan

Study Sponsor:	Philips RS North America LLC f/k/a Respironics, Inc.
Clinical Investigation Plan ID:	SRC-301688-Trillium Pivotal
Clinical Investigation Plan Title:	A Phase III Randomized Crossover Study to Validate the Clinical Safety, Performance, and Benefit of the Trillium Full Face Mask in Patients with Obstructive Sleep Apnea
Clinical Investigation Plan Version:	Version 00

Statistical Analysis Plan Author: Melissa Nelson, MS
Associate Director, Biostatistics
Philips Chief Medical Office

CONFIDENTIALITY STATEMENT

This document and the information contained herein or attached hereto ("Confidential Material") are confidential and proprietary to Philips. This Confidential Material should only be viewed by those individuals or companies that have been given prior authorization to do so by Philips.

TABLE OF CONTENTS

1 List of Abbreviations 5

2 Introduction 6

3 Study Objectives 6

3.1 Clinical Performance Objectives 6

3.2 Clinical Benefit Objectives 6

3.3 Clinical Safety Objectives 6

4 Investigational Plan 7

4.1 Overall Study Design 7

4.2 Selection of Study Population 9

4.3 Treatment Groups 10

4.4 Treatment Randomization 10

4.5 Treatment Blinding 11

4.6 Sample Size 11

5 Extraction of Device Data 12

6 Clinical Performance Measurements 12

6.1 Unintentional Leak 12

6.2 Patient Reported Outcomes Survey 13

7 Clinical Benefit Measurements 13

7.1 Apnea-Hypopnea Index 13

7.2 Adherence 14

7.3 Epworth Sleepiness Scale Questionnaire 14

7.4 European Quality of Life-5D-3L Questionnaire 15

8 Clinical Safety Measurements 16

8.1 Averse Events 16
8.2 Device Deficiencies 18
9 Data Quality Assurance 18
10 General Statistical Methods 18
10.1 Reporting Conventions 18
10.2 Baseline Definition 19
10.3 Analysis Populations 19
10.3.1 Performance Analysis Set 19
10.3.2 Safety Analysis Set 19
10.4 Handling of Dropouts or Missing Data 20
10.5 Cohen's Effect Size Calculation 20
10.6 Linear Mixed Effect Modeling Approach 20
11 Subject Disposition 21
12 Demographics and Baseline Characteristics 22
12.1 Demographic Characteristics 22
12.2 Baseline Characteristics 23
12.3 Medical History 23
12.4 Concomitant Medication 23
13 Protocol Deviations 23
14 Product Accountability 24
15 Clinical Performance Analysis 24
15.1 Unintentional Leak Analysis 24
15.2 Patient Reported Outcomes Analysis 24
16 Clinical Benefit Analysis 25

16.1 Apnea-Hypopnea Index Analysis 25
16.2 Adherence Analysis 25
16.3 Epworth Sleepiness Scale Analysis 26
16.4 European Quality of Life-5D-3L Analysis 27
17 Clinical Safety Analysis 28
17.1 Adverse Events 28
17.2 Device Events/ Deficiencies 29
18 Subgroup Analysis 29
19 Adjustments for Covariates 29
20 Multiple Comparison and Multiplicity 29
21 Changes in the Conduct of the Study or Planned Analysis 29
21.1 Changes in the Conduct of the Study 29
21.2 Changes in the Planned Analysis 29
22 Tables, Figures, and Listings 30
23 References 31
24 Document Revision History 32
25 Document Approval History 33

1 LIST OF ABBREVIATIONS

Abbreviation	Definition
ADE	Adverse Device Effect
AE	Adverse Event
AHI	Apnea-Hypopnea Index
ANOVA	Analysis of Variance
ASP	Average Sleep Propensity
DB	Database
CI	Confidence Interval
CIP	Clinical Investigation Plan
CIR	Clinical Investigation Report
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
EQ-5D-3L	European Quality of Life-5 Dimension 3-level Version
ESS	Epworth Sleepiness Scale
FFM	Full Face Mask
OSA	Obstructive Sleep Apnea
PAP	Positive Airway Pressure
PRO	Patient Reported Outcomes
SADE	Serious Adverse Device Effect
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard Error
TLFs	Tables, Listings, and Figures
UADE	Unanticipated Adverse Device Effect
VAS	Visual Analog Scale

2 INTRODUCTION

This statistical analysis plan (SAP) describes the analyses and data presentations for study SRC-301688-Trillium Pivotal "A Phase III Randomized Crossover Study to Validate the Clinical Safety, Performance, and Benefit of the Trillium Full Face Mask in Patients with Obstructive Sleep Apnea".

The primary purpose of this study is to evaluate clinical safety, clinical performance, and clinical benefit of the Trillium Full Face Mask (FFM) among patients with obstructive sleep apnea (OSA) that are already established on positive airway pressure (PAP) therapy.

Subjects will be exposed to two treatments in a crossover manner. The investigational device is the Philips Trillium mask and the comparator device is the RedMed AirFit F20 mask.

This study is sponsored by Philips RS North America LLC f/k/a Respironics, Inc.

The purpose of this SAP is to ensure credibility of the study findings by pre-specifying the statistical methodology prior to database (DB) lock and final analysis.

The analysis will include one final analysis. There are no interim analyses planned.

