STATISTICAL ANALYSIS PLAN

Protocol title:	**A Phase 3 Open-Label, Multicenter Study of the Safety, Efficacy, and Pharmacokinetics of Intravenous Recombinant Coagulation Factor VIII Fc-von Willebrand Factor-XTEN Fusion Protein (rFVIIIFc-VWF-XTEN; BIVV001) in Previously Treated Patients ≥12 Years of Age With Severe Hemophilia A**
Protocol number:	**EFC16293**
Compound number (INN/Trademark):	**BIVV001**
Study phase:	**Phase 3**
Short title:	**XTEND-1**
Statistician:	[redacted]
Statistical project leader:	[redacted]
Date of issue:	**11-Aug-2021**

Regulatory agency identifier number(s):

IND:	17464
EudraCT:	2019-002023-15
WHO:	U1111-1223-4867

NCT04161495

Total number of pages: 68

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

VERSION HISTORY

This statistical analysis plan (SAP) for Study EFC16293 is based on the protocol dated 13-May-2020.

The first participant was enrolled on 04-Dec-2019.

Table 1 - Major changes in statistical analysis plan

SAP Version	Approval Date	Changes	Rationale
1.0	10-Jun-2021	The change from baseline to Week 52 in Haem-A-QoL physical health score, PROMIS Pain Intensity 3a and HJHS will be analyzed via mixed-effects model repeated measures (MMRM) model. A sensitivity analysis will be performed on the patients who rollover from lead-in observational Study 242HA201/OBS16221 using the same MMRM model.	Specifying the MMRM analysis for Haem-A-QoL physical health score, PROMIS Pain Intensity 3a and HJHS.
1.0	10-Jun-2021	The Haem-A-Qol physical health score change from Baseline to Week 52 in Arm A, the PROMIS pain intensity 3a change from Baseline to Week 52 in Arm A, and the HJHS pain intensity 3a change from Baseline to Week 52 in Arm A will be listed as the 3rd, 4th and 5th testing in the multiplicity part (Section 4.7).	Specifying the newly added statistical inference (ie, MMRM analysis for Haem-A-QoL physical health score, PROMIS Pain Intensity 3a and HJHS) in multiplicity testing. The hierarchy order for the newly added testing is based on clinical and statistical consideration.
1.0	10-Jun-2021	For the key secondary efficacy endpoint analysis (ie, an intra-patient comparison of ABR during the BIVV001 weekly prophylaxis treatment period versus the historical prophylaxis ABR), participants in Arm A need to have at least 6 months of duration in this study and in observational Study 242HA201/OBS16221 to be included in the analysis.	Clarify the participants selection for the key secondary efficacy endpoint analysis.
2.0	Current Version	The change from baseline to Week 52 in PROMIS Pain intensity 3a past 7 days intensity of pain at its worst score (PAINQU6) will be analyzed via mixed-effects model repeated measures (MMRM) model. A sensitivity analysis will be performed on the patients who rollover from lead-in observational Study 242HA201/OBS16221 using the same MMRM model. The PROMIS Pain intensity 3a past 7 days intensity of pain at its worst score (PAINQU6)will be listed as the 4th testing in the multiplicity part (Section 4.7).	Clarify the MMRM analysis for PROMIS Pain intensity 3a past 7 days intensity of pain at its worst score (PAINQU6).

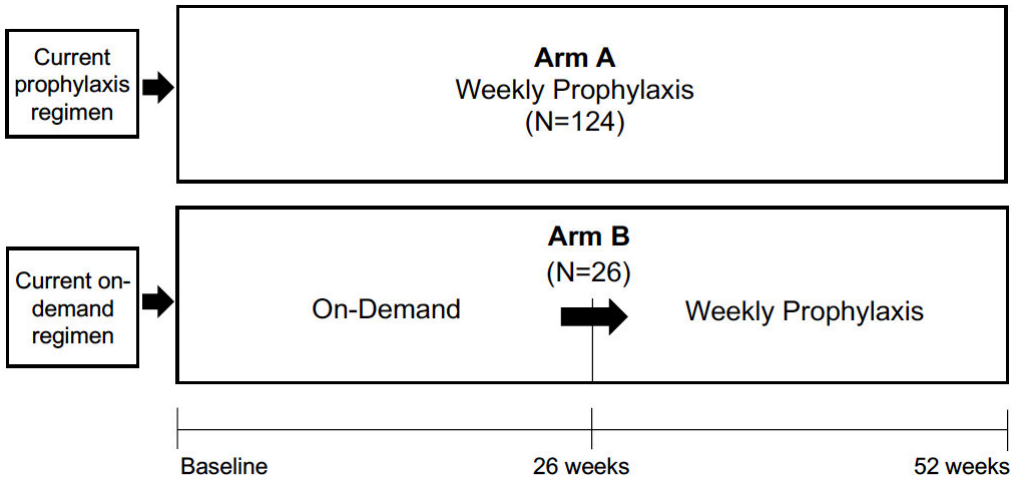
1 INTRODUCTION

1.1 STUDY DESIGN

This is a multinational, multicenter, open-label Phase 3 study of the safety, efficacy and pharmacokinetics (PK) of BIVV001 in previously treated patients (PTPs) ≥12 years of age with severe hemophilia A (defined as <1 IU/dL [<1%] endogenous FVIII).

The study is comprised of two arms: Arm A and Arm B. Approximately 124 participants currently on a prophylaxis regimen with FVIII will enter Arm A and receive BIVV001 at a dose of 50 IU/kg once-weekly (QW) on a prophylaxis regimen for 52 weeks, of which at least 75 participants will have at least 6 months of participation in Study 242HA201/OBS16221 prior to baseline. Approximately 26 participants currently on an on-demand regimen will enter Arm B to receive BIVV001 at a dose of 50 IU/kg on an on-demand basis for 26 weeks, then switch to receive BIVV001 at a dose of 50 IU/kg QW on a prophylaxis regimen for 26 weeks.

Figure 1 - Study schema



All enrolled participants will come in to scheduled visits at Baseline (Day 1), Week 4, Week 13, Week 26, Week 39 and Week 52.

Following a washout period (at least 4 to 5 days, depending on current therapy), all participants (except those in the sequential PK subgroup in Arm A) will undergo abbreviated PK sampling after the first dose of BIVV001 (Baseline). At least 16 participants in Arm A will participate in the sequential PK subgroup and undergo more extensive PK sampling at Baseline and again at Week 26.

In addition, participants from any arm who undergo major surgery during the study will be included in the surgery subgroup to assess control and prevention of bleeding with BIVV001 treatment in the surgical setting. A minimum of 10 major surgeries in at least 5 participants will be targeted to assess control and prevention of bleeding in the surgical setting.

Participants will be offered to enroll in an open-label extension study after completion of this study based on eligibility criteria.

1.2 OBJECTIVE AND ENDPOINTS

Table 2 - Objectives and endpoints

Objectives	Endpoint
Primary Efficacy Objective	**Primary Efficacy Endpoint**
• To evaluate the efficacy of BIVV001 as a prophylaxis treatment	• Annualized bleeding rate (ABR) in Arm A
Secondary Efficacy Objectives	**Secondary Efficacy Endpoints**
• To evaluate the efficacy of BIVV001 as a prophylaxis treatment	• Intra-patient comparison of ABR during the BIVV001 weekly prophylaxis treatment period versus the historical prophylaxis ABR for participants in Arm A who participated in Study 242HA201/OBS16221, an observational study (**key secondary endpoint).** • ABR by type and location for prophylaxis treatment per study arm. • ABR for all bleeding episodes (including untreated bleeding episodes) for prophylaxis treatment per study arm. • Intra-patient comparison of ABR during the QW prophylaxis treatment period versus the ABR during the on-demand treatment period in Arm B. • Percentage of participants who maintain FVIII activity levels over 1%, 5%, 10%, 15%, and 20% in Arm A.
• To evaluate the efficacy of BIVV001 in the treatment of bleeding episodes	• Number of injections and dose of BIVV001 to treat a bleeding episode per study arm and treatment regimen. • Percentage of bleeding episodes treated with a single injection of BIVV001 per study arm and treatment regimen. • Assessment of response to BIVV001 treatment of individual bleeding episodes based on the International Society on Thrombosis and Haemostasis (ISTH) 4-point response scale per study arm and treatment regimen. • Physician's global assessment (PGA) of participant's response to BIVV001 treatment based on a 4-point response scale per study arm and treatment regimen.
• To evaluate BIVV001 consumption for the prevention and treatment of bleeding episodes	• Total annualized BIVV001 consumption per participant per study arm and treatment regimen.
• To evaluate the effect of BIVV001 prophylaxis on joint health outcomes	• Change from Baseline to Week 52 in total score and domain scores (eg, swelling and strength) assessed by the Hemophilia Joint Health Score (HJHS) in Arm A. • Annualized Joint Bleeding Rate (AJBR) per study arm and treatment regimen. • Target joint resolution at Week 52, based on ISTH criteria in Arm A.

Objectives	Endpoint
• To evaluate the effect of BIVV001 prophylaxis on Quality of Life (QoL) outcomes	• Changes in Haem-A-QoL (≥17 years old) total score and physical health score measures from baseline to Week 52 in Arm A. • Changes in PROMIS Pain Intensity 3a from baseline to Week 52 in Arm A. • Changes in PROMIS SF Physical Function (≥18 years old) measures from Baseline to Week 52 in Arm A.
• To evaluate the efficacy of BIVV001 for perioperative management	• Investigators' or Surgeons' assessment of participant's hemostatic response to BIVV001 treatment on the ISTH 4-point response for surgical procedures scale. • Number of injections and dose to maintain hemostasis during perioperative period for major surgery. • Total BIVV001 consumption during perioperative period for major surgery. • Number and type of blood component transfusions used during perioperative period for major surgery. • Estimated blood loss during perioperative period for major surgery.
Safety Objective • To evaluate the safety and tolerability of BIVV001 treatment	Safety Endpoints • The occurrence of adverse events (AEs) and serious adverse events (SAEs). • The occurrence of clinically significant changes from baseline in physical examination, vital signs, and laboratory tests. • Development of inhibitors (neutralizing antibodies directed against factor VIII [FVIII]) as determined via the Nijmegen modified Bethesda assay. • The occurrence of embolic and thrombotic events.
Pharmacokinetic Objective • To assess the PK of BIVV001 based on the one-stage activated partial thromboplastin time (aPTT) and two-stage chromogenic FVIII activity assays	Pharmacokinetic Endpoint • PK parameters including, but not limited to, maximum activity (C_{max}), elimination half-life ($t_{1/2}$), total clearance (CL), total clearance at steady state (CLss), accumulation index (AI), area under the activity time curve (AUC), volume of distribution at steady state (Vss), mean residence time (MRT), incremental recovery (IR), trough activity (C_{trough}), time above predefined FVIII activity levels.
Exploratory Objectives • To evaluate joint-health structural outcomes via ultrasound using the Joint Activity and Damage Exam (JADE) protocol and/or Hemophilia Early Arthropathy Detection with Ultrasound (HEAD-US)	**Exploratory Endpoints** • Changes in anatomical structural joint health outcomes via the JADE and/or HEAD-US protocol for ultrasound imaging from Baseline to Week 52 in a subpopulation at selected sites in Arm A.
• To assess the impact of BIVV001 treatment on patient reported outcome (PRO) measurements, and physical activity (measures per study arm and treatment regimen)	• Changes in PRO measures and physical activity per study arm and treatment regimen. - PROMIS-SF Physical Function (≥18 years old) or PROMIS Pediatric-SF Physical Activity (<18 years old), - Haem-A-QoL (≥17 years old) or Haemo-QoL (<17 years old) total score and subscale scores,

Objectives	Endpoint
	- PROMIS-SF Pain Interference (≥18 years old) or PROMIS Pediatric-SF Pain Interference (<18 years old), - EQ-5D-5L, - Treatment Satisfaction Questionnaire for Medications (TSQM-9), - Patient Global Impression of Severity (PGIS) (activity component), - PGIS (joint component), - Change in physical activity measures (ActiGraph Activity Monitor), - Hemophilia Activities List (HAL) (≥18 years old) or pedHAL (<18 years old). • Change in HJHS total score and domain score (eg, swelling and strength) in Arm B by treatment regimen. • Patient Global Impression of Change (PGIC) at Week 26 and Week 52 per study arm and treatment regimen. • Patient Preference at Week 52 per study arm and treatment regimen. • Exit interview at Week 52 (or at subsequent follow-up visit).

2 SAMPLE SIZE DETERMINATION

The sample size was estimated to rule out a greater-than-acceptable risk of immunogenicity. Assuming a drop-out rate of approximately 15%, a sample size of 124 participants in the prophylaxis arm is expected to provide 104 evaluable participants with at least 50 exposure days (ED). If ≤2 participants out of 104 evaluable participants develop an inhibitor, then the upper bound of an exact 95% confidence interval (CI) would exclude 6.8%, a threshold determined at the FDA Factor VIII Inhibitor Workshop that was held in 2003.

The primary efficacy objective of this study is to evaluate the efficacy of BIVV001 weekly prophylaxis as estimated by the mean annualized bleeding rate (ABR) and one-sided 97.5% CI in Arm A. Based on currently marketed FVIII products, mean ABR during clinical trials typically ranges from approximately 2 to 5 bleeds per year but can be as high as 6 bleeds per year (1, 2, 3, 4). This indicates that while the ABR is typically low with adequate treatment, there is variability in clinical bleeding phenotype with some severe hemophilia A participants having higher bleeding rates than others. In order to show adequate control of bleeding consistent with currently marketed FVIII products, and to account for this variability, a clinically meaningful treatment effect may be claimed if the upper bound of the confidence interval of the estimated ABR is less than or equal to 6. In a phase 3 study of rFVIIIFc, the mean ABR for an individualized prophylaxis arm was 2.9 (3) and the dispersion factor was estimated at 2.3 (data on file). Based on 2000 simulations of a negative binomial regression model with mean ABR of 2.9 and dispersion factor of 2.3, a sample size of 124 participants will provide at least 90% power for the upper bound of the confidence interval to exclude an ABR greater than 6, assuming a 15% drop out rate.

For the key secondary efficacy endpoint, an intra-patient comparison of ABR during the BIVV001 weekly prophylaxis treatment period versus the historical prophylaxis ABR will be performed using non-inferiority testing for participants in Arm A who have at least 6 months of participation in this study and at least 6 months of historical data on prophylaxis treatment from observational study 242HA201/OBS16221. The non-inferiority margin was estimated based on the known treatment effect between on-demand and prophylaxis treatment. A meta-analysis of Phase 3 registrational studies for recombinant FVIII products that include both on-demand and prophylaxis treatment arms estimated an average reduction of 31 bleeds per year between on-demand and prophylaxis treatment regimens. The lower bound of this treatment effect was 27 bleeds per year. Using a fixed margin approach to maintain a substantial amount (85%) of the treatment effect results in a non-inferiority margin of 4. Further details of the meta-analysis and derivation of the non-inferiority margin are specified in Section 5.6 (Appendix 6). For a non-inferiority test of the null hypothesis (median difference in ABR exceeds or is equal to non-inferiority margin) versus the alternative hypothesis (median difference in ABR is less than non-inferiority margin), a sample size of 63 achieves 90% power to detect non-inferiority using a one-sided paired Wilcoxon Signed-Rank test at a 0.025 significance level when the actual mean of paired differences is 0 and the non-inferiority margin is 4. Without prior knowledge of the standard deviation of the paired differences, a conservative estimate of 10 was assumed. In order to account for drop-out and the use of the Per Protocol Set, a total of at least 75 subjects who have completed at least 6 months of participation in observational study 242HA201/OBS166221 will be enrolled in Arm A.

3 ANALYSIS POPULATIONS

The following populations for analyses are defined:

Table 3 - Populations for analyses

Population	Description
All-Enrolled Analysis Set	All participants who were enrolled in the study, regardless of whether they were dosed with study drug or not. Participants will be considered enrolled when the Investigator has verified that they are eligible according to the criteria in Section 5 of the protocol. Participant disposition and enrollment summaries will be based on the All-Enrolled Analysis Set.
Full Analysis Set (FAS)	All participants who receive at least one dose of study drug. All analyses of demographics, baseline characteristics, and efficacy will be based on the FAS, unless otherwise specified.
Per Protocol Set	A subset of the Full Analysis Set including participants who do not have important protocol deviations potentially impacting efficacy. The Per Protocol Set will be utilized for analysis of the key secondary efficacy endpoint, as well as sensitivity analysis of the primary endpoint.
Safety Analysis Set	The safety analysis set is the same as the full analysis set and will include all participants who receive at least one dose of study drug. All analyses of safety will be based on the Safety Analysis Set, unless otherwise specified.
PK Analysis Set (PKAS)	All participants who have completed adequate blood sample collection to assess key PK parameters, as determined by the PK scientist. All analyses of PK will be based on the PKAS, unless otherwise specified.
Sequential Pharmacokinetic Subgroup	All participants who have evaluable PK profiles for both the Baseline and Repeat PK profiles, as determined by the PK scientist.
Surgery Subgroup	All participants who have undergone major surgery after the first dose of study drug.

