NCT03347422

STATISTICAL ANALYSIS PLAN

Protocol title:	**A Phase 3, randomized, double-blind, placebo-controlled study to assess the efficacy and safety of BIVV009 in patients with primary Cold Agglutinin Disease without a recent history of blood transfusion**
Protocol number:	**BIVV009-04 Part A**
Compound number (INN/Trademark):	**BIVV009**
Study phase:	**Phase III**
Short title:	**Cadenza**
Statistician:	[redacted]
Statistical project leader:	[redacted]
Date of issue:	**02-Oct-2020**

Regulatory agency identifier number(s):

Registry:	NA
Enter Registry Name:	NA

Total number of pages: 52

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

VERSION HISTORY

This Statistical Analysis Plan (SAP) for study BIVV-009-04 (Cadenza) is based on the protocol dated [15-Oct-2019]. There are no major changes to the statistical analysis features in this SAP. The first participant was randomized on 16 Apr 2018. This SAP is approved before the first interim analysis is conducted.

Table 1 - Major changes in statistical analysis plan

SAP Version	Approval Date	Changes	Rationale
Version 1 (electronic 1.0)	07-Jul-2020	Not Applicable	Original version
Version 2 (electronic 2.0)	02-Oct-2020	Updated primary endpoint to CMH test (previously Fisher's Exact test) per feedback from FDA.	Feedback from FDA

1 INTRODUCTION

1.1 STUDY DESIGN

This is a randomized, double-blind, placebo-controlled, multicenter study designed to evaluate the efficacy and safety of BIVV009 in symptomatic patients with the complement-mediated disorder, primary CAD, who do not have a recent history of blood transfusion.

During the 6-week Screening/Observation Period, prospective patients will have a detailed medical history documented (including all available transfusion history), physical evaluations, and blood samples collected on 3 occasions approximately every 2 weeks.

The study will enroll approximately 40 primary CAD patients who do not have a recent history of blood transfusion (ie, ≤1 transfusion during the last year and no transfusion during the last 6 months prior to enrollment). Eligible patients should have been diagnosed with primary CAD at least 6 months prior to enrollment and should have had no history of transfusion during this period.

Eligible patients are randomized with a ratio of 1:1 to BIVV009 or Placebo by a permuted block randomization. The block size is fixed at 4. The randomization is performed by Pharm-Olam's Interactive Web Response System (IWRS). Sponsor personnel (except for those who provide drug supply) are blinded to the treatment assignment until the interim analysis (defined below).

Eligible patients will be randomized 1:1 to receive an IV infusion of BIVV009 or placebo over approximately 60 minutes on Day 0, Day 7, and every 14 days thereafter through Week 25 (ie, Days 21, 35, 49, 63, 77, 91, 105, 119, 133, 147, 161, and 175). Patients who weigh less than 75 kg will receive fixed doses of 6.5 g of BIVV009 or placebo. Patients who weigh 75 kg or more will receive fixed doses of 7.5 g of BIVV009 or placebo. Patients who miss a dose (ie, outside the dosing window or >17 days since last dose) should return to the site for an unscheduled visit 1 week prior to the next scheduled dose in order to receive an additional loading dose. Patients will have an EOT visit on Day 182 (Week 26).

During the 6-month treatment period, patients will receive a transfusion, if his or her Hgb level meets either of the following criteria:

- Hgb is <9 g/dL and the patient is symptomatic, or
- Hgb is <7 g/dL and the patient is asymptomatic.

Following completion of dosing in the 6-month treatment period, patients will continue to receive BIVV009 dosing during Part B, the long-term safety and durability of response extension phase. Part B will run for approximately one year following last patient out (LPO) under Part A.

Study primary analysis will be conducted for Part A after all patients have completed Part A and the data are cleaned. Parts A and B will have separate database locks to enable submission of the MAA following completion of Part A. Additional analyses of Part B data will be defined in a separate Part B SAP.

This statistical analysis plan contains information pertaining definitions of analysis sets and derived variables, and statistical methods for the analysis of efficacy, safety, pharmacokinetics (PK), and pharmacodynamics (PD) for Part A of the referenced study.

1.2 OBJECTIVE AND ENDPOINTS

Table 2 - Objectives and endpoints

	Objectives	Endpoints
Primary	• to determine whether BIVV009 administration results in a ≥1.5 g/dL increase in hemoglobin (Hgb) level and avoidance of transfusion in patients with primary CAD without a recent history of blood transfusion.	• responder rate: A patient is considered a responder if he or she did not receive a blood transfusion from Week 5 through Week 26 (end of treatment [EOT]) and did not receive treatment for CAD beyond what is permitted per protocol. Additionally, the patient's Hgb level must meet the following criterion: Hgb increased ≥1.5 g/dL from baseline (defined as the last Hgb value before administration of the first dose of study drug) at treatment assessment timepoint (defined as mean value from Weeks 23, 25, and 26)
Secondary	• To assess the effect of BIVV009 on clinical events and laboratory parameters related to hemolysis and anemia in patients with primary CAD • To assess the effect of BIVV009 on specific complications of CAD (acrocyanosis, Raynaud's syndrome, hemoglobinuria, and thromboembolism) • To assess the effect of BIVV009 on quality of life (QOL) in patients with primary CAD • To evaluate the overall safety and tolerability of BIVV009 in patients with primary CAD	• Mean change from baseline in Hgb at treatment assessment timepoint (mean of values at Week 23, 25, and 26) • Mean change from baseline in bilirubin (excluding patients with Gilbert's Syndrome) at treatment assessment timepoint • Mean change from baseline in QOL, as assessed by the change in Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue scale scores at the treatment assessment timepoint • Mean change from baseline in lactate dehydrogenase (LDH) at the treatment assessment timepoint • Incidence of solicited symptomatic anemia at EOT • Incidence of treatment-emergent AEs (TEAEs) and serious AEs (SAEs) • Change from baseline in clinical laboratory evaluations • Change from baseline in SLE panel • Change from baseline in vital signs • Change from baseline in electrocardiogram (ECG) data • Physical examination findings • Serum disease-related biomarkers

	Objectives	Endpoints
		• Incidence of hemolytic breakthrough (rapid fall in Hgb ≥2 g/dL associated with an increase in LDH/bilirubin and/or decrease in haptoglobin since the last scheduled visit) through the EOT at Week 26 • Incidence of infections of ≥ Grade 3 severity (ie, requiring intravenous [IV] antibiotics) • Incidence of thromboembolic events
Tertiary/exploratory	• To assess the effect of BIVV009 on other quality of life (QOL) measures in patients with primary CAD • To assess the effect of BIVV009 on other laboratory parameters in patients with primary CAD • Pharmacokinetics of BIVV009 • Pharmacodynamics of BIVV009	• Mean change from baseline in QOL, as assessed by the change in the five level EuroQol – five dimensions questionnaire (EQ-5D-5L) scores at the treatment assessment timepoint • Mean change from baseline in QOL, as assessed by the change in the 12-item short form survey (SF-12®) at the end of treatment assessment timepoint • Proportion of patients with Hgb level of ≥12 g/dL at the treatment assessment timepoint • Incidence of thromboembolic events after the first 5 weeks of study drug administration • Median time to normalization of bilirubin • Median time to normalization of LDH • Median time to normalization of haptoglobin • Median time to obtain Hgb level of ≥12 g/dL • Proportion of patients normalizing haptoglobin at the treatment assessment timepoint • Proportion of patients normalizing bilirubin at the treatment assessment timepoint • Proportion of patients with abnormal LDH at baseline who normalize LDH at treatment assessment timepoint • Patient's Global Impression of Change (PGIC) to assess patient's perception of changes in CAgD disease burden at EOT • Patient's Global Impression of [Fatigue] severity (PGIS) to assess the patient's perception of changes in fatigue at EOT • Incidence of disabling circulatory symptoms at treatment assessment timepoint • Total healthcare resource utilization at EOT • Plasma concentrations of BIVV009 • PK parameters: Appropriate exposure parameters (Cmax, AUC) will be derived using a population PK approach • Plasma concentrations of BIVV009 • PK parameters: Appropriate exposure parameters (Cmax, AUC) will be derived using a population PK approach

1.2.1 Estimands

Following International Conference of Harmonisation (ICH) (1) and National Research Counsel (NRC) (2) guidelines, the following estimands are considered for efficacy endpoints to provide clarity of what to be estimated and its connection to the trial objectives. Akacha, et al. (3) proposed a framework for defining an estimand using three attributes:

- The population of interest;
- The variable to be used for the clinical question, which consists of measurements at a specific timepoint. It could be a composite endpoint that incorporates post-randomization events.
- The measure of intervention effect, which accounts for potential effects of post-randomization events.

A composite estimand will be used for the primary endpoint and a hypothetical estimand will be used for lab parameter defined secondary endpoints. Alternative estimands are also proposed to either assist the interpretation of primary estimand or used for exploratory endpoints.

Primary estimand defined for main endpoints are summarized in below Table 3. More details are provided in Section 4.

For all these estimands, the comparison of interest will be the comparison of BIVV009 vs. Placebo control.

Table 3 - Summary of primary estimand for main endpoints

	Estimands			
Endpoint Category (estimand)	**Endpoint**	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
Primary objective: to determine whether BIVV009 administration results in a ≥1.5 g/dL increase in hemoglobin (Hgb) level and avoidance of transfusion in patients with primary CAD without a recent history of blood transfusion				
Primary endpoint (COVID-adjusted Composite estimand)	Responder if: • Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates, and • The change from baseline in Hgb at treatment assessment timepoint >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 5.1.1 within the range of Week	FAS	• Subjects who early discontinue study prior to Week 23, for reasons other than COVID-19, are considered non-responders • Subjects with no Hgb data from Week 23, 25, and 26, for reasons other than COVID-19, are considered as non-responders	• odds ratio of the proportion of responders in BIVV009 and placebo, using stratified Cochran-Mantel-Haenszel test • Multiple imputation will be performed to fill in Hgb data at TAT using an MMRM approach, for subjects missing infusions at Week 23

Endpoint Category (estimand)	Estimands			
	Endpoint	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
	5 and Week 26 visit dates (excluding medications taken during a COVID-related infusion gap or within the following 5 weeks)		• If a subject has a COVID-related infusion gap (defined as 2 consecutive missed infusions due to COVID-19), transfusions received and protocol-prohibited CAD medications taken during the infusion gap and within the 5 weeks following the infusion gap will not be included in the responder derivation.	and 25 due to COVID-19
Primary endpoint (Composite estimand)	Responder if: • Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates, and • The change from baseline in Hgb at treatment assessment timepoint >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 5.1.1 within the range of Week 5 and Week 26 visit dates	FAS	• Subjects who early discontinue study prior to Week 23 are considered as non-responders • Subjects with no Hgb data from Week 23, 25 and 26 are considered as non-responders	odds ratio of the proportion of responders in BIVV009 and placebo, using stratified Cochran-Mantel-Haenszel test
Primary endpoint *(Modified Composite estimand)*	Responder if: • Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates, and • The change from baseline in Hgb at treatment assessment timepoint >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 5.1.1	mFAS	• Subjects who early discontinue study prior to Week 23 are considered as non-responders • Subjects with no Hgb data from Week 23, 25 and 26 are considered as non-responders	Odds ratio of i the proportion of responders in BIVV009 and placebo, using stratified Cochran-Mantel-Haenszel test

Endpoint Category (estimand)	Estimands			
	Endpoint	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
	within the range of Week 5 and Week 26 visit dates			
Primary endpoint *(Completer estimand)*	Responder if: • Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates, and • The change from baseline in Hgb at treatment assessment timepoint >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 5.1.1 within the range of Week 5 and Week 26 visit	A subset of FAS with those who complete treatment at least through Week 23 and have at least 1 evaluable Hgb from Week 23, 25, and 26	N/A	Odds ratio of the proportion of responders in BIVV009 and placebo, using stratified Cochran-Mantel-Haenszel test
Primary endpoint *(Per-Protocol estimand)*	Responder if: • Free of post-baseline transfusion within the range of Week 5 and Week 26 visit, and • The change from baseline in Hgb at treatment assessment timepoint >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 5.1.1 within the range of Week 5 and Week 26 visit dates	PP set	N/A	Odds ratio of the proportion of responders in BIVV009 and placebo, using stratified Cochran-Mantel-Haenszel test
Secondary objective: To assess the effect of BIVV009 on clinical events and laboratory parameters related to hemolysis and anemia in patients with primary CAD				
Secondary endpoint (Hypothetical estimand)	Mean change from baseline in Hgb and FACIT-F Score at treatment assessment timepoint (mean of values at Week 23, 25, and 26)	FAS	• Any value after transfusion and prohibited medication is considered missing	Difference in mean change from baseline at TAT between BIVV009 and Placebo
Secondary endpoint	Mean change from baseline in Hgb and FACIT-F Score at treatment assessment	mFAS	• Any value after transfusion and	Difference in mean change from baseline at

Endpoint Category (estimand)	Estimands			
	Endpoint	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
(modified hypothetical estimand)	timepoint (mean of values at Week 23, 25, and 26)		prohibited medication is considered missing	TAT between BIVV009 and Placebo
Secondary endpoint *(De-facto estimand)*	Mean change from baseline in Hgb and FACIT-F Score at treatment assessment timepoint (mean of values at Week 23, 25, and 26)	FAS	N/A	Difference in mean change from baseline at TAT between BIVV009 and Placebo

2 SAMPLE SIZE DETERMINATION

Approximately 40 patients with primary CAD who do not have a recent history of transfusion will be randomized. There are limited data from the published literature or available natural history data to estimate the placebo response for the transfusion-independent study population. Since patients with evidence of ongoing hemolysis, hemoglobin <10 g/dL, and symptomatic are included in the study, it is expected that few patients would meet the responder definition for the primary endpoint without treatment intervention. The proposed sample size was chosen to provide sufficient power to detect a 50% improvement for BIVV009 over a range of possible placebo responder rates with a significance level of 0.05. A total of 40 patients (20 patients per group) provides a statistical power of greater than 85% to detect a treatment difference of 50% for placebo response rates between the range of 15% to 40%. This calculation assumes a 2-sided 5%-level test comparing the response rates between the BIVV009 and placebo groups. A 50% improvement over placebo is considered clinically relevant.

3 ANALYSIS POPULATIONS

The following analysis populations are defined for Part A of the study. For the remainder of the SAP, the term "subjects" will be used to refer to patients in the study to keep consistent with the term used within table, figure, and listing (TFL) displays.

The following populations for analyses are defined:

Table 4 - Populations for analyses

Population	Description
Full Analysis Set (FAS)	The ITT population consists of all randomized subjects who received at least 1 dose (including partial dose) of study drug. All subjects in the ITT population will be included in the Full Analysis Set. Analyses of efficacy will be performed on the FAS. For analysis, FAS and the ITT population are considered exchangeable in the SAP, but the term FAS will be used in subsequent sections of the SAP and TFLs.
Modified Full Analysis Set (mFAS)	The COVID-19 pandemic occurred during the conduct of Part A study. Following the guidelines from FDA and EMA to assess any potential impact of COVID-19 on efficacy evaluation, we propose a modified full analysis (mFAS) set. The mFAS population consists of a subset of FAS who did not miss any visits or discontinue early due to the COVID-19 pandemic.
Per-protocol Set	The per-protocol (PP) population is defined as a subset of FAS who do not have any important protocol deviations impacting their efficacy assessments (defined in Section 6.3). Selected efficacy endpoints will be analyzed for the PP Population.
Safety Analysis Set	Subjects who received at least 1 dose (including partial dose) of study drug will be included in the Safety Analysis Set. Note that Safety Analysis Set is the same as FAS in this study.
PK Analysis Set	Subjects who received at least 1 dose of study drug and have evaluable PK concentrations will be included in PK analysis set.
PD Analysis Set	All subjects who receive at least 1 dose of study drug and have at least 1 evaluable PD sample during Part A will be included in the PD analysis population.

4 STATISTICAL ANALYSES

Statistical analysis will be performed by Sanofi, using SAS® version 9.3 or higher and, where appropriate, additional validated software. This SAP is based on protocol Version 6, dated 15 October, 2019.

