Statistical Analysis Plan

A Phase III, Randomised, Double-Blind, Placebo-Controlled Trial Evaluating Arinestat® in Adults with Moderate-to-Severe Chronic Obstructive Pulmonary Disease

Protocol Number	ARI-COPD-301
IND Number	123456
EudraCT Number	2024-001234-56
SAP Version	2.1
Effective Date	14 March 2025
Replaces Version	2.0 (08 November 2024)
Document Status	**FINAL — LOCKED PRIOR TO UNBLINDING**
Sponsor	Arion Therapeutics, Inc., 100 Biopark Drive, Cambridge, MA 02139
CRO	Cytel Statistical Services, 675 Massachusetts Ave, Cambridge, MA 02139

Prepared by Biostatistics Department

Lead Statistician: Dr. Meera Krishnaswamy, PhD

Statistical Programmer: James O'Brien, MSc

Medical Monitor: Prof. Alejandro Garza-López, MD, PhD

Contents

1 Introduction

1.1 Purpose of This Document

This Statistical Analysis Plan (SAP) describes in detail the pre-specified statistical methodology, analysis populations, estimands, and output specifications for the primary, secondary, and exploratory analyses of Protocol **ARI-COPD-301**. This document was finalised and locked *prior to database lock and unblinding* in accordance with ICH E9(R1) and the trial's pre-registration commitments.

1.2 Protocol Synopsis

Study Title	CLARITY: A Phase III Trial of Arinestat vs. Placebo in COPD
Design	Randomised, double-blind, parallel-group, placebo-controlled
Indication	Moderate-to-severe COPD (GOLD Stage II–III)
Investigational Product	Arinestat 250 µg inhaled twice daily
Comparator	Matching placebo inhaler twice daily
Randomisation	1:1, stratified by GOLD stage (II vs. III) and smoking status
Sample Size	620 participants (310 per arm)
Treatment Duration	52 weeks (plus 4-week safety follow-up)
Primary Endpoint	Change from baseline in trough FEV_1 at Week 52
Key Sites	48 sites across USA, EU, Canada

1.3 Regulatory Context

Analyses will be conducted in accordance with:

- ICH E9 *Statistical Principles for Clinical Trials* (1998) and ICH E9(R1) Addendum on *Estimands and Sensitivity Analysis* (2019)
- ICH E3 *Structure and Content of Clinical Study Reports* (1995)
- FDA Guidance: *Adjusting for Covariates in Randomised Clinical Trials* (2023)
- CHMP Guideline on *Missing Data in Confirmatory Clinical Trials* (EMA/CPMP/EWP/1776/99 Rev.1, 2010)
- CDISC ADaM Implementation Guide v1.3 (ADSL, ADLB, ADEFF, ADAE)

2 Study Objectives and Estimands

2.1 Primary Objective

To evaluate the efficacy of Arinestat 250 µg BID versus placebo in improving pulmonary function in adults with moderate-to-severe COPD as measured by trough forced expiratory volume in one second (FEV_1) at Week 52.

2.2 Secondary Objectives

1. To assess the effect of Arinestat on health-related quality of life (SGRQ Total Score).
2. To evaluate the rate of moderate-to-severe COPD exacerbations over 52 weeks.
3. To examine the effect on peak FEV_1 (2 h post-dose) at Week 52.
4. To determine the proportion of responders achieving $\geq$100 mL improvement in trough FEV_1 at Week 52.
5. To characterise the pharmacodynamic biomarker profile (serum fibronectin-1 levels).

2.3 Estimand Framework (ICH E9R1)

Population	Adults with GOLD Stage II–III COPD meeting all eligibility criteria (ITT population)
Variable	Change from baseline in trough FEV_1 (litres) at Week 52
Intervention	Arinestat 250 µg BID vs. matching placebo BID
Intercurrent events	**Treatment policy** for discontinuation due to any reason (all post-IC data included); **Hypothetical** strategy for initiation of rescue LABA/ICS (data censored)
Summary measure	Difference in adjusted means (Arinestat minus Placebo)

