16.1.9 Documentation of Statistical Methods

16.1.9.1 Statistical Analysis Plan for Study CT-AMT-061-02

STATISTICAL ANALYSIS PLAN FOR

STUDY CT-AMT-061-02

Protocol Number:	CT-AMT-061-02
Investigational Drug and Drug Number:	AMT-061 (AAV5-hFIXco-Padua); CCI
Indication:	Hemophilia B
Dosage Form/Dose:	2×10^{13} gc/kg AMT-061
Client:	uniQure biopharma B.V.

Protocol Title: Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B

Date of Issue: 10 Jun 2021

Version: Version 4.0

Signed Agreement on Statistical Analysis Plan

FINAL SIGN-OFF SIGNATURES

Author: PPD
PPD, Statistical Operations
Everest Clinical Research

PPD PPD

Signature Date

Peer Reviewer: PPD
PPD
Everest Clinical Research

PPD PPD

Signature Date

Approved by: PPD
PPD, Clinical Development
uniQure Biopharma B.V.

PPD PPD

Signature Date

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
2.0	06-May-2019	Added correlation of Month 6 factor IX activity levels and pre-IMP anti-AAV5 antibody titers as an efficacy endpoint.	1.0
2.0	16-Jun-2019	Made the distinction between the lead-in safety population and the post-treatment safety population.	1.0
2.0	20-Jun-2019	Defined a separate lead-in baseline and post-treatment baseline for safety laboratory and vital sign measurements.	1.0
2.0	20 Jun 2019	Added the requirement that a bleed must be designated as a new and "true" bleed by the investigator (or designee) in order for the bleed to be counted.	1.0
2.0	08-Jul-2019	A single-treatment cumulative responder analysis was added as an additional sensitivity analysis for the primary endpoint.	1.0
2.0	08-Jul-2019	To examine the correlation (and relationship) between factor IX activity levels and bleeding rates, a GAM analysis to graph the estimated ABR and its 95% CI as a function of the mean FIX activity has been added to the SAP.	1.0
2.0	08 Jul 2019	A comparative between-treatments "Cumulative Responder Analysis comparing the Lead-In Period and after Treatment with AMT-061 at Week 26" is being provided as a sensitivity analysis for and as an alternative to the "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 at Week 26"	1.0
2.0	21-Jul-2019	Vital Signs has been removed as an endpoint. Vital signs will still be listed but will not be summarized.	1.0
2.0	30 Mar 2020	To correct some typographical errors.	1.0
2.0	30 Mar 2020	To make some minor clarifications.	1.0
2.0	29 May 2020	Added "Proportion of subjects with zero bleeds in 52-week post-treatment follow-up" as the fourth-to-last endpoint in the list of secondary endpoints.	1.0
2.0	06 Jun 2020	Added annualized infusion rate of factor IX replacement therapy as a secondary efficacy endpoint.	1.0
2.0	06 Jun 2020	Stipulated that subjects with zero post-AMT-061 uncontaminated central-laboratory factor IX activity values will have their planned-assessment post-AMT values imputed with the subject's baseline factor IX activity value (based on historical hemophilia B severity).	1.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
2.0	06 Jun 2020	Added Q1 and Q3 to descriptive statistics.	1.0
2.0	06 Jun 2020	Removed the use of baseline covariates from the multiple-period analyses, because baseline covariates do not add information for a within-subject comparison.	1.0
2.0	06 Jun 2020	Moved "Rate of traumatic bleeding events during the 52-week post-treatment follow-up compared to the lead-in phase" from the position of third-to last secondary efficacy endpoint to the position of last secondary efficacy endpoint.	1.0
2.0	06 Jun 2020	Operationally defined "return to routine prophylaxis".	1.0
2.0	09 Jun 2020	A schema for the assignment of such unplanned assessments to scheduled time points for visit-based endpoint analysis is provided. This schema is called Time Windows for Statistical Analysis. The reasons are as follows. Scheduling difficulties due to the COVID-19 pandemic may result in an increased number of missed, delayed, or unscheduled visits. Scheduling difficulties may also result in the performance of assessments at a scheduled visit where performance of the assessment was not originally planned.	1.0
2.0	16 Jul 2020	Added Section 7.6.6, which defines a new secondary efficacy endpoint: Annualized infusion rate of factor IX replacement therapy during the 52-week post-treatment follow-up, excluding replacement for invasive procedures compared to the lead-in phase.	1.0
2.0	17 Jul 2020	Adaptations of analyses -- for which the main data cut for analysis is at one year -- to the six-month data cut (and sometimes to the 5-year data cut) are provided.	1.0
2.0	17 Jul 2020	Removed statement in Section 7.8.8.1 referring to a separate SAP document addressing the statistical analysis of the PROBE questionnaire.	1.0
2.0	28 Jul 2020	Applied a simplification of the model specified in Section 7.6.8 (due to limitations of the both the GLMMIX procedure and the GENMOD procedure in SAS). The revision specifies a covariance structure using only one level instead of two.	1.0
2.0	16 Aug 2020	Specified that protocol deviations will include those related to COVID-19 in Section 7.1.	1.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
2.0	02 Sep 2020	Added Section 7.6.4.5 detailing a fifth sensitivity analysis for the primary efficacy endpoint evaluating the impact of steroid therapy on factor IX activity levels and specifying the creation of a listing examining subject-specific factor IX activity levels relative to instances of corticosteroid exposure.	1.0
2.0	02 Sep 2020	Added a sensitivity analysis in Section 0 repeating the main analysis while considering only bleeds treated with factor IX that are assessed to be new and true.	1.0
2.0	02 Sep 2020	Added descriptive summaries of bleeding rates in Section 0.	1.0
2.0	02 Sep 2020	Added a descriptive sensitivity analysis in Section 1.1.1 that is a cumulative responder analysis based on subject-specific annualized bleeding rates in the respective treatment periods.	1.0
2.0	03 Sep 2020	Made revisions in Section 7.6.8.1 clarifying that comparisons will be made between treatment periods rather than treatment “groups”.	1.0
2.0	10 Sep 2020	Specified contents of a reference listing defining the duration of contamination due to corticosteroid exposure in Section 7.6.4.5.	1.0
2.0	10 Sep 2020	Added reference to "Bretz et. al. 2015" in Section 7.7 and provided said reference in References section.	1.0
2.0	11 Sep 2020	Added clarification in Section 0 specifying that p-values and confidence intervals provided for data cuts that are not the main data cut for a given endpoint will be considered to be descriptive rather than inferential.	1.0
2.0	11 Sep 2020	Added information to Section 4.3 describing the bleed reporting and assessment process.	1.0
2.0	11 Sep 2020	Removed references to SAP Appendices 5, 6, and 7. Appendices 5 – 7 referred to the Mock tables, listings, and graphs, which will be provided as a single supporting document rather than a formal SAP Appendix or Appendices.	1.0
2.0	05-Oct-2020	Promoted two secondary efficacy endpoints (endogenous Factor IX activity at Week 52 and ABR comparison between AMT-061 and prophylaxis for non-inferiority) to primary efficacy endpoints. Made corresponding revisions in Sections 2.1.1, 2.1.2.1, 2.2.1.1, 2.2.1.2, 7.5, and 0 as a result.	1.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
2.0	05-Oct-2020	Moved description of Type I error control procedures to Section 7.7, a section following the description of primary and secondary endpoints and their respective analyses.	1.0
2.0	05-Oct-2020	Added a clarification indicating only short form CCI data will be considered for descriptive and inferential analyses.	1.0
2.0	05-Oct-2020	Added clarifications that inflammatory marker data will not be available as part of the interim six-month locked data and therefore these data will not be summarized as part of the interim CSR.	1.0
2.0	05-Oct-2020	Added a clarification to the definitions of steatosis subgroups referenced in the subgroup analysis section (Section 7.8.7). Added references supporting these clarifications.	1.0
2.0	05-Oct-2020	Added a clarification indicating that time to resolution is presented relative to the dosing date for target joints in existence at the time of dosing.	1.0
2.0	08-Oct-2020	CCI [redacted]	1.0
2.0	08-Oct-2020	CCI [redacted]	1.0
2.0	08-Oct-2020	CCI [redacted]	1.0
2.0	08-Oct-2020	CCI [redacted]	1.0
2.0	15-Oct-2020	Revised wording and ordering of safety endpoints to match Version 5.0 of the study protocol.	1.0
2.0	15-Oct-2020	Added text addressing the “Change in abdominal ultrasound” safety endpoint.	1.0
2.0	15-Oct-2020	Updated the table summarizing protocol recommended study time intervals for collection of efficacy and safety parameters to match Version 5.0 of the study protocol.	1.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
2.0	15-Oct-2020	Changed p-value threshold for statistical significance from "< 0.025" to "<= 0.025" and "< 0.05" to "<= 0.05".	1.0
2.0	15-Oct-2020	Added text addressing the "Factor IX recovery" safety endpoint.	1.0
2.0	18-Oct-2020	Changed three-month window to two-month window for the endpoint "return to continuous prophylaxis".	1.0
3.0	20 Feb 2021	A period-specific baseline is now being used for CCI [redacted] (for descriptive statistics). The reason is to allow the most recent information to be used as baseline. This does not affect inferential analyses, because baseline is not used as a covariate for inferential analyses.	2.0
3.0	04 Mar 2021	Added clarifying wording about the baseline definition for vital signs.	2.0
3.0	08 Mar 2021	Added clarifying wording about the baseline definition for height, weight, and BMI.	2.0