3 STUDY OBJECTIVES

3.1 Clinical Performance Objectives

The clinical performance objective of this study is:

- To objectively and subjectively evaluate the clinical performance of the Trillium FFM.

3.2 Clinical Benefit Objectives

The clinical benefit objective of this study is:

- To objectively and subjectively evaluate the clinical benefit of the Trillium FFM.

3.3 Clinical Safety Objectives

The clinical safety objective of this study is:

- To objectively evaluate the clinical safety profile of the Trillium FFM.

4 INVESTIGATIONAL PLAN

4.1 Overall Study Design

This is a prospectively planned, single-center, two-treatment, two-period, open-label randomized crossover design, pivotal study to evaluate the clinical safety, clinical performance, and clinical benefits of the Trillium FFM. This study will enroll adult OSA patients who have been established on PAP therapy for at least 3 months and are currently using an AirFit F20 FFM. Approximately 45 participants will be enrolled and randomly assigned in a 1:1 ratio to one of the two treatment sequences:

- Treatment Sequence 1: Trillium mask followed by the AirFit F20 mask
- Treatment Sequence 2: AirFit F20 mask followed by the Trillium mask

There are 3 planned study visits which will range up to approximately 30 days for each participant.

Visit 1 (Day 1)

The rationale for this visit is to obtain informed consent and to assess participant eligibility. Enrolled participants will be randomly assigned to one of the 2 treatment sequences.

Participants randomized to Treatment Sequence 1 (Trillium mask followed by the AirFit F20 mask) will undergo a fitting of the Trillium mask. Once a mask size is selected, the participant will undergo short trials using their own PAP machine at therapeutic pressure for up to 15 minutes, and an assessment of mask fit will be performed. If the participant's PAP device is not available, they will be fitted with the mask and will undergo short trials using a ResMed S9 PAP device provided by the site.

Participants randomized to receive the AirFit F20 mask first will be given a new AirFit F20 mask to take home in the same size that they currently use. Participants will not be fit with the AirFit F20 mask by the clinician or have their therapy pressure applied with their PAP device.

All participants will be instructed to use the assigned mask for the full 15 days of the first trial period.

Visit 2 (Day 15 ± 3 days)

After the first 15-day home use period with the first assigned mask, participants will return to the site for their second study visit. Participants will be instructed to bring their PAP device, the first study mask and any associated parts and/or cushions to return at this visit.

Safety assessments will be conducted.

Participants will complete a Patient Reported Outcomes (PRO) survey providing feedback on attributes of the first mask, including comfort of breathing. Participants will also be asked to complete the Epworth Sleepiness Scale (ESS) to assess daytime sleepiness for the previous 15 days, and the European Quality of Life-5-Dimension three-level version (EQ-5D-3L), which will assess their health-related quality of life.

De-identified data from participants' PAP devices will be extracted from the connected ResMed device software, AirView. Extracted data will include average unintentional leak, Apnea-Hypopnea Index (AHI), usage and pressure data collected during the previous trial period.

Participants will then undergo the fitting and product dispensation procedure for the second mask following the same process as Visit 1.

Visit 3 (Day 30 ± 3 days)

After the second 15-day home use period with the second assigned mask, participants will return to the site for their third study visit. Participants will be instructed to bring their PAP device, the second study mask and any associated parts and/or cushions to return at this visit.

Safety assessments, performance assessments (PRO survey, ESS, EQ-5D-3L) and device data collection will be performed similarly to Visit 2.

The detailed schedule of events is presented in **Table 1**.

Table 1: Schedule of Events

Procedure	Performed/ Supervised by	Pre-Screen [a]	Visit 1 Day 1	Visit 2 Day 15 ±3 days	Visit 3 Day 30 ±3 days
Preliminary eligibility assessment	Investigator/Designee	X			
Consent	Investigator/Designee		X		
Eligibility questionnaire	Study Coordinator		X		
Demographic questionnaire	Study Coordinator		X		
Mask Information Survey	Clinician		X		
Dispense Study Mask	Clinician/Study Coordinator		X	X	
Mask Survey	Clinician/Study Coordinator			X	X
Epworth Sleepiness Scale	Clinician/Study Coordinator			X	X
EQ-5D-3L	Clinician/Study Coordinator			X	X
Device Data Report	Study Coordinator			X	X

Procedure	Performed/ Supervised by	Pre-Screen [a]	Visit 1 Day 1	Visit 2 Day 15 ±3 days	Visit 3 Day 30 ±3 days
Review of potential AEs	PI/Designee Study Coordinator			X	X
Inventory check to ensure product return	Clinician/Study Coordinator			X	X
Compensation	Study Coordinator		X	X	X
[a] Pre-screening procedures will occur at the site. To identify participants who may be eligible for the study, specifically OSA patients using a ResMed therapy device and AirFit F20 mask, the site will utilize their electronic health record to search for potential participants based on available information. Patients will be contacted based on their potential eligibility. This information will not be collected in the study database.					

4.2 Selection of Study Population

This study will enroll adult OSA patients, aged 21-85 years, who have been established on PAP therapy for at least 3 months and are currently using an AirFit F20 FFM and using a ResMed PAP therapy device with a connection to ResMed's AirView system.

