Statistical Analysis Plan (SAP)

Protocol Title:	A Phase 2b Study to Evaluate the Safety and Efficacy of IMR-687 in Subjects with Sickle Cell Disease
Protocol Version No./Date:	5.0/5.1-UK/17-Feb-2022
CRF Version No./Date:	4.0/02-Jul-2021
SAP Version No./Date:	Final 1.0/01-Mar-2022

1.0 Approvals

Sponsor	
Sponsor Name:	Imara, Inc.
Representative / Title:	Constance Barysauskas, MS / Associate Director, Biostatistics
Signature / Date:	
Representative / Title:	Kenneth M. Attie, MD / Senior Vice President and Chief Medical Officer
Signature / Date:	
Everest Clinical Research	
Biostatistician / Title:	Francis Tang, M.Sc. / Associate Director, Statistical Operations
Signature / Date:	
Review Biostatistician / Title:	Christine Kollar, M.Sc. / Principal Biostatistician
Signature / Date:	

2.0 Change History

Version/Date	Change Log
Final 1.0/01-Mar-2022	Created as new

3.0 Table of Contents

4.0 Purpose

The Statistical Analysis Plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Imara Protocol IMR-SCD-301.

5.0 Scope

The Statistical Analysis Plan outlines the following:

- Study Objectives
- Study Design
- Applicable Study Definitions
- Statistical Methods

6.0 Introduction

This SAP should be read in conjunction with the study protocol and case report form (CRF). Any further changes to the protocol or CRF may necessitate updates to the SAP prior to database lock. If the SAP and the final protocol are different, the SAP prevails.

7.0 Study Objectives

7.1 Primary Efficacy Objective

- To evaluate the effect of IMR-687 versus placebo on the annualized rate of vaso-occlusive crises (VOCs)

7.2 Primary Safety Objective

- To evaluate the safety and tolerability of IMR-687 versus placebo.

7.3 Secondary Efficacy Objectives

The key secondary efficacy objective is:

- To evaluate effect of IMR-687 versus placebo on the time to the first occurrence of a vaso-occlusive crisis
- To evaluate the fetal hemoglobin (HbF) response to IMR-687 versus placebo

The other secondary efficacy objectives are:

- To evaluate the effect of IMR-687 versus placebo on other measures of vaso-occlusive crises (VOCs)
- To evaluate the effect of IMR-687 versus placebo on percentage of cells positive for HbF (% F-cells) and total hemoglobin (Hb)
- To evaluate the effect of IMR-687 versus placebo on biomarkers of red blood cell (RBC) hemolysis
- To evaluate the effect of IMR-687 versus placebo on quality of life (QoL) measures
- To evaluate the effect of IMR-687 versus placebo on biomarkers of adhesion, inflammation, and cardiac stress and on RBC indices

7.4 Secondary Pharmacokinetic Objective

The secondary pharmacokinetic (PK) objective is:

- To evaluate the pharmacokinetic (PK) exposure of IMR-687

7.5 Exploratory Efficacy Objective

The exploratory objective is to evaluate the effect of IMR-687 versus placebo on renal function

8.0 Study Design

This is a Phase 2b, randomized, double-blind, placebo-controlled, multicenter study of subjects aged 18 to 65 years with sickle cell disease (SCD; homozygous sickle hemoglobin [HbSS], sickle-β^0 [HbSβ^0] thalassemia, or sickle-β^+ [HbSβ^+] thalassemia) to evaluate the safety and efficacy of the phosphodiesterase type 9 inhibitor, IMR-687, administered once daily (QD) for 52 weeks. This study will enroll approximately 99 subjects with SCD. This study consists of a screening period (up to 4 weeks), a double-blind treatment period (52 weeks), and a safety follow-up period (4 weeks).

The study schematic is provided in Figure 1 below, and the schedule of assessments is provided in Appendix 3 of this document.

Subjects will receive either IMR-687 (lower dose or higher dose) or placebo in a blinded fashion. IMR-687 will be supplied as 100, 150, or 200 mg white tablets. Placebo and the different doses of IMR-687 are visually identical. Subjects will be advised to take 2 tablets orally, once daily (QD). Prior to the Protocol version 4.0, lower dose was defined as an exposure of ≥3.0 to ≤4.5 mg/kg, subjects in the lower dose group weighing <67 kg will be dispensed 100 mg tablets and those weighing ≥67 kg will be dispensed 150 mg tablets. Higher dose was defined as an exposure of >4.5 to ≤6.7 mg/kg, subjects in the higher dose group weighing <67 kg will be dispensed 150 mg tablets and those weighing ≥67 kg will be dispensed 200 mg tablets. In Protocol version 4.0 and later, lower dose was defined as an exposure of ≥3.4 to ≤5.0 mg/kg, subjects in the lower dose group weighing <60 kg will be dispensed 100 mg tablets and those weighing ≥60 kg will be dispensed 150 mg tablets. Higher dose was defined as an exposure of >5.0 to ≤6.7 mg/kg, subjects in the higher dose group weighing <60 kg will be dispensed 150 mg tablets and those weighing ≥60 kg will be dispensed 200 mg tablets.

Initially, subjects will be randomly assigned in a 2:1 ratio to receive either IMR-687 lower dose or placebo. Prior to the introduction of IMR-687 higher dose, the Data Monitoring Committee (DMC) will review safety data for at least 5 subjects who received IMR-687. If the DMC recommends inclusion of the higher dose, randomization will then proceed in a 1:2:1 ratio (IMR-687 lower dose, IMR-687 higher dose, or placebo). During study conduct under Protocol Version 3.0, the DMC approved opening of enrollment to the higher dose IMR-687 group, enrollment to the higher dose went into effect on 12 March 2021.

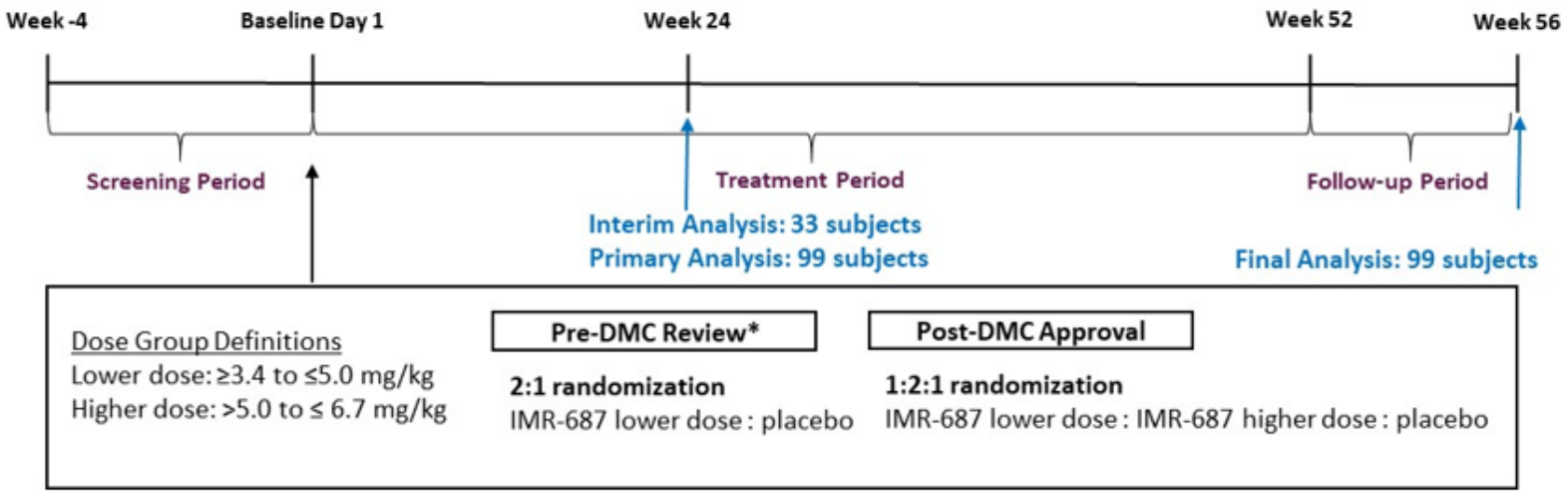
Subjects may or may not be concomitantly receiving a stable dose of hydroxyurea (HU) according to the subject's established treatment plan. Randomization will be stratified by use of HU (yes vs. no) and by region (North America/Europe vs. Africa/Middle East).

Subjects will be directed to take study drug (either IMR-687 or placebo) with food, but if a subject does not do so, it will not be considered a protocol deviation.

An interim analysis (IA) will be conducted when all subjects have completed study assessments (including VOC incidence) at Week 24 or terminated the study early. The IA will be a partial unblinding event with corresponding alpha spending strategies to manage Type 1 error.

A final analysis (FA) will be conducted at study completion (i.e., when all subjects have reached Week 56 or terminated early), following database lock and unblinding. Alpha available at the final analysis will be propagated based on the alpha spent from the IA.

Figure 1: Overview of Study Design and Dose Groups



Abbreviation: DMC = data monitoring committee.
* Prior to the introduction of IMR-687 higher dose, the DMC will review safety data for at least 5 subjects who received IMR-687. If the DMC recommends inclusion of the higher dose, randomization across 3 dose groups will then proceed as shown.
During study conduct under Protocol Version 3.0, the DMC approved the opening of enrollment in the higher dose IMR-687 group, which went into effect on 12 March 2021.

8.1 Sample Size and Power Considerations

8.1.1 Sample Size Determination

The sample size of this study was originally determined based on HbF responder rate as the primary endpoint:

Assuming a 5% HbF response rate for placebo, the study having approximately 31 placebo and 49 higher dose subjects would have > 85% power to detect a targeted difference of 30%, using a Chi-square test for comparison of proportions (with continuity-correction) and a two-sided alpha of 0.050. The assumed responder rate for placebo (5%) and highest tolerated dose subjects (>35%) includes the dropout subjects (regarded as are non-responders at Week 24).

Based on FDA feedback a new primary endpoint was adopted with a corresponding change in statistical methodology as specified in Section 8.1.2 below.

As of completion of this SAP, enrollment was completed in the ongoing study, with a total of 115 subjects randomized between August 2020 and August 2021. Out of the 115 subjects, the first 17 subjects were randomized in a 2:1 ratio to Lower Dose: Placebo, and the remaining 98 subjects proceeded with randomization in a 1:2:1 ratio for Lower Dose: Higher Dose: Placebo.

8.1.2 Power Estimation

Annualized rate of VOCs was changed from a key secondary endpoint in Protocol 4.0/4.1 (April 2021) to the primary efficacy endpoint in this SAP and in Protocol Amendment 5.0/5.1 (February 2022). The power calculation below was based on the projection that there will be approximately 49 randomized subjects in higher dose, 35 subjects in lower dose and 31 subjects in placebo group.

Assuming that the actual data distribution is normal when the significance level (alpha) of the test is 0.050, the standard deviation is 2.7 in both groups, and there are no dropouts, there will be 88% power to detect a difference between the medians of 5.0 and 3.0 for the placebo and the IMR-687 highest tolerated dose groups, respectively (i.e., a 40% difference), using a two-sided Wilcoxon Rank-Sum test (Table 1). Assuming a 25% dropout rate, the power will be 78%.

Table 1 Power Estimation

Scenarios	N for IMR-687 Highest Tolerated Dose	N for Placebo	Significance Level (two-sided alpha)	Standard Deviation	Difference	Power
No Dropout	49	31	0.05	2.7	40%	88%
With 25% Dropout Rate*	37	24	0.05	2.7	40%	78%

* estimated based on Ataga et al.

The power calculation for the Annualized rate of VOCs was carried out in PASS Version 2020 v20.0.3. See Appendix 7 for details.

8.2 Randomization

All subjects who are screened (including screen failures) will be assigned a unique subject identification number.

On Day 1, eligible subjects will be assigned another unique number (randomization number) in sequential order. The randomization number codes the subject's initial treatment assignment according to the

randomization schedule generated prior to the study. Randomization will follow a permuted block design and will be stratified by use of HU (yes vs. no) and by region (North America/Europe vs. Africa/Middle East). Initially, subjects will be randomly assigned in a 2:1 ratio using a block size of 6 to receive either IMR-687 lower dose or placebo respectively. If the DMC recommends inclusion of the higher dose, randomization will then proceed in a 1:2:1 ratio (IMR-687 lower dose, IMR-687 higher dose, or placebo) using a block size of 4.

Randomization numbers will not be re-used once assigned. If a subject is replaced, the replacement subject will receive the same treatment assignment as the replaced subject, and a different leading number will be used for the replacement randomization number.

Randomization schedules were generated by Cytel, a subcontractor of the interactive voice/web response system (IXRS) vendor Suvoda.

9.0 Study Endpoints, Variables and Covariates

9.1 Primary Efficacy Endpoint

The primary endpoint is the annualized rate of VOCs

9.2 Safety Endpoints

Safety endpoints include the number and percentage of subjects with the following:

- At least one treatment-emergent adverse events (AEs)
- At least one treatment-emergent serious adverse events (SAEs)
- At least one clinically significant change in a laboratory test parameter
- At least one clinically significant change in a vital signs parameter
- At least one clinically significant change in an electrocardiogram (ECG) parameter

9.3 Secondary Efficacy Endpoints

- Key secondary efficacy endpoint 1: Time to first VOC
- Key secondary efficacy endpoint 2: Proportion of HbF responders (defined as proportion of subjects with an absolute increase of ≥3% in HbF from baseline) at Week 24
- The study has the following additional secondary endpoints:
    - Proportion of VOC-free subjects
    - Annualized rate of VOC-related hospitalizations
    - Annualized rate of days hospitalized for VOC-related hospitalizations
    - Time to second VOC
    - Proportion of HbF responders (defined as proportion of subjects with an absolute increase of ≥3% in HbF from baseline) at Week 52
    - Proportion of Hb responders (defined as proportion of subjects with an absolute increase of ≥1.0 g/dL in total Hb from baseline) at Week 24 and Week 52
    - Change from baseline in the following at Weeks 24 and 52:
        - HbF (%), F-cells (%), total Hb (g/dL)
        - Biomarkers of red blood cell (RBC) hemolysis (% and absolute reticulocytes, unconjugated [indirect] bilirubin, and lactate dehydrogenase (LDH))

- Each measured subdomain of the Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-Me) questionnaire
- Total preference score and individual domain scores of the Patient-Reported Outcomes Measurements Information System – Preference (PROMIS 29 + 2 Profile v2.1 [PROPr]) questionnaire
- Overall score of the Sickle Cell Self-Efficacy Scale (SCSES)
- Change from baseline in biomarkers of adhesion including soluble E-selectin (E-sel), P-selectin (P-sel), intercellular adhesion molecule 1 (ICAM-1), and vascular cell adhesion molecule 1 (VCAM-1) at Week 24 and Week 52
- Biomarkers of inflammation such as high-sensitivity C reactive protein (hsCRP) and myeloperoxidase [MPO]
- Biomarkers of cardiac stress such as N-terminal prohormone of brain natriuretic peptide (NT-proBNP)
- RBC indices, such as mean corpuscular volume (MCV)

9.4 Secondary Pharmacokinetic Endpoint

PK will be assessed for IMR-687 (where feasible)

- PK profile (concentration-time measurements) and population PK of IMR-687

9.5 Exploratory Endpoints

- Renal Function
  - Change from baseline in renal function as measured by the urine protein-to-creatinine (Pr:Cr) ratio, microalbumin, and urine Cystatin Cat Week 24 and Week 52
- VOCs:
  - Number and percentage of subjects with at least 1 VOC event
  - Number and percentage of subjects by category of number of VOC events: 0, 1, 2-4, 5-7, 8-12, 13+
  - Number of and percentage of subjects with at least one VOC event lasting for ≥ 48 hours
  - Number and percentage of subjects with complicated VOCs (ACS, priapism, hepatic or splenic sequestration)
    - Number and percentage of subjects with at least 1 ACS
    - Number and percentage of males with priapism
    - Number and percentage of subjects with hepatic or splenic sequestration
  - Number of and percentage of subjects with uncomplicated VOCs
  - Time to first event of complicated VOCs
  - Annualized rate of days of VOCs (in days)
  - Annualized rate of VOCs, excluding subjects with a short period of exposure (e.g. <4 weeks)
  - Annualized rate of uncomplicated VOCs

- VOC-Related Hospitalizations:
    - Number and percentage of subjects with at least 1 VOC-related hospitalization
    - Number and percentage of subjects by category of number of VOC-related hospitalizations: 0, 1, 2-4, 5-7, 8-12, 13+
- RBC Transfusions:
    - Time to first RBC transfusion
    - Number and percentage of subjects with at least 1 RBC Transfusion
- Relationships between VOC and PD parameters
    - Correlation of annualized rate of VOCs and PD parameters (e.g., HbF, Hb and F-cells) with each other, at Week 24 and Week 52, for HU and non-HU use subgroups separately and together

10.0 Conventions and Derivations

10.1 Method for Handling Missing Data

There will be no imputation of incomplete or missing data unless specified otherwise in this SAP.

Please refer to Appendix 4 for specific data handling rules, including missing last dose date for IA, incomplete/missing dates for adverse events and prior/concomitant medications, interpretation of non-numeric results (e.g. “< BLQ”, “>x.xx”) from PK and laboratory measures, etc.

Efficacy and PK imputation for missing data will be detailed in Section 13.5.

Imputation will be noted in the footnotes. The original values/results will be presented in the data listings.

10.2 Definitions of Baseline and Change from Baseline Values

Unless otherwise specified, baseline is defined as the last non-missing measurement prior to the first dose of study drug.

All ECG assessments at a visit are to be performed in triplicate at each scheduled visit. The baseline value for ECG parameters will be derived as the average of all non-missing values, at the last visit prior to the first dose of study drug. Post-baseline assessments at a visit will be averaged for analysis.

In analyses of pharmacodynamics (PD) parameters, baseline is defined as the average of non-missing values from the screening and baseline visits. PD assessments include the following laboratory measurements based on blood samples (HbF-associated biomarkers and total Hb; biomarkers of RBC hemolysis, adhesion, inflammation, and cardiac stress; RBC indices).

Change from baseline at a post-baseline visit is defined as follows:

- For the ECG parameters: Post-Baseline Average Value – Baseline Average Value
- For PD parameters: Post-Baseline Value – Baseline Average Value
- For all other parameters: Post-Baseline Value – Baseline Value

10.3 Definitions of Study Days

Study Day 1 is defined as the date on which subjects are administered their first dose of study drug. For visits or events that occur on or after the first dose of study drug, study day is defined as (date of visit [event] – date of first dose of study drug + 1). For visits (or events) that occur prior to Study Day 1, study day is defined as (date of visit [event] – date of first dose of study drug). There is no Study Day 0.