3.0	18 Mar 2021	Added a paragraph about The Use of Lead-In Period Month 6 Visit versus Lead-In Final Visit for Analyses of Factor IX Activity. When present, the Lead-In Month 6 value will be used for analysis. When a Lead-In 6-month value is not available, then the Lead-In Final Visit value (if available) will be used instead. The rationale is that (1) Lead-In Month 6 is a planned assessment (for factor IX activity) and that in the presence of a Month 6 assessment, a Lead-In Final value is not essential for analysis and (2) the planned duration of the Lead-In period is approximately 6 months. The wording is to clarify how the choice between these two visits (for analysis) is being carried out.	2.0
3.0	02 Mar 2021	For certain efficacy endpoints, have provided clarity about how (and under what circumstances) unscheduled-assessment values are to be applied to serve as the value for either the Lead-In Month 6 value or the Lead-In Final value (depending on the endpoint based on the endpoint's schedule of assessments).	2.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
3.0	10 Feb 2021	Have stipulated that for the second primary endpoint "Endogenous Factor IX activity at 52 Weeks after AMT-061 dosing" the period of "contamination" from exogenous factor IX use is to be the time period starting at the time of the infusion (start time) and ending at the time that is 5 half-lives after that. The half-life is medication-dependent. This refinement is being made to improve the accuracy of the contamination period. This contamination-rule refinement is also being applied to other, similar or related endpoints. An analysis using the old 10-day contamination rule (while now taking into account both date and time) is now relegated to a sensitivity analysis.	2.0
3.0	23 Mar 2021	Have added the following paragraph: The primary analysis of the first primary efficacy endpoint "Endogenous Factor IX activity at 26 Weeks after AMT-061 dosing" actually took place using the 26-week data cut and was done according to the existing Statistical Analysis Plan text (at the time). The refinement to the definition of the "contamination period" mentioned in this version of the Statistical Analysis Plan text is stated here (in the 26-week section) only because the second primary endpoint "Endogenous Factor IX activity at 52 Weeks after AMT-061 dosing" makes reference to the analysis methods for the first primary endpoint, and writing the update in this "26-week" section preserves the structure of the SAP text, thus allowing the refinement – pertaining to the definition of the "contamination period" – that will be applied to the 52-week second primary efficacy endpoint analysis – to be more easily and transparently seen in context.	2.0
3.0	08 Mar 2021	To clarify that a conservative penalty that has been stipulated for the primary (factor IX activity) efficacy analysis is also to be applied to other key efficacy endpoints and sensitivity analyses. This conservative penalty is as follows: If a subject has zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that is to be used in the analysis) will be imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity.	2.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
3.0	04 Feb 2021	To stipulate that event-rate-based, consumption-rate-based, and proportion-based endpoints – for the purpose of statistical inference – are to pertain to a time interval time that ends one year after the subject's dosing (with AMT-061). This stipulation is being made to be true to the name of each endpoint, which contains the words "52-week" as a part of the endpoint's name.	2.0
3.0	21 Feb 2021	CCI [redacted]	2.0
3.0	22 Feb 2021	For the endpoint "Annualized infusion rate of factor IX replacement therapy during the 52-week post-treatment (AMT-061) follow-up, excluding replacement for invasive procedures, compared to the lead-in phase", it is now stipulated that the analysis should count the number of infusions rather than the number of days with infusion. This is to account for the possibility that a subject may have more than one factor IX infusion on the same day.	2.0
3.0	24 Feb 2021	For the endpoint "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 at Week 52", it is now specified that "a contrast will be employed to compare the average across the scheduled visits from the Month 6 to the Month 12 post-treatment visits" to the lead-in period. This is to provide a broader time-period base for the post-treatment period than just the 12-month time point alone would have afforded.	2.0
3.0	24 Feb 2021	For the endpoint "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 at Week 52", now more alternatives to the analysis are offered in the case of non-convergence of the statistical model.	2.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
3.0	02 Mar 2021	For the sensitivity analysis "Sensitivity Analysis 1: Cumulative Responder Analysis Comparing the Lead-In Period and after Treatment with AMT-061 at Week 52" – which is a sensitivity analysis for the endpoint "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 at Week 52" –, it is now specified that the average of the scheduled post-treatment time points between Month 6 and Month 12 is to be used for the comparison. This is to provide time-period alignment between the sensitivity analysis and the "<12%" endpoint analysis.	2.0
3.0	19 Feb 2021	Have added an additional sensitivity analysis "Sensitivity Analysis 2: Excluding Contaminated Values Using a 10-day Contamination Rule" for the endpoint "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 at Week 52". The reason is to provide alternative results using the old 10-day contamination rule (while now taking into account both date and time) as an alternative to the more refined 5-half-life contamination rule.	2.0
3.0	21 Mar 2021	For the endpoint "Estimated ABR over the 52-weeks post-treatment follow-up as a Function of Mean Factor IX Activity ("Correlation" Analysis)" it is now stated "Factor IX activity assessments and ABR counts beginning at Day 21 (of the post-treatment period) will be used in the analysis." This is for consistency with the portion of the post-treatment time period being used for the primary ABR analysis.	2.0
4.0	08 Apr 2021	Are making the full analysis set the primary population for the ABR non-inferiority analysis, and relegating the per-protocol population to a sensitivity analysis. The reason is that the FDA statistical team requested this change.	3.0
4.0	08 Apr 2021	The main ABR analysis is hereby to be excluding (from the person-time "at risk") any time intervals (during the post-treatment period) for which there was "contamination" (by the 5 half-life rule) due to exogenous factor IX infusions/use. The ABR analysis that keeps such time intervals (in the time "at risk") is hereby relegated to a sensitivity analysis. The purpose of this change is to increase accuracy.	3.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
4.0	18 Apr 2021	Have clarified which data cut to be doing sensitivity analyses on (Section 7.6.4). Have clarified that for factor IX activity (statistical analysis), the sensitivity analyses to be included in the CSR can all be performed using the 52-week data cut, and such sensitivity analyses do not need to be carried out for the 26-week data cut. That is because the 26-week (visit) time point is one of the time points included in the 52-week-data-cut analysis. This provision allows parsimony of sensitivity analyses.	3.0
4.0	23 April 2021	Added text to show how to employ existing uncontaminated factor IX assessments in any factor IX activity analysis that focuses on a restricted set (or interval) of time points, when there are no available uncontaminated values in that restricted set (or interval). This stipulation does not affect the repeated measures analyses, which already use a wide range of time points as a basis for the model-estimated statistics and inference.	3.0
4.0	05 Jun 2021	ABR has now become the sole primary endpoint. The reason is that the FDA (Food and Drug Administration) statistical and clinical teams requested this.	3.0
4.0	05 Jun 2021	Have changed the data cut for the main CSR to be 52 weeks after stable FIX expression (18 months post-treatment). The reason is that the FDA asked for the main efficacy analysis to pertain to the year after a stable factor IX activity level is reached in the last patient.	3.0
4.0	05 Jun 2021	Have changed the main ABR analysis to count all (unique) bleeds which occur (after the start point for counting bleeds toward ABR) in the total, irrespective of the investigator's designation of newness and trueness. This will be done for all ABR analyses except for a small number of sensitivity analyses. The reason is that the FDA (Food and Drug Administration) clinical team requested this.	3.0
4.0	05 Jun 2021	Have added a sensitivity ABR analysis that excludes person-time during or within 5 half-lives after systemic corticosteroid use. The reason is that the FDA (Food and Drug Administration) clinical team requested this.	3.0
4.0	05 Jun 2021	Have added the following new endpoint: Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis)	3.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
4.0	05 Jun 2021	Have moved the endpoint "Correlation of factor IX activity levels during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)" to be higher in the list of secondary endpoints than the following endpoint "Occurrence of (and resolution of) new target joints during the 52 weeks following stable FIX expression (6-18 months) following AMT-061 dosing and resolution of pre-existing target joints following AMT-061 dosing". The reason is that the FDA (Food and Drug Administration) clinical team requested that this endpoint be raised higher in importance.	3.0
4.0	05 Jun 2021	Have added the following new endpoint "Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis)" and have place this endpoint to be higher in the list of secondary endpoints than the following endpoint: "Correlation of factor IX activity levels during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)". The reason is that the FDA (Food and Drug Administration) clinical team requested this.	3.0

