Precision Interventions for Severe and/or Exacerbation-Prone Asthma

(PrecISE)

National Clinical Trial (NCT) Identified Number: NCT04129931

Investigational New Drug Application (IND) Sponsor: University of North Carolina at Chapel Hill

Funded by: National Heart, Lung, and Blood Institute (NHLBI)

STATISTICAL ANALYSIS PLAN

Protocol Version 3.3

Protocol Version Date: 27 Sep 2024

Statistical Analysis Plan Version: 5.0

14 Oct 2024

Approval Page

Study Title: Precision Interventions for Severe and/or Exacerbation-Prone Asthma (PrecISE)

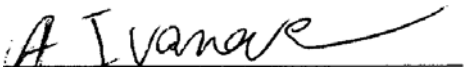
Version: 5.0

Date of Issue: October 14, 2024

Study Sponsor: David Peden, MD, MS
Harry S. Andrews Distinguished Professor of Pediatrics
Senior Assoc. Dean for Translational Research
University of North Carolina School of Medicine
Chapel Hill, NC 27599-7310
Tel: 919 966 0768
Email: david.peden@med.unc.edu

We, the undersigned, have read and approve of this Statistical Analysis Plan and agree on its content.

_____
Sponsor Representative – David Peden, MD, MS



_____
Protocol Statistician – Anastasia Ivanova, PhD

Contents

List of Abbreviations 5
1. Introduction 5
1.1 Study Overview 5
1.2 Changes due to the Covid-19 Pandemic 7
1.3 Other Study Design Changes 7
1.4 Document Disposition 8
2. Objectives, Endpoints, and Estimands 8
3. Statistical Hypotheses and Endpoints 10
3.1 Hypotheses 10
3.2 Primary and Key Secondary Efficacy Assessments 10
3.3 Other Secondary Efficacy Assessments 12
3.4 Exploratory Efficacy Assessments 12
3.5 Study Definitions and Created Variables 12
3.5.1 Visits 12
3.5.2 End of Study 12
3.5.3 Asthma Exacerbations 12
3.5.4 Major Change in Therapy 13
4. Randomization 13
4.1 Two-Step Randomization Procedure 13
4.2 Intervention Assignment 14
4.3 Assignment of Test Treatment versus Placebo and Masking 14
5. Power and Sample Size 15
6. Data Sources, Documentation, and Verification of Results 16
6.1 Clinical Data Entry and Management in CDART 16
6.2 External Data Transfers 17
6.3 Pandemic-Related Changes in Data Sources 18
7. Statistical Analyses 18
7.1 Analysis Populations 18
7.2 Baseline Descriptive Statistics 19
7.3 Unblinding for Interim and Final Analyses of Each Intervention 19
7.4 Analysis of Primary Efficacy Endpoints 20
7.4.1 Statistical Model 20
7.4.2 Intercurrent Events 21
7.4.3 Sensitivity Analyses 21
7.5 Analysis of Secondary Endpoints 22
7.6 Safety Data Analyses 22

7.7 Interim Analysis 23
7.7.1 Futility Interim Analysis 23
7.7.2 Blinded Sample Size Re-Estimation 23
7.7.3 Precision Medicine Interim Analysis 23
7.8 Precision Medicine Final Analysis 24
7.9 Exploratory Analyses 25
7.9.1 Analysis of Monitoring Biomarkers 25
7.9.2 Analysis of Adolescents 25
7.9.3 Subgroup Analysis Using Proximal Biomarkers 25
7.9.4 Precision Medicine Analyses for Interventions with Overlapping Target Subgroups 25
References 27

List of Abbreviations

ACQ-6	Asthma Control Questionnaire
AE	Adverse Event
AQLQ	Asthma Quality of Life Questionnaire
ASSESS	Asthma Severity Scoring System
ATS	American Thoracic Society
BD	Bronchodilator
CDART	Carolina Data Acquisition Reporting Tool
CFR	Code of Federal Regulations
CRF	Case Report Form
CSCC	Collaborative Studies Coordinating Center
DMCC	Data, Modeling and Coordinating Center
DMS	Data Management System
DSMB	Data and Safety Monitoring Board
eCRF	Electronic Case Report Forms
ED	Emergency Department
FDA	US Food and Drug Administration
FeNO	Fractional exhaled Nitric Oxide
ICU	Intensive Care Unit
IND	Investigational New Drug Application
ITT	Intention-To-Treat
MCT	Medium Chain Triglycerides
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intention-To-Treat
MMRM	Mixed Model for Repeated Measures
NCT	National Clinical Trial
NHLBI	National Heart Lung and Blood Institute
NIH	National Institutes of Health
OCS	Oral Systemic Corticosteroids
ODM	Operational Data Model
PARC	Protocol Adaptation Review Committee
PEF	Peak Expiratory Flow
PrecISE	Precision Interventions for Severe and/or Exacerbation-Prone Asthma Network
SAE	Serious Adverse Event
SARP	Severe Asthma Research Program
SD	Standard Deviation
SUSARs	Serious, Unexpected, Suspected Adverse Reactions
WPAI	Work Productivity and Activity Impairment Questionnaire

1. Introduction

1.1 Study Overview

PrecISE is a platform trial conducted under a single Master Protocol to identify new therapies for severe asthma that are effective in biomarker-defined subgroups of participants. Five therapies, clazakizumab, medium chain triglycerides (MCT), imatinib, cavosonstat, and Broncho-Vaxom are currently being

investigated. The trial is designed to meet two primary objectives: (1) Identify novel therapies that work in biomarker-defined subgroups of participants with severe asthma, and (2) Optimize the subgroups targeted for treatment by refining the biomarker and subgroup definitions.

The study population corresponds to adult and pediatric (12-18 years of age) participants who meet modified guideline criteria for severe asthma and who are currently uncontrolled or continue to have exacerbations. Participants must be at least 12 years of age, on a stable regimen of asthma medications prior to enrollment, and satisfy other inclusion/exclusion criteria. Two interventions, MCT and Broncho-Vaxom, are being studied in both adults and adolescents and the other three, clazakizumab, imatinib and cavosonstat, in adults only. The treatment assignment probabilities are restricted accordingly.

For each participant, the study consists of three successive phases: (i) an initial screening phase, (ii) a 2-period crossover phase, and (iii) a single-period crossover phase with successive re-randomizations.

i. Initial screening phase: Inclusion and exclusion criteria for each intervention are applied to determine the set of interventions for which a participant is eligible. Biomarker profiles of the participants are also determined.
ii. 2-period crossover phase: participants are randomly assigned to one of the five interventions based on their biomarker profile and enter a 2-period, double-blind crossover phase consisting of two 16-week treatment periods separated by a washout period of length dependent on the half-life of the intervention (the wash-out length is 16 weeks for clazakizumab and 8 weeks for all other interventions). Treatment sequence is randomly assigned as either test treatment followed by matching placebo (T:P) or vice-versa (P:T).
iii. Single-period crossover phase with repeated re-randomizations: participants enter the single-period crossover phase of the study consisting of successive single 16-week treatment periods followed by washout periods. Randomization occurs in two steps. First, a participant is randomly assigned to an intervention for which they are eligible and have not yet received, based on their biomarker profile. Second, the participant is assigned to receive either test treatment or matching placebo. The probability of receiving a placebo during period 3 or later is controlled such that approximately 25% of participants will have two placebos.

Each intervention has an a priori target subgroup defined as the patients thought most likely to benefit from the intervention based on the literature at the time of study start. Table 1 shows the definition of the a priori target (biomarker positive) subgroups and their prevalence estimated with data from the Severe Asthma Research Program (SARP).[1] The biomarker negative subgroup for each intervention is defined as the complement of the biomarker positive subgroup.

Table 1. Definition of Target Subgroups Based on Biomarkers Assessed at Screening

Treatment	Target Subgroup Definition	Estimated Prevalence
Imatinib	Blood eosinophils < 300	62%
Clazakizumab	Interleukin 6 > 3.1	33%
Cavosonstat	Genotypes rs7669660 TT and rs11547772 AA	64%
Broncho-Vaxom	Blood eosinophils ≥ 300	38%
MCT	Fractional Exhaled nitric oxide ≥ 15 ppb	64%

The participants' biomarker profiles are assessed during screening and used to determine participants' membership in each intervention's target subgroup. Participants' biomarker profiles are not updated

during the study, as participants are exposed to the various interventions. This approach reduces the likelihood of misidentification based on carry over effects from an intervention received.

Throughout the study, interventions are randomly assigned based on participants' biomarker profiles such that participants have a higher likelihood of assignment to an intervention that targets their particular profile.