10.4 Definition of Treatment-Emergent Adverse Event

A treatment-emergent adverse event (TEAE) is defined as an AE that emerges on or after initiation of study drug, or an AE that existed pre-treatment and worsened in severity on or after initiation of study drug, through 30 days after the last dose of study drug.

10.5 Windowing Conventions

All scheduled study visits are defined relative to Study Day 1, the date of first dose. Scheduled visit windows are defined in Appendix 3. A windowing convention will be used to determine the analysis visit value for a given measurement and will be applicable for all by-visit summaries and analyses for efficacy and safety data. Refer to Table 2 below for specific visit windows.

Table 2 Visit Windows for Efficacy and Safety Analyses

For all Efficacy (Except for QoL) and PD Analyses	
Scheduled Study Visit (Protocol Scheduled Day)	**Analysis Visit Window (Study Day)**
Baseline	≤ 1
Week 4 (Day 28 ± 5)	2 to 43
Week 8 (Day 56 ± 5)	44 to 71
Week 12 (Day 84 ± 5)	72 to 99
Week 16 (Day 112 ± 7)	100 to 141
Week 24 (Day 168 ± 7)	142 to 211
Week 36 (Day 252 ± 7)	212 to 309
Week 52 (Day 364 ± 7)	≥ 310
For QoL Analysis	
Scheduled Study Visit (Protocol Scheduled Day)	**Analysis Visit Window (Study Day)**
Baseline	≤ 1
Week 4 (Day 28 ± 5)	2 to 57
Week 12 (Day 84 ± 5)	58 to 127
Week 24 (Day 168 ± 7)	128 to 211
Week 36 (Day 252 ± 7)	212 to 309
Week 52 (Day 364 ± 7)	≥ 310
For Safety Analysis	
Scheduled Study Visit (Protocol Scheduled Day)	**Analysis Visit Window (Study Day)**
Baseline	≤ 1
Week 1 (Day 7 ± 2)	2 to 15
Week 4 (Day 28 ± 5)	16 to 43
Week 8 (Day 56 ± 5)	44 to 71
Week 12 (Day 84 ± 5)	72 to 99
Week 16 (Day 112 ± 7)	100 to 127
Week 20 (Day 140 ± 7)	128 to 155
Week 24 (Day 168 ± 7)	156 to 190
Week 30 (Day 210 ± 7)	191 to 232
Week 36 (Day 252 ± 7)	233 to 274
Week 44 (Day 308 ± 7)	275 to 337

For Safety Analysis	
Scheduled Study Visit (Protocol Scheduled Day)	**Analysis Visit Window (Study Day)**
Week 52 (Day 364 ± 7)	338 to 379
Week 56 (Day 392 ± 7)	≥ 380

Windowing will be applied prior to any missing data calculations. The last non-missing measurement taken prior to the date of first dose (including unscheduled assessments) will be labeled as “Baseline”. Unscheduled or Screening visits that occurred before the baseline visit will not be assigned an analysis visit. The early discontinuation visit will be eligible for allocation to an analysis visit.

If one or more results for a variable are assigned to the same analysis visit, the result with the date closest to the scheduled day per protocol will be used in the analysis. If 2 measurements in the same analysis visit window are equidistant from the scheduled study day, the earliest measurement will be used in the analysis. If multiple assessments are available on the same day (for the same time point), then the average of the assessment will be used in the analysis. If both central and local assessments of the same lab test are available on the same day, the central result will take precedence over the local result.

11.0 Analyses Sets

All analysis sets will include only subjects who gave informed consent and are summarized in Table 3.

Table 3 Analysis Sets and Definitions

Analysis Set	Definition	Use	How Analyzed
Safety	All subjects who have received any amount of study drug	Safety and tolerability	According to actual treatment received
PK (Analysis)	All subjects in the Safety Analysis Set who have received at least 1 dose of IMR-687 and have any measurable post-dose IMR-687 concentration-time data	PK concentrations	According to actual treatment received
PK (Evaluable)	All subjects in the Safety Analysis Set who have received at least 1 dose of IMR-687 and have at least 4 consecutive non-zero post-dose IMR-687 concentration-time data points.	PK parameters	According to actual treatment received
ITT	All randomized subjects	Efficacy/PD	According to randomized treatment assigned
mITT	All randomized subjects who had at least 1 VOC in the 12 months prior to randomization and have received at least one dose of study drug	Efficacy/PD	According to randomized treatment assigned
Per Protocol (PP)	All subjects in the Safety Analysis Set who had at least 1 VOC in the 12 months prior to randomization, completed at least 1 valid clinical outcomes assessment without major protocol deviations or events that would be expected to affect the analysis and did not discontinue prior to Week 24 for reasons other than AEs	Efficacy/PD	According to actual treatment received

11.1 Safety Analysis Set

The safety analysis set will include all subjects who have received any amount of study drug and from whom informed consent has been obtained. The safety analysis set will be used to summarize all safety and tolerability data. In safety data summaries, subjects will be analyzed according to the actual treatment they received.

11.2 PK Analysis Set

The PK Analysis Set, will be defined as a subset of the Safety Analysis Set who are randomized, have received at least 1 dose of IMR-687, and have any measurable post-dose IMR-687 concentration-time data This set will be used to generate the corresponding PK concentration summaries and plots.

11.3 PK Evaluable Set

The PK Evaluable Set will be defined as all subjects in the Safety Analysis Set who are randomized, have received at least 1 dose of IMR-687, and have at least 4 consecutive non-zero post-dose IMR-687 concentration-time data points. The PK Evaluable Set will be used to generate the corresponding PK exposure (parameter) assessments.

11.4 Intent-to-Treat Analysis Set

The intent-to-treat (ITT) analysis set will include all randomized subjects. The ITT set will be used as the primary analysis set to summarize efficacy, PD, and clinical outcome parameters. Subjects will be analyzed according to their randomized treatment assignment.

11.5 Modified Intent-to-Treat Analysis Set

The modified intent-to-treat (mITT) analysis set will include all randomized subjects who had at least 1 VOC in the 12 months prior to randomization and have received at least one dose of study drug. The mITT set will be used for sensitivity analyses of efficacy, PD, and clinical outcome parameters. Subjects will be analyzed according to their randomized treatment assignment.

11.6 Per Protocol Analysis Set

The per protocol (PP) analysis set will include all subjects in the safety analysis set who had at least 1 VOC in the 12 months prior to randomization, completed at least 1 valid clinical outcomes assessment without major protocol deviations or events that would be expected to affect the analysis and did not discontinue prior to Week 24 for reasons other than AEs. Subjects will be analyzed according to the actual treatment they received. The PP set will be used for sensitivity analyses of efficacy, PD, and clinical outcome parameters.

Major protocol deviations and subjects (or data) excluded from the PP Analysis Set will be defined by the Sponsor in a blinded manner prior to database lock.

12.0 Overview of Planned Analyses

This section summarizes the planned analyses for efficacy, PD, PK and safety for the interim analysis (IA) and final analysis (FA). The statistical analyses, which will be performed on locked and cleaned data extracts, will include all data through protocol-specified analysis time points. The data cut-off for the IA are specified below. The DMC may recommend the discontinuation of a dose arm, due to safety, which will inform the highest tolerated dose. Selected personnel involved in the analysis and review of unblinded data for the IA will abstain from attending study conduct and operational meetings after the data has been unblinded for any reason. Detailed statistical methods are defined in Section 13.0.

An IA will be performed after all subjects have reached Week 24 or terminated early. The FA will be performed after all subjects have completed the study. The critical significance levels for the IA and FA to maintain an overall two-sided type I error rate of 5% will be based on an O'Brien-Fleming boundary, calculated from a Lan-DeMets alpha spending function, based on the number of highest tolerated dose and placebo subjects used in the primary efficacy analysis in the ITT analysis set.

A summary of planned analysis for IA/FA are summarized in Table 4 below.

Table 4 Planned Analyses for IA and FA

	IA	FA
Disposition, Baseline and Disease Characteristics	Yes	Yes
Exposure	Yes	Yes
Concomitant and Prior Medications		Yes
Primary Efficacy: Annualized Rate of VOCs	Yes, with Alpha spending - fixed sequence design [a]	Yes, with propagated Alpha from IA [a]
Key Secondary Efficacy: Time to first VOC and HbF Responder rate at Week 24	Yes [b]	Yes [b]
Other Secondary Efficacy: Biomarkers and Specialty Hematology parameters		Yes
Other Secondary Efficacy (Annualized rate of VOC-related hospitalizations, etc.)		Yes
Other Secondary Efficacy: PROs (ASCQ-Me, PROPr, SCSES)		Yes
Exploratory Efficacy		Yes
Subgroup Analysis		Yes
Safety (AEs, Laboratory, Vital Signs, ECGs, etc.)	Yes, selected outputs only	Yes
PK	Yes, selected outputs for plasma concentration only	Yes

[a] Two-sided alpha spending at IA for primary efficacy analysis is 0.002. Propagated Alpha of 0.0494 is to be spent at FA.
[b] Statistical testing for key secondary efficacy will proceed only after the efficacy of the primary efficacy endpoint is demonstrated.

12.1 Hypothesis Testing

For the primary comparisons, the null hypothesis for each pairwise comparison will be that the treatment effect of the highest tolerated dose of IMR-687 is equal to that of placebo; the alternative hypothesis is then that the treatment effect of the highest tolerated dose of IMR-687 and that of placebo are not equal. P-values will thus be reported as two-sided.

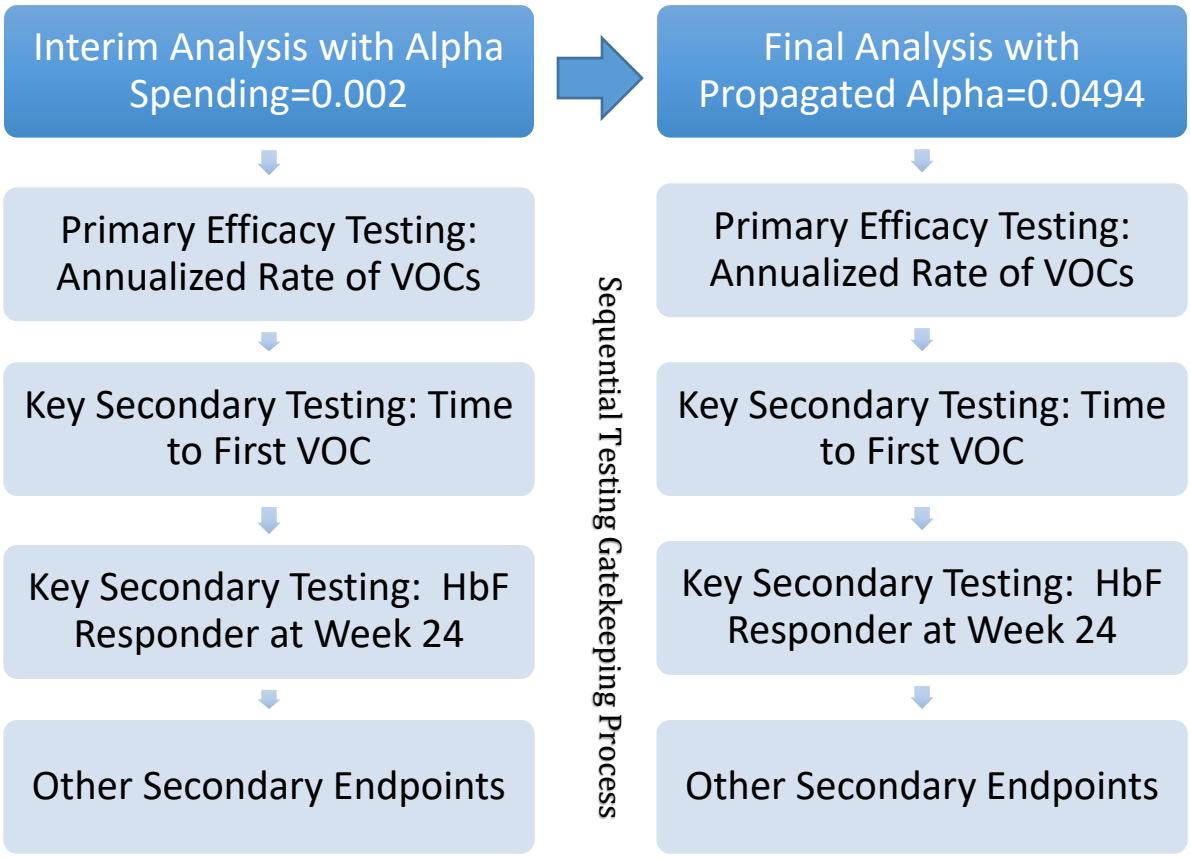
Key secondary and other secondary efficacy analyses will involve the above hypothesis applied to the secondary efficacy endpoints.

12.2 Multiple Comparisons and Control of Type I Error

An IA will be performed after all subjects have reached Week 24 or terminated early. The FA will be performed after all subjects have completed the study at Week 52. As Figure 2 shows, a fixed sequence hierarchical testing strategy will be used to control the Type I error rate across planned analyses (IA/FA) for all efficacy analyses. An O'Brien-Fleming boundary, calculated from a Lan-DeMets alpha spending function, will be used to control the alpha spending for the primary comparison at IA. The details of alpha

spending and the order of testing of the other secondary endpoints are summarized in Sections 12.3 and Section 12.4.

Figure 2: Two-sided Alpha Available for Spending in the Fixed Sequence Hierarchical Testing Procedure



12.3 Interim Analysis

The study will have one unblinded (partial) interim analysis (IA). An unblinded statistical team from Everest will perform the analyses. The IA will be based on the ITT analysis set.

To facilitate the IA, certain sponsor representatives and designees may be unblinded to individual treatment assignments prior to and during the IA (including the unblinded clinical research organization (CRO) biostatistician and external groups for bioanalytical, PK, and PK/PD analyses). Details are provided in the study unblinding plan. The blinded Sponsor and CRO study team members will be unblinded solely to the group analysis results (but not individual subject results) at the conclusion of the IA.

The IA will be performed after all subjects have reached Week 24 or terminated early. The IA will include summaries of disposition, baseline demographics and disease characteristics, safety analysis (summary of adverse events, change from baseline for laboratory [excluding specialty hematology], vital sign and 12-lead ECG parameters), as well as analyses of the efficacy endpoints.

At the IA, all the efficacy endpoints will be tested in a hierarchical order, at the highest tolerated dose of IMR-687 vs placebo. If all the efficacy comparisons are successful, testing of the IMR-687 pooled dose vs. placebo, and of lower dose vs. placebo will be performed. If the first comparison fails (i.e. the null hypothesis is not rejected), the statistical analysis and the p-values from all remaining comparisons will

not be reported at the interim analysis. The hierarchical order of testing of primary, key secondary and other secondary endpoints is:

1. Annualized rate of VOCs
2. Time to first VOC
3. HbF responder rate at Week 24
4. Proportion of VOC-free subjects
5. Annualized rate of VOC-related hospitalization
6. Annualized rate of days hospitalized for VOC-related hospitalizations
7. Time to second VOC
8. Change from baseline in key biomarkers of RBC hemolysis (indirect bilirubin and % reticulocytes) at Week 24
9. Change from baseline in HbF (%) and F-cells (%) at Week 24
10. Hb responder rates at Week 24
11. Change from baseline in total Hb (g/dL) at Week 24
12. Change from baseline in other biomarkers of RBC hemolysis (absolute reticulocytes and lactate dehydrogenase [LDH]) at Week 24
13. Change from baseline in biomarkers of adhesion such as soluble E-selectin (E-sel), P-selectin (P-sel), intercellular adhesion molecule 1 (ICAM-1), and vascular cell adhesion molecule 1 (VCAM-1) at Week 24
14. Change from baseline in biomarkers of inflammation such as high-sensitivity C-reactive protein (hsCRP) and myeloperoxidase (MPO) at Week 24
15. Change from baseline in biomarkers of cardiac stress such as N-terminal prohormone of brain natriuretic peptide (NT-proBNP) at Week 24
16. Change from baseline in RBC indices, such as mean corpuscular volume (MCV) at Week 24
17. Change from baseline in each measured subdomain of ASCQ-Me questionnaire at Week 24
18. Change from baseline in total preference score and individual domain scores of PROPr questionnaire at Week 24
19. Change from baseline in overall score of the Sickle Cell Self-Efficacy Scale (SCSES) at Week 24

The primary efficacy endpoint (annualized rate of VOCs) will be tested using a stratified Wilcoxon rank-sum test. The annualized rate of VOCs will be based on the data collected once all subjects have reached Week 24. Based on the proportion of information available at the IA (24 weeks at IA/52 weeks at FA = 0.4615), the alpha spending and O'Brien-Fleming boundaries as implemented by a Lan-DeMets alpha spending function will be as shown in Table 5 below. The lower bound of O'Brian-Fleming boundary will be used as the critical value to test the primary endpoint at the highest tolerated IMR-687 dose. If the p-value from the primary comparison is equal or below the critical value at the IA, the success of the primary efficacy analysis can be claimed. If not, the comparison will be re-done at the FA, when the information from the whole study is available, using a propagated alpha (also shown in Table 5).