Change Log			
Version No.	**Effective Date (dd-mmm-yyyy)**	**Reason for the Change / Revision**	**Supersedes**
4.0	05 Jun 2021	Have re-defined the efficacy endpoint "return to continuous routine prophylaxis" as having at least 80% of the time being "contaminated" by exogenous factor IX (by the 5 half-life rule) during some contiguous three-month period … . The reason is that the FDA (Food and Drug Administration) clinical team requested that the definition of "return to continuous routine prophylaxis" be independent of the purpose-for-use category assigned in the CRF to instances of exogenous factor IX infusion.	3.0

TABLE OF CONTENTS

List of Tables

List of Figures

GLOSSARY OF ABBREVIATIONS

Abbreviation	Term
AAV	adeno-associated viral
AAV5	adeno-associated viral vector serotype 5
AAV5-hFIXco	recombinant adeno-associated viral vector serotype 5 containing the wild type human FIX gene, codon-optimized for optimal expression in humans, under control of a liver-specific promoter (AMT-060)
AAV5-hFIXco-Padua	recombinant adeno-associated viral vector serotype 5 containing a codon-optimized Padua derivative of human coagulation FIX cDNA (AMT-061)
ABR	annualized bleeding rate
AE	adverse event
AFP	Alpha-fetoprotein
ALP	alkaline phosphatase
ALT	alanine aminotransferase
aPTT	activated partial thromboplastin time
AR	autoregressive
AST	aspartate aminotransferase
ATMP	advanced therapy medicinal product
CCI	CCI
cDNA	complementary deoxyribonucleic acid
CI	confidence interval
COVID-19	Coronavirus Disease (discovered in) 2019
CRP	c-reactive protein
CSR	Clinical Study Report
DMC	Data Monitoring Committee

DNA	deoxyribonucleic acid
eCRF	electronic case report form
ELISA	enzyme-linked immunosorbent assay
CCI	CCI
FAS	Full Analysis Set
FDA	Food and Drug Administration
FIX	coagulation factor IX
GAM	Generalized Additive Model
GGT	gamma-glutamyl transpeptidase
GEE	Generalized Estimating Equations
GTWP	Gene Therapy Working Party
CCI	CCI
HBeAg	hepatitis B extracellular antigen
HBsAg	hepatitis B surface antigen
HBV DNA	hepatitis B virus deoxyribonucleic acid
HCV RNA	hepatitis C virus ribonucleic acid
CCI	CCI
hFIX	human coagulation factor IX
CCI	CCI
IFNγ	interferon gamma
IgG	immunoglobulin G
IgM	immunoglobulin M
IL-1β	interleukin-1beta