The study is adaptive in that interim analyses are planned to assess the futility of each agent, with the goal of discontinuing those with lower probabilities of success to more effectively utilize trial resources for the remaining agents. Additionally, randomization to the biomarker negative subgroup can be stopped at the interim analysis if the pre-defined threshold to continue is not met.

The Data and Safety Monitoring Board (DSMB) established for PrecISE has oversight responsibility for the study. Periodic reviews of safety data will be performed by the DSMB throughout the study, and interventions found to pose significant safety risks in the severe asthma population will be recommended for discontinuation from the study. In addition, the DSMB may make recommendations regarding early stopping of interventions for futility as well as recommendations for modifying the target subgroup for an intervention based on accumulating data.

The Protocol Adaptation Review Committee (PARC) may be asked to review potential design adaptations that are pre-specified in the study protocol and that require the decision-making body to be masked to treatment assignment and comparative analyses (which the DSMB is not).

1.2 Changes due to the Covid-19 Pandemic

Several changes regarding the number of interventions we are able to investigate and the data collection procedures have been made in response to the Covid-19 pandemic. PrecISE enrollment opened in December 2019 but was closed in March 2020 just prior to the first patient's scheduled randomization (following the 8-week screening period). Enrollment re-opened in August 2020 but has proceeded much slower than anticipated due to pandemic related changes in clinic operations (e.g., Covid testing requirements at some clinics) as well as concern that some interventions being studied could impair efficacy of the Covid vaccines, once they were rolled out late in 2020. As a result, the following study changes have been enacted:

- Five interventions, rather than the six originally selected, will be studied in PrecISE. The JAK-2 inhibitor already obtained from the company will not enter the master protocol as planned.
- Procedures have been implemented to allow for home-based data acquisition in the event of clinic closures. Participant-administered questionnaires and eDiaries have been programmed in Qualtrics for web-based administration from the participants' homes.
- The use of a home spirometer for lung function assessments is being implemented as an alternative to in-clinic spirometry
- Lab Corp was contracted to conduct safety labs in the event a participant cannot travel to the clinical site for blood draws required to monitor the safety of an intervention.

1.3 Other Study Design Changes

We performed a planned blinded interim analysis of the within-subject (between-period) correlation (not adjusted for treatment) for the PrecISE crossover trial design. The analysis was performed on 64 pairs of observations from consecutive crossover periods. The estimated within-subject correlation for the CompEx[2] endpoint was 0.33, lower than the value of 0.38 assumed for the power analysis and sample size determination specified in the protocol. This is due, in part, to a lower-than-expected CompEx event rate. The estimated within-subject correlation for the other two primary endpoints, FEV_1 and ACQ-6, was 0.7 or higher. Based on this information and external data on within-subject correlation

for FEV_1 and ACQ, 0.7 was used as the within-subject correlation for re-calculating the required sample size.

CompEx is a novel endpoint which combines exacerbations and deterioration events that are not exacerbations. It relies on study participants to record their peak flow, symptoms and rescue use twice daily. Participant compliance with their daily diaries was lower than expected. This information plus the low correlation and its impact on the power available to detect an intervention effect with respect to CompEx contributed to the decision to remove CompEx as a primary endpoint in the study and consider it a key secondary endpoint.

The target sample size of the study has been revised to 395 participants, due to the change from three to two primary endpoints and change in crossover correlation assumptions for the efficacy analysis. The enrollment period has been extended to a duration of 42 months, compared to the originally planned 30 months. The study period has been extended to a duration of 54 months, with the last patient visit expected in December 2024 and safety follow-up expected to conclude in February 2025. The revised study period and sample size should be sufficient to support our investigation of the five interventions currently active in the PrecISE master protocol. We estimate a total of 1000 treatment periods will be available to contribute to the statistical analyses for evaluation of efficacy and optimization of target subgroups for each intervention. With the current study timelines and enrollment goals, each participant should be able to complete a minimum of two treatment periods and a maximum of six, including at least one placebo period, depending on their time of enrollment.

Changes to the target sample size and study duration created difficulty in accruing to biomarker positive subgroups with small prevalence. This applies to a priori best subgroups and to subgroups that could result after cut-off estimation during the interim analysis. Hence, the planned precision medicine analyses to re-estimate the biomarker cut-off for each intervention has been replaced with a pre-specified analysis where the biomarker positive and negative subgroups are evaluated at the interim with an option of stopping accrual to the biomarker negative group. The precision medicine interim analysis to evaluate biomarker cut-offs and estimate subgroups based on additional biomarkers will be conducted at the end of the study as previously planned.

1.4 Document Disposition

This document describes the planned statistical analyses that will be conducted at interim analysis points and at the completion of the study of each of the interventions. Once the initial unblinded data review has occurred for an intervention, only the blinded statistics team members will be allowed to modify the Statistical Analysis Plan.

2. Objectives, Endpoints, and Estimands

The PrecISE study has two primary objectives:

1. Identify novel therapies that work in biomarker-defined subgroups of severe asthma patients
2. Optimize the subgroups targeted for treatment by refining the biomarkers and subgroup definitions

To achieve the first objective, we will perform hypothesis tests for comparisons of each intervention to placebo, (independently of the other interventions) for each of the two primary endpoints (defined below). An intervention will be considered efficacious if significant benefit is found with respect to either of the two primary endpoints. To achieve the second objective, we will enroll participants inside

and outside the target subgroups and optimize the target subgroup definitions using precision medicine methods.

The study also has two secondary objectives:

1. Gain information about potential monitoring biomarkers for selected therapies
2. Explore the safety and effectiveness of selected therapies in adolescent patients with severe asthma

For the first secondary objective, we will identify which monitoring biomarkers could potentially serve as early or intermediate outcomes for an investigation. These biomarkers may prove useful in predicting whether a drug will be effective with respect to clinical outcomes, which could in turn inform future trials of the drug. For the second, we will describe trends with respect to the primary efficacy outcomes observed in the adolescent subgroup and determine if the risk profile of an intervention appears to differ between adults and adolescents. We do not expect to have sufficient power for formal hypothesis tests of effects in the adolescent subgroup alone or for a test of age by treatment interactions.

For both primary objectives, we have defined two primary endpoints:

1) FEV_1 percent predicted, assessed prior to bronchodilator administration (primary)
2) Asthma symptom control, assessed via the Juniper Asthma Control Questionnaire (ACQ-6) (primary)

An intervention will be considered beneficial if it demonstrates improvement with respect to either of the two primary endpoints relative to placebo. See Section 3 for details about assessments for each of the primary and additional secondary endpoints.

Summaries of the questions of interest and descriptions of the estimand attributes are shown in Table 2, following the framework provided in the International Council on Harmonisation (ICH) guideline, E9(R1).[3]

Table 2. Study Objectives and Estimands for FEV_1 percent predicted, ACQ-6 and CompEx

Question of Interest	Objective Description / Study Population	Endpoint(s)	Intercurrent Events	Population Summary
Does Intervention A improve lung function compared to placebo, for patients suffering from severe and/or exacerbation-prone asthma?	Primary objective of the study / All randomized participants receiving Intervention A during any study period	FEV_1 percent predicted assessed prior to bronchodilation	For analyses of FEV_1 percent predicted, intercurrent events include (i) asthma exacerbations (protocol Section 6.5.3), (ii) major change in therapy (protocol Section 6.5.4); and (iii) treatment discontinuation due to adverse events or lab abnormalities (protocol Section 7.1). Note that study discontinuation when outside of a treatment period is not an intercurrent event, as each new treatment period is associated with a new randomization for treatment assignment.	Mean difference in FEV_1 between test treatment and placebo and its associated confidence interval from a mixed model for repeated measurements (MMRM). Positive values indicate a benefit of Intervention A.

Does intervention A improve asthma symptom control compared to placebo, for patients suffering from severe and/or exacerbation-prone asthma?	Primary objective of the study / All randomized participants receiving Intervention A during any study period	6-item score from the Juniper Asthma Control Questionnaire (ACQ-6)	For analyses ACQ-6, intercurrent events are defined identically to those defined for FEV_1 percent predicted.	Mean difference in ACQ-6 between test treatment and placebo and its associated confidence interval from a MMRM. Positive values indicate a benefit of Intervention A.
Does intervention A reduce asthma exacerbations and deterioration events compared to placebo, for patients suffering from severe and/or exacerbation-prone asthma?	Secondary objective of the study / All randomized participants receiving Intervention A during any study period	CompEx events, which include both asthma exacerbations and clinically-relevant deterioration events	For CompEx, asthma exacerbations and use of rescue medications are incorporated in the CompEx event definition. Major change in therapy (protocol Section 6.5.4) and treatment discontinuation due to adverse events or lab abnormalities are intercurrent events for this primary endpoint.	Incidence ratio of CompEx events per unit time between test treatment and placebo and its associated confidence interval from a log-linear model fit with generalized linear model methods. Values less than 1 indicate a benefit of Intervention A.