Table 5 Alpha Spending Summary

Analysis	One-sided Nominal p-Value (Lower Bound of O'Brien-Fleming Boundary)	Two-sided Nominal p-Value	Lower Bound Alpha Spending
Interim Analysis	0.0010	0.0020	0.001
Final Analysis	0.0247	0.0494	0.024
Total			0.025

12.4 Final Analysis

The database will be locked prior to the final analysis (FA) which will be conducted when all randomized subjects have reached Week 56 or terminated early. As described in Section 12.3, all the efficacy endpoints will be tested in a hierarchical order, at the highest tolerated dose of IMR-687 vs placebo. If all the efficacy comparisons are successful, testing of the IMR-687 pooled dose vs. placebo, and of lower dose vs. placebo will be performed. Propagated alpha from the primary efficacy analysis at the IA will be used in the FA to maintain an overall type I error rate of 5% for 2-sided testing. After the first instance of a failure to reject the null hypothesis, p-values from all remaining comparisons will still be reported, but deemed nominal. The hierarchical order of testing of primary, key secondary and other secondary endpoints is:

1. Annualized rate of VOCs
2. Time to first VOC
3. HbF responder rate at Week 24
4. Proportion of VOC-free subjects
5. Annualized rate of VOC-related hospitalization
6. Annualized rate of days hospitalized for VOC-related hospitalizations
7. Time to second VOC
8. Change from baseline in key biomarkers of RBC hemolysis (indirect bilirubin and % reticulocytes) at Week 24 and Week 52
9. HbF responder rate at Week 52
10. Change from baseline in HbF (%) and F-cells (%) at Week 24 and Week 52
11. Hb responder rates at Week 24 and Week 52
12. Change from baseline in total Hb (g/dL) at Week 24 and Week 52
13. Change from baseline in other biomarkers of RBC hemolysis (absolute reticulocytes and lactate dehydrogenase [LDH]) at Week 24 and Week 52
14. Change from baseline in biomarkers of adhesion such as soluble E-selectin (E-sel), P-selectin (P-sel), intercellular adhesion molecule 1 (ICAM-1), and vascular cell adhesion molecule 1 (VCAM-1) at Week 24 and Week 52
15. Change from baseline in biomarkers of inflammation such as high-sensitivity C-reactive protein (hsCRP) and myeloperoxidase (MPO) at Week 24 and Week 52
16. Change from baseline in biomarkers of cardiac stress such as N-terminal prohormone of brain natriuretic peptide (NT-proBNP) at Week 24 and Week 52
17. Change from baseline in RBC indices, such as mean corpuscular volume (MCV) at Week 24 and Week 52
18. Change from baseline in each measured subdomain of ASCQ-Me questionnaire at Week 24 and Week 52
19. Change from baseline in total preference score and individual domain scores of PROPr questionnaire at Week 24 and Week 52
20. Change from baseline in overall score of the Sickle Cell Self-Efficacy Scale (SCSES) at Week 24 and Week 52

Some of the other secondary endpoints listed above include multiple endpoints at multiple time points. In terms of the hierarchical order of the testing, the endpoint listed earlier will be tested first for the earlier time points, before the endpoint listed later. For example, for point #9, it will be considered as 4 endpoints in the order of:

- Change from baseline in HbF (%) at Week 24
- Change from baseline in F-cells (%) at Week 24
- Change from baseline in HbF (%) at Week 52
- Change from baseline in F-cells (%) at Week 52

For points #18 and #19, the individual domains are in the order as listed in Section 13.5.4.

12.5 Data Monitoring Committee

The specific activities of the DMC will be governed by a charter which will define the DMC's membership, meeting frequency, procedures/conduct, and requirements for reporting its observations to the sponsor.

To ensure safety oversight throughout the trial, the DMC will review safety and preliminary efficacy data and provide recommendations to the Sponsor as described below.

The DMC will convene at the following times and for the following activities during the trial:

- Review safety data for at least 5 subjects who received IMR-687 at the lower dose and make recommendation as to the inclusion of the higher dose. This potential study adjustment makes the design adaptive.
- Review safety data to confirm acceptable safety and tolerability of the higher dose of IMR-687 and make recommendation as to whether the dose levels could be modified.
- Review safety data on a periodic basis.
- At any time during the trial—upon request by the Sponsor, medical monitor, or DMC—should a concern arise from emerging safety data for which DMC review and assessment is desired. This includes the emergence of a frequency or pattern of AEs or SAEs that suggests an unexpected or otherwise concerning safety signal. At any time, the DMC may recommend stopping a dose arm or the study for safety concerns.

The DMC may review the following subject data, depending on the scope of the meeting, and may recommend that the study continues as planned, that the protocol be modified, that a dose level should be stopped (which will inform the highest tolerated dose for analyses), or that the study terminate for safety concerns:

- Unblinded safety, PK, and preliminary efficacy/PD data, which may include TEAEs, SAEs, PK data, clinical laboratory test results, vital signs, and other relevant data for all subjects randomized.
- The DMC chairperson will receive copies of all SAE reports for ongoing review during the study. The DMC chairperson may forward an SAE report to the full DMC if he/she feels that their immediate input on, or awareness of, the SAE would be helpful.
- Additional unblinded summary reports (at group level) for IA and FA.

All assessments, decisions, and recommendations by the DMC will be documented in writing as noted in the DMC charter and prior to any resultant changes to the study unless their immediate implementation is considered necessary for subject safety.

13.0 Statistical Methods

All statistical analyses will be conducted using SAS Version 9.4 or other validated software.

Descriptive summary statistics will be provided for demographics, disposition, and treatment exposure. The number and percentage of subjects who discontinue from the study, along with reasons for discontinuation will be tabulated.

Continuous data will be summarized using descriptive statistics (number of subjects, mean, SD, median, first quartile (Q1), third quartile (Q3), minimum, and maximum, as well as geometric statistics as applicable). Counts and percentages will be used for summarizing categorical (discrete) data. For summaries of categorical variables, counts and percentages are based on the number of subjects in the analysis set unless otherwise specified.

All statistical tests and resulting p-values will be reported as 2-sided. Two-sided confidence intervals, when presented, will generally be constructed at the 95% level. P-values should be presented to 4 decimal places, with values less than 0.0001 presented as <0.0001. Multiplicity adjustment will be used for the IA and FA using the Lan-DeMets alpha spending function, as specified in Sections 12.3 and Section 12.4.

In general, means and medians will be presented to 1 or more decimal places than the recorded data. The SDs will be presented to 2 more decimal places than the recorded data. Confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate. Percentages will be presented to 1 decimal place.

A month is operationally defined to be 30.4375 days. A year is operationally defined to be 365.25 days.

If a subject was assigned the wrong stratum at randomization, their stratum based on the randomization system will be used for the statistical analyses. Analyses may be repeated using actual stratum if mis-randomization occurs in more than 10% of randomized subjects.

Where specified, there will be pair-wise comparisons between IMR-687 higher dose vs placebo, IMR-687 pooled dose vs. placebo, and IMR-687 lower dose vs placebo. Of note, by study design, subjects will be randomly assigned in a 2:1 ratio to receive either IMR-687 lower dose or placebo respectively at phase one of the trial. When the DMC recommends inclusion of the higher dose, randomization will then proceed in a 1:2:1 ratio (IMR-687 lower dose, IMR-687 higher dose, or placebo), as the second phase. The placebo and lower dose from the two phases will be analyzed together without any adjustment in the statistical analysis.

The following labels will be displayed in tabulations, unless otherwise specified:

Table 6 Treatment Groups in Tabulations

Placebo	IMR-687 200/300 mg	IMR-687 300/400 mg	IMR-687 Pooled Dose	Total

The total column will only be included in the concomitant medication summaries and in the summaries pertaining to information collected before the treatment started, such as demographic, baseline and disease characteristics and prior medications.

In general, all clinical data collected from CRF, or laboratory will be stored in SDTM datasets in STDM structure and presented in subject listings.

13.1 Subject Disposition

The number of subjects screened, randomized, and treated in the study will be presented, together with the number and percentage of subjects who discontinued study drug along with the primary reason for study drug discontinuation, discontinued from the study and the primary reason for study discontinuation.

A tabulation of the number and percentage of subjects included in each analysis set will be provided.

Details regarding randomization assignments, subjects who discontinued study drug and/or study, and subject analysis sets will be reported in by-patient listings.

A separate listing will be provided for the replacement subjects and the subjects replaced, including their demographic and stratification factors.

13.2 Important Protocol Deviations

Prior to the database lock for the primary analysis, the Sponsor or designee medical and statistics personnel will identify and review any deviations from the study protocol. Any protocol deviations that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a subject's rights, safety, or well-being will be classified as important protocol deviations (FDA, ICH, 2013, January).

Tabulation of the number and percentage of subjects with important protocol deviations by center and category within each treatment arm will be presented on the Safety Analysis Set. Categories will be sorted alphabetically, and deviations within each category will be sorted alphabetically. Categories will include:

- Deviations related to study inclusion or exclusion criteria
- Receipt of any prohibited therapies as defined in the protocol
- Significant deviations in study drug administration
- Subjects were not withdrawn after developing withdrawal criteria during the study.

Details for protocol deviations for subjects in the Safety Analysis Set will be provided in the protocol deviations data listing.

13.3 Exposure and Drug Accountability

A summary of exposure and compliance, to be reported on the safety analysis set, will include the total number of tablets taken, compliance (%), total exposure (mg), average daily dose (mg/day), duration of exposure (in weeks), and incidence and duration of dose interruption. These parameters will be calculated as follows:

- Duration of exposure (in weeks) will be calculated as [(date of last dose of study medication – date of first dose of study medication + 1)/7]. The date of first dose of study medication will be collected from the Study Drug Administration CRF page and the date of last dose of study medication will be collected from the End of Treatment CRF page. If there is no date for end of treatment on the End of Treatment CRF page, the last available dose date will be used. See Appendix 4 for a complete set of data handling rules.
- Total number of tablets taken = total number of tablets dispensed – total number of tablets returned. The total number of tablets dispensed and returned will be collected from study drug accountability CRF Page. They are to be collected at study visits and will be summed over the whole study treatment period.
- Total exposure (mg) of IMR-687 = Total number of tablets taken * tablet dose level (100/150/200 mg). IMR-687 will be supplied as 100, 150, or 200 mg white tablets in the study. In the lower dose group (≥3.0 to ≤4.5 mg/kg), subjects weighing <67 kg will be dispensed 100 mg tablets and those weighing ≥67 kg will be dispensed 150 mg tablets. In the higher dose group (4.5 to ≤6.7 mg/kg), subjects weighing <67 kg will be dispensed 150 mg tablets and those weighing ≥67 kg will be dispensed 200 mg tablets. The dose regimen for all subjects is 2 tablets once daily. The weight

at the baseline will be used to establish the dose level and the same dose level will be used throughout the trial.

- Incidence and duration of dose interruption: as collected from the IP Interruptions CRF page.
- Overall compliance (%) = [(total number of tablets dispensed – total number of tablets returned) / duration of exposure (days)] x 100.
- Average daily dose (mg/day) = total exposure (mg) / duration of therapy (days).

Additional considerations and data handling in exposure and drug accountabilities will be needed for the DMC and IA when some of drug dosing and accountability information is deemed incomplete. See Appendix 4 for a complete set of data handling rules for drug exposure and accountability.

The same tabulation will be generated for the ITT, mITT and PP analysis set if it is not the same as the safety analysis set.

Details of the administration of study drug, including randomized and actual treatment, start date and time of each dose, dose per administration, and date, time, and composition of last meal will be listed on the safety analysis set.

Details of study drug accountability including date dispensed, date returned, amount returned, and compliance per dosing period (in %, as collected directly from CRF), overall compliance (in %, for the whole duration of study drug exposure, as calculated as above) will be listed based on the safety analysis set.

13.4 Demographic and Baseline Characteristics

Demographics, baseline and disease characteristics will be summarized for all analysis sets. The following are included:

- Age (years) –as a continuous variable
- Sex (Female, Male)
- Ethnicity (Categories as collected from CRF)
- Race (Categories as collected from CRF)
- Region (North America/Europe, Africa/Middle East)
- Height (cm)
- Weight (kg)
- BMI (kg/m2)
- Diagnosis Type (HbSS, HbSB0 thalassemia, HbSB+ thalassemia)
- Baseline HU Use (as collected from IRT)
- Actual Baseline HU Use (as collected from Prior and Concomitant Medications CRF)
- Number of VOCs in the past 12 months (in categories of 1, 2-4, 5-7, 8-12 and as a continuous variable)
- The number of hospitalizations for VOCs in the past 12 months ((in categories of 1, 2, 3, ≥ 4 and as a continuous variable)
- Baseline HbF - as a continuous variable

- Baseline Hb - as a continuous variable

The median value of Age and Baseline HbF/Hb will be determined by the median in the ITT analysis set and displayed as a numeric value in tabulations.

13.4.1 Medical and Surgical history

Medical and Surgical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA, version 23.0 or higher) and will be summarized for the safety analysis set, by treatment group and overall, using System Organ Class (SOC) and MedDRA preferred term (PT). The table will include the number and percentage of subjects and will be sorted alphabetically by SOC and in decreasing frequency by PT and then alphabetically for ties (based on the total count in the analysis set). A subject will only be counted once within a class.

A by-patient listing will show medical and surgical history generated on the safety analysis set.

13.4.2 Prior and Concomitant Medications

Concomitant medications are medications ongoing at the time of the first dose of study drug or medications that started after first dose and within 30 days of the last dose of study drug. Any medication that was used at any time prior to the date of first study drug is a Prior Medication.

Medications will be captured on the Prior and Concomitant Medications CRF and coded using the WHO Drug dictionary (WHODDE 01MAR2020 or later version).

Prior medication and concomitant medications will be summarized separately by Anatomical Therapeutic Chemical (ATC) Level 2 and preferred terms for the safety analysis set. If ATC-2 is missing, the next non-missing level will be used. The summary will be sorted by ATC class alphabetically, then by descending frequency for preferred term in the pooled IMR-687 dose column and then alphabetically for ties. A patient will only be counted once within each ATC class and within each preferred name.

Medications related to sickle cell disease can be identified in the Prior and Concomitant Medications CRF when the Primary Reason Medication = Disease Under Study. Prior medication and concomitant medications related to sickle cell disease will be summarized separately, in a similar manner as for the general prior and concomitant medications.

In addition, opioids and pain medications will be identified via Standardized Drug Groupings (SDGs) and the concomitant use of opioids and the concomitant use of pain medications will be summarized in a similar manner as for concomitant medications.

All prior and concomitant medications data will be presented in a listing for the safety analysis set.

13.4.3 Concomitant Procedures

The number and percentage of subjects having concomitant procedures will be tabulated for the safety analysis set by treatment group using MedDRA (version 23.0 or later). Procedures will be classified by SOC, then by PT. Subjects will only be counted once within each SOC and PT. The SOCs and PTs will be sorted in descending frequency, using the numbers in the higher dose IMR-687 treatment arm, and then alphabetically in the event of any ties.

All concomitant procedures data will be presented in a listing based on the Safety Analysis Set.

13.5 Endpoint Analyses

13.5.1 Summary of Analyses

Primary efficacy endpoint, safety endpoints, secondary efficacy endpoints, and corresponding analysis methods are summarized in Table 7 below. Exploratory analyses are detailed in Section 13.5.5.

Table 7 Summary of Efficacy and Safety Analyses

Endpoints	Analysis Time Points	Analysis Methods	Analysis Set	Time Point for analysis
Primary Efficacy • Annualized rate of VOCs	NA	Comparison of median differences among treatment groups using Stratified Wilcoxon rank-sum tests, with stratification factors of region and HU use. The van Elteren test may be used for the evaluation of treatment difference across strata. P-values for each active group (including pooled IMR-687) vs. placebo	ITT as primary analysis set, mITT, PP as sensitivity	IA, FA
Safety	NA	Descriptive statistics	Safety	IA, FA
Secondary Efficacy				
• Time to first VOC • Time to second VOC	NA	Stratified log rank test, and stratified Cox regression model. p-values from a stratified log rank test and hazard ratios (with 95% CIs) estimated from a stratified Cox regression model with treatment arm and stratification factors of region and HU use as covariates will be provided for comparison of each active group (including pooled IMR-687) vs. placebo	ITT as primary analysis set, mITT, PP as sensitivity	IA* , FA
• Annualized rate of VOC-related hospitalizations • Annualized rate of days hospitalized for VOC-related hospitalizations	NA	Stratified Wilcoxon rank-sum tests, with stratification factors of region and HU use. P-values for each active group (including pooled IMR-687) vs. placebo	ITT as primary analysis set, mITT, PP as sensitivity	IA*, FA
• Proportion of VOC-free subjects • HbF responder rate at Week 24 and Week 52 • Hb responder rate at Week 24 and Week 52	NA	Stratified CMH test adjusted for stratification factors of region and HU use.	ITT as primary analysis set, mITT, PP as sensitivity	IA*, FA
Change from baseline at Week 24 and Week 52 • HbF (%) • F-cells (%) • Total Hb (g/dL)	Baseline to Week 24 and Baseline to Week 52	MMRM analysis with fixed effects for treatment, visit, and treatment-by-visit interaction, and baseline value, region, and HU use as covariates. Mean (95% CI), Mean difference (95% CI), and p-value will be provided (IA/FA). Both change from baseline and percentage change from baseline will be analyzed (IA/ FA).	ITT as primary analysis set, mITT, PP as sensitivity	FA
Change from baseline at Week 24 and Week 52 • Biomarkers of RBC Hemolysis (% and	Baseline to Week 24,	MMRM analysis with fixed effects for treatment, visit, and treatment-by-visit interaction, and baseline value, region, and HU use as covariates. Mean (95% CI), Mean difference (95% CI), and p-	ITT, mITT	FA

Endpoints	Analysis Time Points	Analysis Methods	Analysis Set	Time Point for analysis
absolute reticulocytes, unconjugated [indirect] bilirubin and LDH) • Biomarkers of adhesion such as Soluble E-sel, P-sel, ICAM-1, and VCAM-1 • Biomarkers of inflammation such as hsCRP and MPO • Biomarkers of cardiac stress such as NT-proBNP • RBC indices, such as MCV	Baseline to Week 52	value will be provided. Both change from baseline and percentage change from baseline will be analyzed.		
Change from baseline at Week 24 and Week 52 • Each measured subdomain of Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-Me) questionnaire • Total preference score and individual domain scores of the Patient-Reported Outcomes Measurements Information System – Preference (PROMIS 29 + 2 Profile v2.1 [PROPr]) questionnaire • Overall scores of the Sickle Cell Self-Efficacy Scale (SCSES)	Baseline to Week 24, Baseline to Week 52	ANCOVA analysis with fixed effects for treatment, and baseline value, region, and HU use as covariates. Mean (95% CI), Mean difference (95% CI), and p-value will be provided.	ITT	FA

* Key secondary and other secondary endpoints will be tested in hierarchical order through a gatekeeping procedure.

13.5.2 Primary Efficacy Analysis

The primary efficacy endpoint, annualized rate of VOCs, will be analyzed using a stratified Wilcoxon rank-sum test based on stratification factors of region (North America/Europe vs. Africa/Middle East) and HU use (Yes vs. No) to compare the treatment effect of IMR-687 to placebo using the ITT population.

The van Elteren nonparametric test may be used for the evaluation of treatment difference across strata. The homogeneity of strata results for HU use and region will be investigated prior to the use of the stratified Wilcoxon test. If the data is homogeneous, we will perform the primary analysis stratifying for both region and HU use for the primary analysis and then for HU use only as supplementary analysis at the final analysis.