IL-2	interleukin-2
IL-6	interleukin-6
IMP	investigational medicinal product
INR	International normalized ratio
CCI	CCI
IU	international unit
JADE	Joint Tissue Activity and Damage Exam
LOD	Limit of detection
MET	Metabolic equivalent of task
MCP-1	monocyte chemotactic protein-1
MedDRA	Medical Dictionary for Regulatory Activities
MRE	Magnetic resonance elastography
NAB	Neutralizing antibody
PCS	potentially clinically significant
PP	Per-protocol
CCI	CCI
PROBE	Patient Reported Outcomes, Burdens, and Experiences
Q1	First quartile
Q3	Third quartile
CCI	CCI
rAAV5	recombinant adeno-associated viral vector serotype 5
SAE	serious adverse event
SAP	Statistical Analysis Plan
SDTM	Study Data Tabulation Model

SOC	System Organ Class
SOP	Standard Operating Procedure
SWE	Shear wave elastography
TEAE	treatment emergent adverse event
US	United States
VAS	visual analogue scale
WFH	World Federation of Haemophilia
WHO	World Health Organization
CCI	CCI

Trademark Information

SAS

nQuery Advisor

1. INTRODUCTION

1.1 General Introduction

Congenital hemophilia B is an inherited bleeding disorder characterized by an increased bleeding tendency due to either a partial or complete deficiency of the essential blood coagulation factor IX (FIX). The deficiency is the result of mutations in the factor IX gene. Hemophilia B is an X-linked, recessive condition, since it occurs almost exclusively in males. Females typically are asymptomatic carriers. The number of people with hemophilia B worldwide is approximately 30,000 and in the United States (US) alone is approximately 4,000 (World Federation of Haemophilia [WFH], 2017). Approximately 1 in 20,000 - 50,000 live male newborns has hemophilia B.

The severity of symptoms can vary and the severe forms become apparent early in life. About one-third of individuals with hemophilia B have a severe disorder characterized by factor IX levels that are less than 1% of normal (Kessler & Mariani, 2006). Moderate and mild hemophilia B, with 1 - 5% or 5 - < 40% of normal factor IX level, respectively, are each observed in about one-third of patients (Kessler & Mariani, 2006).

There is no cure for hemophilia B. The primary goals of hemophilia B therapy are the prevention of bleeding episodes, rapid and definitive treatment of bleeding episodes (breakthrough bleeds) that occur even while on a regular prophylactic regimen and provision of adequate hemostasis during surgery and emergencies. Currently, these goals are essentially met for hemophilia B subjects by intravenous (IV) injections of commercially available recombinant- or plasma-derived factor IX products, either at the time of a bleed (on-demand) or by regular infusions up to several times a week (prophylactically).

Somatic gene therapy offers the potential for a shift of the disease severity phenotype from severe to a moderate or mild hemophilia phenotype or complete amelioration through continuous production of stable factor IX levels after a single administration of vector, especially since a small rise in circulating factor IX to at least 1% of normal levels can substantially ameliorate the bleeding phenotype and potentially improve the CCI for patients.

AMT-061 has been developed for the treatment of hemophilia B. AMT-061 is a recombinant adeno-associated viral vector serotype 5 (rAAV5) containing the coding sequence for Padua variant of the human coagulation factor IX (hFIX Padua), codon-optimized, under control of a liver-specific promoter (also known as AAV5-hFIXco Padua).

AMT-061 is a derivative of AMT-060, which has been studied in a Phase I/II clinical trial in humans with severe and moderately severe hemophilia B. Both AMT-060 and AMT-061 have the same rAAV5 containing the codon-optimized wild-type human factor IX gene, but the latter incorporates two-nucleotide change in order to encode the naturally occurring Padua variant of human coagulation factor IX. The factor IX-Padua protein differs from the ‘wild-type’ human factor IX protein by a single amino acid and it is responsible for the observed increased factor IX activity per unit of dose achieved with AMT-061 as compared to its predecessor AMT-060 (Simioni et al, 2009).

This Phase III trial is to demonstrate the efficacy of AMT061 in terms of endogenous factor IX activity and annualized bleed rate (ABR), and to further describe its safety profile. The strong efficacy (protein expression) and safety results obtained during the Phase I/II trial with AMT-060 demonstrate 2×10^{13} gc/kg to be the selected dose for use in future trials. In addition, non-clinical data support the implementation of the prospectively defined product enhancement of AMT-060 (i.e. incorporation of the Padua mutation to form the new construct AMT-061) at a dose of 2×10^{13} gc/kg for this pivotal Phase III trial. The interim data analysis from the dose confirmation trial (CT-AMT-061-01) is now available and the dose for AMT-061 of 2×10^{13} gc/kg has been confirmed for the treatment phase of this Phase III trial.

1.2 Interim Safety Results for CT-AMT-061-01

Enrollment and dosing in an ongoing dose confirmation Phase IIb trial with AMT-061 (CT-AMT-061-01) has been completed and three subjects have been dosed with 2×10^{13} gc/kg of AMT-061. An interim analysis was performed after all subjects had completed 6 weeks of post-treatment follow-up and the interim analysis data set includes a combined 24 weeks of observation. The interim analysis results from CT-AMT-061-01 were reviewed by the Data Monitoring Committee (DMC) who recommended that a single dose of 2×10^{13} gc/kg of AMT-061 be used in the treatment phase of this Phase III trial.

There have been no deaths, serious adverse events (SAEs), or treatment emergent adverse events (TEAEs) resulting in early discontinuation from CT-AMT-061-01 at the time of the interim analysis. Overall, there have been 13 TEAEs reported in 2 of 3 subjects, including 2 treatment-related TEAEs reported in a single subject (C-reactive protein (CRP) increased and headache). All TEAEs were mild in severity and the majority (11/13 events) were assessed as not related to study treatment. According to the system organ class (SOC), the most reported events were Nervous System Disorders (4 events) and the most common TEAE was headache, reported in 2 subjects, though one subject experienced 6 separate events of pain. One subject had a mild, asymptomatic, and transient increase in liver enzyme levels (specifically, elevated aspartate aminotransferase (AST) levels) which resolved without any additional treatment and was not reported as a TEAE. No T-cell response has been observed in any of the 3 subjects. None of the subjects had developed inhibitory antibodies to factor IX.

1.3 Efficacy Results for CT-AMT-061-01

All 3 subjects had been diagnosed with severe hemophilia B with corresponding circulating factor IX activity levels <1% of normal. A single dose of 2×10^{13} gc/kg of AMT061 was demonstrated to result in mean ± SD factor IX activity of 30.6 ± 6.97% (range: 23.9-37.8%) at Week 6 and 40.8 ± 9.45% (range: 31.3-50.2%) at Week 52 (factor IX activity was measured using the activated partial thromboplastin time (aPTT) assay performed at a central laboratory). Factor IX activity at 52 weeks after administration was 50.2% of normal in the first subject, 40.8% of normal in the second subject, and 31.3% of normal in the third subject. The second and third subjects had previously screen-failed another gene therapy study due to pre-existing neutralizing antibodies to a different adeno-associated viral (AAV) vector. All 3 subjects tested

low positive in the luciferase-based assay for pre-existing neutralizing activity against adeno-associated viral vector serotype 5 (AAV5).