The population selected for this trial is sufficiently representative of the product's intended OSA users. Use of patients who are naïve to therapy or those established on therapy but using a different style of FFM (e.g., AmaraView) would introduce confounders (i.e., adjustment to therapy, unfamiliar mask style), making interpretation of results difficult.

It is planned that approximately 45 participants will be enrolled and randomized, with approximately 37 participants expected to complete all study procedures. Complete inclusion/exclusion criteria are provided in the Clinical Investigation Plan (CIP) Section 5.2.

4.3 Treatment Groups

The treatment groups are described below in **Table 2**. Detailed descriptions are provided in CIP Section 3.

Table 2: Treatment Groups

Treatment Group	Description	Device Image
Trillium Mask	The Trillium Full Face Mask (Philips Respironics, Murrysville, PA) is an over-the-nose, full-face mask with the hose connection in the front. The mask is used with a positive airway pressure therapy device and covers both the nose and mouth, ensuring pressurized air is delivered directly into the patient's airways.	
AirFit F20 Mask	The AirFit F20 (ResMed, San Diego, California) is an externally placed vented mask covering the mouth and the nose of the patient. It provides a seal such that pressure from a positive pressure source is directed to the patient's nose and/or mouth. The mask connects via a standard conical connector to a conventional air delivery hose, which in turn connects to the positive pressure source.	

4.4 Treatment Randomization

The randomization scheme is based on a crossover study design. Subjects will use both masks, and the order of usage will be randomized. Subjects will use either the Trillium investigational mask or a new AirFit F20 mask for 15 consecutive days (±3 days) of use. They will then return for a follow-up visit and will be crossed over to the other mask (i.e., AirFit F20 mask or Trillium mask) per the randomization assignment.

Approximately 45 participants will be randomized. Randomization will not be stratified.

Table 3: Treatment Sequences

Treatment Sequence Number	Treatment Sequence Name
1	Sequence 1: Trillium mask followed by AirFit F20 mask
2	Sequence 2: AirFit F20 mask followed by Trillium mask

The Philips' Biostatistics Group is responsible for preparing the randomization schedule specifications. The study randomization schedule will be automatically generated within the electronic data capture (EDC) system with restricted access to the study data manager, clinical study manager, site staff in-charge of randomization, and statistician.

Upon activation/ initiation of the EDC system for the site team member, the site personnel who is responsible for the randomization process will be trained from the sponsor to use the EDC system for randomizing the participants. The randomization schedule should be kept confidential and accessible only to the person responsible for the randomization and the Philips monitors responsible for checking compliance.

4.5 Treatment Blinding

This is an open label study design. Due to the different designs of masks used, the study cannot be blinded.

4.6 Sample Size

This study will randomize approximately 45 participants in order to complete with 37 evaluable participants.

The sample size for this study is based on the primary endpoint of Leak rate (L/min).

The statistical hypothesis is:

H_0: $\mu_{Trillium} = \mu_{Pass}$

H_A: $\mu_{Trillium} < \mu_{Pass}$

Where

$\mu_{Trillium}$ = Mean Leak rate for the Trillium mask

μ_{Pass} = Pass Criteria value = 24 L/min

The sample size is selected based on results of a pilot study, with similar participant enrollment criteria.

Assuming a mean leak rate of 13 L/min, and a SD of 23, a sample size of 37 participants will provide 80% power to show that the Trillium mask does not exceed a leak rate of 24 L/min, with a one-sided significance level of α=0.025 using a one-sample t-test.

This sample size will also provide over 80% power to show that the Trillium mask has a mean comfort of breathing score of at least 7; assuming a mean comfort of breathing score of at least 8.3 and SD no more than 2.25 with a one-sided significance level of α=0.025 using a one-sample t-test.

To account for possible attrition, ~45 participants will be enrolled in order to complete with the required 37 evaluable participants.

5 EXTRACTION OF DEVICE DATA

At the end of each treatment period, data will be extracted from the participant's PAP device via the connected ResMed device software, AirView. The site personnel will extract data for each treatment period based on the dates of participant visits and will enter the device data into the electronic case report form (eCRF). Extracted data will include unintentional leak, AHI, usage and pressure data.

The start date of each treatment period is the device dispense date. The end date of each treatment period is the device return date – 1.

That is, treatment period 1 will include data collected from Visit 1 date to Visit 2 date minus 1 and will be assigned to the first treatment used.

Treatment period 2 will include data collected from Visit 2 date to Visit 3 date minus 1 and will be assigned to the second treatment used.

6 CLINICAL PERFORMANCE MEASUREMENTS

The primary clinical performance measurements will include both objective and subjective assessments. The assessments are described below.

6.1 Unintentional Leak

Mask leaks occur when pressurized air escapes from the mask, bypassing the airway. The PAP device can detect the leak and calculates a leak rate in liters per minute (L/min).

This leak rate data will be extracted from the participant's PAP device at the end of each treatment period (see Section 5). Unintentional leak will be measured as the median mask leak rate (L/min) over the treatment period.

6.2 Patient Reported Outcomes Survey

At the completion of each treatment period, participants will be asked to complete a subjective survey to assess characteristics of the mask usage. The following questions will be asked:

- How would you rate your satisfaction with the seal of this mask?
- How would you rate your satisfaction with the comfort of this mask?
- How would you rate your satisfaction with your comfort of breathing while using this mask?