3. Statistical Hypotheses and Endpoints

3.1 Hypotheses

For the first primary study objective (overall assessment of efficacy), the following hypotheses will be tested for each intervention (independently of the other interventions):

1. Intervention A does not improve FEV_1 percent predicted compared to placebo (the null hypothesis) versus A is better than placebo (the alternative hypothesis)
2. Intervention A does not improve asthma symptom control assessed via ACQ-6 compared to placebo (the null hypothesis) versus A is better than placebo (the alternative hypothesis)

The second primary objective involves estimation of the optimal subgroup for targeting participants for each intervention. No formal statistical hypotheses will be tested for this objective.

3.2 Primary and Key Secondary Efficacy Assessments

Primary Efficacy Assessments: Efficacy of each intervention will be evaluated with respect to lung function and asthma symptom control. Lung function will be measured prior to and post bronchodilator administration (pre-BD and post-BD). The American Thoracic Society (ATS) criteria for assessing lung function will be applied for spirometric maneuvers. A central spirometry reading center will provide quality control over-reads of the measurements and identify any issues requiring retraining of clinical staff. The primary efficacy outcome with respect to lung function is FEV_1 percent predicted pre-BD, measured in the clinic. In the event of a clinic shut-down, home spirometry assessment of the primary endpoint will be used.

Asthma symptom control will be measured using six items from the Juniper Asthma Control Questionnaire (ACQ-6) corresponding to five self-reported symptoms and self-reported rescue BD use. The primary efficacy outcome with respect to asthma control symptoms is the average score of these six items (range 0-6).

Secondary Efficacy Assessments: The use of CompEx events as a key secondary endpoint provides an outcome with statistical properties that approximate exacerbations but with a shorter follow-up time.[2] CompEx is a composite outcome specific to asthma that combines clinically relevant deteriorations captured by diary events with exacerbations, thereby providing an increase in power compared to using exacerbations alone. Exacerbations will be analyzed as another secondary efficacy endpoint.

CompEx events include exacerbations and deterioration events defined based on

1) Daily recordings of peak expiratory flow (PEF) morning/evening (L/min)
2) Reliever use morning/evening (doses)
3) Symptoms morning/evening (score 0–3) assessed from twice-daily diary recordings

The diary event can occur as defined by threshold and slope criteria within a moving window of 5-day length. Evening and morning recordings are treated as separate variables. The baseline levels of deterioration are calculated for each individual as the mean over the 5-10 days ending just before the day of randomization for each of the diary variables. No imputation of missing diary data after the randomization is performed. Deterioration criteria is assessed for each (single) diary variable for thresholds and slopes as follows:

Thresholds: The change from baseline is calculated. If two consecutive days fulfil the chosen threshold limit as defined in Table 3, the deterioration criterion is met.

Slopes: A slope is calculated via univariate linear regression over 5 days. If the slope fulfills the chosen cut-point as defined in Table 3, the deterioration criterion is met.

Diary event start definition: A diary event can occur when (i) the threshold deterioration criterion is met for at least two diary variables, or when (ii) the threshold deterioration criterion is met for one diary variable, and the slope criterion is fulfilled for all included variables. In case of (i), the diary event is defined to start on the first day of the two consecutive deterioration days (event days 0–1). Any missing data in this two-day window will make the event missing. In case of (ii), the event is defined to start on the first of the two days fulfilling the threshold criterion. This means that the slopes are calculated for days -4 to 0 of an event. At least two days with data are needed to calculate slopes in order to qualify as an event. The end of a diary event is the last day that the criteria for a diary event is fulfilled. In order to be counted as a new diary event it must be preceded by at least 7 days in which neither criterion for a diary event is fulfilled.

Table 3. CompEx Threshold and Slope Values

Diary variable	PEF (P) morning/evening	Reliever use (R) morning/evening	Symptoms (S) morning/evening
Threshold type	Decrease from baseline (%)	Increase from baseline (doses)	Increase from baseline (scores) or absolute maximum score
Threshold	15	1.5	1
Slope type	Decrease rate (% per day)	Increase rate (doses per day)	Increase rate (scores per day)
Slope	3	0.3	0.2

3.3 Other Secondary Efficacy Assessments

1. FEV_1, post-BD
2. FVC, pre-BD
3. Exacerbations, time to first (see Section 3.5.3 below)
4. Symptom Free days (defined using diary/CompEx data), where a symptom free day is defined as a symptom score = 0 for both morning and evening in the daily diary
5. Healthcare utilization
    a. Asthma-specific ED visits
    b. Asthma-specific hospital admissions
    c. Asthma-specific Intensive Care Unit (ICU) admissions
6. Asthma free days
   An asthma-free or asthma-control day, as documented in each patient's diary, is a day during which there is
    - no use of albuterol rescue (excluding the use of albuterol as pre-exercise treatment),
    - no daytime or nighttime asthma symptoms,
    - no peak expiratory flow of less than 80% of the predetermined reference value.

   If any of the morning or evening records of rescue use, symptoms and peak flow are missing, that day is considered to be missing and is excluded from analysis.

3.4 Exploratory Efficacy Assessments

1. Rate of recovery from exacerbation, time to improvement after
2. Economic burden of disease assessment (Days missed from work/school, cost of hospitalizations)
3. Asthma Quality of Life Questionnaire (AQLQ)
4. Work Productivity and Activity Impairment Questionnaire (WPAI)[4]
5. Modified Asthma Severity Scoring System (ASSESS)

3.5 Study Definitions and Created Variables

3.5.1 Visits

- Baseline visit is Visit X.1, where X indicates the treatment period and ranges from 1 to 6.
- Endpoint visit is Visit X.5, where X indicates the treatment period and ranges from 1 to 6.
- On-treatment visits are monthly and include Visits X.2, X.3, X.4, and X.5
- Visit windows for baseline and on-treatment visits are ± 5 days

3.5.2 End of Study

The original timeline for randomization and study completion has been extended due to the impact of the Covid-19 pandemic. The last randomization visit is now expected to occur in February 2024 and the last on-treatment visit in December 2024. Safety follow-up will conclude in February 2025, 54 months after enrollment was resumed in August 2020 following the shut-down due to the Covid-19 pandemic. Study closeout, analyses and dissemination of findings will conclude in June 2025.

3.5.3 Asthma Exacerbations

An asthma exacerbation in PrecISE is defined as a worsening of asthma requiring the use of a systemic corticosteroid (at least 3 days of treatment) to prevent a serious outcome (Protocol Section 6.5.3). This definition is consistent with recommendations from the National Institutes of Health (NIH) Outcomes

Workshop.[5] Events meeting one or more of the following criteria will be recorded as asthma exacerbations during the study:

- All worsening asthma events in which systemic corticosteroids were initiated to prevent a serious outcome, including use of systemic corticosteroids in association with any form of healthcare provider encounter
- All asthma-specific emergency department or urgent care visits that involved treatment with systemic corticosteroids (defined as evaluation and treatment for <24 hours in an ED or urgent care center)
- All asthma-specific hospitalizations that involved treatment with systemic corticosteroids (defined as admission to an inpatient facility and/or evaluation and treatment in a healthcare facility for ≥24 hours) (also reported as a serious adverse event (SAE))
- All asthma-specific intensive care unit admissions or intubations (also reported as a serious adverse event)
- All asthma-related deaths (also reported as a serious adverse event)

For the purposes of this study, a course of oral systemic corticosteroids (OCS) is defined as any increase of at least 20 mg prednisone or double the daily maintenance dose for a minimum of 3 days. Note that a single depo-injectable dose of corticosteroids or oral dexamethasone will be considered equivalent to a 3-day course of OCS. If the maintenance dose is <5 mg, an increase must be both a doubling of that dose and at least a 5 mg increase (e.g., 2.5 mg would need to increase to at least 7.5 mg to count as a course).

Two courses of systemic corticosteroids must be separated by at least two weeks to count as two exacerbations. The two-week clock starts when the corticosteroid course for the first exacerbation ends.

3.5.4 Major Change in Therapy

A major change in therapy is defined in Section 6.5.4 of the protocol. If a participant has a major change in therapy during a treatment period, they will discontinue their current assigned treatment and then discontinue the study. Any assessments collected after the major change in therapy will not be included in the efficacy analyses.