13.5.2.1 Derivation of Annualized Rate of VOCs

As described in Protocol Inclusion Criterion 3, a VOC is defined as a documented episode of acute painful crisis (for which there was not an explanation other than VOC) that involves moderate to severe pain lasting for at least 2 hours and at least one of the following:

- Use of escalated analgesia (including healthcare professional-instructed use of an analgesic prescription)
- A hospital, emergency department, or clinic visit and/or healthcare telephone consultation at the time of occurrence
- Diagnosis of acute chest syndrome (ACS) (defined as an acute illness characterized by fever and/or respiratory symptoms, accompanied by a new pulmonary infiltrate on a chest X-ray), hepatic sequestration, or splenic sequestration, or priapism (in males)

VOCs collected during the study are assumed to meet these criteria as well.

VOC information, including start and end date/time, primary setting for treatment (hospitalization/ emergency department, etc.), and treatments are collected on the “VOC” CRF page.

Annualized rate of VOCs is calculated as Total number of VOCs / ([Date of last dose of study drug– date of first dose of study drug + 1]/365.25).

Date of last dose of study drug is collected from the “End of Treatment” CRF page. If that date is not available (which is expected for some of the subjects in the IA), the data cut-off date will be used.

13.5.2.2 Primary Analysis of Primary Efficacy Endpoint

For the primary efficacy analysis, the following hypotheses (H_0 and H_a) will be tested first:

H_0: The annualized rate of VOCs between IMR-687 highest tolerated dose and placebo = 0

H_a: The annualized rate of VOCs between IMR-687 highest tolerated dose and placebo ≠ 0

If the above testing is successful, the following hypotheses (H_1 and H_b) will be tested:

H_1: The annualized rate of VOCs between pooled IMR-687 and placebo = 0

H_b: The annualized rate of VOCs between IMR-687 pooled and placebo ≠ 0

If the above testing is successful, the following hypotheses (H_2 and H_c) will be tested:

H_2: The annualized rate of VOCs between IMR-687 lower dose and placebo = 0

H_c: The annualized rate of VOCs between IMR-687 lower dose and placebo ≠ 0

A stratified Wilcoxon rank-sum test based on stratification factors of region (North America/Europe vs. Africa/Middle East) and HU use (Yes vs. No) will be used to compare the treatment effect of IMR-687 to placebo in the ITT analysis set and a two-sided p-value will be provided. The homogeneity of strata results will be investigated prior to conducting the primary analysis at the end of the study.

The primary endpoint comparison will be performed at both the IA and FA. The critical values for the significance are calculated based on a Lan-DeMets alpha spending function for IA (See Section 12.2 for details), and the propagated alpha will be used for FA. If the two-sided p-value from the primary comparison is below the critical values for the significance in IA or FA (in favor of IMR-687 highest tolerated dose), the study is considered a success in terms of efficacy. If this boundary is crossed during the IA, the study may be unblinded and patients receiving placebo may be offered active treatment. See Section 12.4 for details on the alpha control and hierarchical testing procedures.

13.5.2.3 Missing Value Consideration for the Primary Efficacy Analysis

The value of the primary endpoint, annualized rate of VOCs, will be non-missing by definition. However, subjects included in the ITT analysis may not all complete the 52-week treatment period. In the context of this endpoint, the unobserved VOC events after treatment discontinuation will be considered missing values. This calculation accounts for early treatment discontinuation by extrapolating the VOC rate annualized to one year. The primary analysis assumes the unobserved rate of VOC will be the same for the observed rate for the same subject.

Missing data will not be imputed for primary endpoint analysis methods. Data imputation strategies will be employed for sensitivity analysis.

Understanding the reasons why data are missing is important to correctly interpret and handle the observed data. If the missing data mechanism is independent of missing outcome values, conditional on earlier observed outcome variables and/or covariates, the outcome values are missing at random (MAR). If the probability of missingness depends on an outside variable not in the model or is related to unobserved outcome values at the time of dropout and possibly afterward then the values are missing not at random (MNAR).

Table 8 summarizes the classification with the potential reasons. The missing data pattern will be summarized and the impact of these missing values will be explored in the sensitivity analysis below.

Table 8 Classification of Missing Data Pattern

Missing Data Pattern	Reason for Treatment Discontinuation
MNAR (Missing Not at Random)	Due to efficacy/safety related reasons: • Adverse Event • Death • Lack of Efficacy • Lost to Follow-up • Non-Compliance with Study Protocol • Non-Compliance with Study Drug • Physician Decision • Study Terminated by Sponsor • Subject Participation in Another Investigational Study • Trial Site Terminated by Sponsor • Withdrawal of Consent • Other (efficacy/safety related)
MAR (Missing at Random)	Due to non-efficacy/non-safety related reasons: • Pregnancy • Study Blind Intentionally or Accidentally Broken • Subject Erroneously Included in Study (i.e. did not meet eligibility criteria) • COVID-19 • Other (not efficacy/safety related)

13.5.2.4 Sensitivity Analyses for Primary Efficacy Analysis

A number of sensitivity analyses will be performed for the primary endpoint, annualized rate of VOCs, to examine the robustness of the results:

1. Sensitivity analysis will be conducted on the mITT and PP analysis sets

2. Sensitivity analysis using negative binomial regression (or Poisson regression as appropriate) with total number of VOCs as dependent variable, log treatment exposure as an offset, stratification factors and treatment as independent variables. The analysis will be performed on the ITT, mITT and PP analysis sets.

3. Tipping Point Analysis #1 - P-Value Frontiers

   The primary analysis method for annualized rate of VOCs is stratified Wilcoxon rank-sum test. This sensitivity analysis is to explore the sensitivity of the Wilcoxon rank-sum test on the distribution of observed data.

   In this tipping point analysis, the annualized rate of VOCs for the IMR-687 highest tolerated dose subjects and placebo subjects will be shifted separately (change by -50% to +50% at individual subject level) and the results (median difference, Interquartile range [IQR] and p-value) from stratified Wilcoxon rank-sum test for each combination of the shifts will be tabulated. P-Value Frontiers will be identified when the p-value has changed from non-significant to significant, or vice versa. The tipping point analysis will be performed on the ITT analysis sets.

4. Sensitivity analysis for Missing values due to early discontinuation

   In the primary analysis of annualized rate of VOCs, there is an underlying assumption that the unobserved rate of VOCs after study discontinuation will be the same as the observed rate for the same subject. Several sensitivity analyses will be performed to examine this assumption and its impact. The analysis below will be performed on the ITT analysis set.

   4.1 Completer, non-completer, and pooled analysis

   The primary analysis will be repeated for subjects who completed the study treatment period and who did not separately, and combined subject populations will be analyzed, to explore the differences between the annualized rates of VOCs among completers, discontinued patients, and the combined patient group.

   For the pooled analysis, a jump-to-reference (J2R) method will be applied to non-completers in order to combine them with completers. In the jumped-to-reference method, after dropout, the subject's outcomes (events of VOC) are assumed to “jump” to those of the reference group (completers in the placebo group).

   For a subject in the IMR-687 dose groups who dropout, the annualized rate of VOCs will be recalculated as:

   [Total number of observed VOCs during treatment period + annualized rate from Placebo completers * (Date of expected last study drug if not discontinued - Date of last dose of study drug + 1)]/365.25

   The annualized rate from Placebo completers will be the median (or average when median is 0) rate of all Placebo completers. Date of expected last study drug if a subject is not discontinued is calculated as First Dose Date + 364 days (52 weeks). Dropout subjects in the Placebo group will not require additional imputation.

   4.2 Tipping Point Analysis #2 – MNAR with J2R

   This tipping point analysis will be based on the J2R reference imputation, where the reference refers to the average annualized rate of VOCs among the completers in the placebo group in Section 4.1 Assuming the p-values for the primary comparison (highest tolerated dose vs placebo) are still significant with J2R (using prorogated alpha from IA), incremental changes (delta) in the referenced rate will be applied to the periods deemed to be MNAR, for the dropout subjects in the IMR-687 dose groups. If during the course of incrementing delta, the study conclusions change from favorable to unfavorable for the IMR-687 highest tolerated dose group, a tipping point is reached. The delta change will start from 0, and increment at a rate of 0.5 event/year, but granularity can be adjusted by incrementing by smaller or large amounts as needed.

4.3 Multiple Imputation Under MAR

For the discontinued subjects, alternative to the J2R method, the annualized rate of VOCs after dropout will be imputed using multiple imputation under MAR. MAR assumes the missing value is independent of unobserved outcomes given observed data (i.e., annualized rate of VOCs after dropout can be modeled based on subjects with observed annualized rate of VOCs).

The following steps will be implemented:

Step 1:

A fully conditional specification (FCS) method will be used to impute the missing annualized rate of VOCs after dropout using a regression model. The regression will include the covariates of treatment, observed number of VOCs, duration in the study, region (North America/Europe vs. Africa/Middle East) and HU use (Yes vs. No). Missing data imputation will be performed using the SAS PROC MI procedure. Refer to Appendix 6 for the order of variables in the model.

Step 2:

For each of the repetitions in the datasets, the total number of VOCs for the entire study will be recalculated for the dropouts as:

[Total number of observed VOCs during treatment period + annualized rate of VOCs after dropout * (Date of expected last study drug if not discontinued - Date of last dose of study drug + 1)]/365.25

Step 3:

The negative binomial regression (or Poisson regression as appropriate) with total number of VOCs as dependent variable, log treatment exposure as an offset, stratification factors and treatment as independent variables will be run for each of the 100 datasets to obtain 100 estimators of interest.

Step 4:

Use SAS PROC MIANALYZE to produce an overall pooled estimate (mean of 100 estimates)

The multiple imputation sample code is specified in Appendix 6. However, there could be certain adjustments due to unexpected data issues after unblinding treatment. All post-unblinding modifications to the multiple imputation model or approaches to address missing data will be described in the CSR.

13.5.2.5 Supplementary Analysis for Primary Efficacy Analyses

13.5.3 Subgroup analyses will be performed as supplementary analyses and detailed in Section 13.5.6.Key Secondary Efficacy Analysis

13.5.3.1 Time to first VOC

The time to first VOC event will be analyzed using Kaplan-Meier methods. Time in months to the first VOC event can be calculated as ([date of first VOC event – date of first dose + 1] / 30.4375).

Both start and end date of VOC event will be collected from “VOC” CRF page. The start date of the VOC is not expected to be missing and will be used for the analysis. In the rare event the start date is missing, then the end date can be used. If both start date and end date are missing, no imputation will be made and the event will be excluded from the analysis.

Subjects who experience no VOC events during the study will be censored at their study end date (or date cut-off date if the study end date is not available). Using the ITT set, the number and percentage of subjects with events (subjects who experienced at least 1 VOC on-study) and censored (subjects who did not experience any VOCs on-study) will be presented. The 25th, 50th (median), and 75th percentiles, along

with associated 2-sided 95% CIs based on the Brookmeyer and Crowley method will be presented. The p-values from a stratified log rank test and hazard ratios (and 95% CIs) estimated from a stratified Cox regression model with treatment arm and stratification factors (region and HU use) as covariates will be presented for comparison of each treatment group to placebo. Kaplan-Meier curves will be produced for the time to first VOC with separate curves for each treatment group.

Sensitivity analysis will be performed for Time to first VOC using the mITT and PP analysis sets. Major protocol deviations and subjects (or data) excluded from the PP analysis set will be defined by the Sponsor in a blinded manner prior to database lock. No imputation will be performed for the PP analysis.

13.5.3.2 Other sensitivity analyses may be added. All sensitivity analyses will be performed at FA.HbF responder rate at Week 24

13.5.3.2.1 Analysis and Imputation Methods

HbF responder rate (as defined by the proportion of subjects with an absolute increase of ≥3% from baseline) at Week 24 will be assessed using a stratified Cochran-Mantel-Haenszel (CMH) test adjusted for stratification factors of region (North America/Europe vs. Africa/Middle East) and HU use (Yes vs. No). The ITT set will be used for this analysis.

The frequency and percentage of responders and non-responders at Week 24 will be presented for each treatment group as well as ORs, RRs, 95% Cis, and p-values from the CMH test comparing active groups to placebo.

Bar plots will be provided, and the proportion of subjects with response in HbF in each treatment group, Odds ratios, 95% CI, and p values will be displayed for each treatment comparison.

The following imputation rules will be applied for the key secondary efficacy endpoint (HbF responder at Week 24):

1. Subject will be treated as a non-responder if there is an initiation of HU post randomization and prior to Week 24.
2. If HbF value is missing at Week 24, the HbF value at Week 16 will be used for the analysis. If HbF value at Week 16 is also missing, subject will be treated as a non-responder.
3. Any HbF value within 8 weeks (56 days) of a RBC transfusion will be excluded from all analyses. If the subject does not have a valid HbF value at Week 24 and/or Week 16, subject will be treated as a non-responder.

Number of subjects with imputed values for the key secondary efficacy endpoint and the reason for imputation (imputed as non-responder because of initiation of HU post randomization, imputed with Week 16 value because of Week 24 value is missing, imputed from the last post-baseline assessment prior to RBC transfusion) will also be summarized. The observed and imputed Week 24 HbF values will both be summarized for the by-visit analysis.

13.5.3.2.2 Sensitivity Analysis

Several sensitivity analyses will be implemented to explore robustness of missing value imputation methods for the analysis of HbF responder rate at Week 24:

1. For imputation rule #2 in Section 13.5.3.2.1, use HbF value from Week 12 when Week 24 and Week 16 are both missing, i.e. If HbF value is missing at Week 24, HbF value at Week 16 will be used for the analysis; If HbF value at Week 16 is also missing, HbF value at Week 12 will be used for the analysis; subject will be treated as a non-responder if HbF value is missing for Weeks 24, 16 and 12.
2. For imputation rule #2 in Section 13.5.3.2.1, If HbF value is missing at Week 24, HbF value at Week 16 will not be used for the analysis; i.e. Subject will be treated as a non-responder if HbF value is missing for Week 24.

3. The average value of HbF values at Week 24 and Week 16 will be used to determine if a subject is a responder/non-responder at Week 24. If one of them is missing, the other will be used for the analysis. If HbF value at Week 24 and Week 16 are both missing, the subject will be treated as a non-responder.

In addition, sensitivity analyses for HbF responder rate at Week 24 will be conducted using the mITT and PP analysis sets. Major protocol deviations and subjects (or data) excluded from the PP analysis set will be defined by the Sponsor in a blinded manner prior to database lock. No imputation will be performed for the PP analysis.

All sensitivity analyses above will be performed at FA.

13.5.3.3 Supplementary Analysis

Subgroup analysis will be performed for the key secondary efficacy endpoints and detailed in Section 13.5.6. Other supplementary analysis may be added.

13.5.4 Other Secondary Endpoints

Proportion of VOC-free subjects

VOC-free subjects are defined as these who did not experience a VOC after the first dose of study drug through the end of treatment. Proportion of VOC-free subjects will be analyzed using a stratified Cochran-Mantel-Haenszel (CMH) test adjusted for stratification factors of region (North America/Europe vs. Africa/Middle East) and HU use (Yes vs. No) using the ITT analysis set. Sensitivity analysis will be performed for proportion of VOC-free subjects using the PP analysis set.

Annualized rate of VOC-related hospitalizations

Annualized rate of VOC-related hospitalization is calculated as total number of VOC-related hospitalizations / ([Date of last dose of study drug– date of first dose of study drug + 1] /365.25).

VOC-related hospitalization will be determined from the “Adverse Event” CRF when the AE’s preferred term = “sickle cell anaemia with crisis” and “Did the adverse event result in initial or prolonged hospitalization for the subject?” = “Yes”. Date of last dose of study drug are collected from “End of Treatment” CRF page. If that date is not available (which is expected for some of the subjects in IA), the data cut-off date will be used.

VOC-related hospitalization-free subjects are defined as these who did not experience a VOC-related hospitalization after the first dose of study drug through the end of treatment.

Analyses of annualized rate of VOC-related hospitalizations will be done in the same way as in Section 13.5.2 as for the primary endpoint (annualized rate of VOCs).

Annualized rate of days hospitalized for VOC-related hospitalizations

Annualized rate of days hospitalized for VOC-related hospitalizations is calculated as (sum of total number of days hospitalized for VOC-related hospitalization / ([Date of last dose of study drug – date of first dose of study drug + 1] / 365.25).

Number of days hospitalized for VOC-related hospitalization is calculated as the sum of (Date of Discharge – Date of Admission + 1) for all AEs with preferred term = “sickle cell anaemia with crisis” and “Did the adverse event result in initial or prolonged hospitalization for the subject?” = “Yes”.

Date of last dose of study drug are collected from “End of Treatment” CRF page. If that date is not available (which is expected for some of the subjects in IA), the data cut-off date will be used.

Annualized rate of days hospitalized for VOC-related hospitalization will be analyzed in the same way as annualized rate of VOC-related hospitalization.

Time to first for VOC-related hospitalization

The time to first VOC-related hospitalization event will be analyzed using Kaplan-Meier methods. Time in months to the first VOC-related hospitalization event can be calculated as ([date of first VOC-related hospitalization event – date of first dose + 1] / 30.4375).

Time to second VOC

This analysis of time to second VOC event will be analyzed using Kaplan-Meier methods similar to the analysis of the time to first VOC. A patient with less than two VOCs before withdrawal or completion of the study is considered censored at the time of withdrawal or completion of the study.

A sensitivity analysis will be performed separately, excluding these subjects without any VOCs.

HbF responder rate (as defined by proportion of subjects with an absolute increase of ≥3% in HbF from baseline) at Week 52

This analysis of HbF responder at Week 52 will be similar to the HbF responder rate analysis at Week 24 (Section 13.5.3.2). Similar imputation rules will apply:

1. Subject will be treated as a non-responder if there is an initiation of HU post randomization and prior to Week 52.
2. HbF value from Week 36 (the last scheduled visit before the target analysis visit) will be used when Week 52 value is missing,
3. Any HbF value within 8 weeks (56 days) of RBC transfusion is excluded from all analyses.

Change from baseline in HbF (%) and F-cells (%) at Week 24 and Week 52

Change and percentage change from baseline in HbF (%) and F-cells (%) will be analyzed using mixed models for repeated measures (MMRM) with baseline value as a covariate, fixed effects for region (North America/Europe vs. Africa/Middle East), HU use (yes vs. no), treatment, visit, and treatment-by-visit interaction. Unstructured covariance matrices will be assumed for all models, but auto-regressive 1 (AR1) and compound symmetry (CS) may also be assessed (in that order) if the unstructured covariance matrix does not converge.