Each question will be measured on a 0-10 numerical rating scale, with 0 indicating 'extremely dissatisfied', 5 indicating 'neutral', and 10 indicating 'extremely satisfied'.

At the completion of the second treatment period, participants will be asked the following additional questions:

- Which mask did you prefer?
- Would you switch from your prescribed mask to the Philips investigational study mask?

7 CLINICAL BENEFIT MEASUREMENTS

Clinical benefit measurements will include both objective and subjective assessments. The assessments are described below.

7.1 Apnea-Hypopnea Index

Apnea-Hypopnea Index (AHI) quantifies the number of apneas (complete pauses in breathing) and hypopneas (partial reductions in breathing) per hour of sleep.

AHI data will be extracted from the participant's PAP device at the end of each treatment period (see Section 5).

AHI will be measured as the AHI (events/hr) over the treatment period.

7.2 Adherence

Adherence will be assessed through device usage data, which will be extracted from the participant's PAP device at the end of each treatment period (see Section 5).

The following measurements will be collected:

- Number of Days Device Actually Used
- Number of Days Device Expected to be Used
- Number of Days Device Used for ≥ 4 hours
- Number of Days Device Used for < 4 hours
- Average Time used (hours/Day) (all days)
- Average Time used (hours/Day) (actual use days)

7.3 Epworth Sleepiness Scale Questionnaire

The Epworth Sleepiness Scale (ESS) questionnaire [1] is designed to assess daytime sleepiness. The ESS questionnaire will be collected at the end of each treatment period.

Respondents will be asked to rate their usual chances of dozing off or falling asleep while engaged in eight different activities.

The items are as follows:

- Sitting and reading
- Watching TV
- Sitting inactive in a public place (for example, a theater or meeting)
- As a passenger in a car for an hour without a break
- Lying down to rest in the afternoon when circumstances permit
- Sitting and talking to someone
- Sitting quietly after lunch without alcohol
- In a car, while stopped for a few minutes in traffic

The items are scored using a numerical scale from 0-3:

- 0 = Would never doze
- 1 = Slight chance of dozing
- 2 = Moderate chance of dozing
- 3 = High chance of dozing

7.4 European Quality of Life-5D-3L Questionnaire

The EQ-5D-3L questionnaire is a standardized measure of health status [2]. The EQ-5D-3L will be collected at the end of each treatment period.

The EQ-5D-3L consists of the EQ-5D descriptive system and the EQ visual analogue scale (EQ VAS). The EQ-5D-3L descriptive system consists of 5 dimensions as described in Table 4. Each dimension has three response levels: no problems, some problems, extreme problems.

The EQ VAS records the respondent's self-rated health on a vertical VAS with anchors from 0 = 'the worst health you can imagine to 100 = 'the best health you can imagine'.

Table 4: EQ-5D-3L Description

Dimension	Response Options
Mobility	I have no problems in walking about I have some problems in walking about I am confined to bed
Self-care	I have no problems with self-care I have some problems washing or dressing myself I am unable to wash or dress myself
Usual Activities	I have no problems with performing my usual activities I have some problems with performing my usual activities I am unable to perform my usual activities
Pain/Discomfort	I have no pain or discomfort I have moderate pain or discomfort I have extreme pain or discomfort
Anxiety/Depression	I am not anxious or depressed I am moderately anxious or depressed I am extremely anxious or depressed

8 CLINICAL SAFETY MEASUREMENTS

8.1 Averse Events

Reportable adverse events (AE) (see CIP Section 15.1) will be monitored from the time of signing informed consent until completion of the study or resolution of the event, as appropriate. The eCRF may include but is not limited to the following data elements: event description, dates of onset, resolution, and site awareness, event status (e.g., ongoing), assessment of severity (intensity) and seriousness, relationship to study product or study activity, action taken, and outcome.

The AEs will be classified as:

- Skin Effects (i.e., localized red marks/skin irritation, breakdown, and/or peeling, localized discomfort at the point of mask contact)
- Oronasal Dryness
- Eye-Related Side Effects
- Anxiety
- Allergic Reaction
- Headache
- Other

The severity/intensity of AEs will be assessed by the Principal Investigator or designee using the following levels:

- S0 – Negligible/ No Harm
- S1 – Minor
- S2 – Moderate/ Major
- S3 – Critical/ Serious

Not all severity levels are applicable to all AE categories. The criteria for assigning severity level for each AE classification are collected on the eCRF as follows (also see CIP Section 15.3).

Skin Effects:

- S0 – Red marks, indentations and/or discomfort at the point of mask contact (with or without redness) that go away quickly once the device is removed without any treatment.
- S1 – Red marks, indentations, discomfort at the point of mask contact (with or without red marks), irritation and/or mild peeling that remain after removing mask, but only self-treatment or no intervention required.
- S2 – Red marks, discomfort at the point of mask contact (with or without redness), indentations, irritation, peeling or breakdown that require professional medical intervention.
- S3 – Serious breakdown or pressure necrosis or scarring.

Oronasal Dryness:

- S0 – Temporary discomfort from dryness that resolves when device removed.
- S1 – Dryness or related irritation persists but is treated by patient (i.e., saline spray). May complain of minor cough or sore throat.
- S2 – Dryness that lead to chronic nose bleeds or cough that require professional medical intervention.