4.1 GENERAL CONSIDERATIONS

Safety, efficacy, and PK/PD data will be summarized using standard summary statistics for continuous, categorical, and event-time data. Data will be summarized for the populations defined. All statistical hypothesis testing will be performed at the 2-sided, 5% significance level, unless specified otherwise. All p-values will be presented to 3 decimal places; p-values less than 0.001 will be presented as "<0.001," and p-values greater the 0.999 will be presented as ">0.999."

Unless otherwise specified, analyses will be performed by intervention group (and overall for baseline and demographics characteristics).

Continuous variables will be summarized using descriptive statistics including the number of non-missing values (n), mean, standard deviation (SD), median, minimum, and maximum. Where specified in the table shells, the 25th and 75th percentiles will also be provided. Means, medians, and the 25th and 75th percentiles will be presented to one decimal place beyond that with which the data were captured. SDs will be presented to two decimal places beyond that with which the data were captured. Minimum and maximum will be displayed to the same number of decimal places as that with which the data were captured.

Categorical variables will be summarized by counts and percentages. All percentages will be rounded to one decimal place. The percentage will be suppressed when the count is zero. Unless specified otherwise, the denominator for all percentages will be the number of non-missing data for a given summarization.

Time-to-event variables will be summarized using Kaplan-Meier estimate over time, along with median time to event and probability of an event at specific time points. Partial dates will not be imputed, hence considered unknown. In general, for subjects who do not experience the event, they will be censored at the study day of last assessment within Part A.

Part A Study Treatment Period

The Part A study treatment period starts at the date of first dose of study drug and ends at the date of Week 26 Visit if the subject completes Part A study treatment, or at the date of Early Termination (ET)/Safety Follow-up Visit if the subject discontinues the study treatment prior to Week 25.

Treatment Assessment Timepoint (TAT)

The value of an assessment (eg, Hgb) at the treatment assessment timepoint is defined as the average of the values from Week 23, 25, and 26 visits. In the case of any missing value at any of

these visits, it will be calculated as the average of the available values, unless no value is available from all three visits.

COVID-related infusion gap

COVID-related infusion gap is defined as 2 or more consecutive missed scheduled infusions due to COVID-19. For the primary endpoint analysis, transfusions and protocol-prohibited CAD medications received during a COVID-related infusion gap and within the first 5 weeks of infusion resumption will not preclude a subject from being a responder.

Solicited Symptomatic Anemia

The severity of CAD symptomatic anemia is defined in Table 5. The severity grading for symptomatic anemia is adopted based on Common Toxicity Criteria for Adverse Events (CTCAE) grades and is collected in eCRFs at every visit.

Table 5 - Severity grade for symptomatic anemia related CAD

CAgD Symptomatic Anemia Term	Grade 1	Grade 2	Grade 3	Grade 4
Fatigue	Fatigue relieved by rest	Fatigue not relieved by rest; limiting instrumental ADL	Fatigue not relieved by rest, limiting self care ADL	Not defined
Weakness	Uneasiness or lack of well being	Uneasiness or lack of well being; limiting instrumental ADL	Not defined	Not defined
Shortness of Breath	Shortness of breath with moderate exertion	Shortness of breath with minimal exertion; limiting instrumental ADL	Shortness of breath at rest; limiting self care ADL	Life-threatening consequences; urgent intervention indicated
Palpitations, fast heart beat	Mild symptoms; intervention not indicated	Intervention indicated	Not defined	Not defined
Lightheadedness	Not defined	Present (eg, near fainting)	Not defined	Not defined
Chest Pain	Mild pain	Moderate pain; limiting instrumental ADL	Pain at rest; limiting self care ADL	Not defined

ADL = activities of daily living

Thromboembolic Events

The thromboembolic events prior to study entry are recorded as medical history in the eCRFs, notably pre-specified categories of TE events occurring within 1 year prior to screening are captured, while thromboembolic events on study are captured as adverse events. Both medical history and adverse events will be coded according to Medical Dictionary for Regulatory Activities (MedDRA) v20.1 (or higher).

In general, thromboembolic events may include but are not limited to:

- Any venous events
  - Portal vein obstruction
  - Deep vein thrombosis
  - Pulmonary embolism
  - Mesenteric (abdominal) venous thrombosis
- Any cerebral events
  - Cerebral infarction, occlusion, and stenosis of cerebral and precerebral arteries
  - Vascular syndromes of brain in cerebrovascular diseases, transient cerebral ischemic attacks, and related syndromes
- Any arterial events
  - Myocardial infarction
  - Mesenteric arterial thrombosis
  - Other arterial embolism and thrombosis

Medical review of medical history and AEs will ascertain thromboembolic events for analysis.

Normalization of Bilirubin

The normalization of bilirubin is defined as total bilirubin ≤ upper limit of normal (ULN) (specified by the central lab).

Normalization of LDH

The normalization of LDH is defined as LDH value ≤ ULN (specified by the central lab).

Normalization of Haptoglobin

The normalization of haptoglobin is defined as haptoglobin above the level of detection.

Normalization of Hemoglobin

The normalization of hemoglobin is defined as hemoglobin ≥12 g/dL. Hemoglobin ≥11 g/dL will also be investigated.

Hemolytic Breakthrough

Hemolytic breakthrough <u>at any visit</u> is defined as

- A decrease of 2 g/dL or more in Hgb from the last scheduled visit, and
- One of the followings:
  - An increase of LDH from the last scheduled visit,

- An increase of bilirubin from the last scheduled visit,
- A decrease in haptoglobin from the last scheduled visit.

Protocol Prohibited CAD Medications

The following concomitant medications are prohibited during the study treatment period in Part A:

- Rituximab alone or as part of combination therapy
- Cytotoxic drugs (such as fludarabine, bendamustine, ibrutinib, cyclophosphamide)
- Other investigational drugs

The efficacy endpoint (estimand) accounting for protocol prohibited CAD medication(s) use will be based on the eCRF questionnaire "Did the subject receive any prohibited medications since the last visit?" at post-baseline visits. The medical review for the list of prohibited medications will be reconciled with the eCRF questionnaire.

Pooling Sites for Analysis

Data from all investigational sites will be pooled for the analysis, unless specified otherwise.

4.2 PARTICIPANT DISPOSITIONS

The number (%) of participants included in each of the analysis populations listed in Table 4 will be summarized.

The number of subjects with the status of completed, and discontinued study treatment early, including the primary reason for those who discontinued, will be tabulated for the FAS in Part A. In particular, the number of subjects who early discontinue study and the number of subjects with a missing dose due to COVID-19 will be presented. Subject disposition, including the date of the last visit and the reason for early termination for subjects who did not complete the study treatment, will be provided in a data listing.

Additionally, subject enrollment by country and site as well as the number and percentage of subjects attending each planned visit will be summarized for the FAS.

Protocol deviations

All protocol deviations will be recorded throughout the study and will be medically adjudicated and categorized into major and minor. Deviations due to COVID-19 will be adjudicated and identified as such. Major and minor protocol deviations/violations are to be pre-specified prior to database lock. All major protocol deviations will be summarized for the FAS.

Important protocol deviations (within Part A study treatment period) impacting efficacy assessment are defined as any of the following:

- Not meeting eligibility inclusion criterion #3 as in the protocol,
- Missing at least 2 consecutive doses or 3 intermittent doses,
- Received protocol prohibited CAD treatment (Section 4.1),
- Free of transfusion after Week 5 but missing hemoglobin assessments at all three visits of Weeks 23, 25, and 26.

Important protocol deviations will be summarized for the FAS. Subjects who have had any of the important protocol deviations will be excluded from the PP population.

4.3 PRIMARY ENDPOINT(S) ANALYSIS

4.3.1 Definition of endpoint(s)

The primary endpoint is the responder rate, defined in Section 1.2, and is a composite endpoint of the following components:

1. Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates;
2. Change from baseline in Hgb at treatment assessment timepoint is >= 1.5 g/dL,
3. Receive no protocol prohibited medications (defined in Section 4.1) within the range of Week 5 and Week 26 visit dates

Components 1 and 3 will be determined by the corresponding questions in eCRF at all visits (from Week 5 through Week 26) in Part A.

4.3.2 Main analytical approach

The primary analysis of the primary endpoint is to compare the responder rate of the BIVV009 treatment arm with the placebo treatment arm via the COVID-adjusted Composite estimand as defined in Table 6 below.

To reject the null hypothesis of no treatment difference, the two-sided p-value based on a stratified Cochran-Mantel-Haenszel (CMH) test must be <0.05. The test will be stratified by baseline hemoglobin (< median baseline Hgb vs ≥ median baseline Hgb) and geographic region (Japan/Australia, United States, Europe). In the case of completely unbalanced strata (all records within any strata fall within a single treatment arm), the CMH test will be stratified only by baseline hemoglobin.

Table 6 - COVID-adjusted composite estimand for primary analysis

	COVID-adjusted Composite Estimand
Population	FAS
Response variable	Responder if

	COVID-adjusted Composite Estimand
	• Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates, and • The change from baseline in Hgb at treatment assessment timepoint >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 4.1 within the range of Week 5 and Week 26 visit dates
ICEs handling	• Subjects who early discontinue study prior to Week 23, for reasons other than COVID, are considered non-responders • Subjects with no Hgb data from Week 23, 25, and 26, for reasons other than COVID, are considered as non-responders • If a subject has a COVID-related infusion gap (defined as 2 consecutive missed infusions due to COVID-19), transfusions received and protocol-prohibited CAD medications taken during the infusion gap and within the 5 weeks following the infusion gap will not be included in the responder derivation.
Measure of treatment effect	• Odds ratio of the proportion of responders in BIVV009 and Placebo using the Cochran-Mantel-Haenszel test • Subjects missing infusions at Week 23 and 25 due to COVID will have their Hgb at TAT imputed using multiple imputation with an MMRM approach.

Multiple imputation (n=100 imputations) will first be run on subjects who missed both the Week 23 and 25 study visits due to COVID-19 to fill in their missing Hgb data at TAT. A CMH test will be conducted on each imputation and the chi-square statistics will be pooled using Rubin's Rules, together with the pooled CMH odds ratio and a 95% confidence interval for the odds ratio will be computed.

In addition, the proportion of subjects who meet each of the three response criteria will be summarized by treatment group. A separate table for the number of transfusions by study period (before Week5 and between Week 5 and Week 26) and by treatment arm will be generated.

If there are no patients who missed both the Week 23 and 25 study visits due to COVID-19, the primary efficacy analysis would be completed using the Composite estimand (Section 4.3.3.1).

4.3.3 Sensitivity analysis

All sensitivity analyses are designed to evaluate the robustness of the primary efficacy results in different aspects, aiming to provide additive evidence of the treatment effect of BIVV009.

4.3.3.1 Sensitivity Analysis I - Composite estimand

The purpose of this sensitivity analysis is to assess the primary endpoint without adjusting for COVID-related ICEs. This analysis will compare the responder rate of the BIVV009 treatment arm with the placebo treatment arm via the Composite estimand as defined in Table 7 below.

Table 7 - Composite estimand for primary analysis

	Composite Estimand
Population	FAS
Response variable	Responder if • Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates, and • The change from baseline in Hgb at treatment assessment timepoint >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 4.1 within the range of Week 5 and Week 26 visit dates
ICEs handling	• Subjects who early discontinue study prior to Week 23 are considered as non-responders • Subjects with no Hgb data from Week 23, 25 and 26 are considered as non-responders
Measure of treatment effect	Odds ratio of the proportion of responders in BIVV009 and Placebo using the Cochran-Mantel-Haenszel test

A stratified CMH test will be used to estimate the treatment effect; the p-value for the comparison, together with the CMH odds ratio and a 95% confidence interval for the odds ratio will be reported. As with the primary endpoint, the test will be stratified by baseline hemoglobin (< median baseline Hgb vs ≥ median baseline Hgb) and geographic region (Japan/Australia, United States, Europe). In the case of completely unbalanced strata (all records within any strata fall within a single treatment arm), the CMH test will be stratified only by baseline hemoglobin.

4.3.3.2 Sensitivity Analysis II - Modified composite estimand

The purpose of this sensitivity analysis is to assess the potential impact of the COVID-19 pandemic to this study by analyzing the primary efficacy endpoint with subjects did not miss any visits or discontinue early due to the COVID-19 pandemic.

This sensitivity analysis will be carried out via a modified composite estimand, which is the same as the composite estimand except with the mFAS population. A stratified CMH test will be used to estimate the treatment effect; the p-value for the comparison, together with the CMH odds ratio and a 95% confidence interval for the odds ratio will be reported. As with the primary endpoint, the test will be stratified by baseline hemoglobin (< median baseline Hgb vs ≥ median baseline Hgb) and geographic region (Japan/Australia, United States, Europe). In the case of completely unbalanced strata (all records within any strata fall within a single treatment arm), the CMH test will be stratified only by baseline hemoglobin.

4.3.3.3 Sensitivity Analysis III - Completer estimand

The purpose of this sensitivity analysis is to evaluate the primary endpoint with subjects with evaluable hemoglobin data at treatment assessment timepoint, in other words, eliminating potential bias introduced by early discontinued subjects. This analysis is performed via the completer estimand as defined in Table 8.

Table 8 - Completer estimand for primary analysis

	Completer Estimand
Population	A subset of FAS with those who complete treatment at least through Week 23 and have at least 1 evaluable Hgb from Week 23, 25, and 26
Response Variable	Responder if • Free of post-baseline transfusion within the range of Week 5 and Week 26 visit dates and • The change from baseline in Hgb at treatment assessment >= 1.5 g/dL, and • Receive no protocol prohibited medications defined in Section 4.1 within the range of Week 5 and Week 26 visit dates
Treatment effect	Odds ratio of the proportion of responders in BIVV009 and Placebo using the Cochran-Mantel-Haenszel test

A stratified CMH test will be used to estimate the treatment effect; the p-value for the comparison, together with the CMH odds ratio and a 95% confidence interval for the odds ratio will be reported. As with the primary endpoint, the test will be stratified by baseline hemoglobin (< median baseline Hgb vs ≥ median baseline Hgb) and geographic region (Japan/Australia, United States, Europe). In the case of completely unbalanced strata (all records within any strata fall within a single treatment arm), the CMH test will be stratified only by baseline hemoglobin.

4.3.3.4 Sensitivity Analysis IV - Per protocol estimand

The purpose of this sensitivity analysis is to evaluate the primary endpoint with subjects who have no important protocol-deviations which potential impact efficacy assessment (the per protocol set). Subjects with protocol-prohibited medication use are excluded from the PP set. Subjects with at least 2 consecutive, or at least 3 intermittent missing dose for any reason including COVID-19, are also excluded.

This sensitivity analysis will be carried out via a per protocol estimand, which is the same as the composite estimand except with the PP set. A stratified CMH test will be used to estimate the treatment effect; the p-value for the comparison, together with the CMH odds ratio and a 95% confidence interval for the odds ratio will be reported. As with the primary endpoint, the test will be stratified by baseline hemoglobin (< median baseline Hgb vs ≥ median baseline Hgb) and geographic region (Japan/Australia, United States, Europe). In the case of completely unbalanced strata (all records within any strata fall within a single treatment arm), the CMH test will be stratified only by baseline hemoglobin.

4.3.4 Subgroup analyses

Definitions of subgroups

The following subgroups will be considered for the selected analyses:

- Age (<65, >= 65)
- Gender (Female, Male)
- Baseline weight (<75 kg, >= 75 kg)

- Baseline Hgb level (<median, >= median g/dL)
- Previous rituximab monotherapy use (Yes, No)
- Previous rituximab monotherapy or rituximab and other cytotoxic therapy use (Yes, No)
- Prior thromboembolic events within the last 1 year (Yes, No)
- Previous reculizumab use (Yes, No)

In case of very few subjects (eg, <5) in a category, the cutoff value may be changed to suit the distribution of these factors.

Analysis Methods

The CMH odds ratio and 95% CI will be summarized for the subgroups defined above. As with the primary endpoint, the odds ratios for each subgroup will be stratified by baseline hemoglobin (< median baseline Hgb vs ≥ median baseline Hgb) and geographic region (Japan/Australia, United States, Europe). The CMH test may not be stratified in the case of completely unbalanced strata. . The consistency of the odds ratio among these sub-populations will be examined by forest plots.