3 Analysis Populations

Population	Definition	Primary Use
ITT	All randomised participants, regardless of treatment received or protocol deviations. Participants analysed in the arm to which they were randomised.	Primary efficacy
PP	Subset of ITT excluding participants with major protocol deviations (MPDs) as pre-defined in the MPD charter. Includes participants with $\geq$80% treatment compliance over 52 weeks.	Sensitivity
Safety	All randomised participants who received at least one dose of study treatment. Participants analysed in the arm in which treatment was actually received.	Safety
PK/PD	Participants in the Safety population with at least one evaluable PK sample and baseline biomarker value.	Exploratory

Population assignments are determined prior to database lock by an independent Data Review Committee based on blinded data only.

4 Primary Endpoint Analysis

4.1 Model Specification

The primary analysis will use a Mixed Model for Repeated Measures (MMRM) fitted to change from baseline in trough FEV_1 (L) at scheduled post-baseline visits (Weeks 4, 12, 24, 40, 52).

Model equation:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_1 \cdot \text{BL}_k + \gamma_2 \cdot \text{BL}_k \cdot V_j + \delta_1 \cdot \text{GOLD}_k + \delta_2 \cdot \text{SMOKE}_k + \varepsilon_{ijk}$$

where Y_{ijk} is the change from baseline for participant k in treatment group i at visit j; α_i is the treatment effect; β_j is the visit effect; $(\alpha\beta)_{ij}$ is the treatment-by-visit interaction (the primary contrast at Week 52); BL_k is the baseline trough FEV_1; GOLD_k and SMOKE_k are stratification factors.

4.2 Model Fitting Details

- Covariance structure: unstructured (UN); if convergence fails, fall back to Toeplitz then compound symmetry, in that order.
- Degrees of freedom: Kenward–Roger approximation.
- Software: SAS® PROC MIXED (v9.4).
- All post-baseline assessments included, regardless of adherence status (treatment-policy intercurrent event strategy).

4.3 Hypothesis and Decision Rule

$$H_0\text{: } \mu_{\text{Arinestat}} - \mu_{\text{Placebo}} \leq 0 \qquad \text{vs.} \qquad H_1\text{: } \mu_{\text{Arinestat}} - \mu_{\text{Placebo}} > 0$$

A one-sided $\alpha = 0.025$ will be used (equivalent to two-sided $\alpha = 0.05$). The primary endpoint will be declared statistically significant if the upper bound of the two-sided 95% CI for the treatment difference excludes zero and the one-sided p-value < 0.025.

4.4 Sensitivity Analyses

S1. Per-Protocol MMRM: Repeat primary model in PP population.

S2. ANCOVA at Week 52: Single-visit ANCOVA using LOCF-imputed values as a pre-specified comparison; no alpha used.

S3. Multiple Imputation (MI): Reference-based MI under missing-at-random (MAR) and missing-not-at-random (MNAR, jump-to-reference) assumptions; 1000 imputed datasets, Rubin's rules.