4 STATISTICAL ANALYSES

All summaries and statistical analyses will be generated using SAS version 9.0 or higher.

4.1 GENERAL CONSIDERATIONS

Unless otherwise specified, analyses and summaries will be presented by study arm and treatment regimen as follows:

Arm A	Arm B	
Prophylaxis 50 IU/kg QW	On-demand	Prophylaxis 50 IU/kg QW

In general, summaries of demographic/baseline characteristics and safety data will include a total column in addition to columns for study arm and treatment regimen.

The surgery subgroup will be included in baseline tables as well as the overall AE summary. When the surgery subgroup is included in a summary table, participants who participated in the surgery subgroup will be included in the columns for both the assigned study arm and treatment regimen and the surgery subgroup but will only be counted once in the total column.

Baseline

Baseline is defined as the last non-missing measurement taken prior to the first dose of study drug, excluding measurements taken during a surgical/rehabilitation period. For HJHS, if there are no measurements taken prior to the first dose of the drug, and the first measurement is taken no more than 7 days after the first dosing, then this first measurement will be considered as baseline for HJHS.

Study Day

Study day is defined as the number of days relative to the date of the first dose of study drug. The first dose of study drug is Day 1. Study day will be calculated as (date of event – first dose date +1) if the date of event is on or after the first dose date or (date of event – first dose date) if the date of event is before the first dose date. If the date of event is missing or partial, the corresponding study day will be left blank.

Exposure Day

One exposure day (ED) is defined as a 24-hour period in which a participant receives 1 or more doses of study drug, with the time of the first injection of study drug defined as start of the ED.

4.1.1 Definition of study periods

This section defines the study periods used for different analyses, including treatment regimen period, efficacy period, surgical/rehabilitation period, and safety period.

All analyses and summaries relating to bleeding and consumption will be based on the ***efficacy period*** (Section 4.1.1.2); data collected during the PK and surgical/rehabilitation periods (major and minor) will not be included.

All other efficacy analyses will be based on the ***treatment regimen period*** (Section 4.1.1.1), excluding data that occurs during any major surgical/rehabilitation period to avoid confounding the treatment effect. Analysis of efficacy endpoints that are visit-based will not include visits that are coincidental with a major surgical/rehabilitation period.

Surgical evaluations will be based on the ***surgical/rehabilitation period*** (Section 4.1.1.3).

All safety analyses will be based on the ***safety period*** (Section 4.1.1.4). Unless otherwise specified, adverse events and laboratory evaluations occurred during major surgical/rehabilitation periods will not be included in the safety summaries.

4.1.1.1 Treatment regimen

The treatment regimen is defined as the actual treatment regimen(s) that the participant follows. For example, a participant who starts the study on an on-demand regimen and switches to a prophylactic regimen will appear in the summaries of both the on-demand and the prophylactic regimens based on the time he/she spent in the respective regimens.

The start/end date and time of the treatment regimen is defined as:

- Arm A: the prophylactic treatment regimen starts at the date and time of the first dose of study drug (PK dose) and ends at 23:59 on the day of the last dose of study drug.
- Arm B (on-demand regimen): the on-demand treatment regimen starts at the date and time of the first dose of study drug (PK dose) and ends 1 minute prior to the first prophylaxis dose at the Week 26 visit (or, if the time of the first prophylactic dose is not available, at 23:59 on the day prior to the day of the first prophylactic dose). If the participant withdraws from the study prior to starting the prophylaxis regimen, then the on-demand regimen will end at the last non-safety-follow-up contact or the last electronic Patient Diary (ePD) entry, whichever is later. If there is no time associated with this event then the time will be imputed to be 23:59 of the day.
- Arm B (prophylaxis regimen): the prophylactic treatment regimen starts at the date and time of the first prophylaxis dose of study drug (or, if the time of the first prophylactic dose is not available, at 00:01 on the day of the first prophylactic dose). The prophylactic treatment regimen ends at 23:59 on the day of the last dose of study drug.

The duration of a given treatment regimen is the time period from the start of the treatment regimen to the end of that treatment regimen. The total duration of time allocated to each treatment regimen will be calculated in minutes and converted to days as the number of minutes divided by 1440.

4.1.1.2 Efficacy period

The efficacy period will be used for the evaluation of bleeding and consumption endpoints. For a subject to have an evaluable efficacy period over the duration of study, he/she must have at least 1 day of treatment for an episodic regimen or at least 2 prophylactic injections for a prophylactic regimen. The efficacy period is determined from a combination of dosing dates/times, PK periods, and/or surgical/rehabilitation periods as follows:

- Arm A: The efficacy period for the prophylactic regimen starts with the date and time of the first prophylactic dose following the completed PK sampling period (ie, 168 or 336 hours), and ends with the end of the treatment regimen as defined in Section 4.1.1.1. That is, if the initial PK sampling period is incomplete due to an aborted collection of blood samples following the PK injection (eg, because of a bleeding episode that required treatment), then the efficacy period will begin following a subsequent fully executed PK sampling period.

  The efficacy period will be interrupted for the repeat PK period in the *sequential PK subgroup* as follows: The efficacy period continues up to 1 minute before the repeat PK dose and then continues at the next prophylactic dose following the end of the repeat PK sampling period. Injections or untreated bleeds occurring after the repeat PK sampling and before the next prophylactic dose are not considered part of the efficacy period. Considerations given to an aborted PK sampling period described for the previous PK period will apply here as well. If, after the aborted PK sampling period, the participant returns to his prophylaxis regimen and has at least 2 prophylaxis injections prior to the subsequent PK sampling period, the efficacy period will re-start at the first prophylaxis injection and end 1 minute prior to the subsequent PK dose. The efficacy period will stop and start for each such encounter.

- Arm B (on-demand regimen): The efficacy period for the on-demand regimen starts 1 minute after the date and time of the last sampling time point (ie, 168 hours) for the PK dose and ends with the end of the treatment regimen as defined in Section 4.1.1.1.

- Arm B (prophylaxis regimen): The efficacy period for the prophylactic regimen starts with the date and time of the first prophylactic dose (or, if the time of the first prophylactic dose is not available, at 00:01 on the day of the first prophylactic dose) and ends with the end of the treatment regimen as defined in Section 4.1.1.1.

In addition, efficacy periods are interrupted for surgical/rehabilitation periods (major and minor) and large injection intervals as described below.

Adjustments Due to Surgical/Rehabilitation Periods

For analysis purposes, the efficacy period for a given treatment regimen will be adjusted for all surgical/rehabilitation periods (major and minor). The start and end of the efficacy period for a given treatment regimen are adjusted as follows:

- For all participants, the efficacy period continues up to 1 minute before the start of a surgical/rehabilitation period.

- For participants on a prophylactic regimen following the end of a surgical/rehabilitation period, the efficacy period for the prophylactic regimen starts or re-starts at the first prophylactic dose following the end of the surgical/rehabilitation period.
- For participants on an on-demand regimen following the end of a surgical/rehabilitation period, the efficacy period for the on-demand regimen re-starts at 00:01 on the day following the end of the surgical/rehabilitation period.

Bleeding episodes that occur during the surgical/rehabilitation period will be attributed to the surgical/rehabilitation period and hence not counted towards the ABR.

Adjustments Due to Large Injection Intervals

For analysis purposes, the efficacy period will also be adjusted to account for large intervals between injections resulting from missing data. A large interval is defined as >28 days between any 2 adjacent BIVV001 injections within a prophylactic treatment regimen, and any such intervals will be removed from the efficacy period. The efficacy period prior to each such interval will end at the time of the last injection prior to the interval and restart at the time of the next prophylaxis injection. The efficacy period will be adjusted for each identified interval that is not within a surgical/rehabilitation period. This algorithm applies only to prophylactic regimens and no adjustment for large injection intervals will be made for on-demand treatment regimens.

4.1.1.3 Surgical/rehabilitation period

The broadest span of time for the surgical/rehabilitation period is from the first dose of BIVV001 given for the surgery (ie, the pre-surgery dose) up to 1 minute before the first regular prophylactic dose after the last day of postoperative care/rehabilitation.

Since not all participants will have these events, specific considerations for the start and end of the surgical/rehabilitation period are as follows:

Start of the surgical/rehabilitation period:

- If there is more than one pre-surgical dose, then the first one should be selected (a pre-surgical dose can be administered on the day or the day before surgery).
- If there is no pre-surgical dose but there was a prophylactic dose, or a dose given for "OTHER" reason prior to the surgery on the day of or the day before surgery, then the last of these prophylactic doses should be selected.
- If there is no pre-surgical dose or no prophylactic dose on the day before surgery, then select the start date/time of the surgery. If the time was not recorded, then select the date and impute 00:01 for the time.

End of the surgical/rehabilitation period:

- If the participant is on a prophylactic regimen following the surgical/rehabilitation period, then the end of the surgical/rehabilitation period is 1 minute before the first regular prophylactic dose after the latest date and time among the following (impute 23:59 if time

is not available): 1) discharge from the hospital, 2) end of perioperative period follow-up phone call, and 3) end of a two-week period (one-week for minor surgeries) after the resumption of weekly prophylaxis treatment.

- If the participant is not on a prophylactic regimen following the surgical/rehabilitation period (eg, the participant is on an on-demand regimen after completing the surgical/ rehabilitation period, or the participant simply received no further prophylactic doses), then the end of the surgical/rehabilitation period is imputed as 23:59 on the last date among the following: 1) discharge from the hospital, 2) end of perioperative period follow-up phone call, and 3) end of a two-week period (one-week for minor surgeries) after the last surgical dose.
- If the overall end of study is declared while the participant is still in the surgical/ rehabilitation period, then select the date of the end-of-study visit and impute 23:59 for the time if a time is not provided.

Two exceptions are noted:

- If 2 (or more) major surgeries are performed without an intervening discharge from the hospital, then the first surgical/rehabilitation period will end 1 minute before the start of the next surgery and the second surgical/rehabilitation will end as described above. The same will also apply among overlapping minor surgeries.
- If minor surgery is performed during postoperative care or rehabilitation of a major surgery then the surgical/rehabilitation period for the minor surgery will start and end on the day of the minor surgery, at 00:01 and 23:59, respectively, if times are not otherwise provided or recorded as 00:00. The surgical/rehabilitation period for the major surgery will include the minor surgery (ie, the surgical/rehabilitation period for the major surgery does not stop and restart around the minor surgery) and will end as otherwise defined.

The surgical/rehabilitation period will be determined in the same manner for both major and minor surgeries.

4.1.1.4 Safety period

The overall safety period is defined as beginning at the first dose of study drug. For participants who enroll in the open-label extension study after completion of this study, the safety period ends at the end of their last treatment regimen as defined in Section 4.1.1.1. For participants who do not enroll in the extension study, the safety period ends at the Safety Follow-up Call or Visit. If the Safety Follow-up Call or Visit is missing, then the safety period ends at the last ePD entry, the last study visit, or the last documented participant contact, whichever occurs latest.

The safety period of a specific treatment regimen is as described in Section 4.1.1.1, except that the end of the participant's last treatment regimen is the end of their overall safety period.

4.2 PARTICIPANT DISPOSITIONS

Disposition of participants will be summarized using the All-Enrolled Analysis Set. The number (%) of participants included in each of the analysis populations listed in Table 3 will be summarized by study arm, treatment regimen, and overall. The number of participants who completed the study and the number of participants who discontinued from the study early, including the primary reason for discontinuation, will be tabulated by study arm, treatment regimen, and overall.

Separate summaries of enrollment by country/region and site and the number of participants attending each visit will be provided using the All-Enrolled Analysis Set by study arm, treatment regimen, and overall. Disposition, including date of last visit and reason for early discontinuation, for participants who did not complete the study, will be provided in a listing by participant using the All-Enrolled Analysis Set.

Protocol deviations

The number of patients with major protocol deviations will be summarized by study arm and treatment regimen and overall for the FAS. Major protocol deviations that occurred during a major surgical/rehabilitation period will be presented separately in the summary.

Important protocol deviations impacting efficacy assessment will be identified by the study team from the following protocol deviation categories and will be finalized before the database lock:

- Entered the study even though entry criteria were not satisfied
- Developed withdrawal criteria during the study but were not withdrawn.
- Received the wrong treatment or incorrect dose.
- Received an excluded concomitant treatment.
- Persistent misuse of the ePD device impacting interpretation of data based on medical review.

Important protocol deviations will be summarized by study arm and treatment regimen and overall for the FAS. Subjects who have at least one important protocol deviation will be excluded from the Per Protocol Set.

4.3 PRIMARY ENDPOINT ANALYSIS

4.3.1 Definition of endpoint

The primary efficacy endpoint for this study is annualized bleeding rate (ABR) in Arm A.

Data for bleeding events will be collected by each participant via an ePD or, in the case of bleeds occurring and/or treated at the study site, on the eCRF and will include type of bleed, location of bleed, and treatment dates, if applicable. This information will be used to derive the primary and related secondary efficacy endpoints.

During the course of the study, the Investigator is given the opportunity to disagree with the classification of bleeding episode as given by the participant/caregiver, and the participant/ caregiver is subsequently given the opportunity to agree or disagree with the reclassification (spontaneous, traumatic, or not a bleed). If the participant/caregiver agrees with the Investigator's assessment, then all analyses of bleeding will be based on the Investigator's determination of the bleeding type, whether or not the change was made to the participant's records.

Bleeding episodes of an unknown type will be included in the determination of the annualized bleeding rate and in summaries based on bleeding episodes but, unless specified otherwise, will not be included in summary tables where endpoints are summarized by type of bleed.

Definition of a bleeding episode (based on treated bleeds)

All primary analyses of bleeding endpoints will be based on treated bleeds consistent with ISTH criteria (Blanchette 2014) using the following standardized definition:

In the analysis of treated bleeds, a bleeding episode starts from the first sign of bleeding and ends no more than 72 hours after the last injection to treat the bleed, within which any subsequent bleeding at the same location, injections ≤72 hours apart, are considered the same bleeding episode. Multiple bleeding locations treated with a single injection will also be considered a single bleeding episode. Any injection to treat the bleed, taken >72 hours after the preceding one, will be considered the first injection to treat a new bleeding episode in the same location. Any subsequent bleeding or injection at a different location will be considered a separate bleeding episode, regardless of the time from the last injection.

Annualized bleeding rate

The ABR for each individual participant is calculated using the following formula:

$$\text{ABR} = \frac{\text{Number of treated bleeding episodes during the efficacy period}}{\text{Total number of days during the efficacy period}} \times 365.25$$

A bleeding episode is counted in the primary analysis if it was treated with BIVV001. All types of bleeding episodes (spontaneous, traumatic, and type unknown) will be included in determining the annualized number.

4.3.2 Main analytical approach

The primary analysis of the primary endpoint will estimate the mean ABR and one-sided 97.5% CI using a Negative-Binomial model for the weekly prophylaxis arm (Arm A) based on the FAS. If the upper limit of the CI is less than or equal to 6, the weekly prophylaxis treatment regimen will be considered to provide adequate bleeding control.

4.3.3 Sensitivity analyses

Sensitivity analyses of the primary endpoint will be performed using the Per Protocol Set, as well as using the FAS including participants with an efficacy period of at least 26 weeks.

4.3.4 Subgroup analyses

To assess the homogeneity of the treatment effect across various subgroups, analyses will be performed on the primary endpoint across the following subgroups (categories with fewer than 5 participants may be combined with other categories):

- By age (12-17 years, 18-64 years, ≥65 years).
- By bleeding phenotype at Baseline (estimated bleeds in prior 12 months, 0, >0-5, >5-10, >10).
- By number of target joints (none present, ≤median of number present, >median of number present).
- By dosing and dosing interval compliance (<80%, ≥80%).

The estimated mean ABR and the corresponding 95% CI will be provided, for each subgroup, using the same method as applied to the primary analysis. Forest plots will be provided.

4.4 KEY SECONDARY ENDPOINT ANALYSIS

The key secondary endpoint for this study is an intra-patient comparison of ABR between BIVV001 weekly prophylaxis treatment and historical prophylaxis treatment. Participants in Arm A who have at least 6 months of participation in this study and at least 6 months of historical data on prophylaxis treatment collected in the observational study OBS16221 (defined as efficacy period ≥26 weeks in both studies) will be included for the analysis.

All analyses of the key secondary endpoint will be based on treated bleeds.