4. Randomization

4.1 Two-Step Randomization Procedure

Randomization and re-randomization are implemented in two steps throughout the study, as follows:

1. At the first step, participants eligible for the study are randomly assigned to receive one of the study interventions from among the set of interventions for which they are eligible, after applying intervention-specific exclusion criteria.
2. At the second step, participants are randomly assigned to receive test treatment or placebo with a method that depends on study phase:
    a. 2-period crossover phase: participants beginning period 1 are assigned a crossover sequence (T:P or P:T) with equal probability, to achieve a 1:1 allocation of test treatment versus placebo in each of the 1st two periods.
    b. Single-period crossover phase with repeated re-randomizations: Participants beginning period 3 and higher are assigned either test treatment or placebo with unequal probability, with the goal of 25% of all participants receiving a second placebo.

4.2 Intervention Assignment

Each of the five interventions has a target (biomarker positive) subgroup defined based on a priori information available at the time the intervention enters the trial (Table 1). Each participant's status relative to the target subgroups is based on the biomarker profile determined at screening.

The target subgroups initially defined for each intervention have some degree of overlap, such that participants may be in the target subgroup for more than one intervention. In this sense, the interventions are competing with each other for targeted participants. The 1st-step randomization probabilities are applied to determine which intervention a participant is assigned among those interventions for which they are eligible. These probabilities are adjusted throughout the study to ensure the study objectives can be met. That is, accrued data on intervention eligibility and biomarker profiles (target subgroup membership) are monitored during the study, and randomization probabilities adjusted as needed to ensure a sufficient number of treatment periods (test and placebo) are available to support the planned interim (futility) and final (efficacy and precision medicine) analyses.

The initial randomization algorithm, implemented at study start, was guided by the following principles, given in priority order:

1. Interventions that have not yet reached the required sample size to test for futility are favored over interventions that have already enrolled enough subjects to conduct the futility analysis.
2. Interventions are randomly assigned to favor participants in the target subgroup of an intervention over participants outside that subgroup in 2:1 ratio, until the time of the interim analysis for futility.
3. Intervention assignment takes into consideration the distribution of three prognostic factors for asthma outcomes: (i) past-year exacerbations at the study's baseline (assessed during the initial screening period), (ii) blood eosinophils, and (iii) IL-6 plasma levels. The goal is to ensure that participants contributing to the efficacy analysis (interim and final) of each intervention resemble the severe asthma subpopulation with respect to these three factors. Initial targets for balancing were based on the factor distributions in the SARP population of severe asthma patients and are updated as PrecISE baseline data become available.

Due to delays in certain interventions being available to enter the study in addition to the pandemic shut-down, modifications to the initial randomization algorithm are needed to achieve the target sample sizes for each intervention at interim and final analysis times. Simulations based on participant accrual, eligibility data, and biomarker profiles are conducted periodically during the study (by statisticians masked to treatment assignments and outcome data) to predict the treatment periods that will be available for each intervention and each analysis time point. Randomization probabilities for intervention assignment are then adjusted, as needed, to achieve the required sample sizes for these analyses.

4.3 Assignment of Test Treatment versus Placebo and Masking

Each intervention has its own placebo that matches the test treatment in mode of administration, color, size, and/or taste.

In the first two periods, each participant will receive a test treatment and its matching placebo, with the order being randomly determined (crossover sequence allocation). During subsequent treatment periods (periods 3-6), the placebo assignment probability is initially set to 0.17 but monitored throughout the study and adjusted as needed to yield, on average, approximately 25% of all participants receiving a second placebo. Once a participant receives a 2nd placebo, their probability of receiving placebo in subsequent periods is set to 0 (with some exceptions – see below). This strategy results in

participants enrolled earlier and therefore eligible for more treatment periods having a higher chance of receiving a second placebo than participants enrolled later in the trial.

If a participant discontinues treatment during periods 1 or 2 such that no post-baseline assessments of the FEV_1 percent predicted and ACQ-6 are available, the probability of their receiving placebo in periods 3 and 4 might be adjusted to ensure the participant will have efficacy evaluable data from at least one placebo period during the study. In period 3, the probability of receiving placebo = 0.5, and in period 4, the probability = 0 if placebo is received in period 3, and = 1 otherwise. The participant is also eligible to receive another placebo in periods 5-6 with the same probability as any other participant. A participant may therefore be randomly assigned placebo at most three times during the study and will contribute efficacy data to the statistical analysis from at most two placebo periods.

The initial randomization to a complete 2-period crossover design, with all participants receiving placebo during one of these two periods, allows for efficient estimation of treatment effects to support futility analysis of each test treatment early in the study. The addition of a randomly assigned second placebo allows for the estimation of time differences in placebo response and the potential to adjust the final efficacy analysis for any time trends observed. Because some participants will receive two placebos with different modes of administration (e.g., oral and injection or dietary and oral, etc.), we will also be able to estimate differences in placebo response between different administration modes.

The second placebo also provides some protection against expectation bias (relative to outcome assessments and other reporting behaviors) on the part of both the participants and the investigators/clinical staff, because there is a positive probability of placebo assignment in each treatment period for periods 3-6. Note also that to further maintain masking in periods 3-6, a subject can get a placebo matched to intervention A only if the subject has not yet received intervention A, and a subject who has completed a period with placebo matched to intervention A is not eligible to receive intervention A for the remainder of the study.

Using simulated data for the 2-step randomization scheme described above, we tabulated the number of periods between each active period (in periods 3-6) and the closest placebo period assuming a 16-week treatment period and 8-week washout (all interventions except clazakizumab). As can be seen from the results shown in Table 4 below, active and placebo periods will be adjacent 36% of the time, have one period between them 38% of the time, and separated by three or more periods (more than one year) only 9% of the time.

Table 4. Distribution of Time between Active Periods and the Closest Placebo Periods for Periods 3-6

0 periods	1 period	2 periods	3 periods	4 periods
0 years	0.46 years	0.92 years	1.38 years	1.87 years
36%	38%	17%	7%	2%

5. Power and Sample Size

An intervention is considered efficacious if significant benefit is shown for either of the two primary endpoints (FEV_1, or asthma symptom control score). Because there are multiple chances for an intervention to demonstrate efficacy, multiplicity adjustments are required. We will use the Hochberg method to adjust for multiplicity. According to the Hochberg method, the treatment is declared effective if the treatment-placebo comparison is significant at level α for both outcomes or it is significant at $\alpha/2$ for at least one of the outcomes.

A sample size of 111 participants for each intervention is required to achieve at least 80% power to detect a treatment effect with respect to at least one of the two primary endpoints equal to 0.3 times the standard deviation (SD) of that outcome, taking into account the possibility of stopping for futility with a non-binding stopping rule. If participant accrual in the target subgroup of one or more interventions is such that an interim futility analysis is not possible or practical (e.g., the study will conclude before the recommendations of the interim futility analysis can be implemented), the interim analysis will not be performed. In this case, 93 participants (2:1 ratio of biomarker positive to negative participants) are required to provide similar power for the final analysis.

All participants randomized to an active intervention or matching placebo will be included in the final efficacy analysis, but only participants who do not withdraw from the study prior to the start of period 2 are counted towards the required total sample size in estimating power. We also assume that about 10% of participants will discontinue an active treatment period soon after randomization in the power calculations. Attrition is monitored throughout the study, and the required sample size may be adjusted, if the treatment discontinuation rate is different from 10% or discontinuations occur later in the treatment period. This adjustment, if needed, will take place prior to any unmasking of treatment assignments, that is, prior to the interim futility analysis or, if the futility analysis is not feasible for an intervention, prior to the final analysis.

The Type I error probability will be controlled under the global null hypothesis (no effect on either endpoint) at $\alpha \leq 0.10$ two-sided. Power was estimated under the assumption of a within-participant, between-period correlation of responses in consecutive placebo and test treatment periods of 0.70.

Examples of treatment effects yielding a 0.3 standardized effect size include: a difference of 4.3 percent predicted FEV_1 between intervention and placebo with a SD of 14.5 percent predicted; and a difference of 0.18 in average ACQ-6 symptom scores between the intervention and placebo with a SD of 0.6.

Our analysis method ensures strong control of the Type I error probability at $\alpha = 0.10$ two-sided for each intervention. This value was chosen to reflect the experimental nature of this phase 2 proof-of-concept study. Taking into account both our stopping rule for futility and the Hochberg method for controlling multiplicity due to multiple primary endpoints, the Type I error probability is expected to equal about 0.05. Interventions that demonstrate evidence of efficacy in PrecISE with respect to at least one of the two primary endpoints may be studied further in a more fully powered phase 3 trial to provide substantial evidence of efficacy, controlling the Type I error probability at the more usual level of 0.05. Information learned from PrecISE regarding the optimal target subgroup for the intervention should be useful in developing an enrichment strategy for such a study.

No adjustment for multiplicity with respect to the multiple interventions in the study is planned. Success is defined independently for each intervention, and interventions will not be compared to each other in the efficacy analyses.