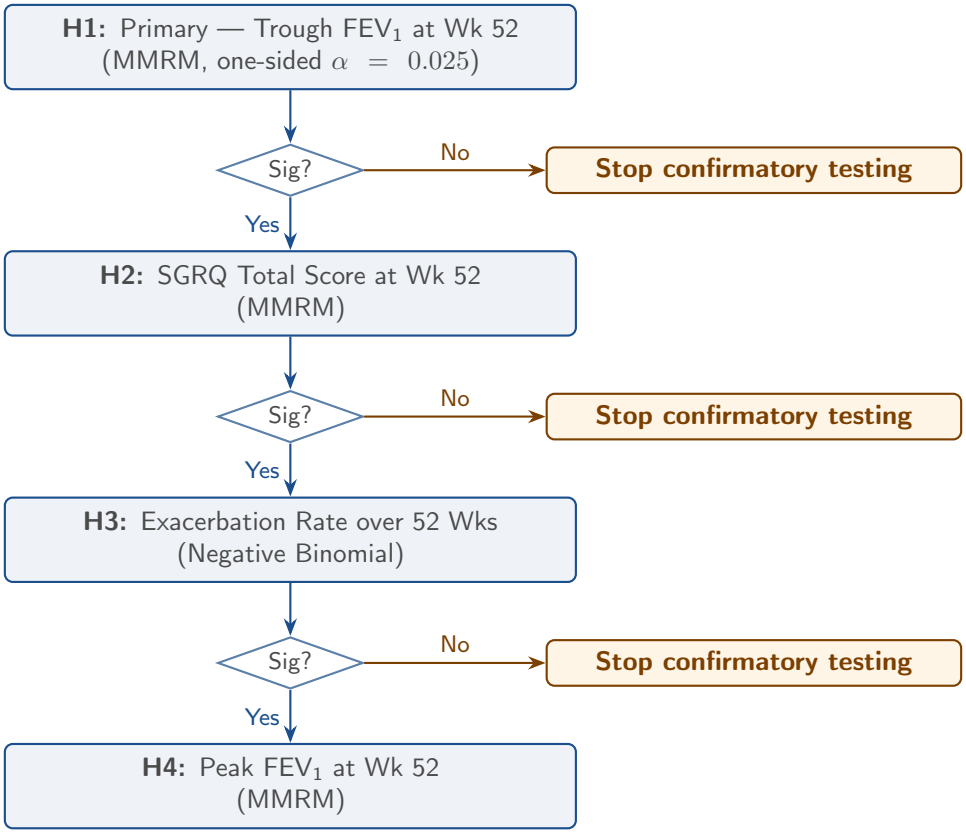
S4. Tipping Point: Compute the degree of MNAR departure needed to reverse significance.

5 Secondary Endpoints and Multiplicity Adjustment

5.1 Hierarchical Testing Procedure

All confirmatory secondary endpoints will be tested using the fixed-sequence (hierarchical) procedure outlined below. Testing proceeds to the next endpoint only if the preceding test is significant at $\alpha = 0.025$ (one-sided).

5.1.1 Multiplicity Adjustment Flowchart



The fixed-sequence procedure controls the family-wise error rate (FWER) at one-sided $\alpha = 0.025$ under arbitrary correlation among test statistics. No Bonferroni, Holm, or Hochberg adjustments are applied within the sequence. Exploratory endpoints (H5 onward) are not multiplicity-adjusted; p-values are nominal.

5.2 Secondary Analysis Methods

H#	Endpoint	Method	TFL Ref
H2	SGRQ Total Score change from baseline at Wk 52	MMRM (same model)	ET-EFF-002
H3	Annualised rate of moderate/severe COPD exacerbations	Negative Binomial GLM, log(time on study) offset	ET-EFF-003
H4	Peak FEV_1 (2h post-dose) change from BL at Wk 52	MMRM	ET-EFF-004
H5	Proportion of FEV_1 responders ($\geq$100 mL) at Wk 52	Logistic regression; OR + 95% CI	ET-EFF-005
H6	Time to first moderate/severe exacerbation	Kaplan–Meier; log-rank; Cox PH	ET-EFF-006

6 Safety Analyses

Safety analyses will be performed in the Safety population using descriptive statistics only; no formal hypothesis testing is pre-specified for safety endpoints.

6.1 Adverse Events

- Adverse events (AEs) are coded using MedDRA v27.0.
- Incidence of Treatment-Emergent AEs (TEAEs) summarised by System Organ Class (SOC) and Preferred Term (PT).
- Participants counted once per PT at the maximum severity and strongest relationship to study drug.
- Separate summaries for: all TEAEs, Grade $\geq$3 TEAEs, SAEs, TEAEs leading to discontinuation, AEs of Special Interest (AESI: cardiovascular, hepatic).

6.2 Clinical Laboratory Parameters

Continuous lab parameters (haematology, biochemistry, urinalysis): descriptive statistics (n, mean, SD, median, Q1/Q3, min, max) at each visit. Shift tables (normal/abnormal at baseline vs. post-baseline) for key parameters. Notable abnormalities flagged using CTCAE v5.0 grading and sponsor-defined reference ranges.

6.3 Vital Signs and ECG

Vital signs (blood pressure, heart rate, respiratory rate, SpO_2) and 12-lead ECG parameters summarised descriptively. QTcF prolongation: proportion of participants with QTcF $>$450, $>$480, $>$500 ms and change from baseline $>$30, $>$60 ms.