In the year prior to screening the number of treatment-requiring bleeding episodes ranged from 1 to 5 spontaneous episodes and there was 1 moderate spontaneous bleeding episode in 1 subject in the period between screening and dosing. After dosing, there were no bleeding episodes and the estimated ABR was 0. Additionally, none of the subjects have required any infusions of factor IX replacement therapy post-treatment with AMT-061 for the selective prevention of a bleed, as of 52 weeks post-dose.

1.4 This Statistical Analysis Plan (SAP)

This Statistical Analysis Plan (SAP) outlines the statistical methods for the display, summary and analysis of data to be performed for the 26-week analysis, 52-week post-dose analysis, , 52-week analysis following the last patient reaching stable FIX ,and CSR addendum covering the long-term follow-up period. The SAP should be read in conjunction with the study protocol. This version of the SAP has been developed using the CT-AMT-061-02 Protocol (Version 5.0 15 Oct 2020) and the CT-AMT-061-02 CRF (Version 12.0 02 Oct 2020).

2. STUDY OBJECTIVES AND ENDPOINTS

2.1 Study Objectives

2.1.1 Primary Objectives

The primary objectives are

- to demonstrate the non-inferiority of AMT-061 (2×10^{13} gc/kg) during the 52 weeks following establishment of stable FIX expression (months 6 to 18) post-treatment (AMT-061) follow-up compared to standard of care continuous routine factor IX prophylaxis during the lead-in phase, as measured by the ABR.

2.1.2 Secondary Objectives

The secondary objective is to demonstrate additional efficacy and safety aspects of systemic administration of AMT-061.

2.1.2.1 Secondary Efficacy Objectives

To investigate the effect of 2×10^{13} gc/kg AMT-061 on the following:

- Endogenous factor IX activity 6 months after a single AMT-061 treatment
- Endogenous factor IX activity 12 months after a single AMT-061 treatment
- Endogenous factor IX activity 18 months after a single AMT-061 treatment
- Annualized consumption of factor IX replacement therapy

- Annualized infusion rate of factor IX replacement therapy
- Discontinuation of previous continuous routine prophylaxis
- Trough factor IX activity
- Prevention of bleedings (comparison for superiority)
- Prevention of spontaneous bleeding
- Prevention of joint bleeding
- Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) – as a function of pre-investigational-medical-product (IMP) anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis)
- Correlation of pre-investigational medicinal product (IMP) anti-AAV5 antibody titers using the luciferase based neutralizing antibody (NAB) assay on factor IX activity levels after AMT-061 dosing
- Occurrence and resolution of target joints
- Proportion of subjects with zero bleeds 6- during the 52 weeks following stable FIX expression (6-18 months) after AMT-061 dosing
- CCI [redacted]
- CCI [redacted].

CCI [redacted]

2.1.3 Safety Objectives

The safety objectives include evaluating the following:

- Adverse Events (AE)
- Changes in abdominal ultrasound
- Formation of anti-AAV5 antibodies (total immunoglobulin M and immunoglobulin G [IgM and IgG], neutralizing antibodies)
- AAV5 capsid-specific T cell response
- Formation of anti-FIX antibodies
- Formation of factor IX inhibitors and recovery
- Hematology and serum chemistry parameters

- AST and alanine aminotransferase (ALT) level increases and use of corticosteroids
- Shedding of vector deoxyribonucleic acid (DNA) in blood and semen
- Inflammatory markers: interleukin-1beta (IL-1β), interleukin-2 (IL-2), interleukin-6 (IL-6), interferon gamma (IFNγ), and monocyte chemotactic protein-1 (MCP-1)
- Alpha-fetoprotein (AFP)

2.2 Study Endpoints

2.2.1 Efficacy Endpoints

2.2.1.1 Primary Efficacy Endpoint

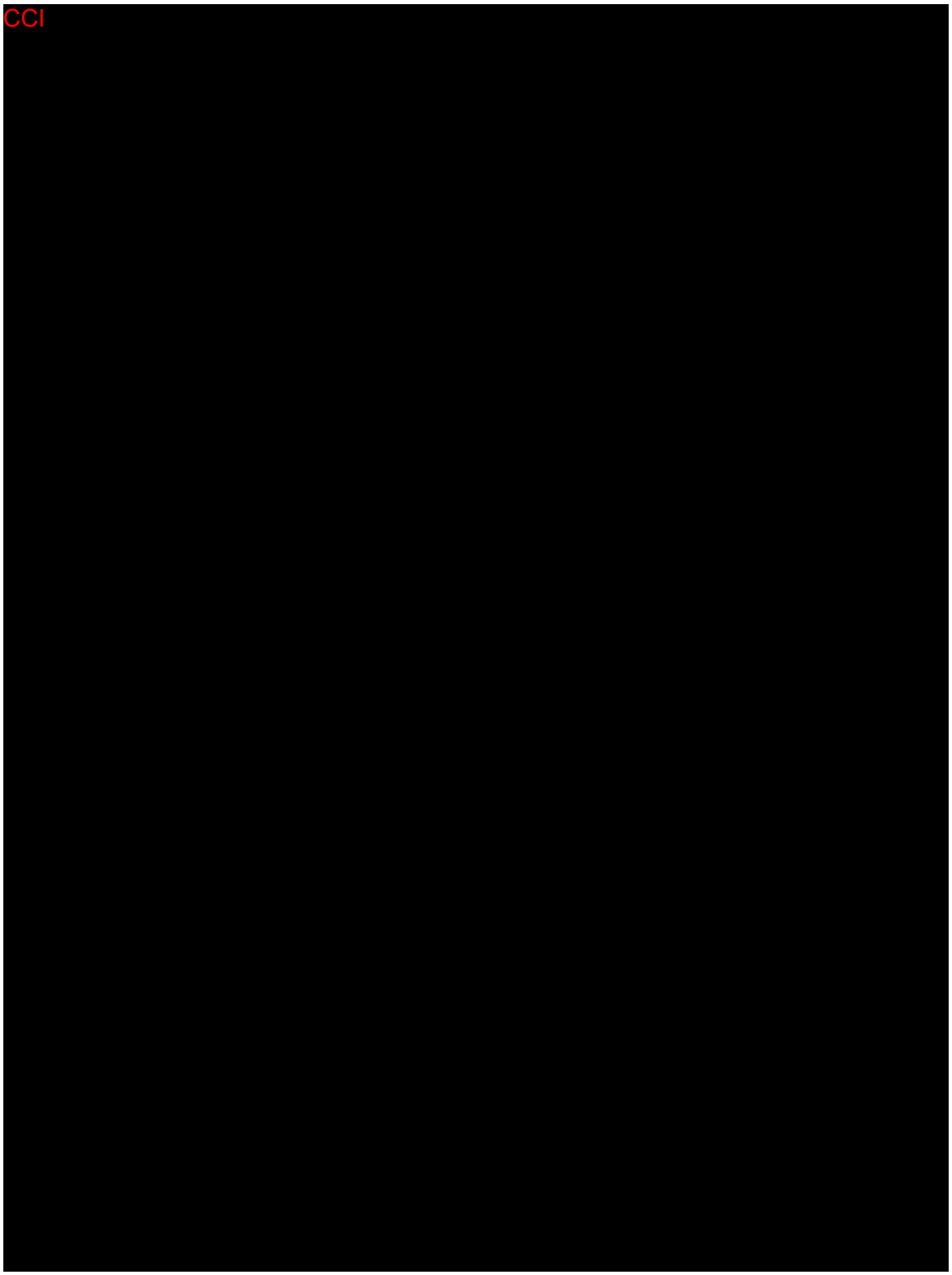
- ABR comparison between AMT-061 and prophylaxis for non-inferiority between the lead-in phase and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up