We are planning to enroll 395 participants. Assuming a study withdrawal rate of 2% a month, we anticipate that each participant will contribute data to two active interventions on average. This will allow us to accumulate required number of active treatment periods to test efficacy in all five interventions.

6. Data Sources, Documentation, and Verification of Results

6.1 Clinical Data Entry and Management in CDART

Clinical data are entered at each study site on electronic case report forms (eCRFs) using the Carolina Data Acquisition and Reporting Tool (CDART), a proprietary web-based data management system (DMS)

developed and maintained by the Collaborative Studies Coordinating Center (CSCC) in the Department of Biostatistics at the University of North Carolina at Chapel Hill. CDART incorporates robust security features, such as authentication log-ins, granular permissions based on user requirement, and encrypted data transmission to protect personal health information (PHI). CDART is compliant with the United States Food and Drug Administration's (FDA) 21 Code of Federal Regulations (CFR) Part 11 guidance (an audit report dated July 2021 confirming 21 CFR Part 11 compliance is available upon request from UNC's Office of Clinical Trials).

Real-time data edits are performed at the time of data entry that include range checks, consistency checks, and routing or skip pattern checks. Cross form checks and other algorithmic data queries are programmed for post-data entry checking in both CDART and SAS and. All SAS queries are sent to the sites for resolution biweekly, and sites are encouraged to run query reports from within CDART at least weekly. Customizable data reports are generated from within CDART (e.g., calendar scheduling for follow-up visits, enrollment reports, etc.). Additionally, data summaries are available through a web-based dashboard with views for report generation outside of CDART (e.g., outstanding query reports by site).

To facilitate data sharing, CDART conforms to the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM) standard. The ODM was designed to support electronic acquisition, exchange and archive of electronic data and meta-data in a standard XML-based format. The ODM defines structures to represent clinical study metadata, clinical study administrative data and clinical study patient data. CDART allows automated data extraction using a RESTful API to provide data and meta-data to authorized callers. SAS and other analysis software packages can easily process the files to define datasets and populate them with study data.

All tables, data listings, and graphical displays summarizing PrecISE study data and all statistical analyses described in this analysis plan will be generated using SAS Version 9.4 or higher (SAS Institute, Cary, NC) and the R language.

Data summaries and analysis results are verified prior to release through the following steps:

- Programs for producing created variables and analysis files; programs for generating data summaries, data listings, and graphical displays; and programs for conducting statistical analysis models are fully specified by project statisticians prior to programming
- Program output is reviewed and verified by comparing against the analysis specifications
- Independent validation is then conducted by team members not involved in the original specifications, for primary and key secondary analyses

6.2 External Data Transfers

Datasets generated from outside vendors are transferred via secure file transfer protocol (FTP) to the PrecISE coordinating center. The data are loaded into the CDART system using an 'Extract, Transform, Load' process. During the load step the data are validated. Any errors in data type, value or expected record structure are reported to the external vendor. External files include data from the central laboratories (PPD and Labcorp), e-diary data (symptoms, controller medication adherence, peak flow) from Propeller Health, data from CT scans of the lung, spirometry data for the assessment of lung function from Vyaire and MIR, results from sputum analyses, and results of other biospecimen analyses. Periodic transfers of data from the respective cores, reading centers, or vendors for each of these sources are processed, and quality control checks performed for data completeness and accuracy. Queries based on comparisons of data from the external data sources and data entered on eCRFs at the clinic are generated, where appropriate, and sent to the sites for resolution. Additional queries to

confirm appropriate follow-up at the clinical sites, when called for, based on external data values (e.g., abnormal lab values) are also generated as quality control checks.

6.3 Pandemic-Related Changes in Data Sources

To accommodate the event of a clinic shutdown during the study, a spirometer that could be used in the home was incorporated for acquisition of lung function data. The study began with the Vyaire spirometer for in-clinic assessments. During the first year of enrollment, participants were provided a MIR spirometer for use in the home. A pilot study was conducted to compare results from the two spirometers, and based on these results, the DSMB approved a switch to the MIR spirometer for assessing lung function at all study visits. The first and last visit in each treatment period (Visit X.1 and X. 5 for period X) are to be conducted with the MIR spirometer in the clinic, and all other visits may use in-clinic or home spirometry at the discretion of the site coordinator, with two exceptions. Note that patients who began the study using the Vyaire spirometer will not switch for their current or any future treatment periods. The methacholine challenge and maximum BD response assessments are to continue to be assessed during screening using the Vyaire spirometer in the clinic. In the event of a clinic shutdown, the Visit X.1 and X.5 spirometry maneuvers are to be conducted in the home using the MIR spirometer already assigned to the participant.

7. Statistical Analyses

7.1 Analysis Populations

Screening Population: Participants who are screened but not randomized will be reported, but not included in any efficacy or safety analyses of study data. SAEs occurring during the screening period are monitored and reported to regulatory agencies each month and will be summarized in final study reports.

Safety population: The Safety population for each intervention consists of all participants randomized to that intervention (active or placebo) who received at least one dose of study medication. This population will form the basis for all safety analyses and selected secondary efficacy analyses of the intervention.

Intention-to-Treat (ITT) Population: The ITT population for each intervention will consist of all participants randomized to that intervention. Participants in the ITT population who are also in the target subgroup for an intervention (in either the a priori target subgroup or in the biomarker negative subgroup) will form the basis for the interim futility analysis.

The primary and secondary efficacy analyses will be conducted using a modified intention-to-treat (mITT) population. The mITT population for an intervention consists of participants who are randomized to the intervention, receive study drug, and contribute at least one post-baseline measurement for each of the primary endpoints during the treatment period in which the participant is randomized to the intervention.

Per-Protocol Population: The per protocol population for each intervention consists of all participants in the mITT population who complete study treatment and are in reasonable compliance with the protocol during the treatment period in which the participant received the intervention. For per protocol analysis purposes, we define compliance as:

- For MCT, clazakizumab, Broncho-Vaxom and cavosonstat, the per-protocol population will include participants completing at least 14 weeks of treatment. In addition, the percentage of intended dose of the study drug taken over the course of follow up is at least 60% for MCT, at

least 75% for clazakizumab (at least 3 out of 4 injections), and 90% for Broncho-Vaxom and cavosonstat.

- For imatinib, the per-protocol population will include participants completing at least 8 weeks of treatment and the end of treatment visit. Additionally, participants should have reached the target dose of 400 mg at least 4 weeks before the end of treatment.

Membership in the per protocol population will be determined prior to unmasking of treatment assignments to avoid the potential for bias. This population will form the basis of selected secondary analyses.

7.2 Baseline Descriptive Statistics

The initial 8-week screening period for each randomized participant is considered to be the baseline evaluation period. Baseline statistical analyses will consist of descriptive statistics (means and standard deviations, medians and first and third quartiles, and minimums and maximums) for continuous baseline measures such as current age, age at first asthma diagnosis, baseline pulmonary function parameters (including methacholine hyperresponsiveness), and asthma symptom severity. Categorical and binary variables such as gender, prior medication history, atopic status, and asthma-relevant genotypes will be summarized as number and percentage for each category. Poisson-type outcomes, such as exacerbations, will be summarized as event rates.

Predictive biomarkers will be assessed at baseline and summarized according to the distributions of values as well as the number and percentage of participants in and out of the target subgroup for each intervention. Predictive biomarkers are re-assessed throughout the study, but the determination of participant biomarker profiles for purposes of treatment assignments are based on the baseline values of predictive biomarkers and not modified during the study.

7.3 Unblinding for Interim and Final Analyses of Each Intervention

Both interim and final analyses for each intervention are conducted independently of other interventions, consistent with the master protocol design. Statistical analyses for Intervention A are based on participants randomized to Intervention A during the trial. The treatment periods in which participants were randomized to A plus matched placebo periods contribute to the analysis. For participants without a matched placebo, all placebo periods will be included in the analysis. For analyses based on the mITT and per protocol populations, the treatment periods included in the analyses are limited to those satisfying the respective mITT and per protocol criteria stated above. Other treatment periods from the participant might be used in the analysis to estimate the within-participant correlation in the crossover if the primary model does not converge.

The target sample size for the interim analysis of Intervention A, with objective to test for futility, is 34 biomarker positive participants and 17 biomarker negative participants.

Once an intervention achieves the target sample size for an interim analysis, an appropriate data cut-off is determined for the interim analysis (sufficient to allow completion of data acquisition for the treatment periods contributing to the analysis), analysis files are generated, and analyses conducted by the independent unblinded project statistician for submission to the DSMB.