Values used at the visit and change from baseline for each parameter will be summarized for each post-baseline timepoint. The LS Mean estimates and 95% CIs will be displayed for all treatment groups. LS Mean difference estimates and 95% CIs with p-values for all IMR-687 treatment groups (relative to placebo) will be provided.

Line plots will be provided for LS mean change from baseline +/- standard error (SE) at each post-baseline visit by treatment group.

Waterfall plots will be provided for change in HbF (%) from Baseline to Week 24 and from Baseline to Week 52 by treatment group, respectively. The plot of individual subjects change from baseline will be displayed with a reference line at 3%.

If the HbF value at Week 24 and Week 52 were imputed from the previous visits in the responder analysis, the imputed value will also be used for the MMRM, summary statistics and waterfall plots. The observed value and imputed value will both be presented in the data listing.

Hb responder rate (as defined by proportion of subjects with an absolute increase of ≥ 1.0 g/dL in total Hb from baseline) at Week 24 and Week 52

This analysis of Hb responder rate at Week 24 and Week 52 will be similar to the HbF responder rate analysis at Week 24 and 52. The same imputation rules will apply.

Change from baseline in total Hb (g/dL) at Week 24 and Week 52

The analysis of total Hb (g/dL) at Week 24 and Week 52 will be similar to Change (and percentage change) from baseline in HbF (%) at Week 24 and 52.

Change from baseline in biomarkers of RBC hemolysis (% and absolute reticulocytes, unconjugated [indirect] bilirubin, and LDH) at Week 24 and Week 52

The analysis of change (and percentage change) from baseline in biomarkers of RBC hemolysis at Week 24 and Week 52 will be similar to the analysis of other change from baseline efficacy variables. The analysis will be performed for each of the following parameters, at Week 24 and Week 52 separately:

- % Reticulocytes
- Absolute Reticulocytes
- Unconjugated [indirect] Bilirubin
- Lactate dehydrogenase (LDH)

Line plots will be provided for LS mean change +/- SE.

Change from baseline in each measured subdomain of the ASCQ-Me questionnaire at Week 24 and Week 52

The Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-Me®) is a systemic, reliable, and valid method for documenting adult SCD patient-reported outcomes. The questions are grouped into 3 domains (physical, social, and emotional) encompassing 7 subdomains, of which the following 5 will be measured: emotional impact, pain impact, sleep impact, social functioning impact, and stiffness impact. Of note, pain episodes and frequency, and SCD medical history checklist will not be evaluated, but the reliability of the instrument should allow for use of only a subset of the validated domains. Each subdomain includes 5 to 9 items. Recall periods vary by subdomain and even question, with most of the periods being from 7 days to 30 days. Subdomain scores are standardized to have a mean of 50 and a standard deviation of 10.

Each of the five subdomains measured in this study consists of five questions. Responses to each question range in numeric value from 1 (Always/Very/Very much) to 5 (Never/Not at all). To calculate each subdomain score, responses to each of the 5 questions will be summed, and then converted to T-scores with standard errors (SE) based on the scales in Appendix 1. If any question within a subdomain is not answered, the subdomain score will be missing for that subject/time point. Untransformed scores range from 5 to 25, and the range of converted scores depends on the subdomain. Higher scores indicate healthier status.

The analysis of change from baseline in each measured subdomain of the ASCQ-Me questionnaire at Week 24 and Week 52 will be based on an analysis of covariance (ANCOVA) model with fixed effects for treatment, visit, and treatment-by-visit interaction, and baseline value, region, and HU use as covariates. Mean (95% CI), Mean difference (95% CI), and p-value will be provided. The T-score analysis will be performed for each of the following parameters (subdomains), at Week 24 and Week 52 separately:

- Emotional Impact
- Pain Impact
- Sleep Impact
- Social Functioning Impact
- Stiffness Impact

Line plots will be provided for LS mean change +/- SE.

Change from baseline in total preference score and individual domain scores of the PROPr questionnaire at Week 24 and Week 52

The PROMIS-29+2 Profile (PROPr) questionnaire (version 2.1) consists of seven domain scores (Physical Function, Anxiety, Depression, Fatigue, Sleep Disturbance, Ability to Participate in Social Roles

and Activities, and Pain Interference), as well as a Cognitive Function short form and a Pain Intensity question. Each domain consists of four questions. Responses to each question range in numeric value from 1 (Unable to do/Never/Not at all/Very poor) to 5 (Without any difficulty/Always/Very much/Very good). The Cognitive Function short form consists of two questions where responses range from 1 (Not at all) to 5 (Very much). To calculate each domain score and cognitive function score, responses to each of the 4 questions (or 2 questions for the cognitive function) will be summed, and then converted to T-scores with SE based on the scales in Appendix 2. If any question within a domain is not answered the domain score will be missing for that subject/time point. Untransformed scores range from 4 to 20, and the range of converted T-score depends on the domain. Higher scores indicate a healthier status.

The total preference (PROPs) score will be calculated using SAS codes provided in Appendix 1 of "The Development of a Preference-based Scoring System for PROMIS (PROPr): A Technical Report (v1.4)" by Hanmer and Dewitt (December 2017). The PROPs score is between 0 and 1 with higher scores indicating better quality of life.

The analysis of change from baseline in total preference score and individual domain T-scores of the PROPr questionnaire at Week 24 and Week 52 will be similar to the analysis of ASCQ-Me. The analysis will be performed for each of the following parameters, at Week 24 and Week 52 separately:

- Total Preference Score
- Physical Function
- Anxiety
- Depression
- Fatigue
- Sleep Disturbance
- Ability to Participate in Social Roles and Activities
- Pain Interference
- Cognitive Function
- Pain Intensity

Line plots will be provided for LS mean change +/- SE.

Change from baseline in overall score of the SCSES at Week 24 and Week 52

The Sickle Cell Self-Efficacy Scale (SCSES) is a 9-item QoL measure for adults with SCD. This measure is a self-appraisal of an SCD patient's ability to engage in daily functional activities and to manage SCD symptomatology. Response choices range from "not at all sure" to "very sure". Responses from individual items are summed to give an overall score with higher scores indicating greater self-efficacy for coping with SCD. If the response for any of the individual questions is missing, the overall score will be missing.

The analysis of change from baseline in overall score of the SCSES at Week 24 and Week 52 will be similar to the analysis of ASCQ-Me. The analysis will be performed for Week 24 and Week 52 separately.

Line plots will be provided for LS mean change +/- SE.

Change from baseline in biomarkers of adhesion at Week 24 and Week 52

The analysis of change (and percentage change) from baseline in biomarkers of adhesion at Week 24 and Week 52 will be similar to the analysis of other change from baseline efficacy variables. The analysis will be repeated for each of the following parameters, at Week 24 and Week 52 separately:

- Soluble E-sel

- P-sel
- ICAM-1
- VCAM-1

Line plots will be provided for LS mean change +/- SE.

Change from baseline in biomarkers of inflammation at Week 24 and Week 52

The analysis of change (and percentage change) from baseline in biomarkers of inflammation at Week 24 and Week 52 will be similar to the analysis of other change from baseline efficacy variables. The analysis will be performed for each of the following parameters, at Week 24 and Week 52 separately:

- hsCRP
- MPO

Line plots will be provided for LS mean change +/- SE.

Change from baseline in biomarkers of cardiac stress at Week 24 and Week 52

The analysis of change (and percentage change) from baseline in biomarkers of cardiac stress at Week 24 and Week 52 will be similar to the analysis of other change from baseline efficacy variables. The following parameter will be analyzed at Week 24 and Week 52 separately:

- NT-proBNP

Line plots will be provided for LS mean change +/- SE.

Change from baseline in RBC indices at Week 24 and Week 52

The analysis of change (and percentage change) from baseline in RBC indices at Week 24 and Week 52 will be similar to the analysis of other change from baseline efficacy variables. The following parameter will be analyzed at Week 24 and Week 52 separately:

- MCV

Line plots will be provided for LS mean change +/- SE.

13.5.5 Exploratory Efficacy Endpoints

The analysis of change (and percentage change) from baseline in renal function at Week 24 and Week 52 will be similar to the analysis of other change from baseline efficacy variables. Change from baseline will be calculated for each of the following renal function parameters at Week 24 and Week 52 separately:

- Urine protein-to-creatinine (Pr:Cr) ratio
    - Urine Microalbumin
    - Urine ACR (albumin to creatinine ratio)
- Urine Cystatin C

Line plots will be provided for LS mean change +/- SE.

Additional exploratory analysis (which are not specified in protocol) will be performed if the data allow, as shown in, but not limited to, the list below. In general, the exploratory endpoints will be analyzed following the same analysis approach specified for similar endpoints specified in primary and secondary analysis sections at FA, unless otherwise specified.

VOC Analyses:

- Number and percentage of subjects with at least 1 VOC event

- Number and percentage of subjects by category of number of VOC events: 0, 1, 2-4, 5-7, 8-12, 13+
- Number of and percentage of subjects with at least one VOC event lasting for ≥ 48 hours
- Number and percentage of subjects with complicated VOCs (ACS, priapism, hepatic or splenic sequestration)
    - Number and percentage of subjects with at least 1 ACS
    - Number and percentage of males with priapism
    - Number and percentage of subjects with hepatic or splenic sequestration
- Number of and percentage of subjects with uncomplicated VOCs
- Time to first event of complicated VOCs
- Annualized rate of days of VOCs (days)
- Annualized VOC rate, excluding subjects with a short period of exposure (< 4 weeks)
- Annualized rate of uncomplicated VOCs

VOC-Related Hospitalizations:

- Number and percentage of subjects with at least 1 VOC-related hospitalization
- Number and percentage of subjects by category of number of VOC-related hospitalizations: 0, 1, 2-4, 5-7, >7.

RBC Transfusions:

- Time to first RBC transfusion
- Number and percentage of subjects with at least 1 RBC Transfusion

Relationships between biomarkers

- Pairwise correlation (Pearson correlation coefficients and p-values) of annualized rate of VOCs and PD parameters (e.g., HbF, Total Hb, F-cells) will be tabulated for Week 24 and Week 52, and for HU and non-HU use separately and together. Corresponding scatter plots will also be provided.

13.5.6 Subgroup Analyses

The following subgroups will be assessed based on the primary, key secondary and PD endpoints (HbF, Hb and F-cells). All subgroup analyses will be performed for each subgroup level. Summary tables and a forest plot will display descriptive statistics for each level of a subgroup for a population.

- Age (≤ median and > median of the total population)
- Sex (Female vs Male)
- Region (North America/Europe vs Africa/Middle East)
- Baseline HU use (Yes vs No)
- Baseline number of VOCs (1, 2-4, 5-12)
- Baseline number of VOCs (1, 2-12) [VOC analysis only]

- Baseline HbF (≤ median and > median of the total population)
- Baseline Hb (≤ median and > median of the total population)

Baseline HU use are based on the randomization stratum used in IRT. The actual baseline HU use from Prior and Concomitant Medications CRF page may be used to perform additional subgroup analysis, if they are not same.

Summary tables and forest plots will be produced to show the treatment difference in the overall population and also within each subgroup. The statistics shown will depend on the type of analysis:

1. Relative risk for the annualized rate of VOCs
2. Log-rank and Cox Proportional hazards from the time to event
3. Percentage of responders from responder analysis (HbF and Hb)
4. LS Mean from MMRM analysis (HbF, Hb, F-cells)

Statistical models will be used to test if the difference between treatments differs between each level of a subgroup (i.e., to test the treatment-by-subgroup interaction).

It should be noted that the study was not designed to detect treatment differences with high statistical power within subgroups. If any subgroup level includes < 5% of all subjects, no inferential statistics will be generated.

13.5.7 Multiplicity Adjustment

A fixed-sequence statistical strategy tests endpoints in a predefined order, all at the same significance level alpha (e.g., α = 0.05), moving to a second endpoint only after a success on the previous endpoint. Such a test procedure does not inflate the Type I error rate as long as there is (1) prospective specification of the testing sequence and (2) no further testing once the sequence breaks, that is, further testing stops as soon as there is a failure of an endpoint in the sequence to show significance at level alpha (e.g., α = 0.05).

Fixed-sequence statistical strategy for multiplicity adjustment will be used in this study. Endpoints will be tested using α = 0.05. The fixed sequence for endpoint testing will follow the order prescribed in section 12.4.

13.5.8 Safety Analyses

In general, safety analyses will be based on the safety analysis set and will be provided by treatment group as defined in Section 11.1. The “Total” column will not be included in the summary tabulations.

Visit windowing (as specified in Section 10.5) will be used to determine the analysis visit for laboratory, vital sign and ECG by-visit analyses.

13.5.8.1 Adverse Events

MedDRA Coding

AEs will be coded using MedDRA (v23.0 or higher).

Treatment-emergent adverse event (TEAE)

A treatment-emergent adverse event (TEAE) is defined as an AE that emerges on or after initiation of study drug, or an AE that existed pre-treatment and worsened in severity on or after initiation of study drug, through 30 days after the last dose of study drug. A pre-treatment adverse event is an AE emerges before the initiation of study drug.

Refer to Appendix 4 for handling of partial dates for AEs for the purpose of identifying treatment-emergent AEs. In the case where it is not possible to define an AE as treatment-emergent or not, the AE will be classified as treatment-emergent.

Severity

Severity is classified as Grade 1: Mild, Grade 2: Moderate, Grade 3: Severe, Grade 4: Life-Threatening, Grade 5: Death Related to AE, using the Common Terminology Criteria for Adverse Events, v5.0 (CTCAE v5.0). For the AE summary by CTCAE severity grade, an AE with missing CTCAE severity grade will not be imputed.

Relationship to Study Treatment

Relationship is classified as “Unlikely/Not related”, “Possible”, “Probably/Likely”, or “Certain/Related” by the Investigator. A “treatment-related TEAE” for the purpose of this summary is defined as a TEAE with relationship to study drug of “Possible”, “Probably/Likely” or “Certain/Related”. For the AE summary by relationship, an AE with a missing relationship to study drug will not be imputed. Imputed values will not be included in data listings.

TEAEs Leading to Discontinuation of Study Treatment

TEAEs leading to discontinuation of study treatment will be identified by action taken with study treatment being recorded as “Drug withdrawn” on the Adverse events CRF.

Serious and Non-Serious TEAEs

Serious adverse events (SAEs) are those events recorded as “Yes” for the question “Was the adverse event serious?” on the Adverse Event CRF.

TEAEs Leading to Death

TEAEs leading to death are those events which are recorded as “Yes” for the question “Did the adverse event result in death?”, or recorded CTCAE Toxicity Grade 5, or with outcome = “Fatal” on the Adverse Events CRF.

TEAEs Leading to Study Discontinuation

TEAEs leading to study discontinuation are those events which are recorded as “Yes” for the question “Did the adverse event cause the subject to be discontinued from the study?” on the Adverse events CRF.

TEAEs Leading to Dose Reduction

TEAEs leading to dose reduction will be identified by action taken with study treatment being recorded as “Dose Reduced” on the Adverse events CRF.

Overall Summary

An overall summary of TEAEs will be presented by treatment group for the safety analysis set. The number and percentage of subjects reporting any of the following categories will be provided:

- Any pre-treatment adverse event
- Any TEAE
- Grade 3 or higher TEAEs
- Treatment-related TEAEs
- Treatment-related grade 3 or higher TEAEs
- Treatment-related grade 2 or higher TEAEs
- SAEs
- Treatment-related SAEs
- Treatment-related grade 3 or higher SAEs
- TEAEs leading to death

- TEAEs leading to discontinuation of study drug
- Treatment-related TEAEs leading to discontinuation of study drug
- TEAEs leading to study discontinuation
- Treatment-related TEAEs leading to study discontinuation
- TEAEs leading to dose reduction
- Treatment-related TEAEs leading to dose reduction

The following AE summaries will be by treatment group, SOC and PT for the safety analysis set. SOCs will be sorted alphabetically and PTs will be sorted by descending frequency in the pooled IMR-687 treatment group column, and then alphabetically for ties.

- The number and percentage of subjects with any TEAE. The AE count will be for subjects not events, and subjects are only counted once within each SOC or PT
- The number and percentage of subjects with Grade 3 or higher TEAEs
- The number and percentage of subjects with Treatment-related TEAEs. Subjects with multiple events within a particular SOC or PT will be counted under the category of their most drug-related event within that SOC or PT (related or not related). AEs will be considered related to study drug if they are reported as certain/related, probably/likely related, or possibly related.
- The number and percentage of subjects with Treatment-related TEAEs
- The number and percentage of subjects with Treatment-related Grade 3 or higher TEAEs
- The number and percentage of subjects with Treatment-related Grade 2 or higher TEAEs
- The number and percentage of subjects with SAEs
- The number and percentage of subjects with Treatment-related SAEs
- The number and percentage of subjects with Treatment-related Grade 3 or higher SAEs
- The number and percentage of subjects with TEAEs leading to death
- The number and percentage of subjects with TEAEs leading to discontinuation of study drug
- The number and percentage of subjects with Treatment-related TEAEs leading to discontinuation of study drug
- The number and percentage of subjects with TEAEs leading to study discontinuation
- The number and percentage of subjects with Treatment-related TEAEs leading to study discontinuation
- The number and percentage of subjects with TEAEs leading to dose reduction
- The number and percentage of subjects with Treatment-related TEAEs leading to dose reduction
- The number and percentage of subjects with TEAEs by severity. Subjects with multiple events within a particular SOC or PT will be counted under the category of their most severe event within that SOC or PT. AE severity (intensity) will be graded using NCI CTCAE, version 5.0.

- The number and percentage of subjects with TEAEs by relationship. Subjects with multiple events within a particular SOC or PT will be counted under the category of their most related event within that SOC or PT.

The following AE tables will be summarized by treatment group (without a "Total" column) and PT for the safety analysis set. PTs will be sorted by descending frequency, by pooled IMR-687 treatment group column, and then alphabetically for ties.

- The number and percentage of subjects reporting each TEAE.
- The number and percentage of subjects reporting common (≥10%) TEAEs (i.e., those occurring in at least 10% of subjects in any IMR-687 treatment group)
- The number and percentage of subjects reporting treatment-related TEAEs
- The number and percentage of subjects reporting treatment-related common (≥10%) TEAEs

All AEs (including non-treatment-emergent events) recorded on the CRF, all SAEs, and all deaths that occur following the first dose of study drug will be listed. All AE tables and listings will be reported using the safety analysis set.