4.4 SECONDARY ENDPOINT(S) ANALYSIS

4.4.1 Key/Confirmatory secondary endpoint(s)

4.4.1.1 Definition of endpoint(s)

Key secondary endpoints are rank ordered by their clinical importance in order to perform a sequential closed testing procedure to protect the type I error. The key secondary endpoints will be tested only when the primary analysis reaches significance. In this case, the key secondary endpoints will be tested using a sequential closed procedure with alpha of 0.05 for each test in the order indicated below:

1. Mean change from baseline in Hgb at the treatment assessment timepoint in BIVV009 arm vs in placebo arm (using Hypothetical estimand)
2. Mean change from baseline in FACIT-F Score at treatment assessment timepoint in BIVV009 arm vs in placebo arm (using Hypothetical estimand)

4.4.1.2 Main analytical approach

1. Mean Change from Baseline in Hgb at the Treatment Assessment Timepoint

The change from baseline in Hgb at the treatment assessment timepoint will be analyzed for the FAS. The primary analysis will be performed according to the Hypothetical Estimand based on the MMRM model, while a sensitivity analysis will be done via De-facto Estimand. Similarly, COVID-specific sensitivity analyses will be performed according to the Modified Hypothetical

Estimand and Modified De-facto Estimand in order to ascertain the results for subjects unaffected by the COVID pandemic. Additional details for these estimands are found below, in Table 9.

Table 9 - Estimands for change from baseline in Hgb

	Hypothetical Estimand	Modified Hypothetical Estimand	De-facto Estimand
Population	FAS	mFAS	FAS
Variable	Change from baseline at treatment assessment time point		
ICE handling	any value after transfusion and prohibited medication is considered missing	any value after transfusion and prohibited medication is considered missing	None
Treatment effect	Difference in mean change from baseline at TAT between BIVV009 and Placebo		

Mixed Model with Repeated Measures (MMRM) will be used for the analysis of such endpoints at study visits (including the treatment assessment time points). The MMRM model will include the baseline value of the endpoint, visit, treatment, and treatment and visit interaction. A heterogeneous Toeplitz (TOEPH) covariance matrix within a subject will be used. For the endpoint at the treatment assessment timepoint (average of Week 23, 25, and 26), the estimate will be calculated as the mean of MMRM estimates at Week 23, 25, and 26 visits.

The pseudo SAS code for MMRM is as follows:

```
Proc mixed data=(ADaM dataset) method=reml;
   class usubjid visit;
   model chg = base visit trt trt*visit/ ddfm = kenwardroger
solution;
   repeated visit / type=TOEPH subject=usubjid;
   estimate 'BIV009 vs. Placebo at treatment assessment
timepoint' trt 1 -1 trt*visit 0 0 0 0 0 0 0 0 0 0 0 0 0.333
0.333 0.333 0 0 0 0 0 0 0 0 0 0 0 -0.333 -0.333 -0.333
/cl;
run;
```

For either estimand, the mean change from baseline in Hgb at treatment assessment timepoint for both treatment arms, the estimated difference in change from baseline at TAT between BIVV009 and placebo, along with its 95% CI, will be estimated by the MMRM model. The significance test for the treatment difference in Hgb at the treatment assessment timepoint will be based on appropriate treatment contrast.

Additional sensitivity analysis based on multiple imputation will be carried out for the Hypothetical Estimand to evaluate the appropriateness of the assumption of MAR. Specifically, multiple imputation will be performed using a pattern-mixture model and the above MMRM model will be run on the imputed data. Non-monotone missing data will first be transformed to a monotone missing data structure. The control-based imputation method assumes subjects in the

BIVV009 treatment arm who discontinued from the study or are missing Hgb values at the TAT will exhibit similar Hgb values to those subjects in the placebo arm (and subjects in the placebo arm will continue to exhibit those Hgb values). Thus, multiple imputation is performed using only information from the placebo arm.

2. Mean Change from Baseline in FACIT-F Score at Treatment Assessment Timepoint

The derivation of FACIT-F score is detailed in Section 4.5.1. The analyses for hemoglobin described above will be performed for the FACIT-F score and is therefore not repeated here.

Additional sensitivity analysis based on multiple imputation will be carried out for the Hypothetical Estimand to evaluate the appropriateness of the assumption of MAR. Specifically, multiple imputation will be performed using a pattern-mixture model and the above MMRM model will be run on the imputed data. Non-monotone missing data will first be transformed to a monotone missing data structure. The control-based imputation method assumes subjects in the BIVV009 treatment arm who discontinued from the study or are missing FACIT-F values at the TAT will exhibit similar FACIT-F values to those subjects in the placebo arm (and subjects in the placebo arm will continue to exhibit those FACIT-F values). Thus, multiple imputation is performed using only information from the placebo arm.

4.4.2 Supportive secondary endpoint(s)

Additional analyses for hemoglobin and bilirubin

To further understand the clinical effect of BIVV009 in increasing hemoglobin level compared with placebo, the number of subjects with an increase of hemoglobin level from baseline ≥1 g/dL, ≥1.5 g/dL, ≥2 g/dL and ≥3 g/dL at TAT will be reported for BIVV009 and for placebo.

Mean Change from Baseline in Bilirubin by visit (including TAT)

Total bilirubin will be summarized by-visit, including change from baseline, using the Full Analysis Set, with subjects with Gilbert's syndrome or unknown Gilbert's test result excluded.

If the number of subjects with unknown Gilbert's syndrome status is greater than 2, additional sensitivity analysis will be performed to include such subjects (unknown Gilbert's syndrome).

Mean Change from Baseline in Lactate Dehydrogenase (LDH) by visit (including TAT)

LDH will be summarized by-visit, including change from baseline, using the Full Analysis Set.

Additional descriptive analyses will be performed for LDH isoforms.

Incidence of Solicited Symptomatic Anemia

- Incidence of solicited symptomatic anemia will be summarized by visit descriptively using the FAS. The summaries will be based on observed data only and missing grades will not be imputed.

- An improvement in each anemia symptom is defined as at least 1 grade reduction from baseline; whereas worsening is defined as at least 1 grade increase from baseline. The percentage of improved and worsened subjects for each symptom will be presented by visit; this percentage is calculated based on the number of subjects with no missing grade at both baseline and each post-baseline visit.
- The improvement (worsening) in all anemia symptoms is defined as at least 1 grade reduction (increase) in at least 1 symptom, and no worsening (improvement) in other remaining symptoms. The proportion of subjects improved and worsened in all anemia symptoms will be summarized by visit.

4.5 TERTIARY/EXPLORATORY ENDPOINT(S) ANALYSIS

The analysis of exploratory endpoints will be based on the Observed Estimand defined in Section 1.2.1. Additionally, summary of hemolytic parameters (Hgb, bilirubin, LDH, haptoglobin, and reticulocyte) over time (by visit) is included to examine the improvement trend over time. Unless specified otherwise, the endpoints will be summarized by visit for continuous and categorical variables, and for time to event variables (eg, time to obtain Hgb level of ≥12 g/dL) for the entire Part A. The analyses for exploratory endpoints will be outlined by the type of endpoints in the following subsection. Details of patient-reported outcomes can be found in Section 4.5.3.

4.5.1 Definition of endpoint(s)

Continuous Endpoints

The continuous endpoints include:

- Mean change from baseline in QOL, as assessed by the change in the five level EQ-5D-5L scores at the treatment assessment timepoint (TAT) and by visit
- Mean change from baseline in QOL, as assessed by the change in the 12-item short form survey (SF-12®) at the end of treatment assessment endpoint and by visit
- Mean change from baseline in hemolytic parameters (Hgb, total bilirubin, LDH, haptoglobin, and absolute reticulocyte) by visit
- Mean change from baseline in FACIT-F by visit

Response (Categorical) Endpoints

The response endpoints include (order different from protocol and not reflecting importance):

- Incidence of thromboembolic events after the first 5 weeks of study drug administration
- Incidence of CAD-related symptoms at last scheduled visit of Part A
- Total healthcare resource utilization at EOT (Part A)
- Proportion of subjects normalizing Hgb (>= 12 g/dL) at TAT and by visit
- Proportion of subjects achieving 1.5 g/dL increase in Hgb at TAT and by visit
- Proportion of subjects normalizing bilirubin at TAT and by visit

- Proportion of subjects normalizing LDH at TAT and by visit (Proportion of subjects with LDH ≤1.5 x ULN at TAT will also be reported)
- Proportion of subjects normalizing haptoglobin at TAT and by visit
- PGIC and change from baseline in PGIS to assess subject's perception of changes in CAgD disease burden at last scheduled visit of Part A

Time-to-Event Endpoints

Time-to-event endpoints include:

- Time to first transfusion after the first 5 weeks of study drug administration from the first dose of study drug
- Time to first transfusion (including the first 5 weeks of study drug administration) from the first dose of study drug
- For subjects who do not receive any transfusion, their time-to-event will be censored at the study day of the last scheduled visit in Part A.
- Time to first normalization of bilirubin from the first dose of study drug
- Time to first normalization of LDH from the first dose of study drug
- Time to first normalization of haptoglobin from the first dose of study drug
- Time to first normalization of Hgb level (≥12 g/dL)
- Time to first achieving 1.5 g/dL increase in Hgb level

4.5.2 Main analytical approach

Continuous Endpoints

The continuous endpoints will be summarized by visit and treatment arm using descriptive statistics for the FAS according to Section 4.1.

For EQ-5D index score and SF-12 component scores (physical component score and mental component score), an ANCOVA (Analysis of Covariance) model will be implemented, where treatment arm is the fixed effect, and the baseline score is the covariate; a 95% confidence interval for the treatment effect difference and the corresponding p-value will be generated. The treatment difference and 95% CI will be estimated by ANOVA with treatment and baseline value as independent variables.

Response (Categorical) Endpoints

The categorical endpoints will be summarized descriptively by visit for the FAS according to Section 4.1 with the exception of total healthcare resource utilization at EOT (Part A), which will be presented in a listing.

All thromboembolic events reported (in Adverse Events) will be reviewed and determined by internal medical staff along with investigators. In case that there are discrepancies between the medical review and investigators' assessment, separate analyses will be carried out for both sets of events. Date of event occurrences will be verified. The detailed listing for all events will be provided including date (study day) of onset, verbatim terms of the event, and the flag for whether the event occurred after the first 5 weeks of study drug administration. Incidence of thromboembolic events will be calculated as the total number of thromboembolic events after the first 5 weeks of study drug administration divided by the total subject-years of time for observation (ie, time from Week 5 through Week 26 or EOT visit for subjects who discontinue early). All events from Week 5 through Week 26 or EOT visit for subjects who discontinue early will be included. A 95% exact CI will be provided.

Time-to-Event Endpoints

All time-to-event endpoints will be analyzed by treatment group for the FAS using the Kaplan-Meier method outlined in Section 4.1. For subjects who receive transfusions after Week 5 infusion, the time to normalization will be censored at the study day of the transfusion visit. For subjects who do not experience normalization, the data will be censored at the study day of the last assessment in Part A.

4.5.3 Details on endpoints based on patient-reported outcomes (PROs)

The FACIT-Fatigue, the EQ-5D-5L, the SF-12, the PGIC and PGIS will be utilized in this study. For continuous variables, the overall score and its change from baseline will be analyzed descriptively for each time point.

When applicable missing items will be handled based on the specific patient-reported outcome (PRO) instrument. Subjects missing baseline evaluations would not be included in the PRO analyses.

Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale

The questionnaires will be analyzed according to the recommendations of the questionnaire authors (see http://www.ser.es/wp-content/uploads/2015/03/FACIT-F_INDICE.pdf). Details of the questionnaire and its algorithm for deriving the score are provided in Section 5.5.1. The questionnaires are considered complete, and the subscale total score can be calculated, if at least 7 of the 13 items were answered by the subject. For each time point, the number and percentage of subjects who complete the questionnaires will be summarized.

The change from baseline in FACIT-F score will be summarized by visit for the FAS.

EQ-5D-5L

The questionnaire will be analyzed based on the EQ-5D-5L scoring manual (see www.euroqol.org). EQ-5D consists of 2 parts – a 5 dimension descriptive system and the EQ visual analogue scale. A copy of the questionnaire is provided in Section 5.5.2.

The descriptive system contains 5 domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The number and percentage of subjects in each response category were tabulated for the FAS. Percentages are based on the number of subjects for whom an assessment is provided at the respective visit.

The EQ visual analogue scale is a visual scale from 0–100 to record a respondent's overall self-rated health state. The respondent is asked to mark an "X" on the scale then record the corresponding number; 0 refers to the worst possible health state, 100 refers to the best possible health state. This will also be presented alongside the 5 dimension descriptive system scores.

The EQ visual analogue scale will be summarized for the observed response and change from baseline for the FAS.

EQ-5D Index Score

The EQ-5D-5L descriptive system can be converted into a single index value, the EQ-5D index score, by using the "EQ-5D-5L Crosswalk Index Value Calculator" provided by the scoring manual (documents and calculation tools will be downloaded from the EuroQol website). The index scores, derived as country specific, are a major feature of the EQ-5D instrument, facilitating the calculation of quality-adjusted life years (QALYs) that are used to inform economic evaluations of health care interventions.

The derived EQ-5D index scores and change from baseline will be summarized by visit for the FAS.

Additionally, the EQ-5D index score, will be summarized using an ANCOVA (Analysis of Covariance) model, where treatment arm is the fixed effect, and the baseline score is the covariate; a 95% confidence interval for the treatment effect difference and the corresponding p-value will be generated. The treatment difference and 95% CI will be estimated by ANOVA with treatment and baseline value as independent variables.

12-Item Short Form Survey (SF-12)

The questionnaire will be analyzed according to the recommendations of the authors (see http://campaign.optum.com/optum-outcomes/what-we-do/health-surveys/sf-12v2-health-survey.html). The SF-12 v2 is a 12-item measure, derived from the original 36-item SF-36, which includes eight (8) health domains which are used to generate a physical component score as well as a mental component score. Each score ranges from 0–100 on a normative scale and may be used to derive a health utility index (SF-6D). A copy of the questionnaire is provided in Section 5.5.3.

The derived scores and change from baseline for the eight health domains, the physical component score, and the mental component score, will be summarized by visit for the FAS.

Additionally, the SF-12 component scores (physical component score and mental component score), will be summarized using an ANCOVA (Analysis of Covariance) model, where treatment arm is the fixed effect, and the baseline score is the covariate; a 95% confidence interval for the treatment effect difference and the corresponding p-value will be generated. The treatment

difference and 95% CI will be estimated by ANOVA with treatment and baseline value as independent variables.

Patient Global Impression of Change (PGIC)

The questionnaire (listed in Section 5.5.4) will be summarized as an ordinal variable by visit for FAS at the end of the study only.

Patient's Global Impression Of [Fatigue] Severity (PGIS)

The change from baseline in questionnaire (listed in Section 5.5.5) will be summarized as an ordinal variable by visit for FAS.

Healthcare Resource Utilization

The number of healthcare visits by type (office visit, hospital ER visit, hospitalization, and ICU stay) will be collected and presented in a listing.

4.6 MULTIPLICITY ISSUES

Primary Endpoint

The study primary objective is to compare the treatment effect of BIV009 with placebo in responder rate (responder definition in Section 7.2).

The study is designed to test the primary hypothesis whether patients treated with BIVV009 will achieve a response rate greater than those treated with Placebo. The primary statistical hypothesis is written as:

H0: response rate (BIVV009) ≤ response rate (Placebo)

H1: response rate (BIVV009) > response rate (Placebo),

where the response rate of a treatment group is the proportion of patients in the intent-to-treat (ITT) population who meet the response criteria defined in of the protocol. This will be tested using a 1-sided test at the 0.025 significance level.

Key Secondary Endpoints

Secondary hypotheses, corresponding to the treatment difference for the key secondary endpoints, will be tested only if the primary null hypothesis is rejected at a 1-sided significance level of 0.025.