Eye-Related Side Effects:

- S0 – Eye dryness that resolves when mask removed and does not impact therapy compliance.
- S1 – Eye redness, dryness, irritation and/or other eye-related effects that is not relieved by removing the mask. Eye drops may be used by patient for self-treatment.
- S2 – Eye redness, dryness, irritation and/or other eye-related effects that persist and require professional medical intervention.

Anxiety:

- S0 – Temporary, able to continue therapy, doesn't impact sleep.
- S1 – Persist throughout night, impact to sleep or therapy compliance.
- S2 – Professional medical intervention needed to continue therapy.

Allergic Reaction:

- S1 – Minor allergy symptoms that resolve with self-treatment or without treatment.
- S2 – Allergic reactions that require professional medical intervention.
- S3 – Anaphylaxis.

Headache:

- S0 – Headache that resolves after mask is removed.
- S1 – Headache that resolves with self-treatment or no treatment.
- S2 – Chronic or severe enough headache that requires professional medical intervention.

Other:

- S0 – Events that do not warrant treatment and do not interfere with the subject's daily activities.
- S1 – Events result in a low level of inconvenience or concern. Events may cause some interference with functioning and may require self-treatment.
- S2 – Events interrupt a subject's usual daily activity and may require systemic drug therapy or other professional medical intervention.
- S3 – Meets the definition of SAE as outlined in Section 15.1.2 of CIP (study protocol).

8.2 Device Deficiencies

Device deficiencies will be monitored throughout the study. The eCRF may include but is not limited to the following data elements: device event type, dates of occurrence and site awareness, action taken with device, and relationship to any AE.

9 DATA QUALITY ASSURANCE

A separate data management plan discusses the steps taken to ensure data quality.

10 GENERAL STATISTICAL METHODS

10.1 Reporting Conventions

All variables will be summarized using descriptive statistics. Results will be summarized by treatment and for the treatment difference (when applicable) using methods appropriate for paired data.

The following reporting conventions apply generally to tables, listings, and figures (TLFs):

- By-subject listings will be created for important variables in each eCRF module and will be presented for all subjects and sorted by subject number, visit, and time (if applicable).
- Summary tables for continuous variables will contain the following statistics: N (number of participants in analysis set), n (number of non-missing observations), mean, standard deviation (SD), median, minimum, maximum, and if applicable, the 95% confidence interval (CI) for the mean.
- Summary tables for categorical variables will include the following statistics: N, n, and percentage. The tables for certain endpoints may also include 95% CIs for the percentage, where appropriate. Unless otherwise indicated, 95% CIs for binary variables will be calculated using the binomial distribution. Missing data will be included as a row in tables where it is appropriate.
- Analyses will be based upon the observed data unless methods for handling missing data are specified. If there is a significant degree of non-normality, analyses may be performed on log-transformed data or nonparametric methods may be applied, as appropriate.
- All analyses will be conducted using SAS® version 9.4 or higher.

10.2 Baseline Definition

The baseline is defined as Visit 1/ Day 1.

10.3 Analysis Populations

10.3.1 Performance Analysis Set

The Performance Analysis Set will include all randomized participants with sufficient data to conduct analyses of performance endpoints. All analyses of clinical performance and clinical benefit endpoints will be conducted on the Performance Analysis Set. Participants will be analyzed according to the actual treatment received.

10.3.2 Safety Analysis Set

The Safety Analysis Set will include all participants who receive treatment. All safety analyses will be conducted on the Safety Analysis Set. Participants will be analyzed according to the actual treatment received.

10.4 Handling of Dropouts or Missing Data

In general, values for missing data will not be imputed unless methods for handling missing data are specified. No data are planned to be excluded due to being an outlier.

10.5 Cohen's Effect Size Calculation

For continuous endpoints, Cohen's d [3] will be calculated as a measure of effect size. The formula for paired data will be used:

d = Mean of paired differences/ SD of the paired differences.

The paired treatment difference will be calculated as:

Treatment Difference = Trillium mask response – AirFit F20 mask response.

Subjects with missing data from one or both treatment periods will be excluded from this calculation. Effect sizes can be interpreted as small (d = 0.2), medium (d = 0.5), or large (d ≥ 0.8).

10.6 Linear Mixed Effect Modeling Approach

Statistical analyses for continuous endpoints will be conducted using a linear mixed effect modeling approach.

The following statistical model will be used:

$Y_{ijkl} = \mu + \beta_{i(k)} + P_j + S_k + T_l + \varepsilon_{ijkl}$

Where:

- Y_{ijkl} denotes the response for Subject i, Period j, Sequence k, Treatment l;
  - Subject $i = 1, 2, \ldots, N_k$ (i^{th} subject in k^{th} sequence).
  - Period $j = 1$ (First Period) or 2 (Second Period).
  - Sequence $k = 1$ (Trillium then AirFit F20) or 2 (AirFit F20 then Trillium).
  - Treatment $l = 1$ (Trillium mask) or 2 (AirFit F20 mask).
- μ is the overall mean.
- $\beta_{i(k)}$ is the random effect for Subject i nested within Sequence k; $\beta_{i(k)}$ are assumed to be independent identically distributed normal random variables, with $\beta_{i(k)} \sim N(0, \sigma_\beta^2)$.
- P_j is the fixed effect for Period.
- S_k is the fixed effect for Sequence.
- T_l is the fixed effect for Treatment.