13.5.8.2 Laboratory Data

Clinical laboratory tests (serum chemistry, hematology, urinalysis, and specialty hematology) will be summarized using the safety analysis set. Results will be evaluated and presented using the International System of Units (SI units) unless otherwise stated. Clinical laboratory values will be graded from Grade 1 to Grade 4 using NCI CTCAE, version 5.0. Post-baseline laboratory values will only be compared to baseline values from the same source. For example, if the baseline value for a parameter came from a local laboratory, post-baseline results from the same local laboratory should be used for the analysis.

Hematology assessments include: RBC count (mean corpuscular volume [MCV], mean corpuscular Hb, and mean corpuscular Hb concentration), hematocrit, Hb, platelets, WBC count (with differential: basophils, eosinophils, neutrophils, monocytes, and lymphocytes), erythrocyte count, and reticulocyte count and reticulocyte percentage. Some of these assessments are also considered PD endpoints, and further detailed in the secondary and exploratory endpoint sections above.

Specialty hematology assessments include: HbF and percent F-cells.

Coagulation assessments include: PT, aPTT, and international normalized ratio (INR).

Serum chemistry assessments include: alanine aminotransferase (ALT), albumin, alkaline phosphatase (ALP), aspartate aminotransferase (AST), bicarbonate, blood urea nitrogen (BUN), chloride, calcium, creatinine, glucose, lactate dehydrogenase (LDH), sodium, potassium, magnesium, phosphate, bilirubin (total, direct, and indirect), creatine kinase, total protein, glomerular filtration rate (GFR), and gamma-glutamyl transferase (GGT). Follicle stimulating hormone (FSH) will be included as well, but this is only collected for post-menopausal women at screening. Some of these assessments are also considered PD endpoints, and further detailed in the secondary and exploratory endpoint sections above.

Urinalysis assessments include: appearance, color, pH, specific gravity, ketones, protein, glucose, bilirubin, and urobilinogen, including occult blood and microscopic examination of sediment (only if occult blood is detected).

The laboratory results and change from baseline will be summarized by visit by treatment group for central laboratory continuous hematology, serum chemistry and urinalysis parameters separately. Note that character urinalysis tests will only be listed.

Shift tables representing the shift from the baseline grade to maximum NCI-CTCAE Grades (Grade I to IV and Missing) will be provided for selected hematology and chemistry parameters by visit by treatment group. See Appendix 5 for details.

Potentially clinically significant (PCS) Criteria for hepatic function and renal function are shown in Table 9 below. The number and percentage of subjects meeting these PCS criteria will be summarized for hepatic function parameters (ALT, AST, ALP, and total bilirubin) and the renal function parameter (serum creatinine) separately. Any subject with any post-baseline result (including "unscheduled" visits) meeting the criteria will be included in the tabulation.

Table 9 PCS Criteria for Hepatic Function and Renal Function

Hepatic Function	Renal Function
○ Post-baseline Alanine Aminotransferase (ALT) ≥ 3 x upper limit of normal (ULN) ○ Post-baseline Aspartate Aminotransferase (AST) ≥ 3 x ULN ○ Post-baseline Direct Bilirubin (BILDIR) ≥ 2 x ULN ○ Post-baseline ALT/AST and BILDIR (ALT≥ 3 x ULN or AST≥ 3x ULN) and BILDIR≥ 2 x ULN (Hy's Law)	○ Post-baseline Creatinine Clearance (CREATCLR) < 0.5 x baseline ○ Post-baseline Serum Creatinine (CREAT) > 2 x baseline

Creatinine clearance are not directly collected and will be calculated using the Crockroft Gault formula:

$$CrCl = \frac{(140 - patient's\ age\ in\ years) \times Body\ weight\ in\ kg}{72 \times serum\ creatinine\ in\ mg/dl}$$

All laboratory parameters (serum chemistry, central laboratory hematology, local laboratory hematology, specialty hematology, coagulation, and urinalysis) will be displayed in individual subject data listings.

13.5.8.3 Vital Signs

Descriptive statistics (N, mean, SD, median, Q1, Q3, minimum, and maximum) will be used to summarize vital sign parameters (body temperature, systolic blood pressure, diastolic blood pressure, heart rate, and respiratory rate) at baseline, each post-baseline visit, and change from baseline to each post-baseline visit, by treatment group. Study baseline as defined in Section 10.2 will be used.

Potentially clinically significant (PCS) post-baseline vital signs can be identified using the criteria shown in Table 10 below. The number and percentage of subjects with any post-baseline treatment-emergent potentially clinically significant vital sign value will be tabulated by treatment group for each criterion. Any subject with a post-baseline result meeting the criteria, including those collected at unscheduled visits, will be counted.

Table 10 PCS Criteria for Vital Signs

Parameter	Criteria
Heart Rate	• ≤ 50 bpm and decrease ≥ 15 bpm • ≥ 100 bpm and increase ≥ 15 bpm • ≥ 120 bpm and increase ≥ 15 bpm
SBP	• ≤ 90 and decrease ≥ 20 mm Hg • < 140 mmHg and Increase ≥ 20 mmHg • ≥ 140 mmHg and Increase ≥ 20 mmHg • ≥ 160 mmHg and increase ≥ 20 mmHg

DBP	• < 100 mmHg and Increase ≥ 20 mmHg • ≥ 100 mmHg and Increase ≥ 20 mmHg • ≤50 mmHg and decrease ≥15 mmHg • ≥100 mmHg and increase ≥15 mmHg

A plot will be presented to show the pattern of the SBP and DBP values over time by treatment group. Mean and SE will be presented in the plot. Additionally, spaghetti plots for SBP and DBP values over time will be provided for individual subjects with 1) maximum post-baseline SBP increased ≥ 20 mmHg and SBP ≥ 150 mmHg and 2) maximum post-baseline DBP increased ≥ 20 mmHg and DBP ≥ 100 mmHg. A shift plot displaying shifts from Baseline heart rate ≤ 100 to maximum heart rate >100 post-Baseline will be provided for the safety analysis set. All vital signs will be presented in a data listing. All vital sign tables and listings will be reported on the Safety Analysis Set.

13.5.8.4 12-lead ECGs

The average of triplicates from 12-lead ECGs will be used to evaluate the change from baseline in ECG parameters (heart rate, PR interval, RR interval, QRS duration, QT interval, and QTcF interval). Descriptive statistics (N, mean, SD, median, Q1, Q3, minimum, and maximum) of ECG parameters will be presented for baseline, each post-baseline visit, and change from baseline to each post-baseline visit by treatment group. Study baseline defined in Section 10.2 will be used. At Baseline, Week 1 and Week 4, descriptive statistics will also be presented for change from pre-dose to post-dose. All measurements, including those during unscheduled visits, will be provided in data listings.

The number and percentage of subjects whose QTcF values have met each of the criteria below, will be summarized by treatment group for baseline and post-baseline assessment. Unscheduled visit results, or any single result from the triplicates meeting criteria criterion will be included.

Table 11 PCS Criteria for ECG

Parameter	Criteria
QTcF Interval	• > 450 msec for males • > 470 msec for females • > 500 msec for males and females
QTcF Interval Increase from Baseline	• ≥ 30 msec • ≥ 60 msec

All ECG tables and listings will be reported using the safety analysis set.

13.5.8.5 Physical Examinations

All physical examination findings will be listed in by-subject listings on the Safety Analysis Set. No tabulation summaries will be provided.

13.5.9 Pharmacokinetic Endpoints

Data from subjects who experience emesis within the PK sampling duration of IMR-687 on Study Day 1 or Week 4 visit, will be evaluated on a case-by-case basis for exclusion from concentration-time data and PK analysis descriptive summaries. The individual subject data will be listed.

If a quantifiable pre-dose concentration of IMR-687 is detected on Study Day 1 and this concentration is greater than 5% of the corresponding Cmax, all subject data will be excluded from all concentration-time and PK analysis descriptive summaries but will be included in the subject data listings.

Any protocol deviations or AEs which could impact PK parameter estimates will be evaluated in consultation with the Sponsor for exclusion from concentration-time data and PK analysis descriptive summaries. All data that are excluded in this manner will be listed and the reason for exclusion noted.

All plasma concentration-time data will be reported to the same number of significant figures as displayed in the bioanalytical data. Lambdaz will be reported up to 4 decimal places and time-related parameters (T1/2, Tmax, and Tlast) will be reported up to 3 decimal places. All other PK parameters will be reported to the same number of significant figures as displayed in the bioanalytical data.

The concentration-time data from subjects with sparse PK sampling collected using the sampling schedule reflected in protocol version 4.0 (08 Apr 2021) will be summarized descriptively and included in subject listings and analyzed for population PK. Noncompartmental analysis will not be performed with the PK data collected from these subjects. The concentration-time data from subjects with PK sampling collected using the sampling schedule reflected in the prior protocol versions (up to version 3.0; 29 Jan 2020) will be summarized descriptively and listed. The PK data from these latter subjects will be used for the estimation of PK parameters using noncompartmental analysis. All available PK data from this study will also be used to explore any relationship between IMR-687 exposure and clinical response, PD endpoints, or AEs, as data permit. These data will be analyzed together with PK data from other clinical studies for a population PK analysis, as appropriate. Population PK analysis and other exploratory PK/PD analyses are not within the scope of this SAP and will be detailed in a separate population PK analysis plan.

No analysis visit windows will be applied for reporting of plasma concentration-time data and PK analysis, nominal visit times will be used (i.e., Study Day 1 or Week 4).

13.5.9.1 Plasma Concentrations

Plasma concentrations of below the limit of quantification (BLQ) will be set to 0.01 ng/mL in the computation of mean concentration values. Descriptive statistics (number of subjects, mean, geometric mean, SD, coefficient of variation (%CV), geometric %CV, median, minimum, and maximum) will be used to summarize the plasma concentrations by dose and by study visit (i.e., Study Day 1 and Week 4 Visit) at each scheduled timepoint.

Linear and semi-logarithmic plots of the arithmetic mean (± SD) and geometric mean (± SD) plasma concentrations by scheduled sampling time will be provided by dose and by study visit. These plots will show time in hours. The plots will present all calculated means and will include a reference line for the lower limit of quantification (LLOQ).

Linear and semi-logarithmic plots of the individual plasma concentration by scheduled sampling time will be provided by subject (one subject per page) and by study visit. These plots will show time in hours. Individual plots will use the BLQ handling procedure described below for “Plasma Pharmacokinetic Parameters”.

All individual subject plasma concentration data will be listed.

13.5.9.2 Plasma Pharmacokinetic Parameters

Plasma PK parameters for IMR-687 will be estimated using noncompartmental methods with Phoenix® WinNonlin® Version 8.1 (Certara, USA). The PK parameters will be estimated from the concentration-time profiles, and AUCs will be calculated using a linear up/log down method using actual sampling times.

The apparent terminal elimination half-life, $T_{1/2}$, where determinable, will be calculated as the natural log of 2 divided by the terminal phase rate constant, $lambda_z$. The number of data points included in the regression will be determined by visual inspection, but a minimum of 3 data points in the terminal phase, excluding C_{max}, is required to estimate $lambda_z$. In order for $lambda_z$ and $lambda_z$ derived parameters (AUC_{inf}, $T_{1/2}$, V_z/F, CL/F (single dose), and V_z/F to be reported, the adjusted r^2 value reported in Phoenix® WinNonlin® must be ≥0.80.

For the purposes of estimating the PK parameters, BLQ values will be set to zero prior to the first measurable concentration and set to missing thereafter, including embedded BLQs. Scheduled sampling times will be used for the interim analyses and actual sampling times will be used for the final analysis.

Missing concentration-time data will be treated as missing except for the pre-dose plasma concentrations where a value of 0 will be imputed for Study Day 1 for the purposes of PK parameter estimation. Similarly for the Week 4 visit, missing concentration-time data will be treated as missing except for the pre-dose concentration following steady state dosing where the lowest concentration value over the dosing interval (C_{min}) will be used for imputation purposes. In addition, subjects must have a 24-hour PK sample collected at Study Day 1 and Week 4 in order to be included in the evaluation of accumulation ratio (RAUC). No extrapolation of AUC_{0-24} for any missing 24-hour PK sample will be performed.

For the final analysis, if the actual time is missing, the scheduled sampling time will be substituted and flagged. Subjects must have at least 4 consecutive non-zero post-dose plasma concentrations in order for their data to be considered evaluable. Profiles that do not meet this criterion will be flagged for exclusion.

Descriptive statistics (number of subjects, mean, geometric mean, SD, %CV, geometric %CV, median, minimum, and maximum) will be used to summarize the PK parameters by dose, timepoint, and by study visit. Additional PK evaluations may be performed as needed.

C_{trough} concentrations for Study Day 1 and Week 4 will be summarized in tables along with C_{trough} concentrations at Week 24 and Week 52.

All individual subject PK parameters will be listed by subject ID and treatment group.

The following PK parameters for IMR-687 may be computed, where feasible. Additional PK parameters may be determined:

Table 12 PK Parameters for IMR-687

Parameter	Description
Study Day 1 (Single Dose)	
C_{max}	Maximum plasma concentration post-dose. Observed peak analyte concentration obtained directly from the experimental data without interpolation, expressed in concentration units.
T_{max}	Time to maximum plasma concentration post-dose. First observed time to reach peak analyte concentration obtained directly from the experimental data without interpolation, expressed in time units.
T_{last}	Time of last measurable plasma concentration.
C_{trough}	Concentration at the end of the dosing interval.
AUC_{last}	Area under the concentration-time curve from time 0 to the last measurable concentration post-dose.
AUC_{inf}	Area under the concentration-time curve from time 0 to infinity, calculated as AUC_{last} + (C_{last}/$Lambda_z$).
$Lambda_z$	Apparent terminal elimination phase rate constant.
$T_{1/2}$	Apparent terminal elimination half-life.
CL/F	Apparent total body clearance from oral administration, calculated as Dose/AUC_{inf}.
V_z/F	Apparent volume of distribution from oral administration, calculated as Dose/$lambda_z$ × $AUC_{inf.}$
Week 4 (Steady State)	
$C_{max,ss}$	Maximum plasma concentration post-dose at steady state. Observed peak analyte concentration obtained directly from the experimental data without interpolation, expressed in concentration units.

Parameter	Description
$T_{max,ss}$	Time to maximum plasma concentration post-dose at steady state. First observed time to reach peak analyte concentration obtained directly from the experimental data without interpolation, expressed in time units.
T_{last}	Time of last measurable plasma concentration.
C_{trough}	Concentration at the end of the dosing interval.
AUC_{0-24}	Area under the concentration-time curve from time 0 to 24 hours post-dose.
$Lambda_z$	Apparent terminal elimination phase rate constant.
CL_{ss}/F	Apparent total body clearance from oral administration at steady state, calculated as $Dose/AUC_{0-24.}$
V_z/F	Apparent volume of distribution from oral administration at steady state, calculated as $Dose/lambda_z \times AUC_{0-24}$
R_{AUC}	Accumulation ratio, calculated as AUC_{0-24} (Week 4)/ AUC_{last} (Study Day 1; AUC_{last} t = 24 hours).

C_{last} = last observed measurable plasma concentration

14.0 References

Ataga et al N Engl J Med 2017;376:429-39. Crizanlizumab for the Prevention of Pain Crises in Sickle Cell Disease.

Edwards R, Telfair J, Cecil H, Lenoci J. Reliability and validity of a self-efficacy instrument

specific to SCD. Behav Res Ther. 2000 Sep;38(9):951-63.

Hanmer J, Dewitt B, Yu L, Tsevat J, Roberts M, Revicki D, et al. Cross-sectional validation of

the PROMIS-Preference Scoring system. PLoS One. 2018 Jul;13(7):e0201093.

Hanmer, Dewitt, The Development of a Preference-based Scoring System for PROMIS (PROPr): A Technical Report (v1.4). December 2017.

Keller SD, Yang M, Treadwell MJ, Werner EM, Hassell KL. Patient reports of health outcome for adults living with sickle cell disease: development and testing of the ASCQ-Me item banks.

Health Qual Life Outcomes. 2014 Aug;12:125. San Keller, Manshu Yang, Christian Evensen, Tamika Cowans, American Institutes for Research. ASCQ-ME User's Manual. December 2017.

Sickle Cell Disease Self-Efficacy Scale Data Collection Worksheet. https://www.phenxtoolkit.org/protocols/view/820401.

U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for Industry, E3 Structure and Content of Clinical Study Reports, Questions and Answers (R1).