The key secondary endpoints will be tested using a sequential closed procedure with alpha of 0.05 for each test in the order indicated below:

1. Mean change from baseline in Hgb at the treatment assessment timepoint in BIVV009 arm vs in placebo arm

2. Mean change from baseline in FACIT-F Score at treatment assessment timepoint in BIVV009 arm vs in placebo arm

The following secondary endpoints will be analyzed descriptively:

- Mean change from baseline in bilirubin at the treatment assessment timepoint in BIVV009 arm vs in placebo arm
- Mean change from baseline in Lactate Dehydrogenase (LDH) at the treatment assessment timepoint
- Incidence of Solicited Symptomatic Anemia

4.7 SAFETY ANALYSES

All safety analyses will be performed on the Safety Analysis Set as defined in Section 3, stratified by treatment arm, unless otherwise specified, using the following common rules:

- The analysis of the safety variables will be essentially descriptive, and no testing is planned.
- Safety data in participants who do not belong to the safety population (eg, exposed but not randomized) will be provided.

4.7.1 Extent of exposure

The extent of IMP exposure will be assessed by the duration of IMP exposure and treatment compliance and summarized for the Safety Analysis Set. Per the protocol, all subjects rolling over to Part B are to receive a loading dose (in a blinded manner). For analysis of Part A exposure, this loading dose will not be included in the Part A database. Thus, Part A exposure will end 1 minute prior to the date/time of the Week 26 loading dose, and Part B exposure will start on the date/time of the Week 26 loading dose.

Duration of IMP exposure

Study drug exposure, measured by the duration of study treatment, number of administrations, and total actual BIVV009 dose will be summarized for the Safety Analysis Set.

The duration of the study treatment (in weeks) in Part A is defined as:

- (Date of Week 26 visit – Date of first dose + 1)/7, if a subject enrolls into Part B;
- (Date of last dose – Date of first dose + 15)/7, if a subject discontinues study treatment early in Part A or does not enroll into Part B.

Treatment compliance

Study drug compliance for a subject is measured by the percent of number of doses received out of the number of protocol-specified doses. For example, for a subject who completes Part A treatment, the number of protocol-specified doses is 14. The compliance of the subject will be the

number of doses received divided by 14 expressed in percentage. For a subject who discontinues early in Part A, the number of protocol specified doses is the number of scheduled doses prior to the date of discontinuation. The proportion of subjects whose compliance is less than 80%, 80% – <100%, and ≥100% will also be summarized. In addition, the number of subjects and number of doses administered out of study windows and number of partial infusion doses will also be summarized.

4.7.1.1 Overall exposure

The actual dose administered at each visit is calculated as:

(Total volume administered/Total volume prepared) * assigned dose.

The total BIVV009 dose is the summation of all actual doses administered in Part A.

The number of subjects with missing infusion due to COVID-19 will be presented.

4.7.2 Adverse events

General common rules for adverse events

AEs will be classified using the MedDRA system organ classes and preferred terms. MedDRA version 20.1 or higher will be used throughout the study.

In general, AEs will be analyzed based on *incidence*, defined as the proportion of subjects who had at least one occurrence of an event out of the number of subjects in the Safety Analysis Set. An adverse event listing will be provided and will include all adverse events reported with the onset and resolution study days relative to Day 0 (Study Day 1).

All summary table analyses will only include treatment-emergent adverse events unless noted otherwise.

Unless specified otherwise, system organ classes (SOCs) and preferred terms (PTs) within each SOC will be presented alphabetically. For the purpose of summarization, a subject is counted once in a SOC or preferred term if the subject reported one or more events in that SOC or preferred term. Unless specified otherwise, percentages will be based on the number of subjects in the Safety Analysis Set.

Overall Summary of Adverse Events

An overall summary of TEAEs will be provided which tabulates the number and percentage of subjects who experienced a TEAE, related TEAE, TEAE grade 3 or higher, TEAE infection of grade 3 or higher, TEAE thromboembolic event, TEAE within 24 hours of the start of infusion, TESAE, related TESAE, and TESAE infection; the number and percentage of subjects who discontinued treatment and/or the study due to a TEAE and the number and percentage of subjects who died on-study.

Treatment-emergent Adverse Events

The incidence of TEAEs will be summarized by SOC and preferred term.

Adverse Events in Descending Order of Incidence

A table will be provided which displays TEAE preferred terms in descending order of incidence. Only preferred terms will be included in this table (ie, the display will not include SOCs).

A similar table will be provided for Grade 3 or higher TEAEs. AEs for which the assessment of severity is missing will be included in this table in the TEAE Grade 3 or higher category.

Severity of Adverse Events

AEs are classified by the Investigator for CTCAE grades v4.03. A summary of TEAEs by system organ class, preferred term, and CTCAE grade will be presented. AEs with missing severity will be excluded from this table and listed separately to supplement the summary table. A subject will be counted once for each SOC and preferred term based on the greatest severity within that SOC and preferred term, respectively.

Relationship of Adverse Events to Study Drug

AEs are classified by the Investigator for relationship to study drug ("Not related," "Possible," and "Probable"). For summary tables, an AE with relationship of "Possible" or "Probable" will be considered "Related." A summary of TEAEs by SOC, preferred term, and relationship ("Related" or "Not related") will be presented. AEs with a missing relationship will be counted as "Related" in the summary table. A subject will be counted once for each SOC and preferred term based on the highest relationship within that SOC and preferred term, respectively.

Serious Adverse Events

Any AE reported as resulting in death, immediate risk of death (life threatening), inpatient hospitalization or prolongation of existing hospitalization, persistent or significant disability/incapacity, or a congenital/anomaly/birth defect will be classified as an SAE by the Investigator. An SAE may also be any other medically important event that, in the opinion of the Investigator, may jeopardize the subject or may require intervention to prevent one of the other outcomes listed above.

Treatment-emergent SAEs and treatment-emergent related SAEs will be summarized by system organ class and preferred term. Treatment-emergent SAEs will also be summarized by preferred term in descending order of incidence.

Adverse Events Leading to Treatment Discontinuation and/or Study Withdrawal

AEs leading to treatment discontinuation and/or study withdrawal will be listed. All AEs reported on the AE log for the item "Action Taken with Study Drug" with a response of "Study drug permanently discontinued" or for the item "Was the subject terminated from this study due to this AE?" with a response of "Yes" will be included.

Deaths on Study

A listing of deaths occurring on study will be provided.

Analysis of adverse events of special interest (AESIs)

AEs within 24 Hours after the Start of Infusion

All AEs within 24 hours after the start of infusion are captured in the eCRF. Incidence of AEs within 24 hours after the start of infusion will be summarized by SOC and preferred term.

Infections and Infections of ≥ Grade 3 Severity

Infections will be captured based on MedDRA SOC of "infections and infestations." Incidence of all infections of Grade 3 or higher will be summarized by preferred term.

Hemolytic Breakthrough

Proportion of subjects with hemolytic breakthrough (defined in Section 4.1) at any visit will be summarized. The date (visit) of onset and hemolytic lab parameters (ie, Hgb, bilirubin, LDH, and haptoglobin) over time will be listed for subjects with haemolytic breakthrough.

Anti-drug Antibody

Subjects who discontinue treatment early in Part A will be evaluated for anti-drug antibodies. Subjects with a positive anti-drug antibody will be listed along with the associated titer, if available.

4.7.3 Additional safety assessments

4.7.3.1 Laboratory variables, vital signs and electrocardiograms (ECGs)

Hematology and Chemistry

All summaries will be structured such that the hematology and chemistry tests are presented in the order shown in the tables below.

Hematology measurements that will be summarized and listed include white blood cell count (WBC), red blood cell count (RBC), differentials (basophils, eosinophils, lymphocytes, monocytes, neutrophils, reticulocytes), hemoglobin, hematocrit, and platelet count.

Chemistry measurements that will be summarized and listed include electrolytes (sodium, potassium, chloride), glucose, total protein, total bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), blood urea nitrogen (BUN), and serum creatinine.

Change from Baseline

Hematology and chemistry results at baseline and post baseline visits, along with change from baseline, will be summarized for the Safety Analysis Set.

Shifts from Baseline

Each subject's laboratory values will be classified according to whether the test result is "low" (below the LLN), "normal" (within the normal range), or "high" (above the ULN). Shift tables will be constructed based on both the minimum and maximum post-baseline values for each subject. Data collected from unscheduled visits will be included in the determination of the per subject minimum and maximum values. A separate table will be provided which summarizes the results of the shift tables in which the number and percentage of subjects with a shift to low (from normal, high, or unknown) and the number and percentage of subjects with a shift to high (from normal, low, or unknown) will be tabulated; percentages will be based on the number of subjects at risk. The number at risk for a shift to low (high) is the number of subjects who have a baseline value that isn't low (isn't high) including unknown, and have at least one post-baseline value. Only directions of change indicating a clinical concern will be included in this table summarizing the shifts. The direction of concern is provided in Table 10.

Table 10 - Direction of change indicating clinical concern for laboratory tests

Laboratory Test	Direction	Laboratory Test	Direction
Chemistry		**Hematology**	
Liver		White blood cells	Low and High
ALT/SGPT	High	Lymphocytes	Low and High
AST/SGOT	High	Neutrophils	Low and High
Total bilirubin	High	Monocytes	Low and High
GGT	High	Eosinophils	Low and High
Renal		Basophils	Low and High
Blood urea nitrogen	High	Red blood cells	Low and High
Creatinine	High	Hemoglobin	Low and High
Electrolytes		Hematocrit	Low and High
Sodium	Low and High	Platelets	Low and High
Potassium	Low and High		
Chloride	Low and High		
Other			
Glucose	Low and High		
Total protein	Low and High		

Potentially Clinically Significant Laboratory Abnormalities

Abnormal laboratory values (excluding Hgb, reticulocytes, haptoglobin, bilirubin, and LDH) will also be evaluated by determining the number and percentage of subjects with at least one potentially clinically significant laboratory abnormality over the course of the study that also represents a worsening from baseline. The potentially clinically significant levels are based on Grade 2 or higher thresholds from the CTCAE v4.03 where possible or are defined by the Sponsor's safety group. Subjects who have a post baseline laboratory value that meets the criteria for being potentially clinically significant but do not have a baseline value will be included in the numerator for determining the percentage of subjects with an abnormality. Data collected from unscheduled visits will be included in this analysis. The laboratory tests and thresholds for determining PCSAs are provided in Table 11.

Table 11 - Thresholds for potentially clinically significant laboratory abnormalities

T	Threshold	Laboratory Test	Threshold
Chemistry		**Hematology**	
ALT (U/L)	>= 3 x ULN	Hematocrit (%)	<= 30% >= 60%
Albumin (g/dL)	<= 2.5 g/dL	Hemoglobin (g/L)	<= 100 g/L >= 190 g/L
Alkaline Phosphatase (U/L)	>2.5 x ULN	MCV (fL)	>= 1.2 x ULN <= 0.8 x LLN
AST (U/L)	>= 3 x ULN	Platelet count (10^9/L)	<= 75 x 10^9/L >= 700 x 10^9/L
Blood urea nitrogen (mmol/L)	>= 10.7 mmol/L	WBC (10^9/L)	<3 x 10^9/L >= 16 x 10^9/L
Chloride (mmol/L)	<= 90 mmol/L >= 118 mmol/L	Basophils (10^9/L)	>1.6 x 10^9/L
Creatinine (umol/L)	>1.5 x ULN	Eosinophils (10^9/L)	>1.6 x 10^9/L
GGT (U/L)	>2.5 x ULN	Lymphocytes (10^9/L)	>0.8 x 10^9/L
Glucose (mmol/L)	<3 mmol/L >8.9 mmol/L	Monocytes (10^9/L)	<2.5 x 10^9/L
Potassium (mmol/L)	<= 3 mmol/L >= 6 mmol/L	Neutrophils (10^9/L)	>1.5 x 10^9/L
Sodium (mmol/L)	<= 126 mmol/L >= 156 mmol/L		
Total bilirubin (umol/L)	>1.5 x ULN		
Direct bilirubin (umol/L)	>1.5 x ULN		
Total protein (g/L)	<= 45 g/L >= 100 g/L		

Systemic Lupus Erythematosus (SLE) Panel

SLE panel will be summarized with descriptive statistics. The number of subjects with all negative or all missing values for all SLE panel parameters prior to receiving study drug and the number of subjects with at least one positive value for at least one SLE panel parameter prior to receiving study drug will be summarized. Further, among subjects who have all negative or all missing values for all SLE parameters prior to receiving study drug and have at least one post-baseline value for at least one parameter, the number of subjects who test negative at all post-baseline assessments for all SLE parameters, and the number of subjects who test positive at any post-baseline assessment for at least one SLE parameter will be summarized.

Similar summaries will be generated for individual SLE parameters. In addition, for a given SLE parameter, among subjects who have at least one positive or all missing values prior to receiving study drug and have at least one post-baseline for this parameter, the number of subjects who test positive at all post-baseline assessments, and the number of subjects who test negative at any post-baseline assessment will be summarized.

Parameters in the SLE Panel are Antinuclear antibody (ANA), multiplex with dsDNA, Anti-La/SSB antibody (SS-B), Anti-riboucleoprotein antibody (RNP), Anti-Smith antibody (Sm), Anti-Ro/SSA antibody (SS A), Anti-scleroderma antibody (Scl-70), Anti-Chromatin antibody, Anti-Jo-1 antibody, Anti-Centromere B antibody, Circulating immune complexes (CIC).

Disease-Related Biomarkers

Disease-related biomarkers include DAT (polyspecific, anti-IgG & anti-C3d), LDH isoforms, cold agglutinin (CAg) titre, IgG subsets (IgA, IgD, IgG, IgM), vaccine titres, CAg thermal amplitude. Actual values and change from baseline for all biomarkers will be summarized by visit by treatment arm.

Electrocardiogram (ECG)

The proportion of subjects with clinically significant change from baseline in ECG findings will be summarized by visit.

Vital Signs

Vital signs (systolic and diastolic blood pressure, pulse, respiratory rate, and temperature) will be summarized for the observed values and change from baseline using descriptive statistics for the Safety Analysis Set.

The number and percentage of subjects with potentially clinically relevant post-baseline abnormalities will be presented. The criteria for clinically relevant post-baseline abnormalities are shown below in Table 12.

Table 12 - Criteria to determine clinically relevant abnormalities in vital signs

Vital Sign	Criteria for Abnormalities
Temperature	>38°C and an increase from pre-dosing of at least 1°C
Pulse	>120 beats per minute post-baseline, or an increase from pre-dosing of more than 20 beats per minute, or <50 beats per minute post-baseline, or a decrease from pre-dosing of more than 20 beats per minute
Systolic Blood Pressure	>180 mmHg post-baseline, or an increase from pre-dosing of more than 40 mmHg, or <90 mmHg post-baseline, or a decrease from pre-dosing of more than 30 mmHg
Diastolic Blood Pressure	>105 mmHg post-baseline, or an increase from pre-dosing of more than 30 mmHg, or <50 mmHg post-baseline, or a decrease from pre-dosing of more than 20 mmHg
Respiratory Rate	>35 breaths per minute post-baseline, or <10 breaths per minute post-baseline

4.8 OTHER ANALYSES

4.8.1 PK Analyses

Plasma BIVV009 concentrations will be listed by subject, nominal visit, date and time of collection and study day. Summary statistics of plasma BIVV009 concentrations will be presented by nominal visit and study day including n, mean, standard deviation, coefficient of variation, geometric mean, median and range stratified by dose cohort (6.5 g vs 7.5 g) for the BIVV009 treatment arm. Individual and mean plasma BIVV009 concentration-versus-time profiles will be plotted.

4.8.2 PD Analyses

Descriptive statistics including number of observations, mean, SD, median, minimum, and maximum will be presented for continuous parameters. Categorical variables will be presented with the number and percentage in each category. Summary descriptive statistics (absolute values and changes from baseline) and individual subject listings will be presented for all PD parameters by time point and study day for both the BIVV009 and Placebo treatment arms.

4.9 INTERIM ANALYSES

For the purposes of regulatory submission, an interim analysis of safety and efficacy data (Part A analysis) will be performed for Part A after all patients have completed Part A and the data are cleaned. Parts A and B will have separate database locks to enable submission of the MAA following completion of Part A. Additional analyses of Part B data will be defined in a separate Part B SAP.

Since Part A analysis constitutes the primary analysis of the study hypotheses, no Type I error adjustment is necessary.