4.4.1 Main analytical approach

The intra-participant comparison of ABR between BIVV001 weekly prophylaxis and historical prophylaxis will be performed using a Negative-Binomial regression model accounting for over-dispersion with the dependent variable as “total bleeding episodes”, covariate as “treatment regimen”, repeated variable as “subject”, and log time as an offset variable. The mean paired difference and 95% CI will be estimated using the Per Protocol Set (as primary analysis) and Full Analysis Set (as supportive analysis).

The null hypothesis that the paired ABR difference (BIVV001 prophylaxis – historical prophylaxis) is ≥M will be tested against the alternative hypothesis that the ABR difference is <M, where M is the non-inferiority margin of 4 bleeds per year (Section 5.6, Appendix 6). Non-inferiority will be declared if the upper bound of the one-sided 97.5% CI is less than 4.

If non-inferiority is achieved, then superiority will be evaluated sequentially using a Negative-Binomial regression model as specified above. The paired ABR ratio and 95% CI will be estimated using the FAS. The null hypothesis that the paired ABR ratio (BIVV001 prophylaxis: historical prophylaxis) is ≥1 will be tested against the alternative hypothesis that the ABR ratio is <1. Superiority will be declared if the upper bound of the one-sided 97.5% confidence interval is less than 1.

4.4.2 Supplementary analysis

As a supportive analysis, the non-inferiority analysis will also be performed using the Wilcoxon signed rank test.

The median of the paired differences in ABR between BIVV001 weekly prophylaxis and historical prophylaxis will be tested using a Wilcoxon Signed Rank test, using the Per Protocol Set. The null hypothesis that the median difference is ≥4 will be tested against the alternative hypothesis that the median difference is <4. The null hypothesis will be rejected if $p<0.025$, establishing non-inferiority of BIVV001 weekly prophylaxis treatment to historical prophylaxis. The 95% CI of the median of the ABR difference will also be calculated using Walsh averages and a Hodges-Lehmann estimate of the median difference (5).

A supportive analysis will also be performed for the superiority test when non-inferiority is achieved. The mean paired ABR difference and 95% CI will be estimated using the same Negative-Binomial regression model as specified above based on the FAS. The null hypothesis that the paired ABR difference (BIVV001 prophylaxis – historical prophylaxis) is ≥0 will be tested against the alternative hypothesis that the ABR different is <0. Superiority can be declared if the upper bound of the one-sided 97.5% confidence interval is less than 0.

4.5 OTHER SECONDARY ENDPOINTS ANALYSES

All secondary efficacy endpoints will be summarized descriptively based on the FAS and presented by study arm and treatment regimen, unless otherwise specified.

All analyses of bleeding endpoints will be based on treated bleeding episodes, except for the summary of ABR for all bleeds which will include both treated and untreated bleeds.

4.5.1 ABR by type and location

ABR will be summarized descriptively by type and location for the following subsets of treated bleeds:

- Type of bleeding (spontaneous, traumatic, unknown).
- Location of bleeding (joint, muscle, internal, skin/mucosa).
- Location and type of bleeding (joint spontaneous, joint traumatic, muscle spontaneous, muscle traumatic, internal spontaneous, internal traumatic, skin/mucosa spontaneous, skin/mucosa traumatic).

The estimated mean ABR and the corresponding 95% CI will be provided, for each subset, using the same method as applied to the primary analysis. These tables will be presented by study arm and treatment regimen for the FAS.

The summary of ABR by type and the summary of ABR by location will also be presented by age groups (12-17 years, ≥18 years) for the weekly prophylaxis arm (Arm A) based on the FAS.

As a description of the raw data collected in this study, the unadjusted number of bleeding episodes per participant will be summarized using categorical (0, 1, 2, 3, 4, 5, and >5) and descriptive statistics overall and by bleed type (spontaneous, traumatic, unknown). The number and percentage of participants experiencing a bleeding episode will be summarized categorically overall and by type of bleed (spontaneous, traumatic, unknown) as well as by bleed location (joint, muscle, internal, skin/mucosa, and unknown) for each bleed type; percentages will be based on the number of participants with an efficacy period in each study arm and treatment regimen in the FAS. The total participant years followed during the efficacy period will be provided in order to put the un-annualized numbers in perspective.

4.5.2 ABR for all bleeding episodes

In addition, ABR will be summarized for all bleeds (treated and untreated) by study arm and treatment regimen. All types of bleeding episodes (spontaneous, traumatic, and unknown) will be included, regardless of whether treatment was used or not. The estimated mean ABR and the corresponding 95% CI will be provided, using the same method as applied to the primary analysis.

Definition of a bleeding episode (based on all bleeds)

The definition of all bleeds will follow the ISTH criteria, accounting for both treated and untreated bleeds, based on the date and time of the bleed as follows:

In the analysis of all bleeds, a bleeding episode starts from the first sign of a bleed (treated and untreated) and ends no more than 72 hours after the last injection to treat the bleed or the last untreated bleed at the same location, within which any sign of bleeding (treated and untreated) at the same location, injections ≤72 hours apart, are considered the same bleeding episode.

Any injection or untreated bleed, occurring >72 hours after the preceding one, will be considered the start of a new bleeding episode. Any subsequent injection or untreated bleed at a different location will be considered a separate bleeding episode, regardless of the time from the last injection or bleed.

4.5.3 Intra-participant comparison of ABR in Arm B

An intra-participant comparison of ABR between weekly prophylaxis and on-demand treatment in Arm B will be performed using a Negative-Binomial regression model accounting for over-dispersion with the dependent variable as “total bleeding episodes”, covariate as “treatment regimen”, repeated variable as “subject”, and log time as an offset variable. A test of the null hypothesis that the paired rate ratio of ABR (weekly prophylaxis: on-demand) is ≥0.5 will be conducted at the 1-sided 0.025 level. Rejection of the null hypothesis ($p \leq 0.025$) will correspond to a clinically important reduction of at least 50% for the weekly prophylactic regimen as compared to the on-demand regimen.

4.5.4 Percentage of participants who maintain FVIII activity levels

The number and percentage of participants achieving steady-state trough FVIII activity levels above 1%, 5%, 10%, 15%, and 20% will be summarized for Arm A. In these summaries, FVIII activity level will be based on the average trough samples (ie, nominal 168-hour time point) from each scheduled visit (Week 4, Week 13, Week 26, Week 39, Week 52) using the aPTT-based one-stage clotting assay. Participants with trough samples that are outside 168 +/- 5 hours from the previous dose will be excluded from this analysis. A listing will be provided for each participant's baseline-corrected FVIII activity levels.

4.5.5 Annualized BIVV001 consumption

The consumption of BIVV001 will be annualized and summarized by study arm and treatment regimen for the FAS. The total annualized BIVV001 consumption (IU/kg) will be calculated for each participant using the following formula:

$$\text{Annualized consumption} = \frac{\text{Total IU/kg of BIVV001 during the efficacy period}}{\text{Total number of days during the efficacy period}} \times 365.25$$

The total amount of BIVV001 received will be the sum of the nominal IU/kg administered for each injection based on the units of BIVV001 as recorded from the participant's ePD and eCRF and the participant's most recent weight.

Total annualized BIVV001 consumption will be determined for the efficacy period, ie, excluding the PK assessments, surgery/rehabilitation periods (major or minor) and large injection intervals.

4.5.6 Number of injections and dose of BIVV001 to treat a bleeding episode

The number of injections and total dose of BIVV001 (IU/kg) to treat a bleeding episode will be determined on both a per-bleeding episode and per-participant basis. A bleeding episode is considered resolved when treatment for the bleeding is no longer needed.

Per bleeding episode: The total number of injections will consider all injections, including initial and follow-up injections, in a bleeding episode. The total dose of BIVV001 will be the sum of these doses. The number of injections to treat a bleeding episode will be summarized across all bleeding episodes both categorically (1, 2, 3, 4, >4; 1, >1; and ≤2, >2) and with descriptive statistics. The total dose of BIVV001 used to treat a bleeding episode will be summarized using descriptive statistics.

Per participant: The number of injections and total dose of BIVV001 to treat each bleeding episode, as determined for the per-bleeding episode summaries, will be averaged across all bleeding episodes for each participant. The average number of injections required for resolution of a bleeding episode will be summarized both categorically (1 to <2, 2 to <3, 3 to <4, and ≥4; and 1 to <2, ≥2) and with descriptive statistics. The averages for the per-participant total dose of BIVV001 used to treat a bleeding episode will be summarized using descriptive statistics.

4.5.7 Assessment of response to BIVV001 treatment of bleeding episodes

Participants will provide an assessment of response to each injection of BIVV001 for treating a bleed using a 4-point scale of excellent, good, moderate, and none, based on ISTH standardized definitions in hemophilia (6). Response categories of excellent and good will be presented combined as well as individually.

The assessment of response will be summarized first on a per injection basis. The number and percentage of injections in each response category will be tabulated based on all injections; percentages will be based on the total number of injections for treating a bleed for which a response was provided.

As a supplemental analysis, the assessment of response will also be summarized on a per bleeding episode basis by presenting the number and percentage of first injections to treat a bleeding episode for which the response to the treatment was categorized as excellent, good, moderate, or none. Percentages will be based on the number of bleeding episodes for which a response was provided for the first injection.

The subject's assessment of response during the efficacy period will be summarized by study arm and treatment regimen for the FAS.

4.5.8 Physician's global assessment of participant's response to BIVV001 treatment

The physician's global assessment of the participant's response to BIVV001 treatment will be summarized by study arm and treatment regimen for each study visit as the number and percentage of participants classified as excellent, effective, partially effective, and ineffective. Percentages will be based on number of participants for whom an assessment was provided at the respective visit.

This table will also include a cumulative tabulation across all scheduled study visits; participants can be included in this tabulation once for each visit. Percentages for this collection of responses throughout the study will be based on the total number of assessments across all visits.

4.5.9 Annualized joint bleeding rate

The annualized joint bleeding rate (AJBR) will be summarized descriptively by study arm and treatment regimen as described in Section 4.5.1. In addition, an intra-participant comparison of AJBR between weekly prophylaxis and on-demand treatment in Arm B will be performed using a Negative-Binomial regression model accounting for over-dispersion with the dependent variable as "total bleeding episodes", covariate as "treatment regimen", repeated variable as "subject", and log time as an offset variable. Statistical significance will be controlled at the 1-sided 0.025 level. The null hypothesis will be rejected if the paired rate ratio of AJBR (weekly prophylaxis: on-demand) is less than 0.5. This corresponds to a clinically important reduction of at least 50% for the weekly prophylactic regimen as compared to the on-demand regimen.

4.5.10 Target joint resolution

Participants will be assessed for target joints at screening by the investigator. A target joint is defined as a major joint (eg, hip, elbow, wrist, shoulder, knee or ankle) into which ≥3 spontaneous bleeding episodes occurred in a consecutive 6-month period. Resolution is achieved when ≤2 bleeds occur into that joint during 12 months of continuous exposure (Blanchette et al. 2014).

The percentage of participants with resolution of at least 1 target joint and the percentage of total target joints that are resolved at 52 weeks will be summarized for Arm A. Determination of resolution will be based on the number of spontaneous bleeding episodes into that joint as well as based on the number of total bleeding episodes into that joint (regardless of type). In this analysis, only participants who have at least 12 months of exposure (defined as duration of dosing ≥52 weeks) to BIVV001 will be included. Target joints on which surgery was performed on or before the end of the participant's 12 months of exposure will be censored.

4.5.11 Hemophilia Joint Health Score

Hemophilia Joint Health Score (HJHS) is a functional measure of joint health using ankle, knee and elbow (to access flexion, extension, range of movement, muscle strength, swelling, duration of swelling, crepitus, gait, pain, and muscle atrophy). The assessment is administered by a healthcare professional trained in the use of anthropometric measures. Ideally HJHS should be performed and assessed by the same Investigator or designee at each time point. HJHS will be assessed at Baseline, Weeks 26 and 52. Details of the questionnaire are provided in protocol Appendix 13 (Section 10.13).

Table 4 - HJHS domains

Range	Left Elbow	Right Elbow	Left Knee	Right Knee	Left Ankle	Right Ankle	Domain Total
Swelling	0-3	0-3	0-3	0-3	0-3	0-3	0-18
Duration	0-1	0-1	0-1	0-1	0-1	0-1	0-6
Muscle Atrophy	0-2	0-2	0-2	0-2	0-2	0-2	0-12
Crepitus on motion	0-2	0-2	0-2	0-2	0-2	0-2	0-12
Flexion loss	0-3	0-3	0-3	0-3	0-3	0-3	0-18
Extension loss	0-3	0-3	0-3	0-3	0-3	0-3	0-18
Joint pain	0-2	0-2	0-2	0-2	0-2	0-2	0-12
Strength	0-4	0-4	0-4	0-4	0-4	0-4	0-24
Joint Total	0-20	0-20	0-20	0-20	0-20	0-20	

The HJHS assessment includes the scoring of six joints (left ankle, right ankle, left elbow, right elbow, left knee, and right knee) on a scale from 0 to 20 according to the following criteria: swelling, duration of swelling, muscle atrophy, crepitus of motion, flexion loss, extension loss, joint pain, and strength (Table 4). Gait will be scored on a scale from 0 to 4 based on the number of skills that are not within the normal limits. The total score will be the sum of scores from all six joints plus the gait score (range from 0 to 124, with 0 being normal and 124 being the most severe disease).

The total joint score (ie, sum of scores from all six joints) will be defined and derived as follows:

- The total joint score will be set as missing if any one of the individual item score at any joint is missing.
- The total joint score will be set as missing if scores are evaluated within 2 weeks after a joint or muscle bleeding episode; and,
- The total joint score will be re-derived if surgery is performed on a joint. The scores for that joint will be replaced with the scores of the same joint at the last visit prior to the surgery using the last observation carried forward (LOCF) technique. Then, the total joint score in the subsequent visits are re-derived using the abovementioned scores.

The individual domain score (ie, sum of scores from all six joints for each domain) will be calculated using the same rules as above for all domains.

The total joint score, gait scores, and total score (calculated as the sum of total joint score and gait score), as well as change from baseline will be summarized by visits and by study arm. In addition, individual domain scores and change from baseline will be summarized by visits and by study arm.

In addition to the above descriptive analyses, the change from baseline to Week 52 in HJHS total score will be analyzed as part of the hierarchical testing procedure highlighted in Section 4.7. Since majority of study participants already had six months to one-year standard-of-care (SOC) prophylactic treatment in the lead-in observational study and all study participants in Arm A were on prior prophylaxis, baseline represents the participants' stable health status on their current SOC prophylactic treatment and thus change from baseline represents further improvement achieved with BIVV001 prophylactic treatment from SOC prophylactic treatment.

A likelihood-based, mixed-effects model with repeated measures (MMRM) will be conducted on the HJHS total score using observed data in Arm A, excluding visits that are coincidental with a major surgical/rehabilitation period. The MMRM model will include the baseline value of the endpoint as a covariate and visit as a fixed effect. A robust sandwich covariance matrix within a subject will be used. The adjusted mean change in HJHS total score from baseline to Week 52, along with its 95% CI, will be estimated by the MMRM model.

A sensitivity analysis will be performed for study participants who rollover from lead-in observational study using the same MMRM model as described above.

4.5.12 Haem-A-QoL and Haemo-QoL

Three HAEM-A-QoL / HAEMO-QoL questionnaires are used in this study and administered at Baseline, Weeks 26 and 52. Details of the questionnaires are provided in protocol Appendix 14 (Section 10.14.6 – 10.14.8).

- Haem-A-QoL (≥17 years old).
- Haemo-QoL (age group III; 13-16 years).
- Haemo-QoL (age group II; 8-12 years).

Each HAEM-A-QoL / HAEMO-QoL questionnaire will be summarized in terms of subscale scores and a total score according to the recommendations of the questionnaire authors. As specified in Section 5.5.1 (Appendix 5.1), questions will either have been coded or will be re-coded to ensure that high scores represent a low quality of life and low scores represent a high quality of life. These scores are then transformed to produce a Transformed Scale Score (TSS). This score is scaled, as a percentage, from 0 to 100%, with higher TSS values representing a worse quality of life for each sub-score and total score summary measurement.

When missing data are presented, a subscale score can only be calculated if at least 50% of questions for that subscale are answered (non-missing and not "Not Applicable").

For each HAEMO-QoL / HAEM-A-QoL questionnaire, the total score and individual subscale scores based on the TSS will be summarized descriptively for the actual value and change from baseline by visit by study arm.

In addition to the above descriptive analyses, the change from baseline to Week 52 in Haem-A-QoL physical health score will be analyzed as part of the hierarchical testing procedure highlighted in Section 4.7. Since majority of study participants already had six months to one-year standard-of-care (SOC) prophylactic treatment in the lead-in observational study and all study participants in Arm A were on prior prophylaxis, baseline represents the participants' stable health status on their current SOC prophylactic treatment and thus change from baseline represents further improvement achieved with BIVV001 prophylactic treatment from SOC prophylactic treatment.