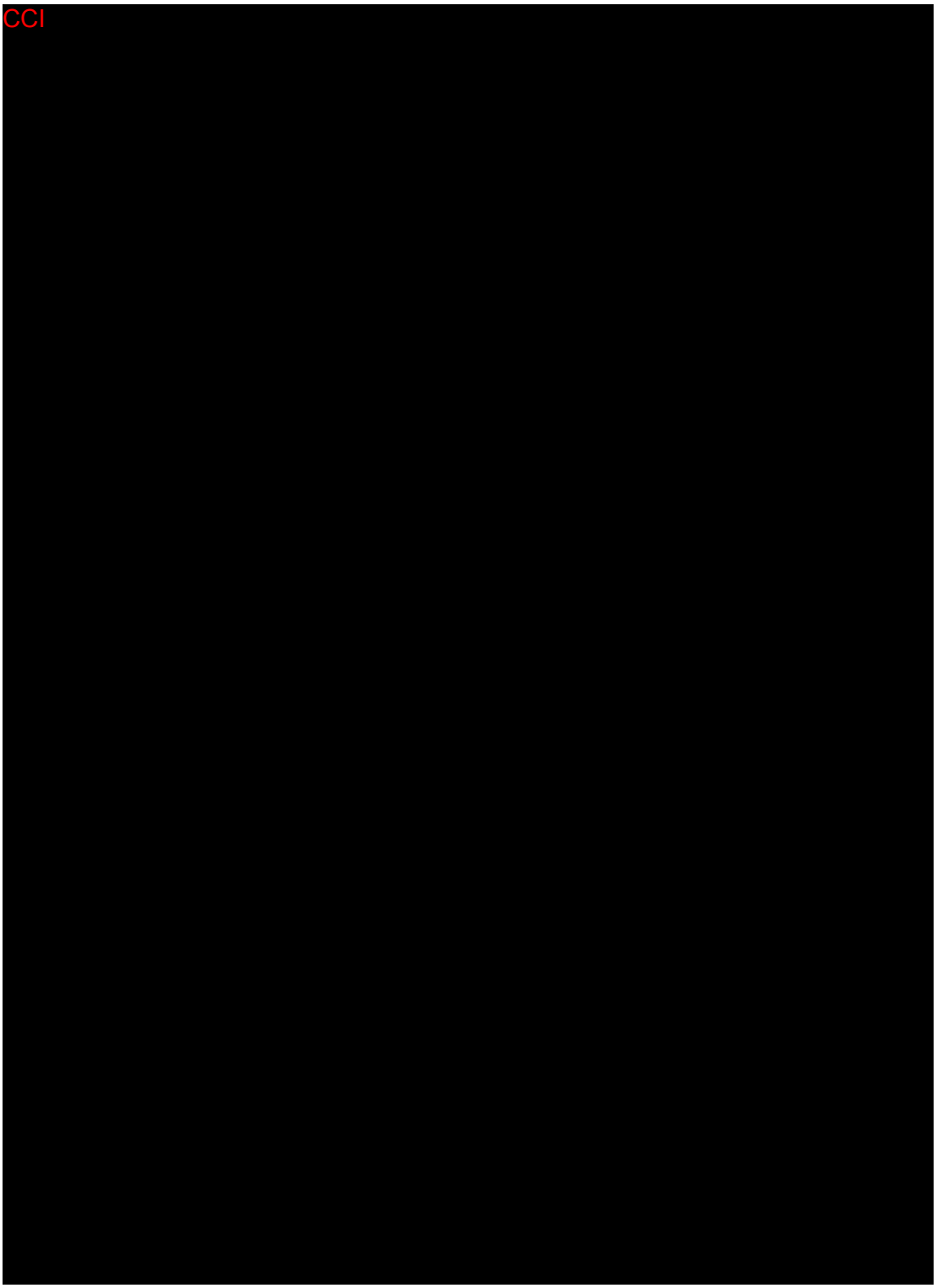
2.2.1.2 Secondary Efficacy Endpoints

- Endogenous factor IX activity at 6 months after AMT-061 dosing
- Endogenous factor IX activity at 12 months after AMT-061 dosing
- Endogenous factor IX activity at 18 months after AMT-061 dosing
- Annualized consumption of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase
- Annualized infusion rate of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase
- Proportion of subjects remaining free of previous continuous routine prophylaxis during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up
- Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 over the 52 weeks following stable FIX expression (6-18 months)

- ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up
- Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to lead-in phase
- Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase
- Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis)
- Correlation of factor IX activity levels during the 6-18 months post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay
- Occurrence of (and resolution of) new target joints during the 52 weeks following stable FIX expression (6-18 months) following AMT-061 dosing and resolution of pre-existing target joints following AMT-061 dosing
- Proportion of subjects with zero bleeds during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up
- CCI [redacted] during the 12 months following AMT-061 dosing compared with the lead-in phase
- CCI [redacted] during the 12 months following AMT-061 dosing compared with the lead-in phase

CCI [redacted]


CCI


CCI

2.2.1.4 Sub-study Endpoints

- Patient Reported Outcomes, Burdens, and Experiences (PROBE) summary scores and individual item responses
- Musculoskeletal Ultrasound Sub-Study Endpoints
    - Joint Tissue Activity and Damage Exam (JADE) scores

2.2.2 Safety Endpoints

All adverse event (AE) data will be collected from signing of the informed consent form until the end of the five-year follow-up. An AE, adverse drug reaction (ADR), and SAE are defined according to the ICH Guidelines E2A.