5 SUPPORTING DOCUMENTATION

5.1 APPENDIX 1 LIST OF ABBREVIATIONS

AE	adverse event
ADL	activities to daily living
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANA	antinuclear antibodies
AST	aspartate aminotransferase
AUC	area under the curve
BLA	Biologics License Applications
BUN	blood urea nitrogen
CAgD	cold agglutinin disease
CI	confidence interval
CIC	circulating immune complexes
C_{max}	maximum concentration
CTCAE	Common Toxicity Criteria for Adverse Events
DBL	database lock
ECG	electrocardiogram
eCRF	electronic Case Report Form
EOT	end of treatment
EQ-5D-5L	EuroQol – five dimensions questionnaire
ET	early termination
FACIT	functional assessment of chronic illness therapy
FAS	full analysis set
GGT	gamma-glutamyl transferase
Hgb	hemoglobin
HUI	health utility index
ICE	intercurrent events
ICH	International Conference of Harmonisation
ITT	intent-to-treat
IV	intravenous
LDH	lactate dehydrogenase
LLN	lower limit of normal
LPO	last patient out
MAA	Marketing Authorization Application
MAR	missing at random
MCMC	Markov chain Monte-Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
MNAR	Missing Not At Random
NRC	National Research Counsel
PD	pharmacodynamics
PGIC	Patient's Global Impression of Change

PGIS	Patient's Global Impression of Severity
PK	pharmacokinetics
PP	per-protocol
PRO	patient-reported outcome
QOL	quality of life
RBC	red blood cell count
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SF-12	12-Item Short Form Survey
SGOT	serum glutamic oxaloacetic transaminase
SGPT	Serum glutamic-pyruvate transaminase
SLE	systemic lupus erythematosus
SOC	system organ class
TEAE	treatment-emergent adverse event
TLF	tables, figures, and listings
TOEPH	heterogeneous Toeplitz
ULN	upper limit of normal
WBC	white blood cell count

5.2 APPENDIX 2 CHANGES TO PROTOCOL-PLANNED ANALYSES

This section summarizes major statistical changes in the protocol amendment(s).

Table 13 - Major statistical changes in protocol amendment(s)

Amendment Number	Approval Date	Changes	Rationale
Version 1 (original)	24-Aug-2017	Not applicable	Not applicable

5.3 APPENDIX 3 DEMOGRAPHICS AND BASELINE CHARACTERISTICS, PRIOR OR CONCOMITANT MEDICATIONS

Demographics and baseline disease characteristics will be summarized for the FAS.

Demography

Demographic characteristics including age (years), age category (<65, >=65), gender, race, ethnicity, country, geographic location, height (cm), weight (kg), weight category (<75 kg, >=75 kg), and body mass index (kg/m^2) will be summarized. Geographic locations are defined as Europe, North America, Asia/Japan, and other.

General Medical History

Medical history by system organ class (SOC) and preferred term (PT) will be summarized. A subject will be counted only once if they reported one or more occurrences in the same SOC or PT.

A summary of prior hospitalizations related to CAD within the last 2 years, prior hematological malignancies within the last 5 years, prior non-malignant hematologic diseases within the last 5 years, and prior thromboembolic events within the last 1 year will be presented based off the "Targeted Medical History" eCRF form.

Baseline Disease Characteristics

Screening and baseline lab parameters, including hematology (such as Hgb, bilirubin, LDH), Gilbert's Syndrome testing, and bone marrow biopsy results will be summarized. Prior transfusion history (eg, number of transfusions within 12 months of study entry) will also be summarized.

Severity of anemia symptoms and other CAD-related disease characteristics (eg, presence of acrocyanosis, Raynaud's syndrome, hemoglobinuria, disabling circulatory symptoms, and major adverse vascular events) within 3 months of screening will be summarized.

The proportion of subjects who had documented vaccines within 5 years and who have received vaccination during the study will be summarized. The vaccines include meningococcal vaccine, S. pneumoniae vaccine, and H. influenza vaccine.

Prior or concomitant medications

Prior medications are those taken before the first dose of study drug in Part A. Concomitant medications are those administered on or after the first dose of study drug, or those administered before the first dose that are ongoing when study treatment begins. Prior and concomitant medications will be coded using World Health Organization drug enhanced dictionary (March 2020 version or higher).

Separate tables will be generated to summarize concomitant medications and concomitant protocol-prohibited medications by ATC level 2 and WHODrug standardized medication text.

Additionally, a summary of targeted prior therapies will be generated based off the responses provided in the "Targeted Prior Therapy" eCRF form.

Concomitant non-drug treatments and procedures

Concomitant non-drug treatments and procedures will be summarized for the FAS.

5.4 APPENDIX 4 DATA HANDLING CONVENTIONS

Study Day

Study Day is defined as days relative to the date of first dose of study drug (BIVV009 or placebo) in Part A (Day 0 as in the protocol). The Study Day of an event is calculated as

- (date of event – date of first dose of study drug in Part A + 1), if the event is on or after the first dose date;
- (date of event – date of first dose of study drug in Part A), if the event is prior to the first dose date.

The date of the first dose of study drug in Part A is Study Day 1.

Baseline

Baseline measures are defined as the last value prior to the first dose of study drug. If a subject receives a transfusion during the screening period, the baseline measure must be at least 7 days after the transfusion.

Visit Windows

Analysis by study visit will utilize the data recorded on the corresponding nominal visit from electronic Case Report Forms (eCRFs). In case of missing data/assessments for a visit, data from an unscheduled visit, or ET/Safety Follow-up Visit will be used, if the Study Day of such a visit fits in the analytic window defined in Table 14, below. If multiple values are identified within the analytic window, the one closest to the target study day will be used.

Table 14 - Analyses window definition

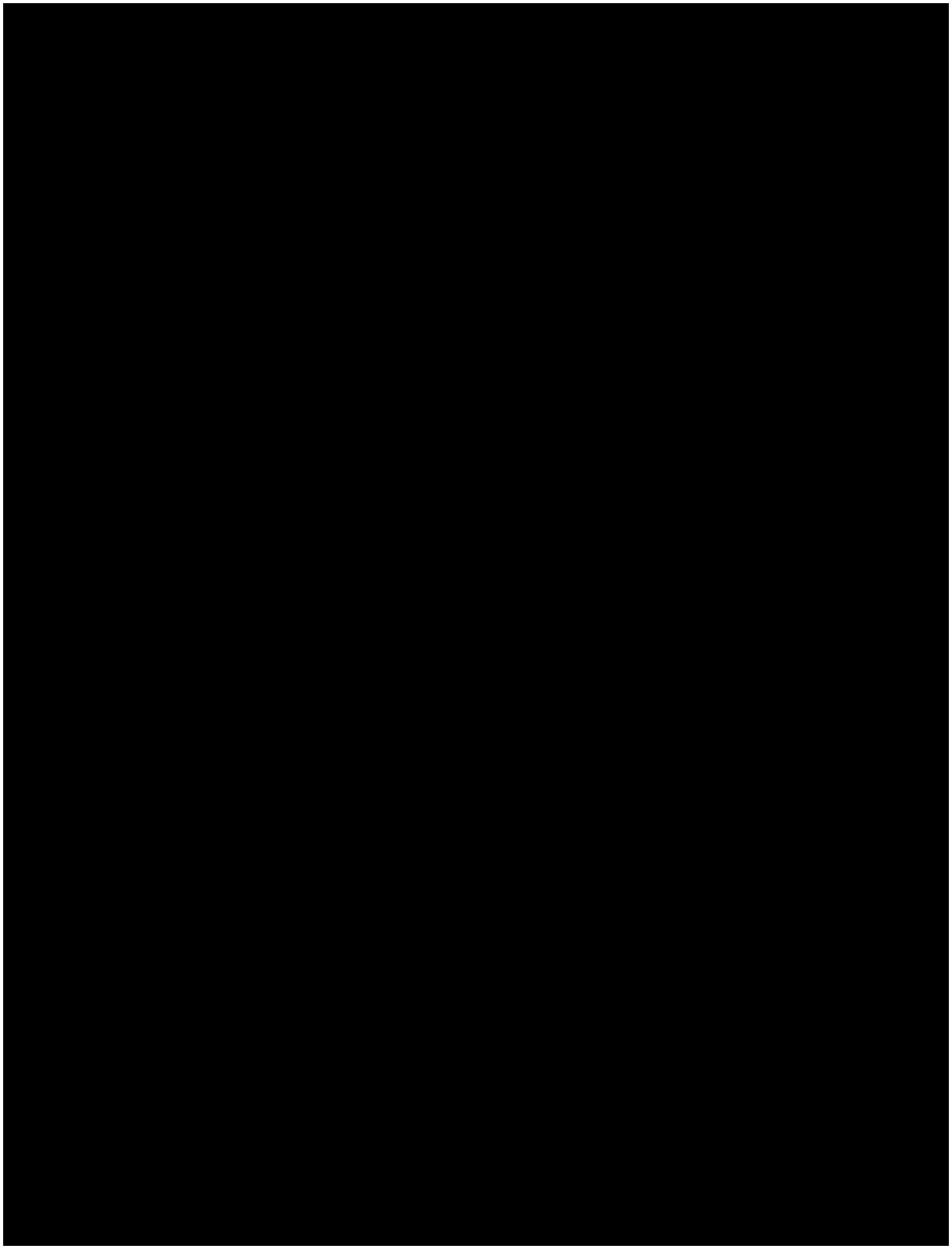
Visit	Target Study Day	Analytic Visit Window
Baseline (Day 0)	1	≤1
Week 1 (Day 7)	8	2 – 15
Week 3 (Day 21)	22	16 – 29
Week 5 (Day 35)	36	30 – 43
Week 7 (Day 49)	50	44 – 57
Week 9 (Day 63)	64	58 – 71
Week 11 (Day 77)	78	72 – 85
Week 13 (Day 91)	92	86 – 99
Week 15 (Day 105)	106	100 – 113
Week 17 (Day 119)	120	114 – 127
Week 19 (Day 133)	134	128 – 141
Week 21 (Day 147)	148	142 – 155
Week 23 (Day 161)	162	156 – 169
Week 26 (Day 175	176	170 – 179
Week 26 (Day 182; EOT Visit in Part A)	183	180 – 189

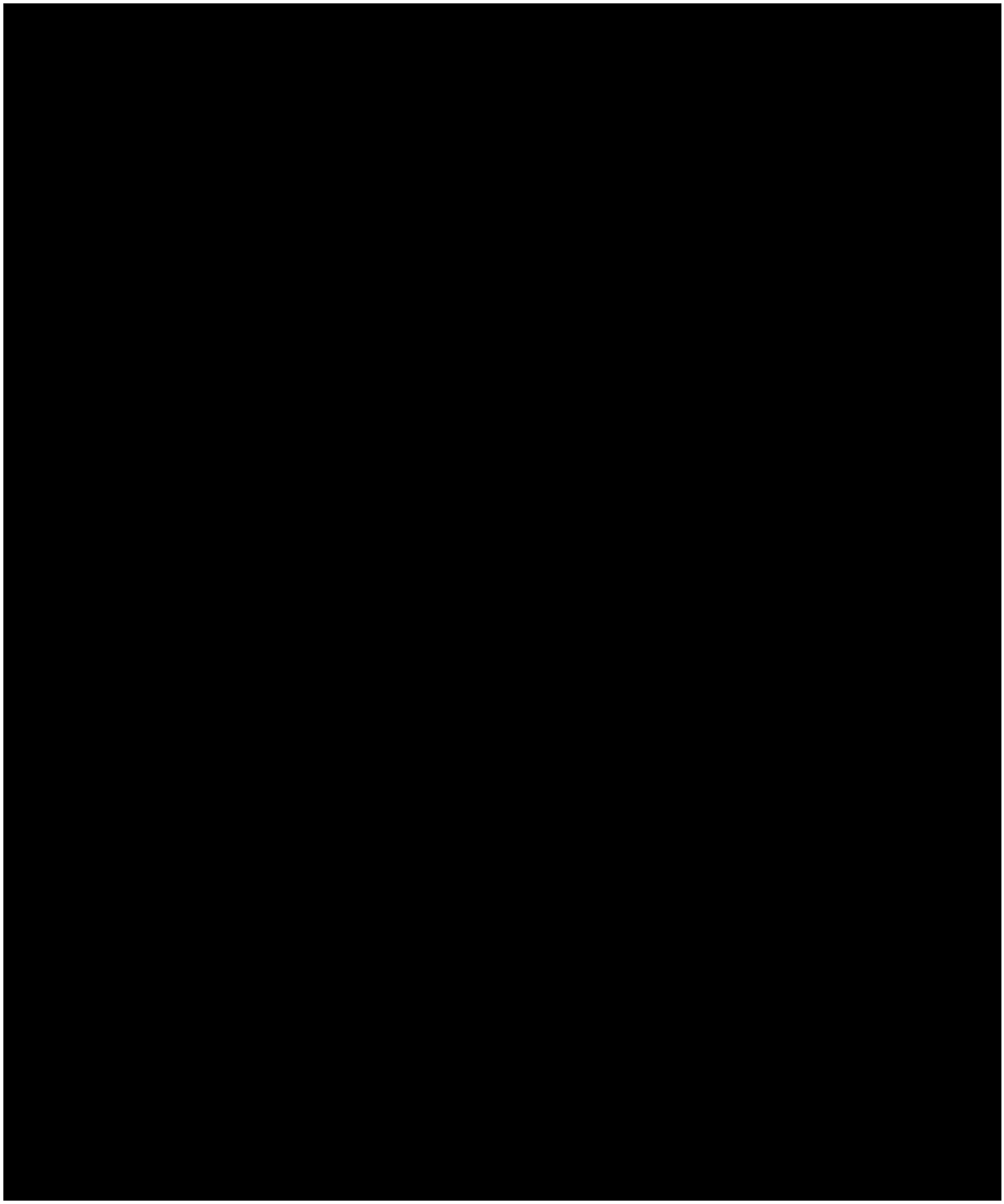
5.5 APPENDIX 5 PATIENT-REPORTED OUTCOMES