- ε_{ijkl} denotes the random error; ε_{ijkl} are assumed to be independent identically distributed normal random variables, with $\varepsilon_{ijkl} \sim N(0, \sigma_e^2)$.
- $B_{i(k)}$ and ε_{ijkl} are assumed to be independent.

Adjusted least squares means, standard errors (SE) of the means, and 2-sided 95% CIs for the means will be presented by treatment, and for the difference in treatments.

It is assumed that there is no carryover effect in this study. The participants enrolled have chronic OSA and have been on PAP therapy for at least 3 months. While there is no wash-out period included in this study, the duration of usage for each treatment is 15 days. This is expected to be sufficient time for any subjective carryover effects to dissipate. While we cannot statistically assess the carryover effect in this study design, the treatment-by-period interaction will be evaluated (i.e., treatment-by-period as an alias for carryover [4]). The other effects (i.e., period effect, and sequence effect) will also be evaluated before testing the treatment effect. If the required assumptions for the mixed model are violated, transformation of the data or non-parametric methods may be used. The model may be adjusted appropriately in the case of non-convergence or other issues. The final model used will be specified in the table output.

11 SUBJECT DISPOSITION

A summary of subject enrollment will be presented based on all screened participants. The summary will present the number and percentage of participants for the following categories:

- Screened (participants who signed informed consent)
- Enrolled (participants who met all eligibility criteria)
- Enrolled and Randomized
- Enrolled but Not Randomized
- Not Enrolled
- Reasons for Not Enrolled
- Failed Eligibility Criteria

The reasons for not enrolling will be collected on the eCRF using the following categories:

- Eligibility not met
- Subject decision

- Investigator decision
- Sponsor decision
- Other

A subject disposition summary will be presented by treatment sequence and overall, based on all randomized participants. The summary will present the number and percentage of participants for the following categories:

- Completed each Visit
- Completed the Study
- Discontinued the Study
- Reasons for Discontinuation

The reasons for discontinuation will be collected on the eCRF using the following categories:

- Physician Decision
- Withdrawal by Subject
- Sponsor Request
- Adverse Event
- Protocol Deviation
- Lost to Follow-up
- Death
- Other

The number and percentage of participants included in each analysis set will be summarized by treatment sequence and overall, based on all randomized participants.

12 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

12.1 Demographic Characteristics

Demographic Characteristics summaries will be based on all enrolled participants, and on the Performance Analysis Set.

Demographic characteristics include age, sex, race, and ethnicity.

Results will be summarized using descriptive statistics by treatment sequence and overall.

For continuous subject characteristics, means will be compared using one way analysis of variance (ANOVA). The incidence of the categorical variables will be compared using the Chi-square test.

12.2 Baseline Characteristics

Baseline characteristics summaries will be based on all enrolled participants, and on the Performance Analysis Set.

Baseline characteristics include the participant's typical mask information (Name of prescribed PAP device, mask size), preferred sleeping position, and presence of facial hair.

Results will be summarized using descriptive statistics by treatment sequence and overall.

For continuous subject characteristics, means will be compared using ANOVA. The incidence of the categorical variables will be compared using the Chi-square test.

12.3 Medical History

The participant's medical history may be screened by the investigator or designee against the eligibility criteria but is not collected in the eCRF.

12.4 Concomitant Medication

Concomitant medications will not be collected in this study.

13 PROTOCOL DEVIATIONS

Protocol deviation summaries will be based on all enrolled participants, and on the Performance Analysis Set.

Deviations that occur between Visit 1 date and Visit 2 date will be assigned to the first treatment group used. Deviations that occur between Visit 2 date + 1 and Visit 3 date will be assigned to the second treatment group used. Any deviations that occur on the date of Visit 2 will be manually reviewed by the study team to confirm that the assignment to the first treatment group is appropriate. If the deviation is more appropriately assigned to the second treatment group, then this will be documented in the analysis dataset. The review will occur in a blinded manner prior to DB lock.

Protocol deviations will be summarized by treatment group and overall. The summary will include the number and percentage of subjects experiencing the following categories:

- Any Protocol Deviation

- Any Major Protocol Deviation

The total number of events in each category will also be included.

The total number of deviations of each type will also be included.

- Deviation Type
  - Eligibility criteria not met
  - Study device not used per protocol
  - Visit completed out of window
  - Other

14 PRODUCT ACCOUNTABILITY

Product accountability information will be included in data listings only.

15 CLINICAL PERFORMANCE ANALYSIS

Statistical analysis of the performance endpoints will be performed for the Performance Analysis Set.

15.1 Unintentional Leak Analysis

The mask leak rate will be summarized descriptively by treatment group and for the treatment difference. Cohen's d effect size will be calculated (See Section 10.5).

The statistical analysis for unintentional leak will be conducted using a linear mixed effect modeling approach (See Section 10.6).

Acceptance Criteria for Leak rate: If the upper limit of the 95% CI for a treatment mean is ≤ 24 L/min, then the leak rate will be considered acceptable.