Safety analyses will be based on the safety population and described as descriptive analyses. Safety set is defined as any study subjects receiving at least some amount of study treatment even when the full dose was not administered (including partial dose).

Safety data will be analysed per study phase i.e., lead in phase, treatment emergent adverse events (TEAEs) during 26 and 52 weeks after study treatment start and during the follow up phase. The overall safety profile of AMT-061 will be assessed given the below safety and tolerability criteria.

Secondary safety endpoints include the following:

- AEs
- Changes in abdominal ultrasound
- Anti-AAV5 antibodies (total [IgM and IgG], neutralizing antibodies)
- AAV5 capsid-specific T cells
- Anti-FIX antibodies
- Factor IX inhibitors and recovery

- Hematology and serum chemistry parameters
- ALT and AST levels, and corticosteroid use for ALT and AST increases
- Vector DNA in blood and semen
- Inflammatory markers:
  - IL-1β
  - IL-2
  - IL-6
  - IFNγ
  - MCP-1
- AFP

Other laboratory evaluations include coagulation and serology parameters. In addition, the following (S)AEs qualify for special notification as they are seen as safety issues of particular concern for Advanced Therapy (ATMP) (ENTR/F/2/SF/dn D(2009) 35810. Brussels, 03/12/2009) and gene therapy medicinal products (EMA / CHMP / Gene Therapy Working Party (GTWP) / 60436/2007):

- AEs related to the IMP administration procedure
- Suspected or confirmed cases of opportunistic or serious infections that in the investigator's opinion might be related to the IMP
- Unexpected reactions (e.g., hypersensitivity, immunological, toxic or other as consequence of a change in the construction or function of the viral vector [e.g., generation of replication competent virus])
- AEs related to product failure (including lack of efficacy), mandatory concomitant medication (e.g., immunosuppression), and medical devices which form part of the product or are used for application of the product
- Development of any new/recurrent cancer

All TEAEs are tabulated displaying the number of subjects (and percentage) experiencing an event and the number of events by SOC and preferred terms within each SOC according to the Medical Dictionary for Regulatory Activities (MedDRA) terminology. TEAEs will also be tabulated by severity and by relationship to trial medication, using frequency counts (number of subjects with

event and number of events) and percentages. Similar tables will be created for TEAEs leading to premature discontinuation or interruption, AESIs, deaths, seriousness, infusion related and hypersensitivity reactions if applicable. These summary tables will be presented by decreasing frequency of occurrence based on SOC and Preferred Term.

The summary tables will be accompanied by individual subject listings of all AEs, including information on AE number, actual AE description, date/time of start and end of AE, preferred term (MedDRA), SOC (MedDRA), severity, relationship/causality, type of AE, seriousness, and outcome. Pre-existing AEs will be flagged. Pre-existing AEs are not considered to be treatment emergent, except in case of worsening during/after trial treatment (to be collected as separate AE). Separate listings will be created for AEs for special notification, deaths, and SAEs, if applicable. Other safety data will be presented using graphical displays, as applicable, descriptive statistics (including change from baseline, if applicable), and/or individual data listing.
The number of days until vector DNA can no longer be detected in semen and blood will be tabulated. The number of days is calculated using the date of collection of the first of three consecutive negative samples for each matrix. (The protocol should have said first of three instead of third of three.)

3. STUDY DESIGN AND ANALYTICAL CONSIDERATIONS

3.1 Study Design

3.1.1 Overall Study Design and Plan

CT-AMT-061-02 is an open-label, single-dose, multi-center, multi-national trial, with a screening period, a lead-in phase, a treatment + post-treatment follow-up phase, and a long-term follow-up phase. Overviews of the trial and its design are presented in Figure 1 and Figure 2.

Figure 1: Study Overview

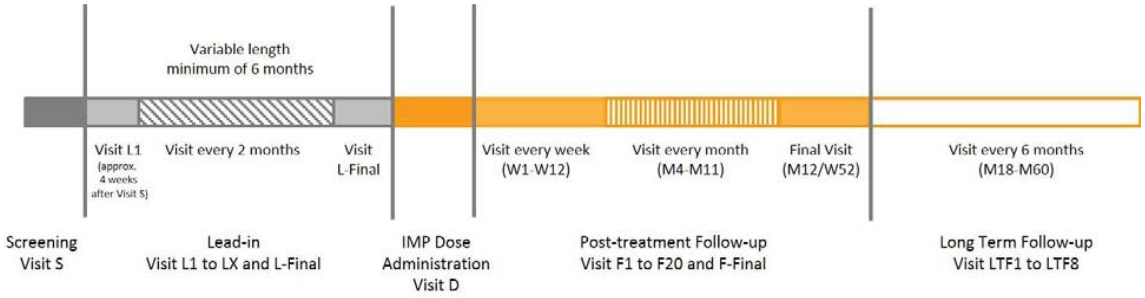


Figure 2: Study Design

Refer to the Protocol for the schedule of events for efficacy and safety evaluation for screening, lead-in, treatment and post-treatment, and follow-up visits. Refer to the Protocol for the schedule of events for laboratory parameters for screening, lead-in, treatment and post-treatment, and follow-up visits. Refer to the Protocol for the schedule of events for efficacy and safety evaluation for long-term follow-up visits. Refer to the Protocol for the schedule of events for laboratory parameters for long-term follow-up visits.

After screening, eligible subjects will enter a lead-in phase prior to the start of AMT-061 treatment. Visits will occur every 2 months during the lead-in phase with the final visit occurring a month prior to dosing. During the post-treatment follow-up, visits will occur weekly up to Week 12 and then monthly up to Month 12, after which subjects will enter a long-term follow-up with visits every 6 months.

Six months after IMP administration, the first secondary endpoint, endogenous factor IX activity at 6 months after IMP administration, will be analyzed and reported via an interim analysis once the last subject has achieved 6 months after AMT-061 treatment. This assessment will be based on clean data and a partially locked database.

Twelve months after IMP administration, the second secondary endpoint, endogenous factor IX activity at 12 months after IMP administration, will be analyzed and reported via another interim analysis once the last subject has achieved 12 months after AMT-061 treatment. This assessment will be based on clean data and a partially locked database.

After 52 weeks following stable FIX expression (18 months post-dose), all available efficacy and safety data collected between screening and 18 months post-treatment will be analyzed and reported in a full CSR, including (but not limited to) the primary ABR endpoint and the third secondary endpoint. The first and second secondary efficacy endpoints (which will have been analyzed in their respective data cuts) will also be added to the full (18-month) CSR. Data up to each analysis time point will be considered locked and will not be changed (with the exception of ending dates and outcomes for continuing events and treatments) without explicit authorization. The subjects will be followed for an additional 3.5 years for evaluation of efficacy parameters and safety. At the end of that 3.5-year period, all safety and efficacy data will be reported in a CSR addendum covering the entire study duration, including the later 3.5-year period.