5.5.1 Scoring algorithm for FACIT-F

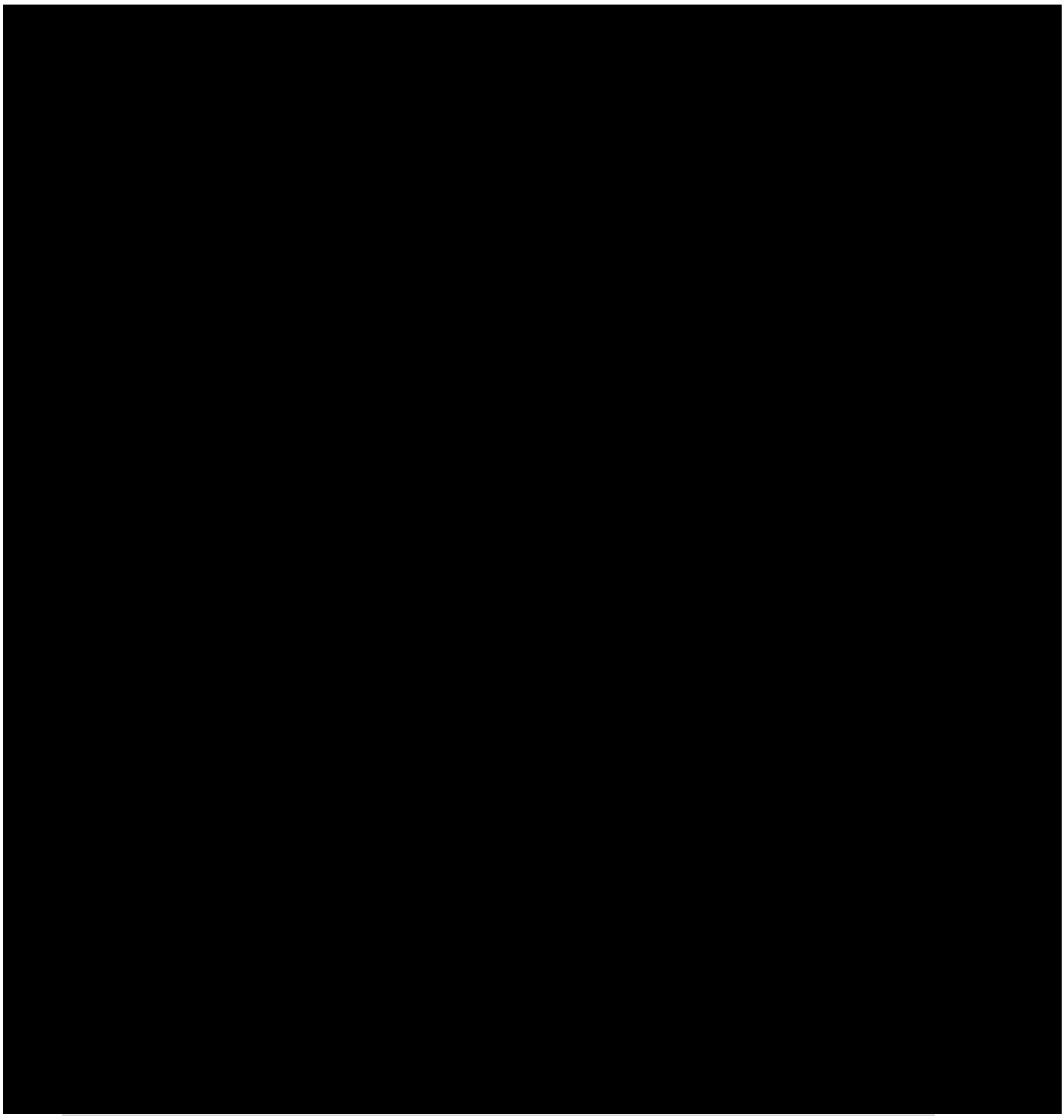
Figure 1 - FACIT-Fatigue Scale

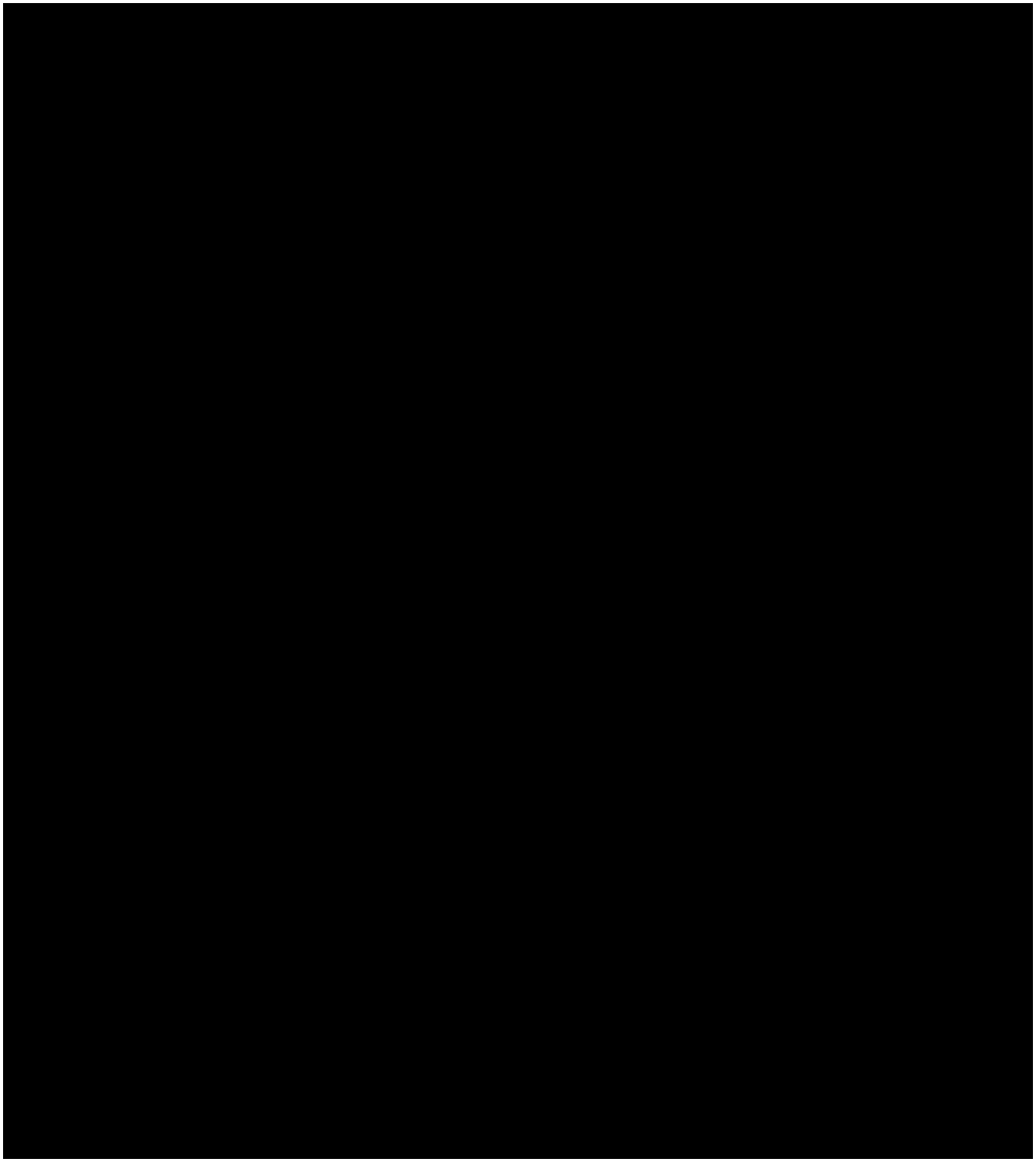
5.5.2 EQ-5D-5L questionnaire score definitions

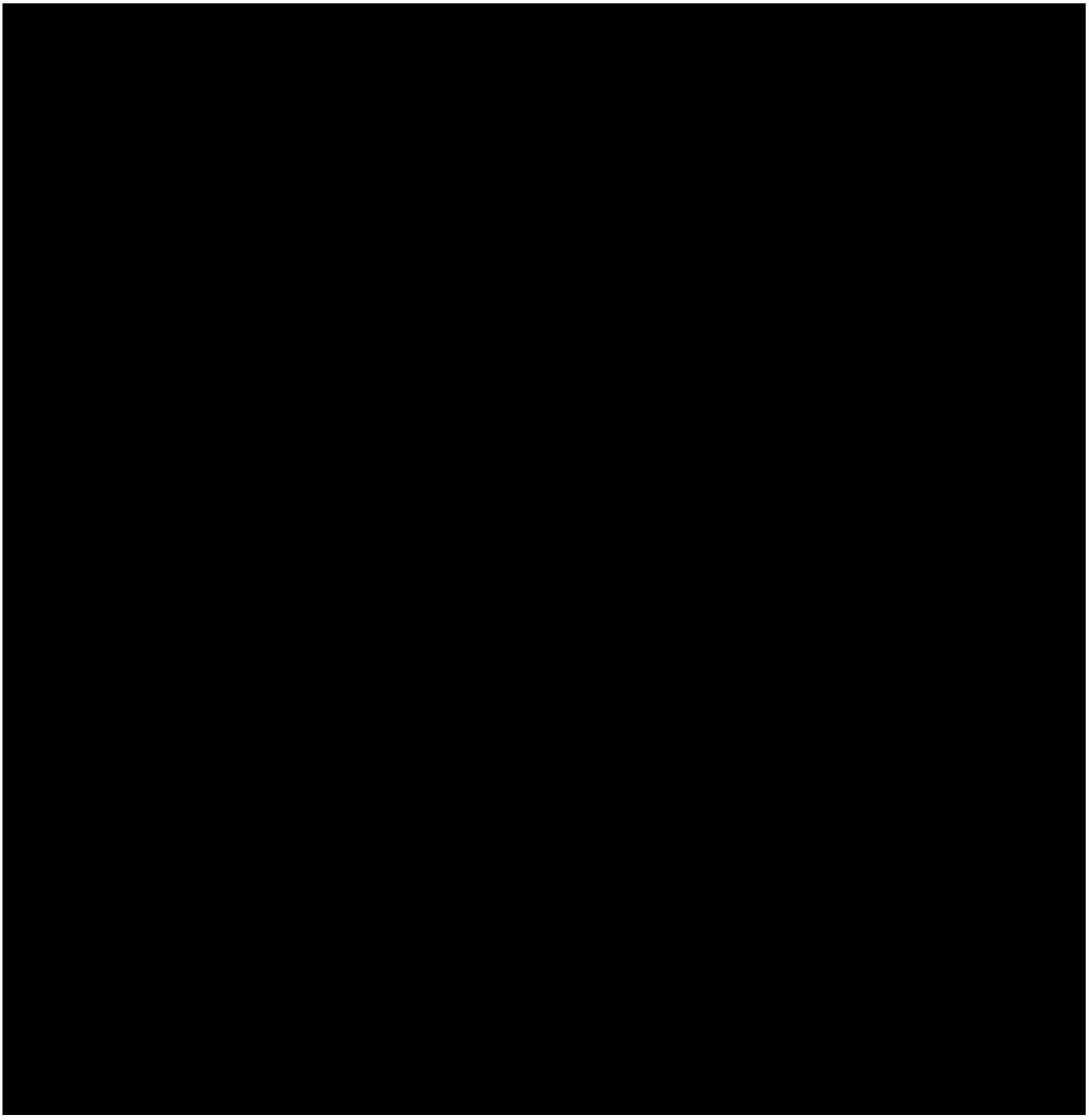




5.5.3 12-item short form survey (SF-12)







5.5.4 Patient's global impression of change (PGIC)

[redacted]

6 REFERENCES

1. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Final Concept Paper: E9(R1): Addendum to Statistical Principles for Clinical Trials on Choosing Appropriate Estimands and Defining Sensitivity Analyses in Clinical Trials. 2014.

2. National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials. Washington DC: National Academies Press. 2010.

3. Akacha M, Bretz F, Ohlssen D, Rosenkranz G, Schmidli H. Estimands and Their Role in Clinical Trials. Stat Biopharm Res. 2017;9(3):268-71.

Signature Page for VV-CLIN-0586110 v2.0

bivv009-04-16-1-9-sap-part-a

Approve & eSign	[redacted]
Approve & eSign	

STATISTICAL ANALYSIS PLAN

Protocol title:	**A Phase 3, randomized, double-blind, placebo-controlled study to assess the efficacy and safety of BIVV009 in patients with primary Cold Agglutinin Disease without a recent history of blood transfusion**
Protocol number:	**BIVV009-04 Part B**
Compound number (INN/Trademark):	**BIVV009**
Study phase:	**Phase III Extension**
Short title:	**Cadenza**
Statistician:	[redacted]
Statistical project leader:	[redacted]
Date of issue:	**11-Aug-2020**

Regulatory agency identifier number(s):

Registry:

Enter Registry Name:

Total number of pages: 29

TABLE OF CONTENTS

LIST OF TABLES

VERSION HISTORY

This statistical analysis plan (SAP) for study BIVV009-04 (Cadenza) Part B is based on the protocol dated 07-Jul-2020. There are no major changes to the statistical analysis features in this SAP.

The first participant was enrolled into Part B on 27-Nov-2018. This SAP is approved before the first interim analysis is conducted.

Table 1 - Major changes in statistical analysis plan

SAP Version	Approval Date	Changes	Rationale
1.0	20-Jul-2020	Not Applicable	Original version

1 INTRODUCTION

1.1 STUDY DESIGN

BIVV009-04 is a Phase 3, randomized, double-blinded, placebo-controlled, multicenter study in patients with primary cold agglutinin disease (CAD) without a recent history of blood transfusion. Eligible patients received study drug (BIVV009 or placebo) and underwent safety and efficacy assessments for 6 months (26 weeks) in Part A of the study.

Following completion of the initial 6-month treatment period, patients will roll into the long-term safety study (Part B) where they will continue to receive the study drug, BIVV009. After completion of all Part A EOT assessments, patients who qualify to roll-over into Part B will receive a cross-over loading dose, in a blinded manner, at the Week 26 visit. Patients who were randomized to placebo during the 6-month treatment period will receive BIVV009; and patients who were randomized to BIVV009 during the 6-month treatment period will receive a placebo dose in order to maintain the blinding. Part B will run for one year following last patient out (LPO) under Part A. In Part B, patients will be dosed with open-label BIVV009 every two weeks, as in Part A. On-site visits will be completed every two weeks when samples for safety and efficacy measures will be collected; samples for pharmacokinetics (PK) and pharmacodynamics (PD) will be collected at visits approximately every three months. Sample for antidrug antibodies will be collected every three months (at a minimum). The study will be complete 12 months following LPO under Part A at which time all patients receiving on-going treatment will proceed to an End-of-Study (EOS) visit 9 weeks after the last administration of study drug.

After Week 39, a subgroup of patients will have their study drug infused at home, every four weeks, alternating with study drug administration during office visits. Additionally, after Week 39, another subgroup of patients will have their study drug administered in the form of infusions undiluted with normal saline.

Study primary analysis will be conducted for Part B after all patients have completed Part B and the data are cleaned.

This statistical analysis plan contains information pertaining to definitions of analysis sets and derived variables, and statistical methods for the analysis of efficacy, safety, PK, and PD for Part B of the referenced study.

1.2 OBJECTIVE AND ENDPOINTS

Table 2 - Objectives and endpoints

	Objectives	Endpoints
Primary	• The primary objective of Part B is to evaluate the long-term safety and tolerability of BIVV009 in patients with primary CAD	Safety Endpoints: • Incidence of treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs) • Change from baseline in clinical laboratory evaluations • Change from baseline in systemic lupus erythematosus (SLE) panel • Change from baseline in vital signs • Change from baseline in electrocardiograms (ECG) data • Physical examination findings • Incidence of infections of ≥ Grade 3 severity (ie, requiring intravenous [IV] antibiotics) • Serum disease-related biomarkers • Incidence of hemolytic breakthrough through EOT • Incidence of thromboembolic events
Secondary	• The secondary objective of Part B is to investigate the durability of response during long-term treatment with BIVV009 in patients with primary CAD.	The following parameters of disease activity will be assessed: • Hemoglobin • Bilirubin • QOL assessments (FACIT-Fatigue, EQ-5D-5L, SF-12, PGIS, and PGIC) • LDH • Transfusion requirements • Haptoglobin • Total healthcare resource utilization at EOT
Tertiary/exploratory	• To describe the safety and patient satisfaction with the convenience of home infusions with BIVV009 in a subset of patients	For subjects receiving home infusions: • Satisfaction with home infusion after first home infusion and after fourth home infusion • Incidence of AEs with onset within 24 hours of a home infusion
	• To describe the safety of undiluted infusions with BIVV009	For subjects receiving undiluted infusions: • Incidence of AEs with onset within 24 hours of an undiluted infusion
	• To evaluate the immunogenicity of BIVV009	• Anti-drug antibodies

Objectives	Endpoints
• Pharmacokinetics of BIVV009	• Plasma concentrations of BIVV009 • PK parameters: Appropriate exposure parameters (Cmax, AUC) will be derived using a population PK approach
• Pharmacodynamics of BIVV009	• Weislab-CP • CH50 • Total C4 • C1q • C1s

1.2.1 Estimands

There are no primary efficacy endpoints in Part B. Thus, estimands will not be utilized in this SAP.

2 SAMPLE SIZE DETERMINATION

Patients in Part B will be a subset of the patients enrolled in Part A. All patients who complete Part A, including those receiving transfusions, will be eligible for Part B. Patients who withdraw from the study in Part A, including those who received prohibited concomitant medications, will not be eligible to participate in Part B. Subject inclusion and exclusion criteria for Part A can be found in Sections 5.1 and 5.2 of the protocol.

No power and sample size analysis was conducted for Part B. Patients from Part A were enrolled into Part B following Part A EOT. The power and sample size calculation for Part A can be found in the Part A SAP.

3 ANALYSIS POPULATIONS

The important protocol deviations defined in Part A do not apply to Part B; and thus, the Per-Protocol Population described in the protocol was determined to have little utility and has not been included in this SAP.

The following populations for analyses are defined:

Table 3 - Populations for analyses

Population	Description
Full Analysis Set	All subjects who enroll into Part B and receive at least 1 dose (including partial dose) of study drug. The Full Analysis Set is synonymous with the Intent-to-Treat Population described in the protocol.
Safety Analysis Set	All subjects who receive at least 1 dose (including partial dose) of study drug in Part B. Note that the Safety Analysis Set is the same as the Full Analysis Set in this study.
PK Analysis Set	All subjects who receive at least 1 dose of study drug in Part B and have evaluable PK concentrations.
PD Analysis Set	All subjects who receive at least 1 dose of study drug in Part B and have at least 1 evaluable PD sample.