An independent Safety Review Committee (SRC) will review unblinded safety data at Weeks 12 and 26. SRC findings may trigger an ad hoc safety analysis per the SRC Charter (Doc. ARI-301-SRC-001). Such analyses are not pre-specified in this SAP and will be reported separately.

7 Missing Data Handling

7.1 Assumed Missing Data Mechanism

For the primary analysis, missing post-baseline FEV_1 values are assumed to be Missing At Random (MAR) conditional on observed history, stratification factors, and baseline covariates. This assumption is untestable but is clinically plausible given the COPD context and observed withdrawal patterns.

7.2 Primary Analysis

The MMRM utilises all available observations; monotone and intermittent missingness are both accommodated by the likelihood-based approach under MAR. No imputation is performed in the primary analysis.

7.3 Sensitivity and Supplemental Analyses

Analysis	Description
Jump-to-Reference (J2R)	Post-withdrawal outcomes for Arinestat participants imputed from the placebo arm conditional mean trajectory. Conservative MNAR assumption.
Copy-Reference (CR)	Post-withdrawal outcomes for all participants imputed from the placebo arm.
LOCF	Last-observation-carried-forward for supplemental ANCOVA; no inferential alpha.
Tipping Point	Vary the degree of MNAR departure (delta) from the MAR estimate; identify the delta value that reverses the primary conclusion.

8 ADaM Dataset Specifications

All analysis datasets will be constructed in accordance with CDISC ADaM v1.3. Source SDTM datasets: DM, AE, LB, EG, VS, CM, EX, QS (SGRQ), CE (exacerbations).

8.1 Dataset Overview

Dataset	Label	Structure	Key Variables
ADSL	Subject-Level Analysis	One record/-subject	USUBJID, TRT01P, SAFFL, ITTFL, PPFL, RANDDT, GOLD, SMOKSTRS
ADEFF	Primary Efficacy	One record/-subject/visit	USUBJID, PARAMCD, AVISIT, AVISITN, AVAL, BASE, CHG, ANL01FL
ADEXAC	Exacerbation Events	One record/event	USUBJID, STARTDT, ENDDT, SEVER, EXACFL, EXACRATE
ADAE	Adverse Events	One record/event	USUBJID, AEDECOD, AESOC, AESEV, AESER, AEREL, TEAEFL
ADLB	Laboratory	One record/-subject/-param/visit	USUBJID, PARAMCD, AVISIT, AVAL, BASE, CHG, ANRIND
ADEG	ECG	One record/-subject/-param/visit	USUBJID, PARAMCD, AVAL, BASE, CHG, QTCFFL
ADQS	Questionnaire (SGRQ)	One record/-subject/-param/visit	USUBJID, PARAMCD, AVAL, BASE, CHG, AVALC
ADPK	Pharmacokinetics	One record/subject/sample	USUBJID, PCTPT, PCORRES, AVAL, DTYPE

8.2 Key ADEFF Derivations

- `BASE`: Trough FEV_1 measured at Day 1 (Visit 2) pre-dose. If Day 1 value is missing, use Screening value.
- `CHG`: `AVAL` − `BASE`.
- `ANL01FL`: 'Y' for records included in the primary MMRM (post-baseline scheduled visits, ITT population, no rescue medication within 48 h of assessment).
- `AVISIT` mapping: Visits within a protocol-defined window are assigned to the nominal visit label (e.g., any assessment at Day 337±14 days → "Week 48"; Day 365±14 days → "Week 52").

All ADaM datasets will be subjected to independent programmer QC (double programming). Validation reports will be filed in the Trial Master File. Discrepancies will be resolved prior to database lock following the Data Management Plan (DMP-ARI-301-v1.4).