A likelihood-based, mixed-effects model with repeated measures (MMRM) analysis will be conducted on the Haem-A-QoL physical health score using observed data in Arm A excluding visits that are coincidental with a major surgical/rehabilitation period. The MMRM model will include the baseline value of the endpoint as a covariate and visit as a fixed effect. A robust sandwich covariance matrix within a subject will be used. The adjusted mean change in Haem-A-QoL physical health score from baseline to Week 52, along with its 95% CI, will be estimated by the MMRM model.

A sensitivity analysis will be performed for study participants who rollover from lead-in observational study using the same MMRM model as described above.

4.5.13 PROMIS instruments

Five PROMIS instruments are used in this study and administered at Baseline, Weeks 26 and 52. Details for each instrument are provided in protocol Appendix 14 (Section 10.14.1 – 10.14.5).

- PROMIS Pain Intensity (v1.0 Pain Intensity 3a; all subjects).
- PROMIS-SF Pain Interference (v1.0 Pain Interference 6a; ≥18 years old).
- PROMIS Pediatric-SF Pain Interference (v2.0 Pain Interference 8a; <18 years old).
- PROMIS-SF Physical Function (v2.0 Physical Function 6b; ≥18 years old).
- PROMIS Pediatric-SF Physical Activity (v1.0 Physical Activity 8a; <18 years old).

For each PROMIS instrument, each question has five response options ranging in value from one to five. To find the total raw score for an instrument with all questions answered, sum the values of the response to each question. All questions must be answered in order to produce a valid total score. The total raw score is then converted into a T-score for each participant using an instrument-specific *Conversion Table* (see Section 5.5.2, Appendix 5.2). The T-score rescales the raw score into a standardized score with a mean of 50 and a standard deviation (SD) of 10. Therefore, a person with a T-score of 40 is one SD below the mean.

A higher PROMIS T-score represents more of the concept being measured. For negatively-worded concepts like Pain Intensity 3a and Pain Interference, a T-score of 60 is one SD worse than average. By contrast, for positively-worded concepts like Physical Function and Physical Activity, a T-score of 60 is one SD better than average.

T-scores for each PROMIS instrument will be summarized descriptively for the actual value and change from baseline by visit by study arm.

In addition to the above descriptive analyses, the change from baseline to Week 52 in PROMIS Pain intensity 3a past 7 days intensity of pain at its worst score (PAINQU6) will be analyzed as part of the hierarchical testing procedure highlighted in Section 4.7. Since majority of study participants already had six months to one-year standard-of-care (SOC) prophylactic treatment in the lead-in observational study and all study participants in Arm A were on prior prophylaxis, baseline represents the participants' stable health status on their current SOC prophylactic treatment and thus change from baseline represents further improvement achieved with BIVV001 prophylactic treatment from SOC prophylactic treatment.

A likelihood-based, mixed-effects model with repeated measures (MMRM) will be conducted on the raw score of the PROMIS Pain Intensity 3a first item (ie, "How intense was your pain at its worst", described in Protocol Appendix 14-Section 10.14.1) using observed data in Arm A, excluding visits that are coincidental with a major surgical/rehabilitation period. The MMRM model will include the baseline value of the endpoint as a covariate and visit as a fixed effect. A robust sandwich covariance matrix within a subject will be used. The adjusted mean change in PROMIS Pain Intensity 3a score from baseline to Week 52, along with its 95% CI, will be estimated by the MMRM model.

A sensitivity analysis will be performed for study participants who rollover from lead-in observational study using the same MMRM model as described above.

4.5.14 Surgery endpoints

An overall summary of surgeries will be provided which summarizes the number of major and minor surgeries and (%) of participants with at least one major surgery and (%) of participants with at least one minor surgery.

4.5.14.1 Investigators' or Surgeons' assessment of participant's hemostatic response to BIVV001 treatment

The Investigators'/Surgeons' assessment of the participant's hemostatic response to BIVV001 treatment will be collected at 24 hours post-surgery based on the ISTH 4-point response scale and will be summarized categorically and using descriptive statistics for all major surgeries for participants in the surgery subgroup. Categorically, the number and percentage of surgeries given each rating will be tabulated. Percentages will be based on the number of surgeries for which a response was provided. Since the response is given as an ordered ordinal scale, the responses have also been given a numeric score (Excellent=1, Good=2, Fair=3, Poor/none=4). A lower average score indicates a better Investigators'/Surgeons' assessment of the participants' response to surgery with BIVV001 treatment. Descriptive statistics will be provided using the numeric value of the 4-point scale.

4.5.14.2 Number of injections and dose to maintain hemostasis for major surgery

The number of injections, mean dose per injection (IU/kg), and total dose (IU/kg) required to maintain hemostasis during surgery will be summarized for all major surgeries for participants in the surgery subgroup. The number of injections per surgery will be summarized categorically (0, 1, 2, 3, 4, >4) and with descriptive statistics. Percentages for the categorical summary will be based on the number of major surgeries. The mean dose per injection and total dose required to maintain hemostasis will be summarized using descriptive statistics. The mean dose per injection will be determined as the average dose across all injections per major surgery (including the loading dose); the total dose will be determined as the sum across all injections (including the loading dose) per major surgery.

4.5.14.3 Total BIVV001 consumption for major surgery

Total consumption (IU/kg) per major surgery on the day of surgery, for the first 2 weeks following surgery (Days 1-3, 4-14, and 1-14), and for the overall surgical/rehabilitation period will be summarized using descriptive statistics for all major surgeries for the surgery subgroup. The day of surgery refers to the calendar day of the surgery and includes the loading dose given for that surgery. The first 2 weeks following surgery begins the day after surgery and extends for 14 calendar days. The overall surgical/rehabilitation period is defined in Section 4.1.1.3. Total BIVV001 consumption will be determined as the sum of all doses administered during the referenced time periods.

4.5.14.4 Estimated blood loss for major surgery

The estimated total blood loss during and post each major surgical procedure will be summarized using descriptive statistics for all major surgeries for the surgery subgroup.

4.5.14.5 Number and type of blood component transfusions for major surgery

The number of transfusions per surgery (regardless of the type of transfusion), the number of transfusions summed across all surgeries for each type of transfusion, and the number of surgeries

requiring each type of transfusion will be summarized categorically for all major surgeries for the surgery subgroup. Percentages in the categorical summaries will be based on the number of major surgeries for which the respective data is available.

In addition, the above data for both major and minor surgeries will be provided in the listings.

4.6 EXPLORATORY ENDPOINTS ANALYSES

4.6.1 HAL and pedHAL

The Hemophilia Activities List (HAL; in patients ≥18 years of age) and pediatric HAL (pedHAL; in patients <18 years of age) questionnaires assess the participant's functional ability to perform activities of daily living. Both questionnaires are administered at Baseline, Weeks 26 and 52. Details of the questionnaires are provided in protocol Appendix 14 (Section 10.14.9 – 10.14.10).

The HAL/pedHAL questionnaire will be summarized in terms of domain scores, component scores, and an overall score. The algorithms for deriving these scores are provided in Section 5.5.3 (Appendix 5.3). The derived scores and their changes from baseline will be summarized by visit by study arm.

4.6.2 Treatment Satisfaction Questionnaire for Medications (TSQM-9)

The 9 items Treatment Satisfaction Questionnaire for Medication (TSQM-9) is a validated psychometric tool that provides a general measure of patient satisfaction with treatment. The TSQM-9 are administered at Baseline, Weeks 26 and 52. A copy of the questionnaire is provided in protocol Appendix 14 (Section 10.14.11).

The TSQM-9 will be summarized in terms of three scale scores. The algorithms for deriving the TSQM-9 Scale scores are provided in Section 5.5.4 (Appendix 5.4). The derived TSQM-9 Scale scores and their changes from baseline will be summarized by visit by study arm.

4.6.3 EQ-5D-5L

The EQ-5D-5L are administered at Baseline, Weeks 26 and 52. The instrument consists of 2 pages: the EQ-5D descriptive system and the EQ visual analogue scale (VAS). A copy of the questionnaire is provided in protocol Appendix 14 (Section 10.14.12).

The questionnaire will be analyzed according to the recommendations of the authors (see www.euroqol.org). The descriptive system contains 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each with 5 response categories (no problems, slight problems, moderate problems, severe problems, and extreme problems). The number and percentage of subjects in each response category will be tabulated by visit by study arm. 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. The EQ VAS will be summarized for the observed response and change from baseline by visit by study arm.

EQ-5D-5L 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 authors (documents and calculation tools will be downloaded from the EuroQol website). The derived EQ-5D-5L index scores and change from baseline will be summarized by visit by study arm.

4.6.4 Patient Global Impression of Severity (PGIS)

Two components of the Patient global impression of severity (PGIS) are used in this study and administered at Baseline, Weeks 26 and 52. A copy of the questionnaire is provided in protocol Appendix 14 (Section 10.14.13).

- PGIS – joint symptoms
- PGIS – physical activity

The PGIS components will be summarized as ordinal variables by visit by study arm.

4.6.5 Patient Global Impression of Change (PGIC)

The Patient Global Impression of Change (PGIC) consists of one item that evaluates patients' overall health in terms of improvement or decline from baseline. The PGIC is administered at Weeks 26 and 52. A copy of the questionnaire is provided in protocol Appendix 14 (Section 10.14.14).

The PGIC will be summarized as an ordinal variable by visit by study arm.

4.6.6 Treatment preference survey

Treatment preference survey consist of 2 questions on perceived impact of treatment at the end of the study. The questionnaire is administered at Week 52/EOS. A copy of the questionnaire is provided in protocol Appendix 14 (Section 10.14.15).

The questionnaire will be summarized as an ordinal variable by study arm.

4.6.7 Physical Activity Monitor (ActiGraph)

Where available, assessments of physical activity (PA) will be done by using a triaxial medical grade accelerometer (ActiGraph Activity Monitor). Patients will be issued the wrist-worn device for continuous monitoring (daily for 8 continuous days) after the scheduled visits at Screening, Baseline, Weeks 4, 13, 26, 39, and 52.

PA measures may include raw acceleration, activity counts, steps taken, physical activity intensity, activity bouts, sedentary bouts, and body position among other biometric measures. The calculation of PA measures during the assessment period requires less than 50% of missing data.

The physical activity monitor analysis is not in the scope of this SAP and will be reported separately. Detail of analysis will be covered in a separate SAP.

4.6.8 Ultrasound measures (if applicable)

A subset of study sites will be selected to perform evaluation of treatment efficacy using the Joint Tissue Activity and Damage Examination (JADE) Protocol in Musculoskeletal Ultrasound (MSKUS) and/or Hemophilia Early Arthropathy Detection with Ultrasound (HEAD-US). Participants at designated and MSKUS-experienced sites who are in Arm A (prophylactic treatment), meet enrollment criteria and are >17 years of age at enrollment will have the option to participate in the JADE and/or HEAD-US protocol, at Baseline and Week 52. Assessments will include the following indicators of joint health:

- Changes in synovial hypertrophy.
- Changes in cartilage thickness.
- Changes in synovial hyperaemia, as observed by power Doppler (JADE).

The US images will be scanned and sent to the third-party imaging CRO for evaluation of the indicators listed above, using two blinded readers with separate adjudication for disagreements.

4.6.9 Healthcare resource utilization (HRU)

The number (0, 1-2, 3-4, >4) and percentage of healthcare visits by type (office visit, hospital ER visit, hospitalization, and ICU stay, etc.) will be summarized by visit and arm for the FAS.

4.6.10 Exit Interview

Qualitative semi-structured interviews will be conducted in a subset of adult participants from select countries within 6 months of exiting the study or after rolling over into the open-label extension study.

The exit interview is not in the scope of this SAP and will be reported separately.

4.7 MULTIPLICITY ISSUES

Type I error for secondary endpoints is controlled through a hierarchical testing framework. The α level is 0.05. Following the estimation approach described in Section 4.3.2 for the primary endpoint (ABR in Arm A), the key and selected secondary endpoints are included in the hierarchy in the following order:

1. Arm A intra-patient comparison non-inferiority (NI): ABR of BIVV001 weekly prophylaxis treatment vs. historical prophylaxis treatment (Section 4.4.1).

2. Arm A intra-patient comparison superiority: ABR of BIVV001 weekly prophylaxis treatment vs. historical prophylaxis treatment (Section 4.4.1).
3. Arm A change from Baseline to Week 52: Haem-A-QoL physical health score (Section 4.5.12).
4. Arm A change from Baseline to Week 52: PROMIS Pain intensity 3a past 7 days intensity of pain at its worst score (PAINQU6) (Section 4.5.13).
5. Arm A change from Baseline to Week 52: HJHS total score (Section 4.5.11).

No multiplicity adjustment will be made on other secondary efficacy variables than mentioned above.

4.8 SAFETY ANALYSES

The summary of safety results will be presented by study arm (and treatment regimen when indicated) as well as overall for the Safety Analysis Set. The analysis of the safety variables will be essentially descriptive, and no systematic testing is planned.

4.8.1 Extent of exposure

The extent of exposure and compliance will be assessed and summarized by study arm and treatment regimen, and overall. In addition, study drug administered will be listed by participant including the reason for administration, date and time of administration, and dose. A listing of lot numbers and nominal potency of the lots will be provided by participant.

Except for PK doses, the unit body weight dose (IU/kg) for analysis of dosing will be calculated as the total IU (nominal dose) for each injection divided by the participant's most recent weight in kg prior to the dose of study drug.

PK doses are calculated using the actual potency of the vial (between 80 to 125% of nominal strength) and used partial vials where necessary. Therefore, PK doses in the analysis are similarly calculated using the exact number of complete and partial vials, the volume and actual potency of each vial and the participant's most recent weight:

$$\text{Dose (IU/kg)} = \frac{\left(\frac{\text{Total volume administered}}{\text{Volume of vial}}\right) \times \text{Actual Potency of Vial}}{\text{Weight (kg)}}$$

4.8.1.1 Overall exposure

The extent of investigational medicinal product (IMP) exposure will be assessed by the number of injections and exposure days to BIVV001 and duration of BIVV001 dosing based on the Safety Analysis Set.

Number of Injections and Exposure Days to BIVV001

For any participant, the total number of days of exposure to BIVV001 will be accumulated from the time of their first on-study injection of BIVV001. An ED is a 24-hour period in which one or more BIVV001 injections are given. The 24-hour window starts from the first injection on study and then for subsequent injections, it starts from an injection taken after/outside of a previously identified ED.

The total number of EDs on BIVV001 for each participant will be summarized categorically (<5, 5-<10, 10-<25, 25-<50 and ≥50; and ≥1, ≥5, ≥10, ≥25, ≥50) and with descriptive statistics for the Safety Analysis Set.

The total number of injections per participant will be summarized overall and by reason for injection (prophylactic regimen, spontaneous bleed, traumatic bleed, follow-up injection, surgical or other) using descriptive statistics for the Safety Analysis Set.

Duration of BIVV001 Dosing

The duration of BIVV001 dosing for a given treatment regimen will be calculated from the start date time of the treatment regimen to the end date time of the treatment regimen, as defined in Section 4.1.1.1. The duration of overall BIVV001 dosing will be calculated from the date time of the first BIVV001 dose to the end date time of the last treatment regimen in the study. Any interruptions to dosing will not be accounted for when calculating this duration.

Duration of BIVV001 dosing (weeks) will be summarized using descriptive statistics for the Safety Analysis Set. Weeks will be represented in the descriptive statistics as if these were data collected with 1 decimal place. The number and percentage of participants whose duration of dosing was at least 13, 26, 39, and 52 weeks will be summarized.

4.8.1.2 Compliance

Compliance will be assessed during the efficacy period and will be summarized for the FAS. Except for doses administered in the clinic, study treatment may be administered by the participant or a caregiver. Data from the eCRF and ePD will be considered for the analysis of treatment received and participants' compliance with the study protocol.

Compliance of prophylaxis injections

The compliance rate of each participant to the study prophylactic dosing regimen during the efficacy period will be calculated in 2 ways: As dose compliance and as dosing interval compliance. Compliance will first be determined on a per-injection basis and then on a per-participant basis. That is, compliance for an individual dose or dosing interval will be determined and then the overall percentage of doses and dosing intervals that were in compliance will be determined for each participant.

For the purpose of evaluating compliance, the following will be considered per injection:

- The nominal dose taken compared to the study dose.

- The actual day of treatment compared to the study day of treatment.

An individual dose will be considered compliant if it is within 80%-125% of the study dose (50 IU/kg). An individual dosing interval will be considered compliant if the time between two prophylactic doses is within 36 hours of the study dosing interval (7 days).