4 STATISTICAL ANALYSES

Part B is an open-label study with no blinding or randomization. No formal statistical hypotheses will be tested. Analyses of efficacy endpoints will be primarily descriptive.

Statistical analysis will be performed by Sanofi, using SAS 9.4 or higher and, where appropriate, additional validated software.

4.1 GENERAL CONSIDERATIONS

Safety, efficacy, and PK/PD data will be summarized using standard summary statistics for continuous, categorical, and time-to-event data. Data will be summarized for the populations defined.

No statistical hypothesis testing is planned.

Continuous Variables

Continuous variables will be summarized using descriptive statistics including the number of non-missing values (n), mean, standard deviation (SD), median, minimum and maximum. Where specified in the table shells, the 25th and 75th percentiles will also be provided. Means, medians, and the 25th and 75th percentiles will be presented to one decimal place beyond that with which the data were captured. SDs will be presented to two decimal places beyond that with which the data were captured. Minimum and maximum will be displayed to the same number of decimal places as that with which the data were captured.

Unless impractical within a given table, statistics will be aligned by the decimal place (or assumed decimal place) in the summary tables.

Categorical Variables

Categorical variables will be summarized by counts and percentages. All percentages will be rounded to one decimal place. The percentage will be suppressed when the count is zero. Unless specified otherwise, the denominator for all percentages will be the number of subjects with non-missing data for a given summarization.

Part B Study Treatment Period

For subjects who enroll into Part B, the Part A study treatment period starts at the date of first study drug administration and ends at the date of Week 26 visit. The Part B study treatment period starts following completion of Part A (ie, immediately following Week 26 visit) and ends at the ET/EOS/Safety follow-up visit, which is 9 weeks after each subject's last dose.

Combined Study Period

The Combined Study Period includes all visits in Part A and Part B and will be utilized for a subset of analyses in order to support regulatory submissions and/or publications.

Thromboembolic Events

Thromboembolic events prior to Part A study entry are recorded as medical history in the eCRFs, while thromboembolic events on Part A and Part B study are captured as adverse events. Both medical history and adverse events will be coded according to Medical Dictionary for Regulatory Activities (MedDRA) v23.0 or higher.

All thromboembolic events reported (as AEs) will be reviewed and determined by internal medical staff via medical adjudication.

Hemolytic Breakthrough

Hemolytic breakthrough at any visit is defined as

- A decrease of 2 g/dL or more in Hgb from the last scheduled visit, and
- One of the followings:
  - An increase of LDH from the last scheduled visit, or
  - An increase of bilirubin from the last scheduled visit, or
  - A decrease in haptoglobin from the last scheduled visit.

4.2 PARTICIPANT DISPOSITIONS

Subject disposition will be summarized for all enrolled subjects in Part B by treatment arm in Part A (ie, BIVV009 and Placebo) and Total. Subject disposition will be summarized including the number of subjects in FAS, Safety Analysis Set, PK Analysis Set, and PD Analysis Set. The number of subjects with the status of completed Part B and discontinued Part B early, including the primary reason for those who discontinued, will be tabulated for all subjects enrolled in Part B.

Protocol deviations

All protocol deviations will be recorded throughout the study. Major and minor protocol deviations/violations are to be pre-specified prior to database lock. All major protocol deviations will be summarized for FAS by treatment arm in Part A (ie, BIVV009 and Placebo) and Total.

4.3 PRIMARY ENDPOINT(S) ANALYSIS

The primary objective of Part B is to evaluate the long-term safety and tolerability of BIVV009 in patients with primary CAD. Thus, all endpoints related to this objective and the analyses for these endpoints will be described, in full, in Section 4.7.

4.4 SECONDARY ENDPOINT(S) ANALYSIS

Unless specified otherwise, all analyses described will be performed using the FAS. All summaries of efficacy endpoints will be descriptive and will not test any formal hypotheses. All summaries will be tabulated by treatment arm in Part A (ie, BIVV009 and Placebo) and Total.

4.4.1 Key/Confirmatory secondary endpoint(s)

4.4.1.1 Definition of endpoint(s)

Continuous endpoints

The continuous endpoints are:

- Mean change from baseline in hemolytic parameters including
  - Hemoglobin
  - Total bilirubin
  - LDH
  - Haptoglobin
- Mean change from baseline in FACIT-Fatigue score
- Mean change from baseline in QOL, as assessed by the change in the five-level EQ-5D-5L scores by visit
- Mean change from baseline in QOL, as assessed by the change in the 12-item short form survey (SF-12 ®)
- Number of transfusions and total transfusion units
- Annualized number of transfusions and annualized transfusion units

Categorical Endpoints

The categorical endpoints include:

- Incidence of thromboembolic events
- Incidence of hemolytic breakthrough
- PGIC
- Change from baseline in PGIS

4.4.1.2 Main analytical approach

Continuous Endpoints

The continuous endpoints will be summarized by visit and treatment arm using descriptive statistics for the FAS according to Section 4.1. Change from baseline will also be summarized by visit.

Line plots of hemoglobin, bilirubin, and FACIT-F score over time, starting at Baseline/Day 0, will be presented. Additionally, line plots of mean change from baseline in hemoglobin, bilirubin, and FACIT-F score over time, starting at Baseline/Day 0, will be presented.

The endpoints related to transfusions and transfusion units will be summarized by study period using the FAS.

Annualized values will be calculated for each patient using the following formula:

$$\text{Annualized value} = \frac{\text{Number of events during the Study Period}}{\text{Total number of days during the Study Period}} \times 365.25$$

Response (Categorical) Endpoints

The categorical endpoints will be summarized descriptively by visit and treatment arm for the using the FAS according to Section 4.1 with the exception of total healthcare resource utilization at EOT (Part A), which will be presented in a listing.

All thromboembolic events reported (in Adverse Events) will be reviewed and determined by internal medical staff along with investigators. In case that there are discrepancies between the medical review and investigators' assessment, separate analyses will be carried out for both sets of events. Date of event occurrences will be verified. The incidence of thromboembolic events and hemolytic breakthrough events will be summarized using all data in Part B (including data from the Week 79 visit and beyond).

4.4.1.3 Details on endpoints based on patient-reported outcomes (PROs)

The patient-reported outcome (PRO) endpoints are as follows:

- Continuous Endpoints
  - FACIT-Fatigue score
  - EQ-5D index score
  - EQ-5D-5L visual analogue scale
  - SF-12 physical component score (PCS) and mental component score (MCS)
  - SF-12 sub-scale scores
- Categorical Endpoints
  - Patient's Global Impression of Change (PGIC)
  - Patient's Global Impression of [Fatigue] Severity (PGIS)
  - EQ-5D-5L Descriptive System

Details of the above PRO instruments and algorithms for deriving the scores can be found in the Part A SAP.

When applicable missing items will be handled based on the specific PRO instrument. Subjects missing baseline evaluations would not be included in the PRO analyses.

Healthcare Resource Utilization

The number of healthcare visits by type (office visit, hospital ER visit, hospitalization, and ICU stay) will be collected and presented in a listing.

4.5 TERTIARY/EXPLORATORY ENDPOINT(S) ANALYSIS

4.5.1 Definition of endpoint(s)

Per the protocol, home infusions will be performed at pre-selected countries/sites for a select number of subject who meet the criteria specified in Appendix K of the protocol. For these subjects, a survey will be conducted after the first and fourth home infusion to assess patient satisfaction with the home infusion.

4.5.2 Main analytical approach

The patient satisfaction survey will be summarized descriptively for the subset of subjects who receive home infusions.

4.6 MULTIPLICITY ISSUES

There are no multiplicity issues for Cadenza Part B.

4.7 SAFETY ANALYSES

All safety analyses will be performed on the Safety Analysis Set, stratified by dose cohort, unless otherwise specified, using the following common rules:

- The analysis of the safety variables will be essentially descriptive, and no testing is planned.
- Safety data in participants who do not belong to the safety population (eg, exposed but not randomized) will be provided.

4.7.1 Extent of exposure

The extent of IMP exposure will be assessed by the duration of IMP exposure and treatment compliance and summarized for the Safety Analysis Set. Per the protocol, subjects rolling over to Part B are to receive a loading dose (in a blinded manner). For analysis of Part B exposure, this loading dose will be included in the Part B database. Thus, Part B exposure will start on the date/time of the Week 26 loading dose.

Duration of IMP exposure

Study drug exposure, measured by the duration of study treatment, the number of BIVV009 administrations and total actual BIVV009 dose, will be summarized by study period for the Safety Analysis Set by treatment cohort in Part A (ie, BIVV009 and Placebo) and Total.

The duration of the study treatment (in weeks) is calculated as follows:

1. Combined Study Period: (date of last dose – date of first dose in Part A + 15)/7
2. Part B Study Period: (date of last dose – Part B start date + 15)/7

Treatment compliance

Study drug compliance for a subject is measured by the percent of number of doses received out of the number of protocol-specified doses. The proportion of subjects whose compliance is less than 80%, 80% - <100%, and ≥100% will also be summarized. In addition, the number of subjects and number of doses administered out of study windows and number of partial infusion doses will be summarized.

4.7.1.1 Overall exposure

The actual dose administered at each visit is calculated as:

Total volume administered / Total volume prepared * assigned dose.

The total BIVV009 dose will be summarized as follows:

1. Combined Study Period: The summation of all actual doses administered in Parts A and B.
2. Part B Study Period: The summation of all actual doses administered in Part B, alone.

The number of subjects with missing infusion due to COVID-19 will be presented.

4.7.2 Adverse events

General common rules for adverse events

AEs will be classified using the MedDRA system organ classes and preferred terms. MedDRA version 23.0 or higher will be used throughout the study.

In general, AEs will be analyzed based on *incidence*, defined as the proportion of subjects who had at least one occurrence of an event out of the number of subjects in the Safety Analysis Set.

All summary tables and listings will only include treatment-emergent adverse events in Part B unless noted otherwise.

Unless specified otherwise, system organ classes and preferred terms within each SOC will be presented alphabetically. For the purpose of summarization, a subject is counted once in a SOC or preferred term if the subject reported one or more events in that SOC or preferred term. Unless specified otherwise, percentages will be based on the number of subjects in the Safety Analysis Set.

Overall Summary of Adverse Events

An overall summary of TEAEs will be provided which tabulates the number and percentage of subjects who experienced a TEAE, related TEAE, treatment-emergent SAE, treatment-emergent related SAE, treatment-emergent Grade 3 or higher AE, treatment-emergent infections of Grade 3 or higher severity, TEAE thromboembolic event, TEAE autoimmune event, TEAE potential hypersensitivity event, and TEAEs within 24 hours after the start of infusion; TEAEs within 24 hours after the start of home infusion of BIVV009 and TEAEs within 24 hours after the start of undiluted infusion of BIVV009; the number and percentage of subjects who discontinued treatment and/or the study due to an AE; and the number and percentage of subjects who died. In addition, the number of TEAEs, TESAEs, TEAE thromboembolic events and TEAEs within 24 hours after the start of infusion will be presented.

Treatment-emergent Adverse Events

The incidence of TEAEs will be summarized by SOC and preferred term.

Adverse Events in Descending Order of Incidence

Tables will be provided which display TEAE and TESAE preferred terms in descending order of incidence. Only preferred terms will be included in these tables (ie, the display will not include SOCs).

A similar table will be provided for Grade 3 or higher TEAEs. AEs for which the assessment of severity is missing will be included in this table.

Severity of Adverse Events

AEs are classified by the Investigator for CTCAE grades v4.03. A summary of TEAEs by system organ class, preferred term, and CTCAE grade will be presented. AEs with a missing severity will be excluded from this table and listed separately to supplement the summary table. A subject will be counted once for each SOC and preferred term based on the greatest severity within that SOC and preferred term, respectively.

Relationship of Adverse Events to Study Drug

AEs are classified by the Investigator for relationship to study drug (“Not related,” “Possible,” and “Probable”). For summary tables, an AE with relationship of “Possible” or “Probable” will be considered “Related.” A summary of TEAEs by SOC, preferred term, and relationship (“Related” or “Not related”) will be presented. AEs with a missing relationship will be counted as “Related” in the summary table. A subject will be counted once for each SOC and preferred term based on the highest relationship within that SOC and preferred term, respectively.

Serious Adverse Events

Any AE reported as resulting in death, immediate risk of death (life threatening), inpatient hospitalization or prolongation of existing hospitalization, persistent or significant disability/incapacity, or a congenital/anomaly/birth defect will be classified as an SAE by the

Investigator. An SAE may also be any other medically important event that, in the opinion of the Investigator, may jeopardize the subject or may require intervention to prevent one of the other outcomes listed above.

Treatment-emergent SAEs and treatment-emergent related SAEs will be summarized by system organ class and preferred term. Treatment-emergent SAEs will also be summarized by preferred term in descending order of incidence.

Adverse Events Leading to Treatment Discontinuation and/or Study Withdrawal

AEs leading to treatment discontinuation and/or study withdrawal will be listed. All AEs reported on the AE log for the item "Action Taken with Study Drug" with a response of "Study drug permanently discontinued" or for the item "Was the subject terminated from this study due to this AE?" with a response of "Yes" will be included.

Death on Study

A listing of deaths occurring on study will be provided.

Analysis of adverse events of special interest (AESIs)

AEs within 24 Hours after the Start of Site Infusion

All AEs within 24 hours after the start of site infusion are captured in the eCRF. Incidence of AEs within 24 hours after the start of site infusion will be summarized by SOC and preferred term.

AEs within 24 Hours after the Start of Home Infusion

All AEs within 24 hours after the start of home infusion are captured in the eCRF. Incidence of AEs within 24 hours after the start of home infusion will be summarized by SOC and preferred term. AEs within 24 hours after the start of home infusion will be flagged on the listing of adverse events.

AEs within 24 Hours after the Start of Undiluted Infusion

All AEs within 24 hours after the start of undiluted infusion are captured in the eCRF. Incidence of AEs within 24 hours after the start of undiluted infusion will be summarized by SOC and preferred term. AEs within 24 hours after the start of undiluted infusion will be flagged in the listing of adverse events.

Infections and Infections of ≥ Grade 3 Severity

Infections will be captured based on MedDRA SOC of "infections and infestations." Incidence of infections of Grade 3 or higher will be summarized by preferred term. Listings of all infections, infections of Grade 3 or higher and serious infections will be provided.

Autoimmune Events

Treatment-emergent autoimmune events will be adjudicated and flagged on the Listing of adverse events.

Hypersensitivity Events

Treatment-emergent hypersensitivity events will be adjudicated and flagged on the Listing of adverse events.

Anti-drug Antibody

Subjects will be evaluated for anti-drug antibodies at Baseline and at the ET/EOS/Safety Follow-up visits in Part B. Subjects with negative or positive anti-drug antibodies at baseline and post baseline, as well as the incidence and prevalence will be summarized for the Combined Study Period by dose cohort in Part B (ie, 6.5 g and 7.5 g) and Total. A listing will be provided for subjects with negative or positive anti-drug antibodies along with the associated titer if available. If applicable, TEAEs for those subjects will be included in the same listing.

4.7.3 Additional safety assessments

4.7.3.1 Laboratory variables, vital signs and electrocardiograms (ECGs)

Hematology and Chemistry

All summaries will be structured such that the hematology and chemistry tests are presented in the order shown in the tables below. All hematology and chemistry lab summaries will be analyzed by dose cohort in Part B (ie, 6.5 g and 7.5 g) and Total using the Safety Analysis Set.

Hematology measurements that will be collected and listed include white blood cell count (WBC), red blood cell count (RBC), differentials (basophils, eosinophils, lymphocytes, monocytes, neutrophils, reticulocytes), hemoglobin, hematocrit, and platelet count.

Chemistry measurements that will be collected and listed include electrolytes (sodium, potassium, chloride), glucose, total protein, total bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), blood urea nitrogen (BUN), and serum creatinine.

Change from Baseline

Hematology and chemistry results at baseline and Part B post-baseline visits, along with change from baseline, will be summarized with descriptive statistics. These by-visit summaries will be presented through the end of Part B.