Upon achievement of the final target sample size for an intervention (111 participants), randomization to that intervention stops. Because interventions may stop at different times in the study, and placebo period data are shared across interventions, care will be taken to observe firewalls that protect confidentiality of data and analysis results for one intervention while other interventions continue to accrue participants. If, for example, data for Intervention A are reported publicly, while the study continues (e.g., A was stopped early for futility), then information about placebo periods included in the

report for A will be limited to comparative group analysis results only. No rare events or within-group event rates that could potentially unblind treatments in the ongoing trial or motivate conjecture about the results of the ongoing interventions will be reported.

Study close-out procedures will not be implemented until all interventions are completed.

7.4 Analysis of Primary Efficacy Endpoints

7.4.1 Statistical Model

The final analysis to assess the efficacy of an intervention will be performed when accrual and follow-up for that intervention are complete and will include biomarker positive and biomarker negative participants randomized to the intervention unless the biomarker negative subgroup is dropped after the interim futility analysis, i.e., cavosonstat. In this case, the primary analysis will include only biomarker positive participants assigned to that intervention. Data from the treatment period in which a participant was randomized to the intervention plus the matched placebo period associated with that participant are included in the final analysis of the intervention. If there is no matched placebo, all placebo periods obtained on that participant are included.

Treatment effects with respect to each of the two primary endpoints (FEV_1 and ACQ-6) will be estimated using a mixed model for repeated measurements (MMRM) that appropriately takes into account the within-participant correlations between periods. The MMRM will include treatment (test vs placebo), time (visit nested within period), subgroup levels and baseline values (FEV_1 and ACQ-6), measured before the first randomization, as fixed effects. The model will also include two-way interactions (fixed effects) between treatment and time and baseline and time. Subgroup-specific treatment effects will be estimated by including subgroup by treatment interaction. Unstructured variance-covariance matrix will be used to model covariances within each treatment period. If the model does not converge with an unstructured variance-covariance matrix, the Topepitz variance-covariance matrix will be used. The model will include a subject-level random intercept to model between-period correlation. If the model fails to converge, other variance-covariance structures will be evaluated until model convergence is achieved. If the variance-covariance matrix in MMRM cannot be estimated well, we will include more data on the participants in the analysis population for that intervention.

For a given intervention and outcome let T^{+}_{k} be the test statistic for the treatment effect in biomarker positive participants and T^{All}_{k} be the test statistic in all participants computed using stage k, k = 1,2, data. The tests are from the corresponding MMRM (subscripts indicating intervention and outcome are omitted here). In the final analysis, the treatment effect of a given intervention with respect to that outcome is performed using the inverse normal combination approach by computing the p-value: $p = 1 - \Phi(\sqrt{0.45}\,\Phi^{-1}(1 - p_1) + \sqrt{0.55}\,\Phi^{-1}(1 - p_2))$. Here p_1 and p_2 are the stage 1 and stage 2 p-values. The first stage p-value, p_1, is computed based on the test statistic $\max(T^{+}_{1}, T^{All}_{1})$ using the skew normal distribution with parameters $(0,1,.32)$[6] as a reference distribution. If the biomarker negative subgroup is not dropped at the interim analysis, the second stage p-value is computed similarly to the stage 1 p-value. If the biomarker negative subgroup is dropped at the interim analysis, the second stage p-value, p_2, is computed based on the test statistic T^{+}_{2} using the standard normal distribution as a reference distribution. The weights $\sqrt{0.45}$ and $\sqrt{0.55}$ in the inverse normal formula correspond to a futility analysis performed at the 0.45 information fraction. If futility is performed at an information fraction xt, the weights $\sqrt{x}$ and $\sqrt{(1-x)}$ will be used. For interventions without an interim analysis, the final analysis for a given outcome is performed as follows. Five possible cutoffs b_1, b_2, b_3, b_4, b_5 are identified for the primary continuous biomarker that defines the biomarker positive subgroup for that intervention. The five potential subgroups are defined as participants with the biomarker values $\geq b_i$ if the treatment effect is assumed to be increasing with the biomarker values, and $\leq b_i$ otherwise. The test statistic $\max(T^{1}, T^{2}, T^{3}, T^{4}, T^{5}, T^{All})$ is compared with 2.163 for a 0.05 one-sided level and with 2.452 for a 0.025 one-sided

level. These critical values were computed based on the joint distribution of the six test statistics under the null hypothesis and assuming that the cutoffs represent the quantiles of the biomarker distribution.[7]

7.4.2 Intercurrent Events

The intercurrent events that could result in missing primary endpoint data (for FEV_1 and ACQ-6) include severe asthma exacerbations that interfere with the 16-week primary outcome assessment, major change in (biologic) therapy, and treatment or study discontinuation due to safety concerns (see Table 2). For the primary efficacy analysis, we will use two approaches to handle missing data depending on the reason for missingness:

1. Missing at random/other: primary endpoint values will be analyzed using the MMRM model described above. Non-informative reasons for missingness include major change in therapy due to change in insurance, availability of a new asthma treatment or due to an addition of a beta blocker; treatment discontinuation, study discontinuations, and loss-to-follow-up due to an administrative event or other event not related to treatment.
2. Informative: primary endpoint values will be handled by fitting the MMRM model described in Section 7.4.1. Informative reasons for missingness include worsening of asthma; major change in therapy due to worsening of asthma; and an adverse event that results in treatment discontinuation.

For intercurrent events that reflect a worsening of severe asthma (i.e., discontinuation due to exacerbations or side effects related to disease worsening), informative missingness (#2 above) will be assumed. For participants who discontinue a treatment period because of a major change in therapy, the reason for the therapy change will influence the approach for handling missing data. If a new asthma therapy becomes available, and the participant withdraws from the study to try the new therapy, the reason for missingness will be assumed to be non-informative (#1). If the reason for change in therapy is determined to be related to a decline in the participant's health status, informative censoring (#2) will be assumed. Treatment or study discontinuation due to adverse events or lab abnormalities will be assumed to be non-informative (#1).

In addition to the intercurrent events defined in Table 2, we anticipate some treatment discontinuations, study discontinuations, and loss-to-follow-up for other reasons that may result in missing primary endpoint data. Missing data resulting from these events will be assumed to be non-informative (#1).

7.4.3 Sensitivity Analyses

Sharing placebo data across interventions: Existing literature and examination of data from prior asthma studies indicate little impact on placebo response from mode of administration,[8-10] supporting the plan to include all placebo data for a participant in the efficacy analysis of an intervention the participant receives. Sensitivity analyses will be conducted on all data using the same MMRM approach as described above but with the addition of a fixed effect for mode of administration of test treatment or placebo (oral medication vs injection vs dietary supplement). A p-value for the mode of administration effect of 0.1 or less would suggest that the response to treatment is affected by the mode of administration.

Carry-over effects: Period-specific baseline values of FEV_1, ACQ-6 and the baseline levels of deterioration for CompEx will be analyzed to look for evidence of carry-over effects from one treatment period to another due to the crossover study design. If there is an indication of carry-over effects, an analysis will be performed using the primary analysis models but with the addition of effects corresponding to the treatment a participant received prior to the current treatment to control for the carry-over effect.

Intercurrent Events: Sensitivity analyses will be conducted to assess the sensitivity of the primary efficacy results to the method of missing data handling.

Temporal effects: Sensitivity analyses will be conducted to examine treatment effects estimated only from contemporaneous comparisons. Data from the first treatment period for all participants will form the basis of comparisons of each intervention versus pooled placebo. A second sensitivity analysis will estimate treatment effects based on comparisons between each intervention and pooled placebo in the first period and in the second period, and then averaged across the two periods. For both analyses, results will be compared to those from the primary analysis to better understand the impact of the crossover design on analysis results.

To assess the sensitivity of the primary analysis results to the potential for a ceiling effect in participants' response to treatment over time, an analysis will be conducted to determine whether participants enrolled in the study improve regardless of the interventions they receive. In this analysis, we will compare response to placebo in periods 1 and 2 to response to placebo in periods 3-6 with adjustment for the previous treatment received. Additionally, we will evaluate the association between the period and subject's baseline by fitting a model similar to the primary analysis model with baseline as the outcome and period and previous treatment as covariates.

We will evaluate the impact of dose received on the treatment effect by fitting the primary analysis model while adjusting for dose intensity measured is proportion of the total intended dose received.

7.5 Analysis of Secondary Endpoints

CompEx event rates will be modeled using a log-linear model for mean event count with treatment and period number and subgroup level as covariates, with log follow-up time as an offset variable, and assuming variance proportional to the mean. Subgroup-specific treatment effects will be estimated by including subgroup by treatment interaction. Unstructured variance-covariance matrix across periods will be used.