The actual dosing intervals will be calculated as the length of time between consecutive prophylactic doses that are not separated by a bleeding episode or surgical/rehabilitation period (date/time of $dose_{x+1}$ – date/time of $dose_x$). The actual time between doses will be determined in minutes and converted to days as the number of minutes divided by 1440. The absolute value of the difference between the actual dosing interval and the study dosing interval of 7 days must be ≤36 hours in order to be compliant.

All prophylactic injections will be used to determine prophylactic dose compliance; only the prophylactic dosing intervals that are not separated by a bleeding episode or surgical/rehabilitation period will be used to evaluate prophylactic interval compliance. Large injection intervals as defined in Section 4.1.1.2 will be included in the interval compliance calculation. Dose and dosing interval compliance rates per participant will be determined as follows:

$$\text{Dose compliance rate} = \frac{\text{Number of doses taken within 80\%-125\% of study dose}}{\text{Total number of doses}} \times 100$$

where the percentage of a study dose is calculated as: (nominal dose taken/ study dose) ×100.

$$\text{Dose interval compliance rate} = \frac{\text{Number of doses taken within +/- 36 hours of study day/time}}{\text{Total number of intervals}} \times 100$$

Both participant dose and dosing interval compliance rates will be summarized as continuous variables using descriptive statistics as well as categorically (<80%, ≥80%) for the FAS.

A participant is considered “dose compliant” or “dosing interval compliant” if his respective rate is at least 80%. Based on this, participants will be further classified into the following mutually exclusive categories for overall compliance to their prophylactic treatment as:

- Both dose and interval compliant.
- Dose compliant or interval compliant (but not both).
- Neither dose nor interval compliant.

Compliance of ePD contemporaneous data entry

Injections (for prophylaxis or a bleed) and untreated bleeds must be entered into the ePD within 7 days from the date of the injection or the untreated bleed. Injections or untreated bleeds entered outside the 7-day window will be reported as protocol deviations. Descriptive statistics of the percentage of the participants with fewer than 80% of their total individual ePD records entered

within this 7-day window and those with ≥80% of records meeting this criterion will be presented for the FAS as well as a summary of % compliance to this criterion by participant.

4.8.2 Adverse events

General common rules for adverse events

Summaries of adverse events will be presented by study arm and treatment regimen as well as overall for the Safety Analysis Set.

All adverse events will be coded using the version of Medical Dictionary for Regulatory Activities (MedDRA) in effect at Sanofi at the time of database lock.

The primary focus of adverse event reporting will be on treatment-emergent adverse events (TEAEs). TEAEs are AEs that developed, worsened or became serious during the treatment-emergent period. Pretreatment adverse events will be listed separately.

An AE with incomplete or missing date/time of onset (occurrence, worsening, or becoming serious) will be classified as a TEAE unless there is definitive information to determine it is a pre-treatment AE. Details on classification of adverse events with missing or partial onset dates are provided in Section 5.3 (Appendix 3).

TEAEs that occurred during major surgical/rehabilitation periods will be included in the overall (top-line) summary of TEAEs but not in any of the other TEAE tables. The exception to this rule is that TEAEs occurring during a major surgical/rehabilitation period with an onset date on the day the surgical/rehabilitation period starts or on the day of the surgery (whichever comes earlier) will be included in the TEAE summaries. Consideration is given to AEs with an onset date at the start of the surgical/rehabilitation period in the event the pre-surgical dose was administered the day before the surgery. TEAE incidence tables will present by system organ class (SOC) and preferred term (PT).

Listings will be provided for all AEs, SAEs, AEs resulting in discontinuation of study treatment and/or from the study, and deaths. AEs that are emergent prior to the first BIVV001 treatment, AEs that are emergent during a major or minor surgical/rehabilitation period, and AEs that are emergent on the day the major surgical/rehabilitation period starts will be flagged.

Analysis of all adverse events

Overall summary of treatment-emergent adverse events

An overall (top-line) summary of TEAEs will be provided which summarizes number of TEAEs, treatment-emergent serious adverse events (TESAEs) and treatment emergent adverse events of special interest (TEAESI) and (%) of participants with any: TEAE, related TEAE, TESAE, related TESAE, TEAESI, related TEAESI**,** TEAE leading to death, and TEAE leading to treatment discontinuation.

Analysis of all treatment-emergent adverse events

The following TEAE summaries will be generated for the safety population.

- TEAEs by SOC and PT.
- TEAEs by PT in descending order of incidence.
- TEAEs by relationship, presented by SOC and PT. AEs with a missing relationship will be counted as "Related" in the summary table, as described in Section 5.3 (Appendix 3). A participant will be counted once for each SOC and PT based on the highest relationship within that SOC and PT, respectively.
- TEAEs by maximal severity, presented by SOC and PT. AEs with a missing severity will be counted as "Severe" in the summary table, as described in Section 5.3 (Appendix 3). A participant will be counted once for each SOC and PT based on the greatest severity within that SOC and PT, respectively.

Analysis of all treatment emergent serious adverse event(s)

- TESAEs by SOC and PT.
- TESAEs by PT in descending order of incidence.
- TESAEs by relationship and SOC and PT.

Analysis of all treatment emergent adverse events of special interest

- TEAESI by SOC and PT.
- TEAESI by relationship and SOC and PT.

Analysis of all treatment-emergent adverse event(s) leading to treatment discontinuation

- TEAEs leading to treatment discontinuation, by SOC and PT.

Analysis of treatment-emergent adverse event(s) occurring during a major or minor surgical/rehabilitation period

- All TEAEs occurring during a major or minor surgical/rehabilitation period will be provided in a listing. The following categories will be flagged separately:
  - TEAEs occurring during a major surgical/rehabilitation period, excluding AEs that occurred on the day the surgical/rehabilitation period starts,
  - TEAEs occurring during a major surgical/rehabilitation period with an onset date on the day the surgical/rehabilitation period starts,
  - TEAEs occurring during a minor surgical/rehabilitation period.

Subgroup analysis

- All TEAEs will also be provided by the factors below:
  - intrinsic factors: age (12–17, 18–64, ≥65 years old), BMI (<30 kg/m^2 and BMI ≥30 kg/m^2), race (White, Black, Asian, Other), HIV (positive, negative), HCV (positive, negative),
  - extrinsic factors: geographic location (Asia/Pacific, Europe, North America, South America).

Analysis of deaths

A listing of AEs with an outcome of death will be provided.

Embolic and Thrombotic Events

The occurrence of embolic and thrombotic events will be described by study arm and treatment regimen. The analysis will consist of a search of TEAE data using the Embolic and Thrombotic Events Standard MedDRA Query (SMQ). Medical adjudication of the search results will also be performed according to the following definition:

- Embolic and thrombotic events are defined as arterial or venous thrombosis, confirmed by imaging.
- Coronary artery thrombosis/occlusion must be confirmed by coronary angiography to be included as part of medical adjudication.
- Thrombosis involving the cerebral vasculature must be confirmed by imaging such as magnetic resonance imaging venogram (MRV), computed tomography venogram (CTV), magnetic resonance angiography (MRA), or computed tomography angiography (CTA) to be included as part of medical adjudication.
- An indwelling central venous access device is a well-established risk factor for thrombosis (7) and thrombotic events associated with such devices will not be included as part of medical adjudication. Occlusion or malfunction of a central venous access device also will not be included as part of medical adjudication.
- Infusion thrombophlebitis is a recognized complication of peripheral vein infusion (8) and will not be included as part of medical adjudication.

4.8.3 Inhibitor

Development of inhibitors will be identified as an inhibitor result of ≥0.6 BU/mL that is confirmed by a second test result of ≥0.6 BU/mL from a separate sample, drawn 2 to 4 weeks following the date when the original sample was drawn. Both tests must be performed by the central laboratory using the Nijmegen-modified Bethesda assay.

The overall incidence of positive inhibitor formation will be calculated as:

$$\frac{\text{Number of participants with an inhibitor}}{\text{Number of participants reaching ED milestone or who have an inhibitor}}$$

In this analysis, any participant who develops an inhibitor following the initial BIVV001 administration will be included in the numerator, regardless of the number of EDs to BIVV001; the denominator will include participants who have an inhibitor as well as participants with a valid inhibitor test following at least 50 EDs of BIVV001.

The calculation will also be performed including all participants with a valid inhibitor test following at least 25 EDs and including all participants with a valid inhibitor test, regardless of how many days they were exposed to BIVV001.

An exact 95% confidence interval for the proportion of participants with a positive inhibitor will be calculated using the Clopper-Pearson method for a binomial proportion.

Results from blood samples collected during surgical/rehabilitation periods for the purpose of determining the presence of an inhibitor will be included in this analysis.

4.8.4 Additional safety assessments

By-visit summaries of laboratory variables and vital signs will be presented by study arm and overall for the Safety Analysis Set. Shifts and potentially clinically significant laboratory abnormality (PCSA) summaries will be presented by study arm and treatment regimen, as well as overall for the Safety Analysis Set.

4.8.4.1 Laboratory variables

Blood samples for clinical laboratories will be taken at all scheduled visits. Clinical laboratory values will be analyzed after conversion into standard international units; international units will be used in all listings and tables. Data collected at local laboratories will be only included in the PCSA analysis and listing, and shift summary.

The laboratory parameters will be classified as follows:

- Hematology
  - Red blood cell (RBC) count,
  - White blood cell (WBC) count and differential,
  - Hemoglobin (Hgb),
  - Hematocrit (HCT),
  - Platelet count.

- Clinical chemistry
  - Alanine aminotransferase (ALT),
  - Aspartate aminotransferase (AST),
  - Alkaline phosphatase (ALP),
  - Gamma glutamyl transferase (GGT),
  - Bilirubin,
  - Blood Urea Nitrogen (BUN),
  - Creatinine,
  - Glucose,
  - Total protein,
  - Potassium,
  - Sodium,
  - Chloride.
- von Willebrand Comprehensive panel
  - VWF ristocetin cofactor activity,
  - VWF antigen.

Laboratory evaluations will be summarized for the Safety Analysis Set. Laboratory evaluations taken during major surgical/rehabilitation periods will not be included in the summaries. In the event of retests or repeat assessments at the same time point, the last non-missing evaluable measurement will be used for the purpose of analysis. Laboratory values of the form "<x" (ie, below the lower limit of quantification [LLOQ]) or ">x" (ie, above the upper limit of quantification [ULOQ]) will be imputed as "x" in the calculation of summary statistics but displayed as "<x" or ">x" in the listings.

Change from baseline

The summary statistics (including number, mean, median, standard deviation, minimum and maximum) of all laboratory variables (central laboratory values and changes from baseline) will be calculated for each visit by study arm and overall.

Shifts

Each participant's laboratory values will be classified according to whether the test result is "low" (below the lower limit of normal [LLN]), "normal" (within the normal range), or "high" (above the upper limit of normal [ULN]). Shift tables will be constructed based on both the minimum and maximum post baseline values for each participant. Data collected from unscheduled visits will be included in the determination of the per participant 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 participants with a shift to low (from normal, high, or unknown) and the number of participants with a shift to high (from normal, low, or unknown) will be tabulated; percentages will be based on the number of participants at risk. The number at risk for a shift to low (high) is the number of participants 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 5.

Table 5 - Direction of change indicating clinical concern for laboratory tests

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

Potentially Clinically Significant Laboratory Abnormalities

Abnormal laboratory values will also be evaluated by determining the number and percentage of participants with at least one PCSA over the course of the study that also represents a worsening from baseline. The PCSA values are defined as abnormal values considered medically important by the Sponsor according to predefined criteria/thresholds based on literature review and defined by the Sponsor for clinical laboratory tests and vital signs (Section 5.4, Appendix 4).

Participants 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 participants with an abnormality. Percentages will be based on the number of participants with at least one post baseline value for the given laboratory test. Data collected from unscheduled visits will be included in this analysis.

4.8.4.2 Vital signs and physical examination

Vital signs will be taken at Baseline, Weeks 4, 26 and 52. Vital signs include temperature, pulse rate, respiratory rate and blood pressure. All vital sign evaluations will be summarized for the Safety Analysis Set. Data from unscheduled visits and vital sign measurements taken during major surgical/rehabilitation periods will not be included in any summary but will be included in the listings and flagged.

Change from baseline

The summary statistics (including number, mean, median, standard deviation, minimum and maximum) of all vital sign variables (and changes from baseline) will be calculated for each visit by study arm and overall for the Safety Analysis Set. Evaluations taken during major surgical/rehabilitation periods will not be included in any summary.

Potentially Clinically Significant Vital Sign Abnormalities (PCSA)

Frequencies and percentages of participants with potentially clinically significant abnormalities will be summarized by study arm and treatment regimen, and overall. Percentages will be based on the number of participants with a baseline and at least one post baseline value. Vital sign abnormality thresholds are defined in Section 5.4 (Appendix 4).

All physical examination findings are presented in a by-patient data listing.

4.9 OTHER ANALYSES

4.9.1 Pharmacokinetic analyses

All PK analyses will be performed based on the PK analysis set (PKAS).

If a repeat PK profile is required due to bleeding (or other reason) during the PK evaluations, then only the repeat profile will be used for the summary of the FVIII activity levels and in the analysis of PK summary parameters. Data from incomplete profiles will be included in data listings.

Actual sampling times, doses, and injection durations will be used for PK analysis. Nominal sampling times and doses will be used for the creation of tables, figures and listings.

FVIII activity values for each timepoint will be corrected for baseline activity at Day 1 using the residual decay function. The detailed methodology used to correct baseline activity will be provided either in the clinical study report or PK report. PK parameters will be computed using baseline corrected FVIII activity values.

PK parameters will be estimated using FVIII activity values measured by aPTT-based one-stage clotting assay, and values measured by chromogenic coagulation assay.

For participants with abbreviated PK sampling at Day 1, for each participant using noncompartmental methods the following pharmacokinetic parameters will be included but not

limited to: maximum FVIII activity (Cmax), area under the plasma FVIII activity-time curve (AUC), terminal half-life ($t_{1/2}$), volume of distribution at steady state (Vss), total clearance (CL), clearance at steady state(CLss), mean residence time (MRT), incremental recovery (IR), time of maximum concentration observed (T_{max}), trough activity (C_{trough}), time above predefined FVIII activity levels, in vivo recovery (IVR).

For participants in the sequential PK subgroup, selected pharmacokinetic parameters will be determined at Day 1 and Week 26 for each participant using noncompartmental methods. Time to percent FVIII activity will be calculated by simulation from the final population PK model using the post-hoc random effects from each of the study participants.

The detailed methodology for computing PK parameters using noncompartmental method will be described in the CSR or PK report. The detailed methodology for computing PK parameters using population PK method will be described in the population PK report.

The population PK analysis will be presented separately from the main CSR. Population PK analysis will be described fully in the population PK analysis plan and will be finalized before database lock.

For the PKAS the following analyses will be performed:

- For each assay type, the summary tables of FVIII activity levels will be provided for BIVV001 at scheduled PK visits and time points for Arm A Sequential PK (Baseline BIVV001, Repeat BIVV001), Abbreviated PK (Baseline BIVV001) by arm and overall, and Abbreviated PK (Baseline BIVV001) by age category, respectively.
- For each assay type, the peak and trough concentrations will be summarized by visit by arm and overall for the PKAS. The incremental recovery will be also summarized by visit by arm and overall for the PKAS. Summary descriptive statistics will include the number of non-missing values, mean, geometric mean, median, Q1, Q3, standard deviation, percent coefficient of variation, minimum, and maximum. In these summaries, values below the LLOQ will be imputed as zero.
- For each assay type, individual PK parameters will be listed for each participant and summarized descriptively for Arm A Sequential PK (Baseline BIVV001, Repeat BIVV001), Abbreviated PK (Baseline BIVV001) by arm and overall, and Abbreviated PK (Baseline BIVV001) by age category. Any FVIII assessments flagged by the PK scientist as being implausible and excluded from computation of PK parameters will be excluded from summaries but included in listings. The listings will indicate which values, if any, were excluded from summaries. A separate listing of excluded values, together with the reason for exclusion will be provided.

For the Sequential PK Subgroup the following analyses will be performed:

- For each assay type, FVIII activity versus time profiles will be plotted for each participant in both the original and log scale. In these plots, values below LLOQ will not be included.

- For each assay type, mean FVIII activity versus time profiles will be plotted by timepoint (Baseline BIVV001 and Repeat BIVV001) in both the original and log scale. In these plots, values below LLOQ will be imputed as zero in the calculation of the mean.
- The geometric mean of the intra-participant ratio of repeated BIVV001 to baseline BIVV001 to for selected PK parameters will be tabulated with 95% CI as necessary.

4.9.2 Additional Immunogenicity analysis

4.9.2.1 Anti-drug antibody

Participant's ADA status (see definitions below) will be summarized on the safety analysis set. A patient with at least one anti-drug antibody (ADA) sample (positive, negative or inconclusive) during the treatment emergent period is considered evaluable.