15.2 Patient Reported Outcomes Analysis

Numerical PRO responses will be summarized descriptively by treatment group and for the treatment difference. Cohen's d effect size will be calculated (See Section 10.5).

The statistical analysis for numerical PRO responses will be conducted using a linear mixed effect modeling approach (see Section 10.6).

Acceptance Criteria for Comfort of Breathing: If the lower limit of the 95% CI for a treatment mean is ≥ 7.0, then the satisfaction will be considered acceptable.

Categorical PRO responses will be summarized using descriptive statistics for all participants combined. Additionally, to assess the impact of order of treatments, results will also be summarized by treatment sequence.

16 CLINICAL BENEFIT ANALYSIS

Statistical analysis of the clinical benefit endpoints will be performed for the Performance Analysis Set.

16.1 Apnea-Hypopnea Index Analysis

AHI will be summarized descriptively by treatment group and for the treatment difference. Cohen's d effect size will be calculated (See Section 10.5).

The statistical analysis for AHI will be conducted using a linear mixed effect modeling approach (see Section 10.6).

Acceptance Criteria for AHI: If the upper limit of the 95% CI for the mean treatment difference is < 5, then the AHI for the Trillium mask will be considered non-inferior to the AirFit F20 mask.

Additionally, AHI will be categorized by severity as follows:

- Normal: AHI < 5 events/hr
- Mild OSA: AHI ≥ 5 to < 15 events/hr
- Moderate OSA: AHI ≥ 15 to < 30 events/hr
- Severe OSA: AHI ≥ 30 events/hr

The proportion of responses in each severity level will be summarized by treatment group. Contingency tables will be used to compare the distribution of severity across the 2 treatment groups.

16.2 Adherence Analysis

In addition to the values collected from the device, the following variables will be calculated programmatically:

- Percent of Days Device Actually Used – calculated as 100*Number of days device actually used /Number of days device expected to be used.

- Percent of Days Device Used for < 4 hours – calculated as 100*Number of days device actually used for < 4 hours / Number of days device expected to be used.
- Percent of Days Device Used for ≥ 4 hours – calculated as 100*Number of days device actually used for ≥ 4 hours / Number of days device expected to be used.
- Participant meets adherence criteria, defined as having average time used ≥ 4 hours/day.

Adherence responses will be summarized descriptively by treatment group and for the treatment difference. Cohen's d effect size will be calculated (See Section 10.5).

The statistical analysis for numerical outcomes will be conducted using a linear mixed effect modeling approach (see Section 10.6).

Acceptance Criteria for Average Time used (hours/Day) (all days): If the lower limit of the 95% CI for the mean treatment difference is > -0.5, then the usage for the Trillium mask will be considered non-inferior to the AirFit F20 mask.

16.3 Epworth Sleepiness Scale Analysis

The ESS score will be calculated as the sum of all 8 item scores, and ranges from 0 – 24. The higher the ESS score, the higher that person's average sleep propensity (ASP) in daily life, or their 'daytime sleepiness'.

The ESS score will be summarized descriptively by treatment group and for the treatment difference. Cohen's d effect size will be calculated (See Section 10.5).

The statistical analysis for ESS score will be conducted using a linear mixed effect modeling approach (see Section 10.6).

Acceptance Criteria for ESS score: If the upper limit of the 95% CI for the mean treatment difference is < 2, then the ASP for the Trillium mask will be considered non-inferior to the AirFit F20 mask.

Additionally, the ESS score will be categorized to assess ASP as follows [5]:

- 0 – 5: Lower Normal Daytime Sleepiness
- 6 – 10: Higher Normal Daytime Sleepiness
- 11 – 12: Mild Excessive Daytime Sleepiness
- 13 – 15: Moderate Excessive Daytime Sleepiness
- 16 – 24: Excessive Daytime Sleepiness

The proportion of responses in each severity level will be summarized by treatment group. Contingency tables will be used to compare the distribution of severity across the 2 treatment groups.

16.4 European Quality of Life-5D-3L Analysis

Each of the 5 dimensions of the EQ-5D descriptive system (see Section 7.4) has three response levels. Responses will be coded as single-digit numbers expressing the severity level as follows:

- Level 1 = no problems
- Level 2 = some problems
- Level 3 = extreme problems

For each dimension, the proportion of responses in each severity level will be summarized by treatment group. Contingency tables will be used to compare the distribution of severity across the 2 treatment groups.

The severity level digits for the five dimensions can be combined in a 5-digit code that describes the participant's health state [2]; for instance, 21133 indicates some problems in the Mobility dimension, no problems in the Self-care, no problems in the Usual Activities dimensions, extreme problems in the Pain/Discomfort, and extreme problems in the Anxiety/Depression dimension.

The EQ-5D-3L summary index will be derived to reflect how good or bad a health state is according to the preferences of the general population of the country/region.

An EQ-5D summary index will be derived by applying a formula that attaches values (weights) to each of the levels in each dimension. The index is calculated by deducting the appropriate weights from 1, the value for full health (i.e., state 11111).

The collection of index values (weights) for all possible EQ-5D health states is called a value set. The most current United States value sets will be used [6] [7].

The EQ-5D summary index ranges from -0.59 to 1; where 1 represents the best possible health state, 0 represents death, values between 0 and 1 represent varying degrees of health problems, and values < 0 represents health states worse than death.