15.0 Glossary of Abbreviations

Glossary of Abbreviations:	
ACS	Acute Chest Syndrome
AE	Adverse event
ANCOVA	Analysis of Covariance
ASCQ-Me	Adult Sickle Cell Quality of Life Measurement Information System
AR1	Auto-regressive-1 covariance matrix
ATC	Anatomic Therapeutic Classification
AUC_{0-24}	Area Under the Concentration-Time Curve from Time 0 to 24 Hours Post-dose
AUC_{inf}	Area Under the Concentration-Time Curve from Time 0 to Infinity
AUC_{last}	Area Under the Concentration-Time curve from Time 0 to the Last Measurable Concentration Post-dose
BLQ	Below the Limit of Quantification
CI	Confidence Interval
C_{last}	Last Observed Measurable Plasma Concentration
CL/F	Apparent Total Body Clearance from Oral Administration after Single Dose
CL_{ss}/F	Apparent Total Body Clearance from Oral Administration at Steady State
C_{max}	Maximum Plasma Concentration Post-dose after Single Dose
$C_{max,ss}$	Maximum Plasma Concentration Post-dose at Steady State
C_{min}	Minimum Plasma Concentration over the Dosing Interval
C_{trough}	Concentration at the End of the Dosing Interval
CMH	Cochran-Mantel-Haenszel
CRF	Case Report Form
CS	Compound symmetric covariance matrix
CSR	Clinical Study Report
DMC	Data Monitoring Committee
ECG	Electrocardiogram
Hb	Hemoglobin
HbF	Fetal Hemoglobin
HU	Hydroxyurea
IA	Interim Analysis
$Lambda_z$	Apparent Terminal Elimination Phase Rate Constant
mITT	Modified intent-to-treat
IRT	Interactive Response Technology
IXRS/IVRS	Interactive Voice Response System

Glossary of Abbreviations:	
J2R	Jump to Reference
LLOQ	Lower Limit of Quantification
LSMean	Least squares mean
MAR	Missing at Random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
MNAR	Missing not at Random
PCS	Potentially Clinically Significant
PD	Pharmacodynamic(s)
PK	Pharmacokinetic(s)
PP	Per Protocol
PROMIS	Patient-Reported Outcomes Measurements Information System
PT	Preferred Term
QD	Once Daily
QoL	Quality of Life
R_{AUC}	Accumulation Ratio using AUC_{0-24} data
RBC	Red Blood Cells
SAP	Statistical Analysis Plan
SAE	Serious Adverse Event
SCD	Sickle Cell Disease
SCSES	Sickle Cell Self-Efficacy Scale
SD	Standard Deviation
SE	Standard Error
SI	International System of Units
SOC	System Organ Class
$T_{1/2}$	Apparent Terminal Elimination Half-life
T_{last}	Time of Last Measurable Plasma Concentration
T_{max}	Time to Maximum Plasma Concentration after Single Dose
$T_{max,ss}$	Time to Maximum Plasma Concentration at Steady State
TEAE	Treatment-Emergent Adverse Event
VOC	Vaso-Occlusive Event
V_z/F	Apparent Total Body Clearance from Oral Administration after Single Dose or Steady State
WBC	White Blood Cells

16.0 Appendices

Appendix 1 ACSQ-Me Scoring Tables

Emotional Impact Short Form Conversion Table		
Raw Score	T-Score	SE[a]
5	26.8	4.5
6	30.8	3.5
7	33.3	3.1
8	35.3	2.9
9	37.0	2.8
10	38.5	2.7
11	39.9	2.6
12	41.2	2.6
13	42.5	2.6
14	43.7	2.6
15	44.9	2.6
16	46.2	2.7
17	47.4	2.7
18	48.7	2.8
19	50.1	2.8
20	51.5	3.0
21	53.3	3.3
22	55.2	3.6
23	57.3	3.8
24	60.5	4.4
25	65.6	5.8

[a]SE = Standard Error for T-Score

Social Functioning Impact Short Form Conversion Table		
Raw Score	T-Score	SE[a]
5	26.0	4.3
6	29.8	3.2
7	32.5	2.8
8	34.7	2.8
9	36.8	2.7
10	38.7	2.7
11	40.4	2.7
12	42.1	2.7
13	43.9	2.6
14	45.6	2.6
15	47.2	2.6
16	48.8	2.6
17	50.5	2.6
18	52.2	2.5
19	54.0	2.5
20	55.8	2.5
21	57.7	2.5
22	59.8	2.6
23	62.1	2.7
24	64.9	3.1
25	69.8	4.6

[a]SE = Standard Error for T-Score

Pain Short Form Conversion Table		
Raw Score	T-Score	SE[a]
5	24.8	3.9
6	28.8	2.5
7	31.0	2.2
8	33.0	2.2
9	34.9	2.2
10	36.7	2.2
11	38.3	2.2
12	39.9	2.1
13	41.5	2.1
14	43.0	2.1
15	44.4	2.1
16	45.7	2.1
17	47.1	2.1
18	48.5	2.0
19	49.9	2.0
20	51.2	2.0
21	52.5	2.0
22	54.0	2.1
23	55.8	2.3
24	58.0	2.8
25	63.8	5.2

[a]SE = Standard Error for T-Score

Stiffness Short Form Conversion Table		
Raw Score	T-Score	SE[a]
5	24.9	4.0
6	29.0	2.8
7	31.5	2.5
8	33.5	2.4
9	35.3	2.4
10	36.9	2.3
11	38.4	2.3
12	39.9	2.3
13	41.3	2.3
14	42.7	2.3
15	44.0	2.3
16	45.4	2.3
17	46.7	2.3
18	48.1	2.3
19	49.5	2.3
20	51.0	2.5
21	52.7	2.7
22	54.7	2.9
23	57.0	3.3
24	59.9	3.8
25	65.4	5.4

[a]SE = Standard Error for T-Score

Sleep Short Form Conversion Table		
Raw Score	T-Score	SE[a]
5	27.9	4.4
6	32.3	3.1
7	35.1	2.7
8	37.3	2.6
9	39.5	2.6
10	41.4	2.6
11	43.2	2.6
12	45.0	2.6
13	46.7	2.5
14	48.2	2.5
15	49.7	2.4
16	51.1	2.4
17	52.5	2.4
18	53.9	2.4
19	55.3	2.4
20	56.7	2.4
21	58.2	2.5
22	59.9	2.7
23	61.9	3.0
24	64.4	3.4
25	69.1	4.8

[a]SE = Standard Error for T-Score

Appendix 2 PROMIS-29+2 Profile v2.1 (PROPr) Scoring Tables

Adult v2.0 - Physical Function 4a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
4	22.5	4.0
5	26.6	2.8
6	28.9	2.5
7	30.5	2.4
8	31.9	2.3
9	33.2	2.3
10	34.4	2.3
11	35.6	2.3
12	36.7	2.3
13	37.9	2.3
14	39.2	2.4
15	40.5	2.4
16	41.9	2.5
17	43.5	2.6
18	45.5	2.8
19	48.3	3.3
20	57.0	6.6
*SE = Standard Error on T-score metric		

Adult v1.0 - Anxiety 4a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
4	40.3	6.1
5	48.0	3.6
6	51.2	3.1
7	53.7	2.8
8	55.8	2.7
9	57.7	2.6
10	59.5	2.6
11	61.4	2.6
12	63.4	2.6
13	65.3	2.7
14	67.3	2.7
15	69.3	2.7
16	71.2	2.7
17	73.3	2.7
18	75.4	2.7
19	77.9	2.9
20	81.6	3.7
*SE = Standard Error on T-score metric		

Adult v1.0 - Depression 4a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
4	41.0	6.2
5	49.0	3.2
6	51.8	2.7
7	53.9	2.4
8	55.7	2.3
9	57.3	2.3
10	58.9	2.3
11	60.5	2.3
12	62.2	2.3
13	63.9	2.3
14	65.7	2.3
15	67.5	2.3
16	69.4	2.3
17	71.2	2.4
18	73.3	2.4
19	75.7	2.6
20	79.4	3.6
*SE = Standard Error on T-score metric		

Adult v1.0 - Fatigue 4a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
4	33.7	4.9
5	39.7	3.1
6	43.1	2.7
7	46.0	2.6
8	48.6	2.5
9	51.0	2.5
10	53.1	2.4
11	55.1	2.4
12	57.0	2.3
13	58.8	2.3
14	60.7	2.3
15	62.7	2.4
16	64.6	2.4
17	66.7	2.4
18	69.0	2.5
19	71.6	2.7
20	75.8	3.9
*SE = Standard Error on T-score metric		

Adult v1.0 - Sleep Disturbance 4a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
4	32.0	5.2
5	37.5	4.0
6	41.1	3.7
7	43.8	3.5
8	46.2	3.5
9	48.4	3.4
10	50.5	3.4
11	52.4	3.4
12	54.3	3.4
13	56.1	3.4
14	57.9	3.3
15	59.8	3.3
16	61.7	3.3
17	63.8	3.4
18	66.0	3.4
19	68.8	3.7
20	73.3	4.6
*SE = Standard Error on T-score metric		

Adult v1.0 – Ability to Participate in Social Roles and Activities 4a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
4	27.5	4.1
5	31.8	2.5
6	34.0	2.3
7	35.7	2.2
8	37.3	2.1
9	38.8	2.2
10	40.5	2.3
11	42.3	2.3
12	44.2	2.3
13	46.2	2.3
14	48.1	2.2
15	50.0	2.2
16	51.9	2.2
17	53.7	2.3
18	55.8	2.3
19	58.3	2.7
20	64.2	5.1
*SE = Standard Error on T-score metric		

Adult v1.0 - Pain Interference 4a		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
4	41.6	6.1
5	49.6	2.5
6	52.0	2.0
7	53.9	1.9
8	55.6	1.9
9	57.1	1.9
10	58.5	1.8
11	59.9	1.8
12	61.2	1.8
13	62.5	1.8
14	63.8	1.8
15	65.2	1.8
16	66.6	1.8
17	68.0	1.8
18	69.7	1.9
19	71.6	2.1
20	75.6	3.7
*SE = Standard Error on T-score metric		

Cognitive Function Short Form v2.0 - Abilities 2a (part of PROMIS-29+2 Profile)		
Raw Summed Score	**T-score**	**SE***
2	29.5	6.4
3	34.4	5.9
4	38	5.7
5	41.2	5.7
6	44.3	5.8
7	47.3	5.8
8	50.5	5.7
9	54.7	5.9
10	61.2	6.9
*Standard Error on T-score metric		

Note: No conversion to T-score for Pain Intensity (0 -10).

Appendix 3 Schedule of Assessments

	Screening Period	Baseline	Treatment Period													Follow-up Period
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Study Week	-4 to 0	NA	1	2	4	6	8	12	16	20	24	30	36	44	ET/EOT 52	EOS 56
Study Day	-28 to -1	1[a]	7 ± 2	14 ± 3	28 ± 5	42 ± 5	56 ± 5	84 ± 5	112 ±7	140 ± 7	168 ± 7	210 ± 7	252 ± 7	308 ± 7	364 ± 7	392 ± 7
Administrative Procedures																
Informed consent	X															
Demographic information	X															
Medical/disease history	X															
Inclusion/exclusion criteria[b]	X	X														
Randomization		X														
Telephone visits[c]				X		X										
Safety Assessments																
Vital signs[d]	X	X	X		X		X	X	X	X	X	X	X	X	X	X
Weight	X	X			X			X			X		X		X	X
Height	X														X	
Physical examination[e]	X	X	X		X		X	X	X	X	X	X	X	X	X	X
12-lead ECG[f]	X	X	X		X		X	X	X	X	X	X	X	X	X	X
Hematology and serum chemistry	X	X	X		X		X	X	X	X	X	X	X	X	X	X
Coagulation studies	X							X			X				X	
Urinalysis with urine creatinine, Pr:Cr ratio, and microalbumin	X	X	X		X		X	X	X	X	X	X	X	X	X	

	Screening Period	Baseline	Treatment Period													Follow-up Period
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Study Week	-4 to 0	NA	1	2	4	6	8	12	16	20	24	30	36	44	ET/EOT 52	EOS 56
Study Day	-28 to -1	1[a]	7 ± 2	14 ± 3	28 ± 5	42 ± 5	56 ± 5	84 ± 5	112 ±7	140 ± 7	168 ± 7	210 ± 7	252 ± 7	308 ± 7	364 ± 7	392 ± 7
Serum pregnancy test[g]	X															
Urine pregnancy test[g]		X	X		X		X	X	X	X	X	X	X	X	X	X
Clinically indicated virology[h]	X															
Adverse events[i]	Continuous															
Concomitant medications[i]	Continuous															
Efficacy and Pharmacodynamic Assessments																
Specialty hematology[j]	X	X			X		X	X	X		X		X		X	
PD biomarkers[k]	X	X			X		X	X	X		X		X		X	
QoL tools[l]		X			X			X			X		X		X	
CV marker	X	X						X			X		X		X	
Pharmacokinetic Assessments																
IMR-687 plasma PK sampling[m]		X			X						X				X	
Study Drug Procedures																
Study drug dispensing		X			X		X	X	X	X	X[n]	X[n]	X[n]	X[n]		
Study drug administration[o]		Oral administration IMR-687 or matching placebo qd														

Abbreviations: AE = adverse event; ASCQ-Me = Adult Sickle Cell Quality of Life Measurement Information System; BP = blood pressure; CV = cardiovascular; ECG = electrocardiogram; EOS = end of study; EOT = end of treatment; ET = early termination; F-cells = cells positive for HbF; HbF = fetal hemoglobin; HBsAg = hepatitis B surface antigen; HCV = hepatitis C virus; HR = heart rate; HU = hydroxyurea; ICAM-1 = intercellular adhesion molecule 1; IgM = immunoglobulin M;

MPO = myeloperoxidase; NA = not applicable; PD = pharmacodynamic; PE = physical examination; PK = pharmacokinetics; Pr:Cr = protein-to-creatinine (ratio); PROPr = Patient-Reported Outcomes Measurements Information System – Preference; qd = once daily; QoL = quality of life; SCSES = Sickle Cell Self-Efficacy Scale; TEAE = treatment-emergent adverse event; VCAM-1 = vascular cell adhesion molecule 1.

Note: Unscheduled visits may occur, limited to visits resulting from potential TEAEs, drug dispensation, or other urgent study-related procedures.

[a] Day 1 assessments should be performed prior to study drug administration.

[b] All inclusion and exclusion criteria will be assessed at the Screening Visit; continued eligibility based on these criteria will be confirmed on Day 1.

[c] Qualified site personnel will contact the subject at Week 2 and Week 6 to capture potential TEAEs and concomitant medications. Subjects will also be reminded of compliance with drug and the next visit schedule. If any AEs of significant clinical concern are identified during the telephonic visit, the subject will be requested to come into the site to be assessed.

[d] Vital signs include HR, respiratory rate, BP, and body temperature and should be consistently measured in either the sitting or semi-supine position. At the Day 1 and Week 4 visits, vital signs will be taken pre-dose and 2 hours (±20 minutes) post-dose, during the PK assessments. At all other timepoints, vital signs can be taken irrespective of taking study drug.

[e] Complete PEs will be performed at Screening and at Weeks 12, 24, 36, 52, and 56; these consist of a general examination of the body, including the abdomen, heart, lungs, lymph nodes, back/neck, neurological system, skin, extremities, head, eyes, nose, and throat. At all other visits, symptom-directed PEs will be obtained after identification of AEs deemed by the investigator to be of significant clinical concern.

[f] All ECGs to be performed in triplicate. At the Baseline, Week 1, and Week 4 visits, ECGs will be obtained pre-dose and 2 hours (±30 minutes) post-dose. At all other timepoints, ECG will be taken pre-dose.

[g] Females of childbearing potential only. A serum pregnancy test will be performed during screening via a central laboratory. All subsequent pregnancy tests will be urine pregnancy tests performed locally (with test kits provided by the central laboratory). If a urine pregnancy test is positive, the result must be confirmed with a serum pregnancy test.

[h] Serum virology (screening assessment) will be conducted only if clinically indicated. Testing will be performed through a central laboratory and may include HBsAg, hepatitis A IgM, and HCV antibody.

[i] AEs and concomitant medications, including opioids and HU use, will be recorded at each visit throughout the study from screening through the EOS (Week 56) follow-up.

[j] Includes HbF and % F-cells. Blood samples should be obtained prior to administration of study drug.

[k] PD biomarkers that are not already assessed as part of standard hematology or serum chemistry testing may include, but are not limited to, the following: soluble E-selectin, P-selectin, ICAM-1, VCAM-1, and MPO.

[l] QoL assessments include the specified components of the ASCQ-Me®, PROPr, SCSES. The translated QoL assessments from PROPr will be used when available.

[m] At the Day 1 and Week 4 visits, serial blood samples for IMR-687 (including metabolites and HU, if applicable) plasma concentrations will be drawn pre-dose (within 30 minutes) and at 15 minutes (±5 minutes), 30 minutes (±5 minutes), and 3 hours (±20 minutes) after administration of study drug. A trough blood sample will be drawn pre-dose at Week 24 and on the last day of dosing (Week 52). The date/time of the last administered dose will be recorded. Patients dosed prior to protocol version 4.0 (08 Apr 2021) used the following PK sampling schedule: At the Day 1 and Week 4 visits serial blood samples for IMR-687 plasma concentrations were drawn pre-dose (within 30 minutes) and at 30 minutes (±5 minutes), 1.5 hours (±15 minutes), 6 hours (±1 hour), and 24 hours (±2 hours) after administration of study drug. On these full profile PK days, food details were also recorded at the study sites. A trough blood sample was drawn pre-dose at Week 24 and on the last day of dosing (Week 52).

[n] Each study drug bottle will contain 84 tablets. The indicated visits will require 2 bottles to be dispensed.

[o] On days when study drug is taken in the clinic, food details will also be recorded.

Appendix 4 Data Handling Rules

Category	Description	Data Handling Rule
1. Age (years)	Age (years)	Age is collected from Demographics CRF page.
2. First Dose Date/Time of Study Drug	date/time of first dose of study drug	The date and time (24 hr. clock) of the first dose of study drug will be taken from the Study Drug Administration CRF.
3. Last Dose Date of Study Drug	date of last dose of study drug	The date of the last dose of study drug will be the Date of last dose from the End of Treatment CRF. If it is missing for the IA, and DMC analyses, use the data cutoff date. If it is missing at the database lock, the date of the last dose of study drug will be imputed as the later of last visit date, last dispensing date (from Drug Accountability CRF), date of Completion/Discontinuation, and death date.
4. Last Visit Date	Date of Last Visit	Date of last visit according to the Clinic Visit and Telephone CRF.
5. Last Study Participation Date (STDM variable, typically named RFPENDTC)	Last Study Participation Date (STDM variable, RFPENDTC), where SDTM denotes Study Data Tabulation Model	Last study participation date is defined as last known date of contact which would be the later of the following dates: last visit date, date of the last dose, date of Completion/Discontinuation, date of clinical study follow-up, or death date.
6. Study Day Definitions	**Study Day** for an assessment/event which occurred on or after the start of study drug	Study Day = Date of assessment/event – date of the first dose of study drug + 1.
	Study Day for an assessment/event which occurred on a day prior to the first dose of study drug in the study	Study Day = date of assessment/event – first dose date of study drug in the study.
	Study Day of Randomization	Study Day of Randomization = date of randomization – date of the first dose of study drug in the study + 1. Study Day is 1 if baseline day is on the day of randomization.
	First Dose Day	**First Dose Day** in the study is defined as the study day of the first dose of study drug in the study (Study Day 1).
	Last Dose Day	**Last Dose Day** in the study is defined as the study day of the last dose of study drug in the study.
7. Duration of an event	The duration of any event	The duration of an event is defined as (stop date – start date + 1).
8. Multiple assessments for the same visit	Vital Sign, ECG and Laboratory assessments	● All ECG assessments are to be performed in triplicate. Valid assessments will be averaged to get the analysis value for that visit.