Participant's ADA status

- Pre-existing ADA: participant with at least one ADA positive prior to receiving BIVV001.
  - Treatment-boosted ADA: participant with pre-existing ADA positive and the ADA titer level anytime post-baseline is significantly higher than that at baseline. The post-baseline titer value that is at least 4-fold of pre-existing ADA titer value is considered significant.
  - Unclassified ADA: participant with pre-existing ADAs that cannot be classified as treatment-boosted ADA because of missing titer(s).
- Treatment-induced ADA: participant with all ADA negative (or missing) prior to receiving BIVV001 and at least one ADA positive at any time post-baseline.
- Treatment-emergent ADA: participant with treatment-induced or treatment boosted ADA.

ADA kinetics:

ADA response duration is defined as the date of last treatment-induced or treatment-boosted ADA sample minus date of first treatment-induced or treatment-boosted ADA sample +1. ADA response duration will be calculated only for participants with at least two treatment-induced or treatment-boosted ADA samples, irrespective of negative samples.

In case of participants with only one treatment-induced or treatment-boosted ADA, the ADA duration will be imputed to 0.

ADA response will be classified in the following categories for each participant:

- Persistent ADA response: an ADA response duration greater or equal than 16 weeks.
- Transient ADA responses: an ADA response duration less than 16 weeks and the last sample in the study is not treatment-induced nor treatment-boosted.
- Indeterminate ADA response: ADA response that is neither persistent nor transient.

The number and percentage of participants with pre-existing ADA, treatment-boosted, unclassified, treatment-induced ADA, and treatment emergent ADA will be summarized by study arm and overall for the Safety Analysis Set. The number and percentage of participants with different ADA kinetic status (ie, persistent, transient, indeterminate) will be summarized by study arm and overall for the Safety Analysis Set. Median and $25^{th}/75^{th}$ quantiles, minimal and maximal of the ADA titer for pre-existing ADA, treatment-boosted, unclassified, treatment-induced ADA, and treatment emergent ADA will be provided.

A listing will be provided for confirmed positive anti-BIVV001 antibody information, including sample collection date and time, and ADA characterization testing.

To evaluate the potential impact of ADAs and PK parameters, PK parameters for participants who have incidence of ADA positive (ie, treatment-boosted or treatment-induced), along with the group mean of participants who are ADA negative at all times will be tabulated. As needed, FVIII activity level over time for participants who have incidence of ADA positive (ie, treatment-boosted or treatment-induced), along with the group mean of participants who are ADA negative at all times may be evaluated.

4.10 INTERIM ANALYSES

No formal interim analyses are planned.

5 SUPPORTING DOCUMENTATION

5.1 APPENDIX 1 LIST OF ABBREVIATIONS

ABR: annualized bleeding rate
ADA: anti-drug antibody
AE: adverse event
AESI: adverse events of special interest
AI: accumulation index
AJBR: annualized joint bleeding rate
ALP: alkaline phosphatase
ALT: alanine aminotransferase
aPTT: activated partial thromboplastin time
AST: aspartate aminotransferase
AUC: area under the curve
BMI: body mass index
BUN: blood urea nitrogen
CI: confidence interval
CL: total clearance
CLss: total clearance at steady state
Cmax: maximum activity
CRO: contract research organization
CSR: clinical study report
C_{trough}: trough activity
dL: deciliter
eCRF: electronic case report form
ED: exposure days
EOS: end of study
ePD: electronic patient diary
ET: early termination
FAS: full analysis set
FDA: Food and Drug Administration
FVIII: Factor VIII
GGT: gamma glutamyl transferase
HAL: hemophilia activities list
HBV: Hepatitis B virus
HCT: hematocrit
HCV: Hepatitis C virus
HEAD-US: hemophilia early arthropathy detection with ultrasound
Hgb: hemoglobin
HIV: human immunodeficiency virus
HJHS: hemophilia joint health score
HRU: healthcare resource utilization
IMP: investigational medicinal product

IR:	incremental recovery
ISTH:	International Society on Thrombosis and Haemostasis
IU:	international unit
IVR:	in vivo recovery
JADE:	joint activity and damage exam
kg:	kilogram
LLN:	lower limit of normal
LLOQ:	lower limit of quantification
LOCF:	last observation carried forward
MedDRA:	medical dictionary for regulatory activities
MMRM:	mixed-effects model with repeated measures
MRT:	mean residence time
MSKUS:	musculoskeletal ultrasound
NI:	non-inferiority
PA:	physical activity
PCSA:	potentially clinically significant laboratory abnormality
PGA:	physician's global assessment
PGIC:	patient global impression of change
PGIS:	patient global impression of severity
PK:	pharmacokinetic
PKAS:	pharmacokinetic analysis set
PRO:	patient report outcome
PROMIS:	patient-reported outcomes measurement information system
PT:	preferred term
PTPs:	previously treated patients
QoL:	quality of life
QW:	once-weekly
RBC:	red blood cell
SAE:	serious adverse event
SAP:	statistical analysis plan
SD:	standard deviation
SMQ:	standard MedDRA query
SOC:	system organ class
$t_{1/2}$:	elimination half life
TEAEs:	treatment-emergent adverse events
Tmax:	time of maximum concentration observed
TSQM:	treatment satisfaction questionnaire for medications
TSS:	transformed scale score
ULN:	upper limit of normal
ULOQ:	upper limit of quantification
US:	Ultrasound
VAS:	visual analogue scale
Vss:	volume of distribution at steady state
WBC:	white blood cell
WHO-DD:	World Health Organization-drug dictionary

5.2 APPENDIX 2 DEMOGRAPHICS AND BASELINE CHARACTERISTICS, PRIOR OR CONCOMITANT MEDICATIONS

Demographic characteristics

Demographic variables include age, height, weight, body mass index (BMI), race, ethnicity and geographic location. Age at informed consent will be summarized both as a continuous variable using descriptive statistics and categorically (12-17, 18-64, and ≥65 years).

Demographic characteristics will be summarized by study arm and treatment regimen, including the surgery subgroup, and overall for the FAS using descriptive statistics.

Hemophilia history

The following hemophilia history variables will be summarized by study arm and treatment regimen, including the surgery subgroup, and overall for the FAS using descriptive statistics:

- Age at diagnosis of severe hemophilia (years).
- Family inhibitor history (Yes, No).
- Blood type (A, B, AB, O).
- Rh factor (Positive, Negative).
- Lowest documented historical FVIII level (<1%, ≥1%).
- FVIII genotype.
- Human immunodeficiency virus (HIV) status.
- Hepatitis C (HCV) status.
- Hepatitis B (HBV) status.
- Vaccination history in past year (Yes, No).
- Types of FVIII product previously administered at study entry.
- Age at start of first prophylaxis regimen (years; <6, 6 to <10, 10 to <18, and ≥18 years).
- Number of prior exposure days to FVIII (<150, ≥150).
- Number of bleeds in the past 12 months.
- Number of joint bleeds in the past 12 months.
- Target joint at baseline.

In addition, the most recent pre-study FVIII regimen and other prior FVIII regimens used in the past 12 months will be collected and summarized including:

- Categories for most recent pre-study FVIII regimen (prophylaxis, on-demand).
- Time on pre-study regimen (>12 months, 6-12 months, <6 months).
- Frequency of injections for those participants who indicated prophylaxis as their most recent pre-study regimen.

Medical and surgical history

Medical (or surgical) history will be collected at screening and will be coded using the version of Medical Dictionary for Regulatory Activities (MedDRA) in effect at Sanofi at the time of database lock.

The number and percentage of participants with any medical and surgical history will be summarized for each body system by study arm and treatment regimen, including the surgery subgroup, and overall using the Safety Analysis Set. A participant will be counted only once if the participant reported one or more occurrences of the same body system.

Prior or concomitant medications

All medications taken in the 30 days prior to the time the participant enrolls in the study and until the Safety Follow Up Call or Visit are to be reported on the eCRF.

Medications will be identified as being prior and/or concomitant based on the start and stop dates compared to the first dose of study drug. Prior medications are those taken before the first dose of study drug. Concomitant medications are those administered during or after the first dose of study drug. A medication that started prior to the first dose of study drug and was ongoing during and/or after the first dose of study drug will be classified as both prior and concomitant.

No imputation of medication start/end dates will be performed. If the start/end date 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 the medications as prior and/or concomitant as follows.

If a concomitant medication start day is missing then the medication will be assumed to be both a prior and a concomitant medication unless the start month and/or year or medication stop date can be used to determine if a medication is concomitant or prior, as follows:

- If the medication start day is missing, but the month and year precede the month and year of the first treatment and the medication stop date is before the date of first treatment, then the medication will be classified as prior only.
- If the medication start month is missing, but the year precedes the year of first treatment and the medication stop date is before the date of first treatment, then the medication will be classified as prior only.
- If the medication start day is missing and the month and/or year are on or after the month and year of the first treatment, then the medication will be classified as concomitant only.

All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD) using the version in effect at Sanofi at the time of database lock.

Separate summaries will be provided for prior and concomitant medications. Prior and concomitant medications will be presented by study arm and treatment regimen and overall for the Safety Analysis Set. Summaries will be based on the number and percentage of participants taking medications by WHO-DD standardized medication text. Within each WHO-DD standardized medication text a participant will be counted once even if he/she reported taking the medication more than once.

Concomitant medications taken for hemophilia-related pain

Concomitant medications taken for hemophilia-related pain will be summarized by study arm and treatment regimen and overall for the Safety Analysis Set. Summaries will be based on the number and percentage of participants taking pain medications by class (acetaminophen/ paracetamol, NSAIDs, opioids, other).

In addition, the number of days pain medication was used for hemophilia-related pain in the past 2 weeks will be summarized categorically (0 [None], 1-3, 4-7, 8-14) for each study visit by study arm and treatment regimen and overall.

5.3 APPENDIX 3 DATA HANDLING CONVENTIONS

Electronic Patient Diary (ePD) Data

Participant recorded diary data will be handled taking into consideration previous audit findings and industry standards to control changes to participant reported data. Only changes defined as not requiring participant confirmation and that are supported by site source documentation will be allowed.

The following programming algorithms are allowable to address record consolidation and true duplicate removal.

1. Record consolidation:

When a dose requires more than one vial and these vials are erroneously recorded in the ePD/eCRF as separate injections, albeit within a short time window, the change to consolidate these multiple records into one record is called single vial consolidation. The programming algorithm to identify and consolidate the records is as follows:

- Identify all injection records within 60 minutes of one another where the variables (injection date/times, lot numbers, number of vials, and vials strength in nominal IU) are not exactly the same on each record. There are four scenarios:
    - Date and time of injections are not exactly same, and lot numbers, number of vials, and vials strength in nominal IU are not the same.
    - Date and time of injections are not exactly same, but lot numbers, number of vials, and vials strength in nominal IU are the same.
    - Date and time of injections are exactly same, but lot numbers, number of vials, and vials strength in nominal IU are not the same.
    - Date and time of injections are exactly same, and lot numbers, number of vials, and vials strength in nominal IU are the same and these duplicates occurred for reasons which cannot be attributed to administrative issues as specified below.
- Consolidate injections as follows:

Record with earliest injection date/time retained with corresponding contextual information.

Combine/sum values related to the dose (eg, lot numbers, number of vials, volume injected, nominal and actual dose) into the single retained record.

If injections identified with the above algorithm have distinctly different reasons (eg, one injection is recorded as bleeding or surgery and another is recorded as prophylaxis, or OTHER), then the records should NOT be consolidated. However, if one record has reason or bleed information missing, then consolidation can be performed.

2. True duplicates removal:

When injections are identified in the ePD with exactly the same date/times, bleed information if applicable (type, location, and sublocation), lot numbers, number of vials, and vials strength in nominal IU, these may have resulted due to administrative issues and are therefore true duplicates. There are three types of administrative issues as follows:

Technical transmission issue.

Entry of same record into 2 different devices.

Records duplicated in ePD and eCRF which remain despite attempts to correct the data via the query process.

The programming algorithm to identify and delete the duplicates is as follows:

Identify all injections that have exactly the same date/times, reasons for injection, bleed information (if applicable), lot numbers, number of vials, and vials strength in nominal IU.

Remove the duplicates and keep a single record.

As with single vial consolidation, if injections identified as duplicates have distinctly different reasons (eg, one injection is recorded as bleeding or surgery and another is recorded as prophylaxis or OTHER), then these cannot be considered duplicates and removed. However, if one reason is missing, then they can be considered duplicates and removed.

Algorithm for determining analysis visit windows for PRO and HJHS observations

A scheduled measurement will be used if it is available. Otherwise, the following analysis window will decide how the unscheduled and EOS/ET visits will be used in the analyses of PRO and HJHS variables.

A measurement (unscheduled or EOS/ET) will be used if it is available and measurement date is within the analysis window.

After applying these time windows, if multiple assessments are associated to the same time point, the closest from the targeted study day will be used. If the difference is a tie, the value after the targeted study day will be used. If multiple valid values exist within a same day, then the first value of the day will be selected.

Table 6 - Analyses window definition

Scheduled visit post baseline	Target study day	Derived Window in study days	
		Lower bound	Upper bound
Baseline	1	-INF	1(7 for HJHS)
Week 26	182	2 (8 for HJHS)	272
Week 52	364	273	-

Handling of adverse events with missing or partial date/time of onset

No imputation of adverse event onset/end dates will be performed. If the onset/end date 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 the adverse events as treatment-emergent or not as follows:

- If the onset day of an adverse event is missing and the onset month and year of the AE are either the same as or later than the month and year of the first treatment, the AE will be considered a TEAE.
- If the onset day of an adverse event is missing and the onset month and year of the AE precede the month and year of the first treatment, the AE will not be considered a TEAE.
- If the onset month of an adverse event is missing and the onset year of the AE is either the same as or later than the year of first treatment, the AE will be considered a TEAE.
- If the onset month of an adverse event is missing and the onset year of the AE precedes the year of first treatment, the AE will not be considered a TEAE.
- If the onset day, month, and year of an adverse event are missing, the AE will be considered a TEAE.
- If start date is partial but the stop date can be determined to be before the start of the first dose of study drug, then the AE will not be considered a TEAE.

Handling of adverse events with missing relationship to investigational product

If the assessment of relationship to study drug is missing, the event will be counted as “related” in the frequency tables of treatment-emergent adverse events by relationship to study drug.

Handling of adverse events with missing severity

If the assessment of severity is missing, the event will be counted as “severe” in the frequency tables of treatment-emergent adverse events by severity.

Unscheduled visits

Unscheduled visit measurements of laboratory data and vital signs will not be included in the by-visit summaries but will be used for computation of baseline, maximum or minimum on-treatment values, and PCSAs.