The EQ-5D summary index and the EQ VAS will be summarized descriptively by treatment group and for the treatment difference. Cohen's d effect size will be calculated (See Section 10.5).

The statistical analysis for EQ-5D summary index and the EQ VAS will be conducted using a linear mixed effect modeling approach (see Section 10.6).

Acceptance Criteria: For EQ-5D summary index: If the lower limit of the 95% CI for the treatment difference is > -0.08, then the summary index for the Trillium mask will be considered non-inferior to the AirFit F20 mask.

17 CLINICAL SAFETY ANALYSIS

All analyses of safety data will be based on the Safety Analysis Set.

17.1 Adverse Events

Adverse events will be summarized in terms of 'reportable' AEs, as defined in CIP Section 15.1.

Adverse events that occur between Visit 1 date and Visit 2 date will be assigned to the first treatment group used. Adverse events that occur between Visit 2 date + 1 and Visit 3 date will be assigned to the second treatment group used. Any AEs that occur on the date of Visit 2 will be manually reviewed by the study team to confirm that the assignment to the first treatment group is appropriate. If the AE is more appropriately assigned to the second treatment group, then this will be documented in the analysis dataset. The review will occur in a blinded manner prior to DB lock.

An overall summary of AEs will be presented by treatment group. The summary will include the number and percentage of participants experiencing the following categories:

- Any AE
- Any Serious AE (SAE)
- Any AE Possibly Related or Causally Related to Study
- Any Adverse Device Effect (ADE; AE Causally Related to Device)
- Any Serious Adverse Device Effect (SADE; SAE Causally Related to Device)
- Any AE Leading to Study Device Removal (temporarily or permanently)
- Any AE Leading to Study Discontinuation
- Any AE Leading to Death

The total number of events in each category will also be included.

For each AE category defined in Section 8.1, the number and percentage of participants experiencing each category and severity level will be summarized by treatment group.

Acceptance Criteria: The acceptance criteria vary by AE category and severity level as described in CIP Section 4. Acceptability will be based on the observed percentage of participants, in consideration of other factors such as device usage and subjective PROs.

17.2 Device Events/ Deficiencies

Device events/ deficiencies will be summarized descriptively by treatment group. The summary will include the number and percentage of participants experiencing the following categories:

- Any Device Event/ Deficiency
- Device Event/ Deficiency Related to Reportable AE

18 SUBGROUP ANALYSIS

If demographic or baseline characteristics show imbalances which may affect interpretation, subgroup analyses accounting for such factors may be conducted.

Additional exploratory subgroup analyses may be conducted as needed.

19 ADJUSTMENTS FOR COVARIATES

There are no adjustments for covariates planned.

20 MULTIPLE COMPARISON AND MULTIPLICITY

No adjustments for multiplicity will be made in this study.

All analyses will be performed using an unadjusted two-sided significance level of α=0.05.

21 CHANGES IN THE CONDUCT OF THE STUDY OR PLANNED ANALYSIS

21.1 Changes in the Conduct of the Study

Any changes in the conduct of this study will be reported in the Clinical Investigation Report (CIR).

21.2 Changes in the Planned Analysis

There are no changes in the planned analysis described in the CIP.

22 TABLES, FIGURES, AND LISTINGS

The shells for Tables, Listings, and Figures will be presented in a separate stand-alone document.

23 REFERENCES

[1] M. Johns, "A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale," *Sleep,* vol. 14, no. 6, pp. 540-545, November 1991.

[2] EuroQol Research Foundation, "EQ-5D-3L User Guide," 2018. [Online]. Available: https://euroqol.org/publications/user-guides.

[3] J. Cohen, Statistical power analysis for the behavioral sciences, 1988.

[4] C.-Y. Lim and J. In, "Considerations for crossover design in clinical study," *Korean J Anesthesiol,* vol. 74, no. 4, pp. 293-299, 2021.

[5] "About the ESS," [Online]. Available: https://epworthsleepinessscale.com/about-the-ess/. [Accessed 04 June 2025].

[6] "EuroQol Value Sets," [Online]. Available: https://euroqol.org/information-and-support/resources/value-sets/. [Accessed 04 June 2025].

[7] J. Shaw, J. Johnson and S. Coons, "US Valuation of the EQ-5D Health States: Development and Testing of the D1 Valuation Model," vol. 43, no. 3, pp. 203-220, 2005.

24 DOCUMENT REVISION HISTORY

Version	Date	Author	Description of Change	Reason for Change
1.0	22APR2026	M Nelson	Not Applicable – Initial Release	

25 DOCUMENT APPROVAL HISTORY

Approved by	Role / Function	Signature & Date
Melissa Nelson, MS Associate Director, Biostatistics	Project Statistician (SAP Author)	[SIGNATURE DETAILS REDACTED]
Sara Garbin Senior Clinical Development Scientist	Clinical Development	[SIGNATURE DETAILS REDACTED]
Kristi Seese Clinical Study Manager	Clinical Study Manager	[SIGNATURE DETAILS REDACTED]
Kevin Najarian, MS Global Director of Biostatistics and Data Management	Director of Biostatistics	[SIGNATURE DETAILS REDACTED]