Category	Description	Data Handling Rule
		• Other than for ECG assessments, the last non-missing measurement taken prior to the date/time of first study drug (including unscheduled assessments) will be the analysis baseline. • Windowing will be applied for by-visit analyses. Unscheduled or Screening visits that occurred before the Baseline visit will not be assigned an analysis visit. The early discontinuation visit will be eligible for allocation to an analysis visit. • If one or more results for a variable are assigned to the same analysis visit, the result with the date closest to the protocol scheduled day will be used in the analysis. If 2 measurements in the same analysis visit window are equidistant from the protocol scheduled study day, the earliest measurement will be used in the analysis. • If multiple assessments are available on the same day (or for the same time point for ECG assessments), then the average of the assessments will be used in the analysis. • If both central and local assessments of the same lab test are available on the same day, the central result will take precedence over the local result. • All data will be listed in data listings.
9. Special Lab Value Handling	Lab values with a prefix such as: '>', '<', '+' and 'Less than' Etc.	• '>': use the available original value +0.001 in the analyses. • '<': use the available original value –0.001 in the analyses. • '+': use the available original value without the prefix in the analyses. • '≥': use the available original value in the analyses. • '≤': use the available original value in the analyses.
10. Prior and concomitant, medication / treatment	Prior, concomitant, and post-treatment medication/treatment	• Any medication that was used at any time prior to the date of first study drug is a Prior Medication. • Any medication that was used at any time on or after the date of first study drug and within 30 days of the last dose of study drug is a Concomitant Medication. • If the date of first study medication is missing, then the date of randomization is used (as long as the subject is known to have received study

Category	Description	Data Handling Rule
		drug on at least one day during the study). Otherwise, in the case of missing dates, any period (Prior and Concomitant) that is possible based on the available information should be applied.
11. Adverse event	Missing CTCAE severity (toxicity grade)	An AE with missing CTCAE severity grade will not be imputed.
	Missing relationship to study drug	An AE with a missing relationship to study drug will not be imputed.
	Treatment-emergent adverse event	An adverse event is considered treatment-emergent if emerges on or after initiation of study drug, or an AE that existed pre-treatment and worsened in severity on or after initiation of study drug, through 30 days after the last dose of study drug. An adverse event that begins on the same date as the first dose of study drug is treatment-emergent if the AE begins after the time of first dose or if the time of AE onset is unknown.
		If the AE start date is partial/missing, then ● If AE start date is completely missing, then the AE is considered as treatment-emergent. ● If both AE start month and day are missing and AE start year is the same or after the first dose year, then the AE is considered as treatment-emergent. ● If AE start day is missing and AE start year and month are the same or after the first dose year and month, then the AE is considered as treatment-emergent. Missing/incomplete (partial) AE start and end dates will not be imputed for data listings.
12. Treatment Duration (days)	Treatment Duration (days)	Treatment duration is defined as Date of last dose of a study drug - Date of first dose of a study drug +1.
13. Duration of Exposure (days)	Duration of Exposure (days)	Exposure is defined as (Date of the last dose of the study drug – Date of first dose of the study drug + 1).
14. General	Conversion of Duration in Days to Duration in Years/Months	Duration in Years = Duration in Days/365.25 Duration in Month = Duration in Days/30.4375
15. Hard coding	Hard coding for data analysis	Hard Coding is not allowed during data analysis unless if agreed in writing by Imara.

Appendix 5 CTCAE Grade Criteria for Selected Hematology and Chemistry parameters

Investigations				
	Grade			
Laboratory Analyte	1	2	3	4
aPTT (activated partial thromboplastin time) prolonged	>ULN - 1.5 x ULN	>1.5 – 2.5 x ULN	>2.5 x ULN	-
Alanine aminotransferase increased	>ULN – 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 – 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 – 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Alkaline phosphatase increased	>ULN – 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 – 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 – 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Aspartate aminotransferase increased	>ULN – 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 – 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 – 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Blood bilirubin increased	>ULN – 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 – 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 – 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal
Creatinine increased	>ULN -1.5 x ULN	>1.5 – 3.0 x baseline; >1.5 -3.0 x ULN	>3.0 baseline; >3.0 – 6.0 x ULN	>6.0 x ULN

Investigations				
	Grade			
Laboratory Analyte	1	2	3	4
GGT (gamma-glutamyl transferase) increased	>ULN – 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 – 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 – 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Blood bicarbonate decreased	<LLN and no intervention initiated			
Blood lactate dehydrogenase increased	>ULN			
Hemoglobin increased (a)	Increase in >0 – 2 gm/dL above ULN (a)	Increase in >2 – 4 gm/dL above ULN (a)	Increase in >4 gm/dL above ULN (a)	n/a
Anemia (hemoglobin decreased)	LLN- 10g/dL; <LLN – 6.2 mmol/L; <LLN – 100 g/L	<10.0 – 8.0 g/dL; <6.2 – 4.9 mmol/L; <100 – 80g/L	<8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated
INR increased	>1.2 – 1.5; >1 -1.5 x baseline if on anticoagulation (use of anticoagulation identified using the B01A ATC code)	>1.5 – 2.5; >1.5 – 2.5 x baseline if on anticoagulation (use of anticoagulation identified using the B01A ATC code)	>2.5; >2.5 x baseline if on anticoagulation (use of anticoagulation identified using the B01A ATC code)	-
Eosinophilia	>ULN and >Baseline		Steroids initiated	
Neutrophil count decreased	<LLN - 1500/mm3; <LLN - 1.5 x 10e9 /L	<1500 - 1000/mm3; <1.5 - 1.0 x 10e9 /L	<1000 - 500/mm3; <1.0 - 0.5 x 10e9 /L	<500/mm3; <0.5 x 10e9 /L
Lymphocyte count decreased	<LLN - 800/mm3; <LLN - 0.8 x 10e9/L	<800 - 500/mm3; <0.8 - 0.5 x 10e9 /L	<500 - 200/mm3; <0.5 - 0.2 x 10e9 /L	<200/mm3; <0.2 x 10e9 /L
Lymphocyte count increased	-	>4000/mm3 - 20,000/mm3	>20,000/mm3	-

Investigations				
	Grade			
Laboratory Analyte	1	2	3	4
Leukocytosis (White blood cell increased)			>100,000/mm^3	Clinical manifestations of leucostasis (sic); urgent intervention indicated Note that the spelling is often "leukostasis" in the literature.
Platelet count decreased	<LLN - 75,000/mm3; <LLN - 75.0 x 10^9/L	<75,000 - 50,000/mm3; <75.0 - 50.0 x 10^9/L	<50,000 - 25,000/mm3; <50.0 - 25.0 x 10^9/L	<25,000/mm3; <25.0 x 10^9/L
White blood cell decreased	<LLN – 3000/mm^3; <LLN – 3.0 x 10^9/L	<3000 – 2000/mm^3; <3.0 – 2.0 x 10^9 /L	<2000 – 1000/mm^3; <2.0 – 1.0 x 10^9 /L	<1000/mm^3; <1.0 x 10^9 /L
(a) Definition: A finding based on laboratory test results that indicate increased levels of hemoglobin above normal – i.e. above ULN.				

Metabolism and Nutrition Disorders				
	Grade			
Adverse Event	1	2	3	4
Hypercalcemia (Calcium Increased)	Corrected serum calcium of >ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L; Ionized calcium >ULN - 1.5 mmol/L	Corrected serum calcium of >11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L; Ionized calcium >1.5 - 1.6 mmol/L; symptomatic	Corrected serum calcium of >12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L; hospitalization indicated	Corrected serum calcium of >13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L; life-threatening consequences
Hyperkalemia (Potassium Increased)	>ULN - 5.5 mmol/L	>5.5 - 6.0 mmol/L; intervention initiated	>6.0 - 7.0 mmol/L; hospitalization indicated	>7.0 mmol/L; life-threatening consequences
Hypermagnesemia (Magnesium Increased)	>ULN - 3.0 mg/dL; >ULN - 1.23 mmol/L	n/a	>3.0 - 8.0 mg/dL; >1.23 - 3.30 mmol/L	>8.0 mg/dL; >3.30 mmol/L; life-threatening consequences
Hypernatremia (Sodium Increased)	>ULN - 150 mmol/L	>150 - 155 mmol/L; intervention initiated	>155 - 160 mmol/L; hospitalization indicated	>160 mmol/L; life-threatening consequences
Hypoalbuminemia (Albumin Decreased)	<LLN - 3 g/dL; <LLN - 30 g/L	<3 - 2 g/dL; <30 - 20 g/L	<2 g/dL; <20 g/L	Life-threatening consequences; urgent intervention indicated
Hypocalcemia (Calcium Decreased)	Corrected serum calcium of <LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L; Ionized calcium<LLN - 1.0 mmol/L	Corrected serum calcium of <8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L; Ionized calcium <1.0- 0.9 mmol/L; symptomatic	Corrected serum calcium of <7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L; Ionized calcium <0.9- 0.8 mmol/L; hospitalization indicated	Corrected serum calcium of <6.0 mg/dL; <1.5 mmol/L;Ionized calcium <0.8 mmol/L;life-threatening consequences
Hypoglycemia (Glucose Decreased)	<LLN - 55 mg/dL; <LLN - 3.0 mmol/L	<55 - 40 mg/dL; <3.0 - 2.2 mmol/L	<40 - 30 mg/dL; <2.2 - 1.7 mmol/L	<30 mg/dL; <1.7 mmol/L; life-threatening consequences; seizures
Hypokalemia (Potassium Decreased)	<LLN - 3.0 mmol/L	Symptomatic with <LLN - 3.0 mmol/L; intervention indicated	<3.0 - 2.5 mmol/L; hospitalization indicated	<2.5 mmol/L; life-threatening consequences
Hypomagnesemia (Magnesium Decreased)	<LLN - 1.2 mg/dL; <LLN - 0.5 mmol/L	<1.2 - 0.9 mg/dL; <0.5 - 0.4 mmol/L	<0.9 - 0.7 mg/dL; <0.4 - 0.3 mmol/L	<0.7 mg/dL; <0.3 mmol/L; life-threatening consequences
Hyponatremia (Sodium Decreased)	<LLN - 130 mmol/L	125-129 mmol/L and asymptomatic	125-129 mmol/L symptomatic; 120-124 mmol/L	<120 mmol/L; life-threatening consequences

Metabolism and Nutrition Disorders				
	Grade			
Adverse Event	1	2	3	4
			regardless of symptoms	

Renal and Urinary Disorders				
	Grade			
Adverse Event	1	2	3	4
Chronic kidney disease	eGFR (estimated Glomerular Filtration Rate) or CrCl (creatinine clearance) <LLN - 60 ml/min/1.73 m^2 or proteinuria 2+ present; urine protein/creatinine >0.5	eGFR or CrCl 59 - 30 ml/min/1.73 m^2	eGFR or CrCl 29 - 15 ml/min/1.73 m^2	eGFR or CrCl <15 ml/min/1.73 m^2; dialysis or renal transplant indicated

Appendix 6 SAS Codes for Statistical Analysis

Test	**Template SAS Code for Modeling** (SAS Version 9.4)
Determination of lower O'Brien-Fleming stopping boundary for the primary efficacy endpoint at the interim analysis	Proc Seqdesign errspend plots=all boundaryscale=pvalue; PrimaryEfficacy: design alpha = 0.05 beta = 0.20 alt = twosided /* two-sided test*/ nstages = 2 method (reject)= errfuncobf /* O'Brien-Fleming as implemented by Lan-DeMets */ info = cum(24 52) /* information proportion are based on 24 wks out of 52 wks*/; run;
Primary Analysis: Stratified Wilcoxon rank-sum test	proc npar1way data=ITT; Title 'Primary Endpoint – Comparison of Annualized VOC Rates'; Title2 'Intent To Treat Population'; class TRT; var ARVOC; strata <stratification factors: REGION HU >; run;
Negative binomial regression	proc genmod data=ITT; Title 'Sensitivity Analysis for Primary Endpoint – Negative Binomial Regression'; class TRT <stratification factors: REGION HU >; *TRT is treatment group; model Y = TRT <stratification factors: REGION HU> / dist = negbin link = log offset = LTIME; *Y is the count of VOC, LTIME is the log of exposure time in years; lsmeans TRT / cl diff exp;

Time to Event Analysis	proc lifetest data=ITT plots=(survival(nocensor atrisk cb));
	Title 'Secondary Endpoint – Time to Event Analysis Time to first VOC';
	time TIME*VOC(0);
	Strat TRT;
	run;

Mixed model repeated measures (MMRM)	proc mixed data=ITT;
	class SUBJECT TRT VISIT HU REGION;
	model CFB = TRT VISIT TRT*VISIT BASELINE HU REGION; *CFB=change from baseline;
	repeated VISIT/ TYPE = UN ddfm=kr sub=SUBJECT; /*Type=CS and then Type AR(1) will be used if UN does not work*/
	lsmeans TRT*VISIT/cl diff alpha=0.05;
	lsmeans TRT/cl diff alpha=0.05;
	run;

Multiple Imputation for the annualized rate of VOCs	proc mi; class TRT REGION HU ; fcs reg(AVOC = TRT REGION HU VOC_O DUR); *AVOC= annualized VOC rate; VOC_O =Number of observed VOCs; DUR = Duration in the study. run;

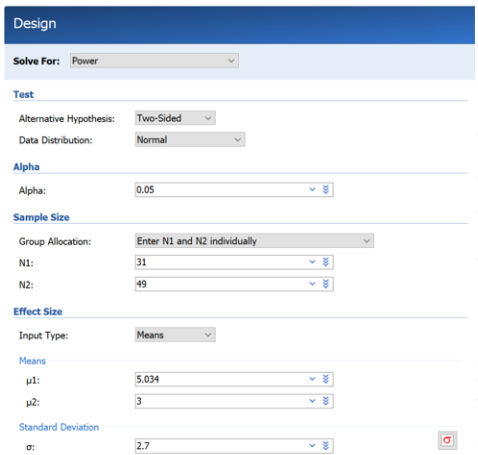
Test	**Template SAS Code for Modeling** (SAS Version 9.4)
with FCS Method	

Negative binomial regression based on Multiple Imputed dataset	proc genmod data=MIDATA;
	by _imputation_;
	class TRT REGION HU ; *TRT is treatment group;
	model Y = TRT REGION HU / dist = negbin link = log offset = LTIME;
	*Y is the count of VOC, LTIME is the log of exposure time in years;
	lsmeans TRT / cl diff exp;
	run;

Combine Parameter Estimates	proc mianalyze data=outreg; modeleffects TRT; /*TRT is the variable for mean treatment differences*/ stderr TRTERR; /*TRTERR is the standard errors for treatment differences*/ run;

Appendix 7 Power Calculation for the Primary Endpoint (Annualized Rate of VOCs)

Wilcoxon 40% Effect Size - 0% Drop Out



Mann-Whitney U or Wilcoxon Rank-Sum Tests

Numeric Results

δ = μ1 - μ2
Hypotheses: H0: δ = 0 vs. H1: δ ≠ 0
Data Distribution: Normal

Power	N1	N2	N	μ1	μ2	δ	σ	Alpha
0.88002	31	49	80	5.0	3.0	2.0	2.7	0.050

References

Al-Sunduqchi, Mahdi S. 1990. Determining the Appropriate Sample Size for Inferences Based on the Wilcoxon Statistics. Ph.D. dissertation under the direction of William C. Guenther, Dept. of Statistics, University of Wyoming, Laramie, Wyoming.
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.
Julious, S. A. 2010. Sample Sizes for Clinical Trials. Chapman & Hall/CRC. Boca Raton, FL.
Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.
Zar, Jerrold H. 1984. Biostatistical Analysis (Second Edition). Prentice-Hall. Englewood Cliffs, New Jersey.

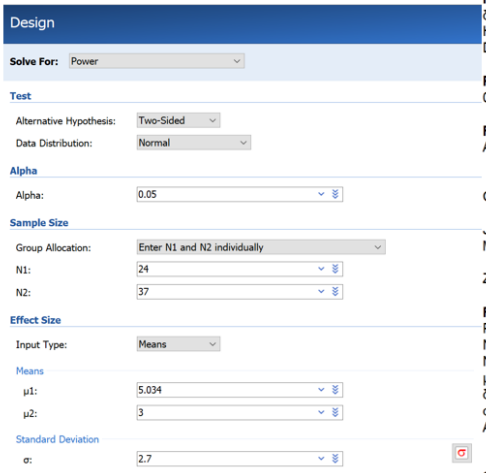
Report Definitions

Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of items sampled from each population.
N = N1 + N2 is the total sample size.
μ1 and μ2 are the assumed population means.
δ = μ1 - μ2 is the difference between population means at which power and sample size calculations are made.
σ is the assumed population standard deviation for each of the two groups.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements

Group sample sizes of 31 and 49 achieve 88.002% power to detect a difference of μ1 - μ2 = 5.0 - 3.0 = 2.0 using a two-sided Mann-Whitney U or Wilcoxon Rank-Sum test assuming that the actual data distribution is normal when the significance level (alpha) of the test is 0.050 and the standard deviation is 2.7 in both groups.

Wilcoxon 40% Effect Size - 25% Drop Out



Mann-Whitney U or Wilcoxon Rank-Sum Tests

Numeric Results

δ = μ1 - μ2
Hypotheses: H0: δ = 0 vs. H1: δ ≠ 0
Data Distribution: Normal

Power	N1	N2	N	μ1	μ2	δ	σ	Alpha
0.77644	24	37	61	5.0	3.0	2.0	2.7	0.050

References

Al-Sunduqchi, Mahdi S. 1990. Determining the Appropriate Sample Size for Inferences Based on the Wilcoxon Statistics. Ph.D. dissertation under the direction of William C. Guenther, Dept. of Statistics, University of Wyoming, Laramie, Wyoming.
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, Third Edition. Taylor & Francis/CRC. Boca Raton, Florida.
Julious, S. A. 2010. Sample Sizes for Clinical Trials. Chapman & Hall/CRC. Boca Raton, FL.
Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.
Zar, Jerrold H. 1984. Biostatistical Analysis (Second Edition). Prentice-Hall. Englewood Cliffs, New Jersey.

Report Definitions

Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of items sampled from each population.
N = N1 + N2 is the total sample size.
μ1 and μ2 are the assumed population means.
δ = μ1 - μ2 is the difference between population means at which power and sample size calculations are made.
σ is the assumed population standard deviation for each of the two groups.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements

Group sample sizes of 24 and 37 achieve 77.644% power to detect a difference of μ1 - μ2 = 5.0 - 3.0 = 2.0 using a two-sided Mann-Whitney U or Wilcoxon Rank-Sum test assuming that the actual data distribution is normal when the significance level (alpha) of the test is 0.050 and the standard deviation is 2.7 in both groups.

Appendix 8 Mock-up Tables, Listings, and Figures

Mock-up tables, listings, and figures are presented in a separate document.