5.4 APPENDIX 4 POTENTIALLY CLINICALLY SIGNIFICANT ABNORMALITIES CRITERIA

Table 7 - Threshold levels for potentially clinically significant hematology abnormalities

Parameter Name	Age	Low	High
White Blood Cells	≥18 years	<3.0 * 10^9/L (non-African American) <2.0 * 10^9/L (African American)	≥16 * 10^9/L
	12-<18 years	<4.5 * 10^9/L or 5000/mm3	>13.5 * 10^9/L or 17 000/mm3
Neutrophils	≥18 years	<1.5 * 10^9/L (non-African American) <1.0 * 10^9/L (African American)	N/A
	12-<18 years	<1.2 * 10^9/L or 1200/mm3	>1 ULN
Lymphocytes	≥18 years	N/A	>4.0 * 10^9/L
	12-<18 years	<0.6 * 10^9/L or 600/mm3	>6.0 * 10^9/L or 6000/mm3
Monocytes	≥18 years	N/A	>0.7 * 10^9/L
	12-<18 years	N/A	N/A
Eosinophils	≥18 years	N/A	>0.5 * 10^9/L or >ULN (if ULN≥0.5 * 10^9/L)
	12-<18 years	N/A	>0.5 * 10^9/L or >ULN (if ULN≥0.5 * 10^9/L)
Basophils	≥18 years	N/A	>0.1 * 10^9/L
	12-<18 years	N/A	N/A
Hemoglobin	≥18 years	≤115 g/L (male), ≤95 g/L (female); decrease from baseline ≥0.31 mmol/L or 20 g/L	≥185 g/L (male), ≥165 g/L (female)
	12-<18 years	<1.55 mmol/L or 10 g/dL; Decrease from baseline ≥0.31 mmol/L or 20 g/L	N/A
Hematocrit	≥18 years	≤0.37 v/v or 37% (male) ≤0.32 v/v or 32% (female)	≥0.55 v/v or 55% (male) ≥0.50 v/v or 50% (female)
	12-<18 years	<0.32 v/v or 32%	>0.47 v/v or 47%
Platelet count	≥18 years	<100 * 10^9/L	≥700 * 10^9/L
	12-<18 years	<100 * 10^9/L	≥700 * 10^9/L

N/A=not applicable

Table 8 - Threshold levels for potentially clinically significant chemistry abnormalities

Parameter Name	Age	PCS Low	PCS High
Alanine Aminotransferase (ALT)	≥18 years	N/A	>3 x ULN
	12-<18 years	N/A	>3 x ULN
Aspartate aminotransferase (AST)	≥18 years	N/A	>3 x ULN
	12-<18 years	N/A	>3 x ULN
Alkaline phosphatase (ALP)	≥18 years	N/A	>1.5 x ULN
	12-<18 years	N/A	>1.5 x ULN
Total Bilirubin	≥18 years	N/A	>1.5 x ULN
	12-<18 years	N/A	>1.3 x ULN
Blood urea nitrogen (BUN)	≥18 years	N/A	≥17 mmol/L or 47.6 mg/dL
	12-<18 years	N/A	≥6.4 mmol/L or 18 mg/dL
Creatinine	≥18 years	N/A	≥150 µmol/L or 1.70 mg/dL ≥ 30% change from baseline
	12-<18 years	N/A	≥132 µmol/L or 1.5 mg/dL
Glucose	≥18 years	≤3.9 mmol/L and <LLN; or ≤70 mg/dL and <LLN	≥11.1 mmol/L or 200 mg/dL (unfasted); ≥ 7 mmol/L or 126 mg/dL (fasted)
	12-<18 years	<2.7 mmol/L; or ≤ 50 mg/dL	≥7 mmol/L or 120 mg/dL (fasted); ≥10.0 mmol/L or 180 mg/dL (unfasted)
Potassium	≥18 years	<3 mmol/L or 3 mEq/L	≥5.5 mmol/L or 5.5 mEq/L
	12-<18 years	≤3.5 mmol/L or 3.5 mEq/L	≥5.5 mmol/L or 5.5 mEq/L
Sodium	≥18 years	≤129 mmol/L or 129mEq/L	≥160 mmol/L or 160mEq/L
	12-<18 years	≤129 mmol/L or 129mEq/L	≥150 mmol/L or 150mEq/L
Chloride	≥18 years	<80 mmol/L or 80mEq/L	>115 mmol/L or 115 mEq/L
	12-<18 years	<80 mmol/L or 80mEq/L	>115 mmol/L or 115 mEq/L

N/A = not applicable, ULN = upper limit of normal

Table 9 - Threshold levels for potentially clinically significant coagulation abnormalities

Parameter Name	Age	Low	High
VWF antigen	All	< LLN	N/A
VWF:RCo activity	All	< LLN	N/A

N/A = not applicable, ULN = upper limit of normal

Table 10 - Threshold levels for potentially clinically significant vital sign abnormalities

Variable	Age	Low	High
Systolic Blood Pressure	≥18 years	≤95 mmHg and decrease from baseline ≥20 mmHg	≥160 mmHg and increase from baseline ≥20 mmHg
	12-<18 years	≤90 mmHg and decrease from baseline ≥20 mmHg	≥119 mmHg and increase from baseline ≥20 mmHg
Diastolic Blood Pressure	≥18 years	≤45 mmHg and decrease from baseline ≥10 mmHg	≥110 mmHg and increase from baseline ≥10 mmHg
	12-<18 years	≤54 mmHg and decrease from baseline ≥10 mmHg	≥78 mmHg and increase from baseline ≥10 mmHg
Heart Rate	≥18 years	≤50 bpm and decrease from baseline ≥20 bpm	≥120 bpm and increase from baseline ≥20 bpm
	12-<18 years	≤50 bpm and decrease from baseline ≥20 bpm	≥120 bpm and increase from baseline ≥20 bpm
Temperature	12-<18 years	N/A	Ear/temporal artery: ≥100.4°F/38.0°C Oral: ≥99.5°F/37.5°C Axillary: ≥99°F/37.2°C
Respiration Rate	12-<18 years	<12 per minute	>20 per minute

5.5 APPENDIX 5 QUESTIONNAIRE SCORING

5.5.1 Appendix 5.1 Haem-A-Qol and Haemo-Qol

The HAEM-A-QoL questionnaire (for adults ≥17 years) and HAEMO-QoL questionnaires (for children and teenagers aged 8-12 years and 13-16 years) will be summarized in terms of subscale scores and a total score according to the recommendations of the questionnaire authors.

When missing data are presented, a subscale score can only be calculated if at least 50% of questions for that subscale are answered (non-missing and not "Not Applicable"), as specified in the tables below.

Haem-A-QoL (≥17 years old)

The Haem-A-Qol questionnaire for participants with hemophilia who are ≥17 years of age consists of 46 items pertaining to 10 subscales:

Time Period	Subscales for Haem-A-QoL	No. of questions	No. of non-missing items required to calculate subscale
In the past 4 weeks	1. Physical health	5	3
	2. Feeling	4	2
	3. View of yourself	5	3
	4. Sports and leisure	5	3
	5. Work and school	4	2

Time Period	Subscales for Haem-A-QoL	No. of questions	No. of non-missing items required to calculate subscale
	6. Dealing with hemophilia	3	2
	7. Treatment	8	4
Recently	8. Future	5	3
	9. Family planning	4	2
	10. Partnership and sexuality	3	2
	Total Score	46	23

Haemo-QoL (age group III; 13-16 years)

The Haemo-QoL questionnaire for participants with hemophilia who are 13-16 years of age consists of 77 items pertaining to 12 subscales:

Time Period	Subscales for Haemo-QoL	No. of questions	No. of non-missing items required to calculate subscale
In the past month	1. Physical health	7	4
	2. Feeling	8	4
	3. View of yourself	10	5
	4. Family	8	4
	5. Friends	4	2
	6. Perceived support	4	2
	7. Other people	6	3
	8. Sports and school	9	5
	9. Dealing with hemophilia	7	4
	10. Treatment	8	4
Recently	11. Future	4	2
	12. Romantic relationships	2	1
	Total Score	77	39

Haemo-QoL (age group II; 8-12 years)

The Haemo-QoL questionnaire for participants with hemophilia who are 8-12 years of age consists of 64 items pertaining to 10 subscales:

Time Period	Subscales for Haemo-QoL	No. of questions	No. of non-missing items required to calculate subscale
In the past month	1. Physical health	7	4
	2. Feeling	7	4
	3. View of yourself	9	5
	4. Family	5	3
	5. Friends	4	2
	6. Perceived support	4	2
	7. Other people	6	3
	8. Sports and school	8	4

Time Period	Subscales for Haemo-QoL	No. of questions	No. of non-missing items required to calculate subscale
	9. Dealing with hemophilia	7	4
	10. Treatment	7	4
	Total Score	64	32

Scoring algorithm

Per the algorithm on http://www.haemoqol.de/manual.htm, to correctly score the questionnaires, each version has to be identified and the appropriate scoring list has to be selected. High score represent low quality of life and scoring involves the following steps:

1. Assigning numbers to the response scale

1= never, 2=rarely, 3=sometimes, 4=often, 5=all the time

For negatively worded items, the above classification can be applied in which higher values represent a lower quality of life. For positively worded items, the score has to be recoded (see below).

2. Recoding positively worded items

Positively worded items in the subscales table have to be recoded so that numeric values assigned are reversed:

1=all the time, 2=often, 3=sometimes, 4=rarely, 5= never

By recoding, high scores in positively worded items reflect not higher but lower quality of life. The then unidirectional values can subsequently be added to yield the summed scores according to the scoring list.

3. Producing the raw score

A raw subscale score is produced by summing up all items within a subscale. Its range lies between the lowest possible (number of items (n) x 1) and highest possible (number of items (n) x 5) value of the respective scale.

The raw total score is produced by the addition of all items (instead of the subscale items only) of the questionnaire (again paying attention to the recoding procedure – see Steps 1 and 2).

4. Transferring a raw score to a Transformed Scale Score (TSS) between 0 and 100

Creating a Transformed Scale Score (TSS) makes it possible to express the score as a percentage between the lowest (0) and the highest (100) possible value. To obtain the TSS the following transformation rule has to be applied:

TSS = 100 x ((raw score – minimal possible raw score) / possible range of raw scores) where:

- Minimal possible raw score = number of questions answered × 1
- Maximal possible raw score = number of questions answered × 5
- Possible range of raw scores = maximum possible raw score - minimal possible raw score

Example: A raw score of 20 on the "Physical Health" scale is to be transformed:

- When all 7 questions were answered:
  - minimal possible score=7
  - maximal possible score=35
  - range of scores=28
  - TSS=20-7/28 = 46.4
- When 6 of the 7 questions were answered:
  - minimal possible score=6
  - maximal possible score=30
  - range of scores=24
  - TSS=20-6/24 = 58.3

Transformed scores can be produced from raw subscale and raw total scores.

The total and subscale scores based on TSS will be presented in outputs.

5.5.2 Appendix 5.2 PROMIS instruments

Five PROMIS instruments are used in this study:

A) PROMIS Pain Intensity (v1.0 Pain Intensity 3a; all subjects).
B) PROMIS-SF Pain Interference (v1.0 Pain Interference 6a; ≥18 years old).
C) PROMIS Pediatric-SF Pain Interference (v2.0 Pain Interference 8a; <18 years old).
D) PROMIS-SF Physical Function (v2.0 Physical Function 6b; ≥18 years old).
E) PROMIS Pediatric-SF Physical Activity (v1.0 Physical Activity 8a; <18 years old).

Within each PROMIS instrument, each question has five response options ranging in value from one to five. To find the total raw score for an instrument with all questions answered, sum the values of the response to each question. All questions must be answered in order to produce a valid total score. The total raw score is then converted into a T-score for each participant using an instrument-specific *Conversion Table* as provided in Table 11. The T-score rescales the raw score into a standardized score with a mean of 50 and a standard deviation (SD) of 10. Therefore, a person with a T-score of 40 is one SD below the mean.

Table 11 - Conversion table for PROMIS instruments

(a) PROMIS Pain Intensity (all subjects)

Adult v1.0 - Pain Intensity 3a		
Short Form Conversion Table		
Raw Score	T-Score	SE*
3	30.7	4.5
4	36.3	3.1
5	40.2	3.0
6	43.5	3.0
7	46.3	3.0
8	49.4	2.9
9	52.1	2.8
10	54.5	2.9
11	57.5	3.1
12	60.5	3.1
13	64.1	3.8
14	67.4	4.2
15	71.8	5.0
*SE = Standard Error on T-score metric		

(b) PROMIS-SF Pain Interference (≥18 years)

Pain Interference 6a - Adult		
Short Form Conversion Table		
Raw Score	T-Score	SE*
6	41.1	6.0
7	48.6	2.4
8	50.7	1.8
9	52.2	1.6
10	53.4	1.6
11	54.5	1.6
12	55.6	1.5
13	56.6	1.5
14	57.6	1.5
15	58.6	1.5
16	59.5	1.5
17	60.4	1.4
18	61.2	1.4
19	62.1	1.4
20	63.0	1.5
21	63.8	1.5
22	64.8	1.5
23	65.7	1.5
24	66.7	1.5
25	67.6	1.5
26	68.7	1.5
27	69.8	1.5
28	71.0	1.6
29	72.6	2.0
30	76.3	3.6

*SE = Standard Error on T-Score

(c) PROMIS-SF Pain Interference (<18 years)

Pain Interference 8a - Pediatric v2.0		
Short Form Conversion Table		
Raw Score	T-Score	SE*
8	34.0	5.6
9	38.7	4.4
10	40.6	4.2
11	42.7	3.8
12	44.3	3.7
13	45.8	3.4
14	47.1	3.3
15	48.4	3.2
16	49.5	3.2
17	50.6	3.1
18	51.7	3.1
19	52.7	3.1
20	53.7	3.0
21	54.7	3.0
22	55.7	3.0
23	56.6	3.0
24	57.6	3.0
25	58.5	3.0
26	59.5	3.0
27	60.4	3.0
28	61.4	3.0
29	62.4	3.0
30	63.4	3.0
31	64.4	3.0
32	65.4	3.1
33	66.5	3.1
34	67.6	3.2
35	68.8	3.2
36	70.1	3.3
37	71.5	3.4
38	73.2	3.7
39	75.0	3.8
40	78.0	4.3

*SE = Standard Error on T-Score

(d) PROMIS-SF Physical Function (≥18 years)

Adult v2.0 – Physical Function 6b		
Short Form Conversion Table		
Raw Summed Score	T-score	SE*
6	21.0	3.8
7	25.0	2.7
8	27.1	2.4
9	28.8	2.2
10	30.1	2.1
11	31.3	2.0
12	32.3	2.0
13	33.2	1.9
14	34.2	1.9
15	35.0	1.9
16	35.9	1.9
17	36.8	1.9
18	37.6	1.9
19	38.5	1.9
20	39.3	1.9
21	40.2	1.9
22	41.2	1.9
23	42.1	1.9
24	43.2	2.0
25	44.3	2.0
26	45.6	2.2
27	47.1	2.3
28	48.9	2.7
29	51.3	3.0
30	59.0	6.2
*SE = Standard Error on T-score metric		

(e) PROMIS Pediatric-SF Physical Activity (<18 years)

Physical Activity 8a - Pediatric v1.0		
Short Form Conversion Table		
Raw Score	T-Score	SE*
8	28.8	4.8
9	32.6	3.8
10	34.5	3.5
11	36.4	3.1
12	37.9	2.8
13	39.2	2.6
14	40.4	2.5
15	41.4	2.4
16	42.4	2.3
17	43.4	2.3
18	44.3	2.3
19	45.2	2.3
20	46.1	2.3
21	47.0	2.3
22	47.8	2.3
23	48.7	2.3
24	49.6	2.3
25	50.5	2.3
26	51.4	2.3
27	52.3	2.4
28	53.3	2.4
29	54.3	2.4
30	55.3	2.4
31	56.3	2.4
32	57.3	2.4
33	58.4	2.4
34	59.5	2.5
35	60.8	2.5
36	62.1	2.7
37	63.7	2.8
38	65.5	3.1
39	67.8	3.5
40	71.7	4.6
*SE = Standard Error on T-Score		

5.5.3 Appendix 5.3 HAL and pedHAL

HAL scores can be calculated for each of the seven domains of the HAL. Additionally, three component scores will be calculated (Activities involving the Upper Extremities, Basic activities involving the Lower Extremities, and Complex activities involving the Lower Extremities) as well as an overall score.

Before summarizing the individual item scores, recoding is required (see Table 12), so that a higher score represents more functional limitations. Possible scoring ranges for the domain, component, and overall scores are given in Table 13. Normalized scores for the domains, components, and the overall questionnaire will be obtained using Table 14. Missing values are

controlled for and the possible scores range from 0 to 100, where 0 represents the worst possible functional status and 100 the best possible functional status.

Table 12 - Scoring HAL/pedHAL: recoding

Score	Recode	Meaning
8	0	N/A
1	6	Impossible
2	5	Always problems
3	4	Mostly problems
4	3	Sometimes problems
5	2	Rarely problems
6	1	Never problems

Table 13 - Scoring HAL: possible scoring ranges

Score		Items	Score range
Lying / sitting / kneeling / standing	LSKS	1-8 (8)	8 - 48
Functions of the legs	LEGS	9-17 (9)	9 - 54
Functions of the arms	ARMS	18-21 (4)	4 - 24
Use of transportation	TRANS	22-24 (3)	3 - 18
Self-care	SELFC	25-29 (5)	5 - 30
Household tasks	HOUSEH	30-35 (6)	6 - 36
Leisure activities and sports	LEISPO	36-42 (7)	7 - 42
Upper Extremity Activities	UPPER *	(9)	9 - 54
Basic Lower Extremity Activities	LOWBAS **	(6)	6 -36
Complex Lower Extremity Activities	LOWCOM ***	(9)	9 - 54
Sum score	SUM 1-42	(42)	42 - 252

* Items for UPPER-component: 18, 19, 20, 21, 25, 26, 27, 28, 29. (9 items)
** Items for LOWBAS-component: 8, 9, 10, 11, 12, 13. (6 items)
*** Items for LOWCOM-component: 3, 4, 6, 7, 14, 15, 16, 17, 22. (9 items)

Table 14 - Scoring HAL: normalization

Score	Normalization
LSKS	100 - ((Σ1−8 - valid) * (100/(5 * valid)))
LEGS	100 - ((Σ9−17 - valid) * (100/(5 * valid)))
ARMS	100 - ((Σ18−21 - valid) * (100/(5 * valid)))
TRANS	100 - ((Σ22−24 - valid) * (100/(5 * valid)))
SELFC	100 - ((Σ25−29 - valid) * (100/(5 * valid)))
HOUSEH	100 - ((Σ30−35 - valid) * (100/(5 * valid)))
LEISPO	100 - ((Σ36−42 - valid) * (100/(5 * valid)))
UPPER	100 - ((Σ18−21;25−29 - valid) * (100/(5 * valid)))
LOWBAS	100 - ((Σ8−13 - valid) * (100/(5 * valid)))
LOWCOM	100 - ((Σ3−7;14−17;22 - valid) * (100/(5 * valid)))
SUM	100 - ((Σ1−42 - valid) * (100/(5 * valid)))

"valid" = number of items scored within the specific domain/component.
Items with "n/a"-response are to be considered NOT valid

If more than half of the items of a domain/component are missing or scored “N/A”, no valid domain score will be calculated. If more than half of the items are missing or scored “N/A”, no valid total score will be calculated.