Secondary endpoints such as pre-BD FVC, post-BD FEV_1, symptom free days (defined using diary/CompEx data), and asthma free days will be analyzed using MMRM as described for the primary efficacy endpoints. Secondary endpoints that correspond to event counts, including asthma-specific ED visits, asthma-specific hospital admissions, and asthma-specific ICU admissions, will be analyzed using a log-linear model for mean event count as described for the CompEx endpoint above, with log follow-up time as an offset variable, and assuming variance proportional to the mean. Generalized estimating equations will be used to estimate regression parameters and compute a robust (sandwich) covariance matrix estimator. In addition, time to first exacerbation will be analyzed using a Cox proportional hazards model.

7.6 Safety Data Analyses

Safety data are evaluated on an ongoing basis. Monthly reports of serious adverse events (SAEs) are provided to all DSMB members, and more comprehensive reports of safety data are provided to the DSMB members for review and discussion at their quarterly meetings. These reports include incidence of treatment-emergent adverse events by assigned treatment group. Adverse events are coded using the Medical Dictionary for Regulatory Activities (MedDRA). Adverse event rates (and severity of non-serious adverse events) are compared between active and placebo periods for each intervention. Note that for a particular intervention, all placebo periods for participants receiving the intervention are included, and not just those where the matching placebo was received.

Based on the results of their review, the DSMB will recommend to the NHLBI whether to continue or discontinue each intervention for safety reasons. In addition, the DSMB monitors all safety data

throughout the course of the trial and will be notified as soon as possible of any SAE that is deemed both unexpected and probably related to treatment (serious, unexpected, suspected adverse reactions or SUSARs).

The PrecISE Medical Monitor reviews lab abnormalities on an ongoing basis that are potentially unblinding to clinical personnel at the site and advises the clinical investigator and staff if follow-up testing or other procedures are needed. Summaries of lab abnormalities are reviewed by the DSMB at their quarterly meetings.

Final safety analyses will be performed for each intervention, once the intervention completes and all follow-up data are obtained. Descriptive summaries (differences from placebo and corresponding 95% confidence intervals, tables and graphs) of safety endpoints, including adverse events, laboratory parameters, and other safety biomarkers, will be generated by test treatment versus placebo.

7.7 Interim Analysis

7.7.1 Futility Interim Analysis

When data are available from approximately 34 biomarker positive participants and approximately 17 biomarker negative participants, an interim analysis will be performed. An intervention will be stopped for futility if the futility boundary is crossed for both primary endpoints in biomarker positive participants and in the combined group of biomarker positive and negative participants. The probability to declare futility if at least one of the two endpoints has the true effect size of 0.3 (the effect size is defined as the treatment effect divided by the standard deviation of the outcome) in the biomarker positive subgroup is at most 0.16 To test for futility, we compute the test of a treatment effect for FEV_1 percent predicted, and ACQ-6. Larger values of the test statistics correspond to more favorable treatment responses. Futility is declared for an endpoint if test statistics for biomarker positives and for the combined sample of biomarker positives and negatives, are both less than 1.29. The intervention must show futility on both endpoints in both samples (biomarker positives and the combined biomarker positive and negative group) to be dropped from the study. Assuming mutual independence of the outcomes, the probability to stop for futility is 0.74 if there is no effect on any of the two endpoints.

Futility is also tested on the biomarker negative sample alone. Enrollment of biomarker negative participants to the intervention is stopped if futility is established. The cutoff for futility in this analysis is 1.08. If the biomarker negative group is stopped for futility, 77 more biomarker positive participants are enrolled. The above sample size is for interventions with a proportion of biomarker positives among all enrolled to the intervention of 67%, assuming the interim analysis is performed at the 0.45 information fraction and a 10% active treatment discontinuation rate is observed. Adjustments might be made if these parameters are different from the above.

7.7.2 Blinded Sample Size Re-Estimation

The pre-specified blinded interim analysis to estimate the crossover correlation coefficient was conducted, and changes were made to the protocol, as described in Section 1.3 above.

7.7.3 Precision Medicine Interim Analysis

Each of the interventions selected for investigation in PrecISE is hypothesized to work in a specific patient subgroup defined based on one or more predictive biomarkers, and the intervention's effect on patients outside the subgroup is suspected to be smaller. The biomarker positive subgroup is specified as a function of the primary predictive biomarker for that intervention before the first unblinded analysis (futility analysis or the final analysis for interventions where futility analysis is not feasible). A precision medicine interim analysis was originally planned to revise the subgroup definitions for any intervention not stopped for futility by evaluating alternate cut-points delineating biomarker -positive

and negative patients as well as determining whether pre-specified secondary biomarkers were useful in defining alternate biomarker positive subgroups. The precision medicine interim analysis is now planned to be performed at the end of the study, due to sample size and feasibility concerns. The precision medicine aspect of the interim analysis now planned is to determine whether biomarker negative participants (using the pre-specified definition of biomarker positive and negative participants) should continue to be assigned to the intervention (see Section 7.7.1).

7.8 Precision Medicine Final Analysis

The monotonicity assumption that the treatment effect is increasing with the value of the primary biomarker will be tested by fitting a model with treatment and treatment by biomarker interaction. If the interaction term is significant at 0.1 level, we will conclude that the treatment effect is increasing or decreasing with the biomarker value depending on the sign of the estimated interaction. If this test yields results different from our initial assumption, the multiple cutoff analysis described in Section 7.4.1 will be repeated according to the adjusted assumption of monotonicity. We will explore a possibility that the treatment effect has a U-shape as a function of a biomarker and consider subgroups defined as the biomarker $\geq b_i$ and the biomarker $\leq b_j$ for $i < j$, $i = 1,\ldots,5$, $j = 1,\ldots,5$. These analyses will be repeated when the best subgroup is defined as the one maximizing the weighted average of the test statistics for the three outcomes with equal weights. These analyses are of an exploratory nature.

To evaluate the predictive effect of the biomarker, if such an effect is monotone, we will evaluate the correlation between the treatment effect and the biomarker. For each intervention with a continuous primary predictive biomarker, we will compute ρ_A, the correlation between the patient's response to active treatment and the patient's baseline (prior to period 1) biomarker value, and ρ_P, the correlation between the patient's response to placebo and the baseline biomarker. If the participant does not have a matching placebo, we will use response data on all (one or two) placebos on the participant. If there are two placebos, we will take the average. This analysis will be repeated using the biomarker value measured right before the corresponding treatment period.

The best subgroup for an intervention is defined as the subgroup that maximizes the power of detecting a treatment effect under the global null hypothesis relative to its alternative, namely, the null hypothesis that the intervention has no effect on the two primary endpoints, FEV_1 and ACQ-6, and the key secondary endpoint, CompEx, versus the alternative hypothesis that there is an effect on at least one of these endpoints. For a single outcome, this definition is equivalent to maximizing the square root of the size of the subgroup measured by its prevalence in the relevant patient population multiplied by the average treatment effect in the subgroup. This approach to defining the best subgroup reflects not only the treatment effect in the subgroup but also the size of the subgroup. Potential target subgroups with prevalence smaller than 16% will not be considered.

The primary predictive biomarkers are: blood eosinophils for imatinib, interleukin 6 for clazakizumab, genotypes rs7669660 and rs11547772 for cavosonstat, blood eosinophils for Broncho-Vaxom, and fractional exhaled nitric oxide (FeNO) for MCT. For each of the three outcomes, FEV_1 and ACQ-6 and CompEx, a model like the primary analysis model is fit that includes fixed effects for each biomarker, two-way interactions among the biomarkers, the biomarkers' interactions with treatment, and the biomarkers' two-way interactions with treatment. The one (1) standard deviation method[10] and the overlapping group exponential Lasso penalty approach[11] with a penalty term that leads to more likely selection of subgroups defined by a smaller number of biomarkers will be used. Three models will be fitted, one for each outcome. From the fitted model for outcome Y_j, $j = 1,2,3$, a set of estimated linear predictors, $\hat{\eta}_1, \ldots, \hat{\eta}_n$, is obtained for the set of biomarker vectors through plugging in estimated model coefficients. The target subgroup is then estimated based on a weighted average of three test statistics,

one for each outcome. The best subgroup is the subgroup that maximizes the power of the hypothesis test.