Shifts from baseline

Each subject's laboratory values will be classified according to whether the test result is "low" (below the LLN), "normal" (within the normal range), "high" (above the ULN). Shift tables will

be constructed based on both the minimum and maximum post-baseline values for each subject. A separate table will be provided which summarizes the results of the shift tables in which the number and percentage of subjects with a shift to low (from normal, low, or unknown) and the number of subjects with a shift to high (from normal, low, or unknown) will be tabulated; percentages will be based on the number of subjects at risk. The number at risk for a shift to low (high) is the number of subjects whose baseline value was not low (high), including unknown, who had at least one post-baseline value. Only directions of change indicating a clinical concern will be included in this table summarizing the shifts. The direction of concern is provided in Table 4.

All data collected during Part B, including data collected from unscheduled visits, will be included in the above-mentioned analyses.

Table 4 - Direction of change indicating clinical concern for laboratory tests

Laboratory Test	Direction	Laboratory Test	Direction
Chemistry		**Hematology**	
Liver		White blood cells	Low and High
ALT/SGPT	High	Lymphocytes	Low and High
AST/SGOT	High	Neutrophils	Low and High
Total bilirubin	High	Monocytes	Low and High
GGT	High	Eosinophils	Low and High
Renal		Basophils	Low and High
Blood urea nitrogen	High	Red blood cells	Low and High
Creatinine	High	Hemoglobin	Low and High
Electrolytes		Hematocrit	Low and High
Sodium	Low and High	Platelets	Low and High
Potassium	Low and High		
Chloride	Low and High		
Other			
Glucose	Low and High		
Total protein	Low and High		

Potentially Clinically Significant Laboratory Abnormalities

Abnormal laboratory values (excluding LDH, haptoglobin, and reticulocytes) will also be evaluated by determining the number and percentage of subjects with at least one potentially clinically significant laboratory abnormality over the course of Part B of the study that also represents a worsening from baseline. The potentially clinically significant levels are based on Grade 2 or higher thresholds from the CTCAE v4.03 where possible, or are defined by Sanofi's safety group. Subjects who have a post baseline laboratory value that meets the criteria for being potentially clinically significant but do not have a baseline value will be included in the numerator for determining the percentage of subjects with an abnormality. The laboratory tests and thresholds for determining PCSAs are provided in Table 5.

Table 5 - Thresholds for potentially clinically significant laboratory abnormalities

Laboratory Test	Threshold	Laboratory Test	Threshold
Chemistry		**Hematology**	
ALT (U/L)	>= 3 x ULN	Hematocrit (%)	<= 30% >= 60%
Albumin (g/dL)	<= 2.5 g/dL	Hemoglobin (g/L)	<= 100 g/L >= 190 g/L
Alkaline Phosphatase (U/L)	> 2.5 x ULN	MCV (fL)	>= 1.2 x ULN <= 0.8 x LLN
AST (U/L)	>= 3 x ULN	Platelet count (10^9/L)	<= 75 x 10^9/L >= 700 x 10^9/L
Blood urea nitrogen (mmol/L)	>= 10.7 mmol/L	WBC (10^9/L)	<3 x 10^9/L >= 16 x 10^9/L
Chloride (mmol/L)	<= 90 mmol/L >= 118 mmol/L	Basophils (10^9/L)	> 1.6 x 10^9/L
Creatinine (umol/L)	> 1.5 x ULN	Eosinophils (10^9/L)	> 1.6 x 10^9/L
GGT (U/L)	> 2.5 x ULN	Lymphocytes (10^9/L)	> 0.8 x 10^9/L
Glucose (mmol/L)	<3 mmol/L > 8.9 mmol/L	Monocytes (10^9/L)	< 2.5 x 10^9/L
Potassium (mmol/L)	<= 3 mmol/L >= 6 mmol/L	Neutrophils (10^9/L)	> 1.5 x 10^9/L
Sodium (mmol/L)	<= 126 mmol/L >= 156 mmol/L		
Total bilirubin (umol/L)	> 1.5 x ULN		
Direct bilirubin (umol/L)	> 1.5 x ULN		
Total protein (g/L)	<= 45 g/L >= 100 g/L		

All data collected during Part B, including data collected from unscheduled visits, will be included in this analysis.

Systemic Lupus Erythematosus (SLE) Panel

SLE panel will be summarized with descriptive statistics. The number of subjects with all negative or all missing values for all SLE panel parameters prior to receiving study drug and the number of subjects with at least one positive value for at least one SLE panel parameter prior to receiving study drug will be summarized. Further, among subjects who have all negative or all missing values for all SLE parameters prior to receiving study drug and have at least one post-baseline value for at least one parameter, the number of subjects who test negative at all post-baseline assessments for all SLE parameters, and the number of subjects who test positive at any post-baseline assessment for at least one SLE parameter will be summarized.

Similar summaries will be generated for individual SLE parameters. In addition, for a given SLE parameter, among subjects who have at least one positive or all missing values prior to receiving study drug and have at least one post-baseline for this parameter, the number of subjects who test positive at all post-baseline assessments, and the number of subjects who test negative at any post-baseline assessment will be summarized.

Parameters in the SLE panel include Antinuclear antibody (ANA), multiplex with dsDNA, Anti-La/SSB antibody (SS-B), Anti-riboucleoprotein antibody (RNP), Anti-Smith antibody (Sm), Anti-Ro/SSA antibody (SS A), Anti-scleroderma antibody (Scl-70), Anti-Chromatin antibody, Anti-Jo-1 antibody, Anti-Centromere B antibody, Circulating immune complexes (CIC).

ECG and Vital Signs

The proportion of subjects with specific clinical interpretations (Normal; Abnormal, not an AE; Abnormal, AE) will be summarized by visit, through the end of Part B.

The proportion of subjects with a shift in clinical interpretation from baseline will be summarized. All data collected during Part B, including data collected from unscheduled visits, will be included in this analysis.

Vital signs

Vital signs (systolic and diastolic blood pressure, pulse, respiratory rate, and temperature) will be summarized by visit, through the end of Part B, for the observed values and change from baseline using descriptive statistics for the Safety Analysis Set.

The number and percentage of subjects with potentially clinically relevant post-baseline abnormalities will be presented. All data collected during Part B, including data collected from unscheduled visits, will be included in this analysis. The criteria for potentially clinically relevant post-baseline abnormalities are shown below in Table 6.

Table 6 - Criteria to determine clinically relevant abnormalities in vital signs

Vital Sign	Criteria for Abnormalities
Temperature	>38°C and an increase from pre-dosing of at least 1°C
Pulse	>120 beats per minute post-baseline or an increase from pre-dosing of more than 20 beats per minute; <50 beats per minute post-baseline or a decrease from pre-dosing of more than 20 beats per minute
Systolic Blood Pressure	>180 mmHg post-baseline or an increase from pre-dosing of more than 40 mmHg; <90 mmHg post-baseline or a decrease from pre-dosing of more than 30 mmHg
Diastolic Blood Pressure	>105 mmHg post-baseline or an increase from pre-dosing of more than 30 mmHg; <50 mmHg post-baseline, or a decrease from pre-dosing of more than 20 mmHg
Respiratory Rate	>35 breaths per minute post-baseline; <10 breaths per minute post-baseline

4.8 OTHER ANALYSES

4.8.1 Pharmacokinetic analysis

Plasma BIVV009 concentrations will be listed by subject, nominal visit, date and time of collection and study day. Summary statistics of plasma BIVV009 concentrations will be presented by nominal visit and study day including n, mean, standard deviation, coefficient of variation, geometric mean, median and range. All summaries will be tabulated by treatment cohort in Part B (ie, 6.5 g and 7.5 g) and Total. Individual and mean plasma BIVV009 concentration-versus-time profiles will be plotted. The above analyses will be conducted using the PK Analysis Set for the Cumulative Study Period.

4.8.2 Pharmacodynamic analysis

Descriptive statistics including number of observations, mean, SD, median, minimum, and maximum will be presented for continuous parameters. Categorical variables will be presented with the number and percentage in each category. Summary descriptive statistics (absolute values and changes from baseline) and individual subject listings will be presented by time point and study day. All summaries will be tabulated by treatment cohort in Part A (ie, BIVV009 and Placebo) and Total. The above analyses will be conducted using the PD Analysis Set for the Cumulative Study Period.

4.9 INTERIM ANALYSES

For purposes of regulatory submission, an interim analysis of safety and efficacy data will be performed after all subjects have completed Part A and the data are cleaned. Part B data, up to the date of Part A database lock, will be summarized descriptively.

5 SUPPORTING DOCUMENTATION

5.1 APPENDIX 1 LIST OF ABBREVIATIONS

AE	adverse event
ADL	activities to daily living
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANA	antinuclear antibodies
AST	aspartate aminotransferase
AUC	area under the curve
BLA	Biologics License Applications
BUN	blood urea nitrogen
CAD	cold agglutinin disease
CI	confidence interval
CIC	circulating immune complexes
C_{max}	maximum concentration
CTCAE	Common Toxicity Criteria for Adverse Events
DBL	database lock
ECG	electrocardiogram
eCRF	electronic Case Report Form
EOT	end of treatment
EQ-5D-5L	EuroQol – five dimensions questionnaire
ET	early termination
FACIT	functional assessment of chronic illness therapy
FAS	full analysis set
GGT	gamma-glutamyl transferase
Hgb	hemoglobin
HUI	health utility index
ICH	International Conference of Harmonisation
ITT	intent-to-treat
IV	intravenous
LDH	lactate dehydrogenase
LLN	lower limit of normal
LPO	last patient out
MAA	Marketing Authorization Application
MAR	missing at random
MCMC	Markov chain Monte-Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
MNAR	Missing Not At Random
NCMV	Neighboring-Case Missing Value
PD	pharmacodynamics
PGIC	Patient's Global Impression of Change
PGIS	Patient's Global Impression of Severity

PK	pharmacokinetics
PP	per-protocol
PRO	patient-reported outcome
QOL	quality of life
RBC	red blood cell count
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SF-12	12-Item Short Form Survey
SGOT	serum glutamic oxaloacetic transaminase
SGPT	serum glutamic-pyruvate transaminase
SLE	systemic lupus erythematosus
SOC	system organ class
TEAE	treatment-emergent adverse event
TLF	tables, figures, and listings
TOEPH	heterogeneous Toeplitz
ULN	upper limit of normal
WBC	white blood cell count

5.2 APPENDIX 2 CHANGES TO PROTOCOL-PLANNED ANALYSES

This section summarizes major statistical changes in the protocol amendment(s).

Table 7 - Major statistical changes in protocol amendment(s)

Amendment Number	Approval Date	Changes	Rationale
Version 1 (original)	24-Aug-2017	Not applicable	Not applicable

5.3 APPENDIX 3 DEMOGRAPHICS AND BASELINE CHARACTERISTICS, PRIOR OR CONCOMITANT MEDICATIONS

Demographics and baseline disease characteristics will be summarized for the FAS.

Demography

Demographic characteristics including age (years), age category (<65, >= 65), gender, race, ethnicity, country, geographic location, height (cm), weight (kg), weight category (<75 kg, >= 75 kg), and body mass index (kg/m^2) will be summarized. Geographic locations are defined as Europe, North America, Asia/Japan, and other.

General Medical History

Medical history by system organ class (SOC) and preferred term (PT) will be summarized. A subject will be counted only once if they reported one or more occurrences in the same SOC or PT.

A summary of prior hospitalizations related to CAD within the last 2 years, prior hematological malignancies within the last 5 years, prior non-malignant hematologic diseases within the last 5 years, and prior thromboembolic events within the last 1 year will be presented based off the "Targeted Medical History" eCRF form.

Baseline Disease Characteristics

Screening and baseline lab parameters, including hematology (such as Hgb, bilirubin, LDH), Gilbert's Syndrome testing, and bone marrow biopsy results will be summarized. Prior transfusion history (eg, number of transfusions within 12 months of study entry) will also be summarized.

Severity of anemia symptoms and other CAD-related disease characteristics (eg, presence of acrocyanosis, Raynaud's syndrome, hemoglobinuria, disabling circulatory symptoms, and major adverse vascular events) within 3 months of screening will be summarized.

The proportion of subjects who had documented vaccines within 5 years and who have received vaccination during the study will be summarized. The vaccines include meningococcal vaccine, S. pneumoniae vaccine, and H. influenza vaccine.

Prior or concomitant medications

Prior medications are those taken before the first dose of study drug in Part A. Concomitant medications are those administered on or after the first dose of study drug, or those administered before the first dose that are ongoing when study treatment begins. Prior and concomitant medications will be coded using World Health Organization drug enhanced dictionary (March 2020 version or higher).

Separate tables will be generated to summarize concomitant medications and concomitant protocol-prohibited medications by ATC level 2 and WHODrug standardized medication text.

Additionally, a summary of targeted prior therapies will be generated based off the responses provided in the "Targeted Prior Therapy" eCRF form.

Concomitant non-drug treatments and procedures

Concomitant non-drug treatments and procedures will be summarized for the FAS.

5.4 APPENDIX 4 DATA HANDLING CONVENTIONS

Study Day

Study Day is defined as days relative to the date of first dose of BIVV009 or placebo in Part A, Day 0. The start/stop day of events will be calculated as (date of event – date of Day 0 +1).

Incomplete dates will be imputed for the calculation of study days, unless specified otherwise:

- If missing day only, the start date will be imputed as the first day of the month, while the end date will be imputed as the last day of the month;
- If missing day and month, the start date will be imputed as the first day of the year, while the end date will be imputed as the last day of the year;

Baseline

For purposes of specified analyses, baseline measures are defined as the last value prior to the first dose of study drug in Part A.

Analysis Windows for time points

Analysis by study visit will utilize the data recorded on the corresponding nominal visit from electronic Case Report Forms (eCRFs). In case of missing data/assessments for a visit, data from an unscheduled visit, or ET/Safety Follow-up Visit will be used, if the Study Day of such a visit fits in the analytic window defined below. If multiple values are identified within the analytic window, the one closest to the target study day will be used.

Starting at Week 27 and every 2 weeks thereafter, analytic visit windows in Part B will be -6 days to +7 days from the Target Study Days of the protocol-defined visits. An example of the first few visits in Part B can be seen below .

Table 8 - Analytic visit window

Visit	Target Study Day	Analytic Visit Window (Study Day)
Week 27	189	183-196
Week 29	203	197-210
Week 31	217	211-224
Week 33	231	225-238
Etc.		

Unscheduled visits

Unscheduled visit measurements of clinical laboratory data, ECG and vital signs will be used for computation of baseline, the last on-treatment value, analysis according to PCSAs, and the shift summaries for safety. They will also be included in the by-visit summaries if they are re-allocated to scheduled visits.

Pooling Sites for Analysis

Data from all investigational sites will be pooled for the analysis, unless specified otherwise.

Handling of Missing Data

Best efforts will be utilized to minimize missing values. If a subject discontinues study treatment early, the subject will be asked to complete study assessments at ET/safety follow up visit.

For the analysis of both continuous and categorical data, summary statistics will be computed based on observed data. Since the duration of treatment period varies in Part B, efficacy endpoints by visit will be summarized through 12 months during Part B. Additional summaries beyond 12 months may be computed depending on the enrollment pattern, ie, the number of subjects attending visits after the Week 77 visit. AE endpoints by visit will be summarized through LPLV in Part B.

For a QOL endpoint, in the event of missing data, the total score will be estimated according to the provision for missing domains in the calculation algorithm of the score. For example, if a subject has missing data in FACIT-F questionnaire at a visit, the prorated FACIT-F score will be calculated if more than 50% of the items (a minimum of 7 out of 13 items) are available. Otherwise, the score is missing.

For the analysis of AEs and concomitant medications/procedures, if the start/stop date of an AE/concomitant medication/procedure is missing or partial, the corresponding study day will be left blank. However, inferences will be made from the partial and missing dates to classify modifications as prior and/or concomitant and AEs as treatment emergent or not treatment emergent.

6 REFERENCES

Not applicable.

Signature Page for VV-CLIN-0587710 v1.0
bivv009-04-16-1-9-sap-part-b

Approve & eSign	[redacted]
Approve & eSign	