9 Tables, Figures, and Listings (TFL) Shells

9.1 TFL Naming Convention

Prefix	Content Type
DT-DM-	Demographic and baseline tables
ET-EFF-	Efficacy tables
ET-SAF-	Safety tables (AE, lab, ECG, vital signs)
EF-	Efficacy figures
EL-	Listings

9.2 Selected TFL Shells

9.2.1 Table DT-DM-001: Demographic and Baseline Characteristics

[DT-DM-001]

Characteristic	Arinestat 250 µg (N=310)	Placebo (N=310)	Total (N=620)
Age (years)			
Mean (SD)	64.2 (8.7)	63.9 (9.1)	64.1 (8.9)
Median [Q1, Q3]	65.0 [58, 70]	64.0 [57, 71]	64.5 [57, 70]
Sex, n (%)			
Male	178 (57.4)	181 (58.4)	359 (57.9)
Female	132 (42.6)	129 (41.6)	261 (42.1)
GOLD Stage, n (%)			
Stage II	155 (50.0)	155 (50.0)	310 (50.0)
Stage III	155 (50.0)	155 (50.0)	310 (50.0)
Baseline Trough FEV_1 (L)			
Mean (SD)	1.43 (0.31)	1.41 (0.33)	1.42 (0.32)
Median [Q1, Q3]	1.42 [1.21, 1.64]	1.40 [1.18, 1.62]	1.41 [1.20, 1.63]
Current Smoker, n (%)			
Yes	93 (30.0)	96 (31.0)	189 (30.5)
No (Ex-smoker)	217 (70.0)	214 (69.0)	431 (69.5)

SHELL ONLY — values are illustrative of expected format. Final values from ADSL.

9.2.2 Table ET-EFF-001: Primary Efficacy — Trough FEV_1 at Week 52

[ET-EFF-001]

Parameter	Arinestat (N=310)	Placebo (N=310)	Difference (95% CI)
Baseline mean (SD)	1.43 (0.31) L	1.41 (0.33) L	—
LS Mean change at Wk 52 (SE)	+0.112 (0.018)	−0.007 (0.018)	—
Treatment difference (LS Mean)	—		+0.119 L
95% Confidence Interval	—		(0.074, 0.164)
One-sided p-value	—		< 0.001

MMRM with UN covariance, Kenward–Roger df. ITT population. SHELL ONLY — illustrative.

9.2.3 Table ET-SAF-001: Overview of Adverse Events

[ET-SAF-001]

Category	Arinestat 250 µg n (%) [N=310]	Placebo n (%) [N=310]
Any TEAE	xxx (xx.x)	xxx (xx.x)
Any TEAE Grade $\geq$3	xxx (xx.x)	xxx (xx.x)
Any Serious AE (SAE)	xxx (xx.x)	xxx (xx.x)
Any TEAE leading to discontinuation	xxx (xx.x)	xxx (xx.x)
AESI: Cardiovascular	xxx (xx.x)	xxx (xx.x)
AESI: Hepatic	xxx (xx.x)	xxx (xx.x)
Deaths	xxx (xx.x)	xxx (xx.x)

Safety population. MedDRA v27.0. SHELL ONLY — xxx = actual data.

10 Sample Size and Power

10.1 Primary Endpoint Power Calculation

Sample size was computed to detect a treatment difference of **80 mL** (0.080 L) in change from baseline trough FEV_1 at Week 52, assuming:

Parameter	Assumption
Clinically meaningful difference	80 mL (0.080 L)
SD (within-group)	0.290 L
Correlation (baseline, endpoint)	0.55
Type I error (α, one-sided)	0.025
Power ($1-\beta$)	90%
Dropout rate (Week 52)	20%
Required per arm (completers)	248
Required per arm (randomised)	310
Total sample size	620

Computed using nQuery Advisor v8.5 (MMRM module, ANCOVA proxy).

10.2 Interim Analyses

No formal efficacy interim analyses are planned. One pre-specified futility-only interim analysis will occur at approximately 50% of events (approximately Week 52 of enrolment). Futility boundary: conditional power $< 15\%$ under the assumed effect size. The interim is conducted by the independent Data Safety Monitoring Board (DSMB) under the DSMB Charter (Doc. ARI-301-DSMB-001). The sponsor remains blinded.

11 Subgroup Analyses

Subgroup analyses of the primary endpoint (trough FEV_1 at Week 52) are **exploratory**; no multiplicity adjustment is applied. Treatment-by-subgroup interaction p-values are presented for hypothesis generation only.