The resulting subgroup definition will be reported out as the best estimate of the types of patients who should be targeted by the intervention, satisfying the second primary study objective. These results can guide the design of further studies of the interventions found to have benefit in PrecISE. We will estimate the best subgroup for each intervention using primary predictive biomarkers. Then the analysis will be repeated with primary and available secondary predictive biomarkers. The secondary predictive biomarkers include: serum tryptase, sputum polymorphonuclear neutrophils (PMNs), sputum tryptase, and Urinary PGD-2-metabolite for imatinib; high sensitivity C-reactive protein (HS-CRP) for clazakizumab; change in maximal BD FEV_1 for cavosonstat; and urine bromotyrosine for assessment of TH2 inflammation, and assessment of T cell profile in circulation for MCT. The target subgroup will be estimated using the same methodology as specified above. The treatment effect in the estimated subgroup for each intervention will be tested using a 2-fold cross-validation with bootstrap methodology. This methodology preserves the Type I error probability of the testing procedure.

7.9 Exploratory Analyses

7.9.1 Analysis of Monitoring Biomarkers

Each intervention included in the PrecISE master protocol has specified one or more monitoring biomarkers hypothesized to provide evidence of early clinical response to that intervention. Exploratory analyses will be conducted at the end of the study to investigate the utility of these biomarkers as potential early or intermediate outcomes (efficacy or safety) for future clinical trials of the interventions. To this end, models like those described above for the primary efficacy analysis will be fit with a particular intervention's monitoring biomarker as outcome to determine if treatment effects are detected. We will evaluate the correlation between each monitoring biomarker and each of the two primary outcomes.

7.9.2 Analysis of Adolescents

The efficacy of the subset of interventions available for testing in adolescents (MCT and Broncho-Vaxom) will be analyzed using the same approach as for the overall population (i.e., the MMRM and log linear models described in Section 7.4) but with the analysis population restricted to adolescents. We anticipate these analyses to be descriptive in nature, due to the lack of power for the limited number of adolescents expected to be enrolled (~40).

Safety data will also be described and analyzed separately for the adolescent population using the same tables and graphics described for the overall population.

7.9.3 Subgroup Analysis Using Proximal Biomarkers

We will repeat the precision medicine analysis described in Section 7.7.3 using primary predictive biomarkers assessed right before the start of each intervention. Using the power for the treatment comparison in the best subgroup, we will compare the predictive ability of baseline biomarkers versus more proximal biomarkers. If additional potentially predictive biomarkers emerge by the time of this final analysis, the analysis will be repeated after addition of these biomarkers to the set of primary and secondary predictive biomarkers.

7.9.4 Precision Medicine Analyses for Interventions with Overlapping Target Subgroups

Once all interventions not stopped early for futility have been completed and data are available, an exploratory precision medicine analysis to estimate individualized treatment rules will be performed for severe asthmatics with different biomarker profiles. As noted above, the primary efficacy analyses for the PrecISE study involve comparisons of each intervention to placebo, conducted independently of the other interventions, with no comparisons of one intervention to another. In an exploratory fashion,

however, interventions that were found to be significantly better than placebo and whose estimated best target subgroups overlap will be analyzed to identify the best intervention in each overlapping region. For example, suppose the target subgroups for two interventions both include patients with high IL-6 plasma levels and high blood eosinophils. In that case, exploratory analyses will be conducted to determine if either appears to provide superior benefit in that overlapping subgroup region.

Results of the study's primary analyses will include estimated decision rules for each participant identifying which interventions' target subgroups the participant belongs to. That is, a set of functions mapping the patients' biomarkers to a set of indicators on whether or not each intervention is better than placebo will be determined for each participant in the study. All biomarkers involved in defining the target subgroups will be combined with standard demographic variables to define a vector of patient tailoring variables $\boldsymbol{X}$ to ensure comparability and compatibility across successful treatments. If there are any intersections of these subgroups which are non-empty for two or more treatments, an exploratory precision medicine analysis will further refine the subgroups to assign patients to the one treatment (or multiple essentially equivalent treatments) that dominates all other treatments. Potential methods for this exploratory analysis include random forests,[12] linear regression models with quadratic terms and interactions, and outcome weighted learning models.[13,14] .Equivalence of two treatments for a given value of $\boldsymbol{X}$ will be defined as a difference in expected outcome Y for those two treatments of less than 10% of the overall standard deviation of outcome.

Formally, two treatments A_j and A_j, are defined for $j \neq j'$ to be equivalent if $\left|E\left[Y\middle|\boldsymbol{X}=\boldsymbol{x}, A=a_j\right] - E\left[Y\middle|\boldsymbol{X}=\boldsymbol{x}, A=a_{j'}\right]\right| < 0.1\hat{\sigma}_Y$, where $\hat{\sigma}_Y$ is the overall standard deviation of the outcome Y. This analysis will be done separately for each of the three outcomes. Results from this analysis will allow future treating clinicians to determine which intervention or interventions are best for each of the three outcomes for a patient with feature value $\boldsymbol{X}=\boldsymbol{x}$.

One challenge with the above approach is that a treating clinician will need to prioritize among the three outcomes (FEV_1 asthma symptoms, and loss of asthma control), for a given patient. To address this, precision medicine analysis will also be conducted combining the three outcomes. Let $\boldsymbol{Y}$ be the vector of the three outcomes, rescaled so that larger values correspond to better outcomes. Specifically, for a given outcome, if larger values indicate a better outcome, the outcome remains as is; otherwise, if larger is worse, then the negative of the outcome will be used. Let $\boldsymbol{W}$ be a vector of three weights between 0 and 1 whose total adds up to 1. For each possible value of $\boldsymbol{W}$ (over a fine grid), the outcome weighted analysis described in the previous paragraph is repeated, including placebo as a treatment. All treatments are simultaneously compared by dividing the feature space, $\boldsymbol{X}$, into regions wherein each of the treatments is optimal for the outcome $\boldsymbol{W}'\boldsymbol{Y}$, the inner product of $\boldsymbol{W}$ and $\boldsymbol{Y}$, with null regions allowed. The expected value of the resulting treatment rule is then estimated for each of the three outcomes separately using the value estimation procedure[13], and this is repeated over all distinct values of $\boldsymbol{W}$ on the grid, where *m* denotes the number of distinct values. A table is then created with three columns, one for each outcome, and *m* rows, one for each distinct value on the grid, where each entry in the table corresponds to the estimated value function for the given outcome and grid value. For each of the three columns, value function estimates are replaced by their percentiles across the m rows within each column. Now for each row, the minimum percentile across the three columns is determined, and the value of $\boldsymbol{W}$ corresponding to the row that maximizes this minimum is called $\boldsymbol{W}_*$. This is the best performing combination overall in the sense that every other value of $\boldsymbol{W}$ will yield a lower percentile for the worst performing outcome. This is the "minimax" optimal choice of $\boldsymbol{W}$, and all of the analyses of the previous paragraph are repeated, including the equivalence determinations, for the weighted outcome $\boldsymbol{Y}_* = \boldsymbol{W}_*{}'\boldsymbol{Y}$. This will provide an alternative clinical decision which combines the three outcomes.

References

1. Severe Asthma Research Program. (Accessed November 15, 2021, at http://www.severeasthma.org/.)

2. Fuhlbrigge AL, Bengtsson T, Peterson S, et al. A novel endpoint for exacerbations in asthma to accelerate clinical development: a post-hoc analysis of randomised controlled trials. Lancet Respir Med 2017;5:577-90.

3. US Food and Drug Administration. Guidance for Industry: E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials. In: (CDER) CfDEaR, (CBER) CfBEaR, eds.: U.S. Food and Drug Administration; 2021.

4. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. PharmacoEconomics 1993;4:353-65.

5. Fuhlbrigge A, Peden D, Apter AJ, et al. Asthma outcomes: exacerbations. J Allergy Clin Immunol 2012;129:S34-48.

6. Azzalini A, Capitanio A. The skew-normal and related families. Cambridge: Cambridge University Press; 2014.

7. Stallard N. Adaptive enrichment designs with a continuous biomarker. Biometrics 2023;79:9-19.

8. de Wit HM, Te Groen M, Rovers MM, Tack CJ. The placebo response of injectable GLP-1 receptor agonists vs. oral DPP-4 inhibitors and SGLT-2 inhibitors: a systematic review and meta-analysis. British journal of clinical pharmacology 2016;82:301-14.

9. Dutile S, Kaptchuk TJ, Wechsler ME. The placebo effect in asthma. Current allergy and asthma reports 2014;14:456.

10. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning : data mining, inference, and prediction : with 200 full-color illustrations. New York: Springer; 2001.

11. Breheny P. The group exponential lasso for bi-level variable selection. Biometrics 2015;71:731-40.

12. Zhu R, Zeng D, Kosorok MR. Reinforcement Learning Trees. Journal of the American Statistical Association 2015;110:1770-84.

13. Murphy SA. An experimental design for the development of adaptive treatment strategies. Statistics in medicine 2005;24:1455-81.

14. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating Individualized Treatment Rules Using Outcome Weighted Learning. Journal of the American Statistical Association 2012;107:1106-18.