The pedHAL domain scores and the overall score will be calculated in a similar manner as the HAL. Responses will be recoded (as described in Table 12) before summing the domain and overall scores (Table 15). Normalized scores for the domains and the overall questionnaire will be obtained using Table 16. The normalized scores range from 0 to 100, where 0 represents the worst possible functional status and 100 the best possible functional status.

Table 15 - Scoring pedHA: possible scoring ranges

Score		Items	Score range
Sitting / kneeling / standing	LSKS	1-10 (10)	10 - 60
Functions of the legs	LEGS	11-21 (11)	11 - 66
Functions of the arms	ARMS	22-27 (6)	6 - 36
Use of transport	TRANS	28-30 (3)	3 - 18
Self-care	SELFC	31-39 (9)	9 - 54
Household tasks	HOUSEH	40-42 (3)	3 - 18
Leisure activities and sports	LEISPO	43-53 (11)	11 - 66
Sum score	SUM 1-53	(53)	53- 318

If more than half of the items of a domain are missing or scored “N/A”, no valid domain score will be calculated. If more than half of the items are missing or scored “N/A”, no valid total score will be calculated.

Table 16 - Scoring pedHAL: normalization

Score	Normalization
LSKS	100 - ((Σ1−10 - valid) * (100/(5 * valid)))
LEGS	100 - ((Σ11−21 - valid) * (100/(5 * valid)))
ARMS	100 - ((Σ22−27 - valid) * (100/(5 * valid)))
TRANS	100 - ((Σ28−30 - valid) * (100/(5 * valid)))
SELFC	100 - ((Σ31−39 - valid) * (100/(5 * valid)))
HOUSEH	100 - ((Σ40−42 - valid) * (100/(5 * valid)))
LEISPO	100 - ((Σ43−53 - valid) * (100/(5 * valid)))
SUM	100 - ((Σ1−53 - valid) * (100/(5 * valid)))

“valid” = number of items scored within the specific domain/component.
Items with “n/a”-response are to be considered NOT valid

5.5.4 Appendix 5.4 TSQM-9

The TSQM Scale scores will be computed by adding the items loading on each factor. The lowest possible score is subtracted from this composite score and divided by the greatest possible score minus the lowest possible score. This provides a transformed score between 0 and 1 that should be multiplied by 100; see details in Table 17. Note that only one item may be missing from each scale before the subscale should be considered invalid for that respondent.

Table 17 - TSQM Scale Scores

Scale	Calculation	In case of missing
Effectiveness	([(Item 1 + Item 2 + Item 3) – 3] divided by 18) * 100	If one item is missing: ([(Sum(Item 1? + Item 2? + Item 3?)) – 2] divided by 12) * 100
Convenience	([Sum(Item 4 to Item 6) – 3] divided by 18) * 100	If one item is missing: ([(Sum(Item4? to Item6?)) – 2] divided by 12) * 100
Global Satisfaction	([Sum(Item 7 to Item 9) – 3] divided by 14) * 100	If either Item 7 or 8 is missing: ([(Sum(Item7? to Item9?)) – 2] divided by 10) * 100 If Item 9 is missing: ([(Sum(Item7 and Item8)) – 2] divided by 8) * 100

5.6 APPENDIX 6 META-ANALYSIS AND DERIVATION OF NON-INFERIORITY MARGIN

For the key secondary efficacy endpoint, an intra-patient comparison of ABR during the BIVV001 weekly prophylaxis treatment period versus the historical prophylaxis ABR will be performed using non-inferiority testing for participants in Arm A who have at least 6 months of participation in this study and at least 6 months of historical data on prophylaxis treatment from observational study 242HA201/OBS16221.

The non-inferiority margin was estimated based on the known treatment effect between on-demand and prophylaxis treatment. A meta-analysis of Phase 3 registrational studies for recombinant FVIII products that include both on-demand and prophylaxis treatment arms was performed using the following search criteria:

1. Hemophilia A population
2. Recombinant FVIII product
3. Phase 3 registrational study
4. Adult population
5. Non-inhibitor population
6. Previously treated patient (PTP) population
7. Includes both prophylaxis and on-demand treatment arms

A total of 12 studies were found meeting the first six search criteria, and 5 studies were excluded that did not include an on-demand treatment arm and therefore did not meet the seventh search criteria. A summary of the search results is provided in Table 18 below.

The 7 selected studies were further evaluated for consistent inclusion/exclusion criteria, dosing regimens, and constancy of treatment effect (Table 19). All 7 studies were deemed appropriate to include in the meta-analysis.

Table 18 - Meta-analysis search results

No.	Product	FDA Approved	Prophylaxis		On-demand		Reason for exclusion	Reference
			N	ABR Mean (SD)	N	ABR Mean (SD)		
1	Kogenate	2000	42	2 (4.5)	42	30.5 (19.5)		Manco-Johnson 2013 (9)
2	Advate	2003	111	6.1 (8.2)	-	-	No on-demand arm	US Prescribing Information
3	Xyntha	2008	94	3.9 (6.5)	-	-	No on-demand arm	Recht 2009 (10)
4	NovoEight	2013	150	6.5	-	-	No on-demand arm	Lentz 2013 (11)
5	Eloctate	2014	117	2.9 (3.9)	23	37.3 (20.6)		Mahlangu 2014 (3)
6	Nuwiq	2015	32	2.3 (3.7)	-	-	No on-demand arm	Lissitchkov 2016 (12)
7	Adynovate	2015	101	3.7 (4.7)	17	40.8 (16.3)		Konkle 2015 (2)
8	Kovaltry (LEOPOLD I)	2016	62	3.8 (5.2)	-	-	No on-demand arm	Saxena 2016 (13)
9	Kovaltry (LEOPOLD II)	2016	59	4.9 (6.8)	21	57.7 (24.6)		Kavakli 2015 (1)
10	Afstyla	2016	146	3.1 (5.1)	27	31.1 (35.6)		Mahlangu 2016 (4)
11	Jivi	2018	43	3.3 (4.3)	20	28.8 (17.8)		Reding 2017 (14)
12	Esperoct	2019	175	3.0 (4.7)	12	31.9 (19.1)		Giangrande 2017 (15)

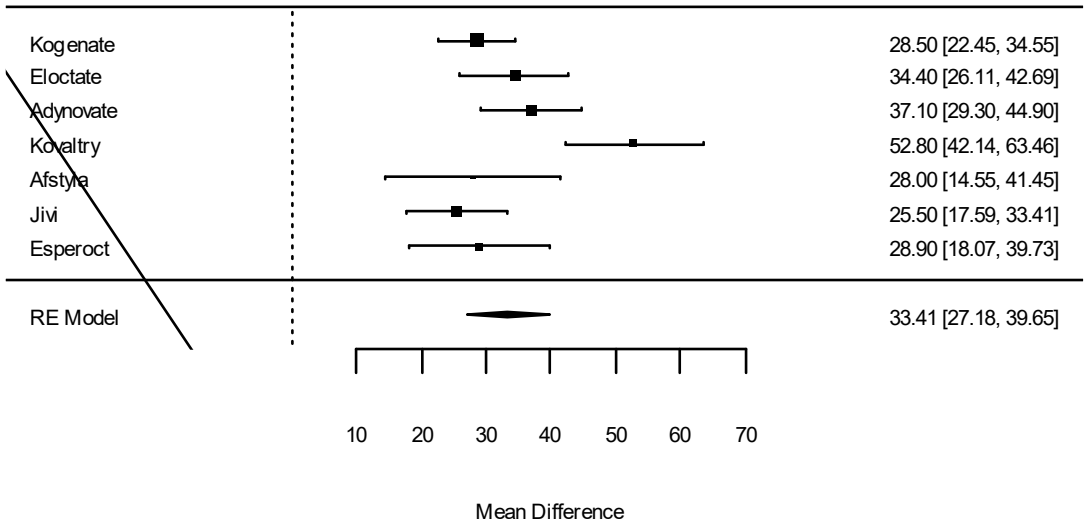
Table 19 - Summary of meta-analysis studies

No.	Product	Inclusion Criteria	Exclusion Criteria	Prophylaxis Dosing Regimen
1	Kogenate	- Males 12-50 years - Severe hemophilia A (<1% FVIII) as confirmed by central lab - Subjects with no inhibitor activity by Nijmegen-modified Bethesda assay, either by positive (>0.6 BU) or borderline (>0.3 and <0.6 BU)	- Subjects with any other bleeding disease besides hemophilia A - Subjects with known hypersensitivity to rFVIII, mouse or hamster proteins	25 IU/kg three times a week
2	Eloctate	- Male >12 years - Severe hemophilia A (<1% FVIII) - History of at least 50 documented prior exposure days to FVIII - No current, or history of, inhibitor development to FVIII	- Other coagulation disorders in addition to hemophilia A - History of anaphylaxis associated with any FVIII or IV immunoglobulin administration	Individualized prophylaxis 25-65 IU/kg every 3-5 days
3	Adynovate	- Males 12-65 years - Severe hemophilia A (<1% FVIII) documented or confirmed by central lab - Previously treated with plasma-derived FVIII concentrates or recombinant FVIII for ≥150 exposure days - Currently receiving prophylaxis or on-demand therapy with FVIII	- Detectable FVIII inhibitory antibodies (≥0.6 BU) as confirmed by central lab - History of FVIII inhibitory antibodies -Diagnosis of inherited or acquired hemostatic defect other than hemophilia A	45 ± 5 IU/kg two times per week
4	Kovaltry (LEOPOLD II)	- Males 12-65 years - Severe hemophilia A - History of more than 150 exposure days with clotting factor concentrates - Currently receiving episodic treatment with FVIII; no regular prophylaxis for more than 6 consecutive months in the past 5 years - No current FVIII inhibitor or history of inhibitor	- Presence of another bleeding disease that is different from hemophilia A - Known hypersensitivity to FVIII	20-30 IU/kg two times per week or 30-40 IU/kg three times per week
5	Afstyla	- Males 12-65 years - Diagnosis of severe hemophilia defined as <1% FVIII documented in medical records - Received or currently receiving FVIII products (plasma-derived and/or recombinant) and have had at least 150 exposure days with a FVIII product	- Any history of or current FVIII inhibitors - Any first order family history of FVIII inhibitors - Known hypersensitivity (allergic reaction or anaphylaxis) to any FVIII product or hamster protein - Any known congenital or acquired coagulation disorder other than congenital FVIII deficiency	20-40 IU/kg every second day or 20-50 IU/kg 2-3 times per week
6	Jivi	- Males 12-65 years - Severe hemophilia A - Previously treated with FVIII for a minimum of 150 exposure days	- Inhibitors to FVIII (current evidence or history) - Any other inherited or acquired bleeding disorder in addition to hemophilia A	25 IU/kg two times per week to 45-60 IU/kg every 5 days

No.	Product	Inclusion Criteria	Exclusion Criteria	Prophylaxis Dosing Regimen
7	Esperoct	- Males with severe congenital hemophilia A (<1% FVIII, according to medical records) - At least 12 years and body weight at least 35 kg - Documented history of at least 150 exposure days to other FVIII products	- Any history of FVIII inhibitors - FVIII inhibitors above or equal to 0.6 BU/mL at screening - Congenital or acquired coagulation disorders other than hemophilia A	A single-bolus dose of 50 IU/kg every fourth day

Using metafor package in R, the combined mean difference (on-demand – prophylaxis) can be calculated from a random effects model (16).

Figure 2 - Forest plot for mean difference in annualized bleeding rates between on-demand and prophylaxis treatment regimens



Using the lower bound treatment effect of 27 bleeds per year and a fixed margin approach to maintain a substantial amount (85%) of the treatment effect (17) results in a non-inferiority margin of 4 (ie, 27 *0.15 = 4).

6 REFERENCES

1. Kavakli K, Yang R, Rusen L, Beckmann H, Tseneklidou-Stoeter D, Maas Enriquez M, et al. Prophylaxis vs. on-demand treatment with BAY 81-8973, a full-length plasma protein-free recombinant factor VIII product: results from a randomized trial (LEOPOLD II). J Thromb Haemost. 2015;13(3):360-9.

2. Konkle B, Stasyshyn O, Chowdary P, Bevan DH, Mant T, Shima M, et al. Pegylated, full-length, recombinant factor VIII for prophylactic and on-demand treatment of severe hemophilia A. Blood. 2015;126(9):1078-85.

3. Mahlangu J, Powell JS, Ragni MV, Chowdary P, Josephson NC, Pabinger I, et al. Phase 3 study of recombinant factor VIII Fc fusion protein in severe hemophilia A. Blood. 2014;123(3):317-25.

4. Mahlangu J, Kuliczkowski K, Abdul Karim F, Stasyshyn O, Kosinova MV, Lepatan LM, et al. Efficacy and safety of rVIII-SingleChain: results of a phase 1/3 multicenter clinical trial in severe hemophilia A. Blood. 2016. 2016;128(5):630-7.

5. Han L, Merck. Calculating the point estimate and confidence interval of Hodges-Lehmann's median using SAS® software. The Proceedings of the SouthEast SAS Users Group 2008 (pp. ST 154).

6. Blanchette VS, Key NS, Ljung LR, Manco-Johnson MJ, van den Berg HM, Srivastava A, et al. Definitions in hemophilia: communication from the SSC of the ISTH. J Thromb Haemost. 2014;12(11):1935-9.

7. Baskin JL, Pui CH, Reiss U, Wilimas JA, Metzger ML, Ribeiro RC, et al. Management of occlusion and thrombosis associated with long-term indwelling central venous catheters. Lancet. 2009;374(9684):159-69.

8. Tagalakis V, Kahn SR, Libman M, Blostein M. The epidemiology of peripheral vein infusion thrombophlebitis: a critical review. Am J Med. 2002;113:146-51.

9. Manco-Johnson MJ, Kempton CL, Reding MT, Lissitchkov T, Goranov S, Gercheva L, et al. Randomized, controlled, parallel-group trial of routine prophylaxis vs. on-demand treatment with sucrose-formulated recombinant factor VIII in adults with severe hemophilia A (SPINART). J Thromb Haemost. 2013;11:1119-27.

10. Recht M, Nemes L, Matysiak M, Manco-Johnson M, Lusher J, Smith M, et al. Clinical evaluation of moroctocog alfa (AF-CC), a new generation of B-domain deleted recombinant factor VIII (BDDrFVIII) for treatment of haemophilia A: demonstration of safety, efficacy, and pharmacokinetic equivalence to full-length recombinant factor VIII. Haemophilia. 2009;15:869-80.

11. Lentz SR, Misgav M, Ozelo M, Salek SZ, Veljkovic D, Recht M et al. Results from a large multinational clinical trial (guardian™ 1) using prophylactic treatment with turoctocog alfa in

adolescent and adult patients with severe haemophilia A: safety and efficacy. Haemophilia. 2013;19: 691-7.

12. Lissitchkov T, Hampton K, von Depka M, Hay C, Rangarajan S, Tuddenham E, et al. Novel, human cell line-derived recombinant factor VIII (human-cl rhFVIII; Nuwiq) in adults with severe haemophilia A: efficacy and safety. Haemophilia. 2016;22:225-31

13. Saxena K, Lalezari S, Oldenburg J, Tseneklidou-Stoeter D, Beckmann H, Yoon M, et al. Efficacy and safety of BAY 81-8973, a full-length recombinant factor VIII: results from the LEOPOLD I trial. Haemophilia. 2016;22(5):706-12.

14. Reding MT, Ng HJ, Poulsen LH, Eyster ME, Pabinger I, Shin HJ, et al. Safety and efficacy of BAY 94-9027, a prolonged half-life factor VIII. J Thromb Haemost. 2017;15(3):411-9.

15. Giangrande P, Andreeva T, Chowdary P, Ehrenforth S, Hanabusa H, Leebeek FGW, et al. Clinical evaluation of glycoPEGylated recombinant FVIII: efficacy and safety in severe haemophilia A. Thromb Haemost. 2017;117:252-61.

16. Normand ST. Meta-analysis: Formulating, evaluating, combining, and reporting. Stat Med. 1999;18(3):321-59.

17. Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. Trials. 2011;12(1):1.

Signature Page for VV-CLIN-0614313 v2.0
efc16293-16-1-9-sap

Approve & eSign	[redacted]
Approve & eSign	[redacted]