Subgroup	Levels	Interaction Test
GOLD Stage	II vs. III	LRT p-value
Sex	Male vs. Female	LRT p-value
Age	< 65 vs. ≥ 65 years	LRT p-value
Smoking Status	Current vs. Ex-smoker	LRT p-value
Baseline FEV_1	$\leq$ vs. $>$ median	LRT p-value
Geographic Region	North America vs. Europe	LRT p-value
Prior ICS use	Yes vs. No	LRT p-value

A forest plot of subgroup-specific LS Mean differences with 95% CIs will be produced `[EF-EFF-010]`. The overall treatment effect estimate is shown at the bottom for reference. Point estimates and CIs are from the primary MMRM with a treatment-by-subgroup interaction term added.

12 SAP Amendment Log

Ver.	Date	Summary of Changes	Author / Approver
1.0	15 Apr 2024	Initial SAP issued for regulatory submission. Defines primary estimand, MMRM specification, ADaM structures, and TFL shells.	M. Krishnaswamy / A. Garza-López
2.0	08 Nov 2024	**Amendment 1:** Added tipping-point sensitivity analysis (Section 7); refined ADEFF ANL01FL flag definition to exclude assessments within 48 h of rescue bronchodilator; clarified Kenward–Roger df method for MMRM. No change to primary analysis model, endpoint, or decision rule.	M. Krishnaswamy / A. Garza-López
2.1	14 Mar 2025	**Amendment 2:** Updated MedDRA coding dictionary from v26.1 to v27.0 (Section 6); corrected minor typographical errors in Table ET-SAF-001 shell column header. No change to analysis methodology. Locked prior to database lock.	M. Krishnaswamy / A. Garza-López

This SAP (Version 2.1) was finalised and electronically locked on **14 March 2025**, prior to database lock (planned 28 March 2025) and prior to unblinding of any participant-level data. Any post-lock modifications require a protocol deviation notification and written justification filed in the Trial Master File.

Lead Statistician	Dr. Meera Krishnaswamy, PhD	_____
Medical Monitor	Prof. Alejandro Garza-López, MD, PhD	_____
Clinical Lead	Dr. Priya Ramachandran, MD	_____

13 Programming and Quality Control

Aspect	Specification
Statistical software	SAS® 9.4 (TS1M7) — primary analysis. R 4.4.1 (sensitivity analyses, forest plots).
ADaM programming	All datasets double-programmed by independent programmers. Comparison using PROC COMPARE (SAS) and `diffdf` (R).
TFL QC	All tables, figures, and listings quality-controlled by an independent statistician against the SAP specifications.
Randomisation code	Provided by IXRS vendor (Medidata Rave) and stored in sealed envelope; not accessible to sponsor biostatistics team.
Operating system	Red Hat Enterprise Linux 8.6 (64-bit).
Code archival	All programs and logs archived in the sponsor DMS (Veeva Vault) under study folder ARI-COPD-301/Biostatistics.

14 Abbreviations

Term	Definition
ADaM	Analysis Data Model (CDISC)
AE	Adverse Event
AESI	Adverse Event of Special Interest
ANCOVA	Analysis of Covariance
BID	Twice Daily
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
CR	Copy-Reference
CRO	Contract Research Organisation
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	Data Safety Monitoring Board
FEV_1	Forced Expiratory Volume in 1 Second
FWER	Family-Wise Error Rate
GOLD	Global Initiative for Chronic Obstructive Lung Disease
ICS	Inhaled Corticosteroid
ICH	International Council for Harmonisation
IND	Investigational New Drug
ITT	Intent-to-Treat
IXRS	Interactive X-act Randomisation System
J2R	Jump-to-Reference
LABA	Long-Acting Beta-Agonist
LOCF	Last Observation Carried Forward
LRT	Likelihood Ratio Test
MAR	Missing At Random
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MMRM	Mixed Model for Repeated Measures
MNAR	Missing Not At Random
MPD	Major Protocol Deviation
PP	Per-Protocol
PT	Preferred Term
QC	Quality Control