The overall trial participation will be approximately five and a half years.

3.2 Interim Analysis

There will be an interim efficacy analysis of the first secondary endpoint, 6-month endogenous factor IX activity levels, after all subjects have completed the 6-month assessment and the database is (partially) locked.

There will be an interim efficacy analysis of the second secondary endpoint, 12-month endogenous factor IX activity levels, after all subjects have completed the 12-month assessment and the database is (partially) locked.

3.3 Sample Size

The study sample size is constrained by the non-inferiority analysis of the primary endpoint, ABR.

Based on a literature search of trials in a similar clinical setting and the same underlying disease, as well as the previous AMT-060 Phase I/II trial, a non-inferiority margin of 1.8 is assessed for the rate ratio of ABR between AMT-061 (post-treatment) and factor IX prophylaxis (lead-in). For establishing the non-inferiority margin, an ABR of 2.4 between factor IX prophylaxis and placebo treatment has been assumed. Via simulation of ABR under a negative binomial distribution with a yearly rate of 2.4 events for lead-in and 1.9 for post-treatment, with a Pearson correlation of 0.05 for the number of events between the two periods, and with a common negative binomial dispersion parameter of 1.5, a sample size of N=50 will demonstrate non-inferiority with a non-inferiority margin of 1.8 and a power of 82.0%. Therefore, the study should consist of at least 50 analyzable subjects.

Given the sample size needed for ABR, this will produce a power >95% for the secondary statistical analysis of endogenous factor IX activity. For the secondary statistical analyses of factor IX activity at 6, 12, and 18 months, assuming a mean of 30.6 percent of normal (as observed at 6 weeks in study CT-AMT-061-01) and assuming a standard deviation of 6.97 (as observed at 6 weeks in study CT-AMT-061-01), assuming conservatively that the baseline factor IX activity is 2%, and assuming that the sample size is 50 subjects, for a one-sample t-test at the 0.025 one-sided level of significance to test whether the change from baseline is > 0, the statistical power is > 99%. Alternatively assuming that the standard deviation is 6.95, which is half of the range of factor IX activity values (23.9 to 37.8) observed at 6 weeks in study CT-AMT-061-01, the statistical power is still > 99%. The nQuery Advisor software was employed for this power calculation.

3.3.1 Non-inferiority Margin for Rate Ratio Analysis

Recent studies in similar Hemophilia B populations have shown the results – presented in the table below – relative to on-demand treatment. In the Idelvion publication, Group 2 (n=19) was followed for 6 months using on-demand treatment and then for 6 months using 7-day prophylaxis treatment. The ABR rates were estimated as 365.25 x (number of bleeding

episodes)/ (number of days in the observed treatment period of interest). The rate reduction, 0.11, was estimated from a Poisson distribution, with a corresponding two-sided, 95% confidence interval (CI) for the rate reduction of (0.051, 0.238).

Table 1: Recent Results of Prophylaxis Compared to On-demand

Publication	ABR On-demand	ABR Prophylaxis	Rate Reduction
Alprolix (Powell, 2013)	18.67 (N=27)	3.12 (N=61)	0.17
Idelvion (Santagostino, 2016)	20.09 (N=19)	2.22 (N=19)	0.11
Nonacog (Collins, 2014)	15.58 (N=15)	40 IU: 2.51 (N=29)	0.16

Using a rate ratio analysis, as opposed to a difference in rates, results in an evaluation that is relatively independent of the magnitude of the baseline ABR. Thus, a rate reduction of 0.50 for a subject with 20 events during lead-in has the same meaning as for a subject with 4 events during lead-in. However, the difference in rates is quite different between such subjects (10 events and 2 events, respectively).

As currently proposed, the null and alternative hypotheses can be written as:

$$H_0: \frac{\lambda_{AMT-061}}{\lambda_{Prophy}} \geq M \; vs \; H_1: \frac{\lambda_{AMT-061}}{\lambda_{Prophy}} < M$$

where M represents the non-inferiority margin and λ_{XXX} represents the rate of events in the XXX group (AMT-061 representing the post-treatment period, Prophy representing the lead-in period). This can be equivalently rewritten as the difference between the rates on the natural logarithmic (base e) scale:

$$H_0: \log \lambda_{AMT-061} - \log \lambda_{Prophy} \geq \log M \;\; vs \; H_1: \log \lambda_{AMT-061} - \log \lambda_{Prophy} < \log M$$

If M1 represents the entire effect of prophylactic treatment compared to on-demand treatment on the natural logarithmic scale, then M1 = log(0.238) = -1.4354846, the upper limit of the CI from the Idelvion publication. If M2 represents the percentage of treatment effect relative to on-demand to be preserved, then M2 can be stated as p*M1 = p*log(0.238). Then, the hypotheses of interest can be restated as:

$$H_0: \frac{\lambda_{AMT-061}/\lambda_{on-d}}{\lambda_{Prophy}/\lambda_{on-d}} \geq M \; vs \;\; H_1: \frac{\lambda_{AMT-061}/\lambda_{on-d}}{\lambda_{Prophy}/\lambda_{on-d}} < M$$

or equivalently as differences on the natural logarithmic scale, where "on-d" represents on-demand treatment. Substitution of M1 in the denominator and M2 in the numerator gives the following equation:

$$(p-1)M1 \geq \log M$$

$$(1-p)x1.4354846 \geq \log M$$

This can then be solved for M or for p. This approach will be called Approach (1).

Alternatively, if M1 represents the entire effect of prophylactic treatment on the efficacy scale, then M1 = 1 – 0.238 = 0.762, using the relationship that efficacy = 1 – rate ratio, and M2 = pM1. Substituting 1-M2 into the numerator and 1-M1 into the denominator (RR = 1-efficacy), yields the following equation:

$$\frac{1-p*0.762}{1-0.762} < M$$

This approach will be called Approach (2).

The two approaches provide slightly different interpretations.

Table 2: Non-inferiority Margin (for the ABR rate ratio) for Retention of p%

	Percentage Retention				
	60%	**70%**	**75%**	**80%**	**90%**
Approach (1)	2.28	1.96	1.80	1.64	1.32
Approach (2)	1.776[a]	1.538	1.43	1.33	1.15

a. A margin of 1.8 using difference in log values corresponds to 59% retention of treatment effect.

Hypothetical scenarios illustrating the upper limit of the non-inferiority margin under observations of varying pre-treatment annualized bleed rates are shown in the table below.

Table 3: Relation of pre-treatment ABR and NI margin using a 1.8 rate ratio

ABR observed during lead-in phase	2.0	2.5	5.0	10
Maximum ABR post-treatment to maintain NI (bleeds)	<3.6	<4.5	<9.0	<18
Permissible increase in number of bleeds	<1.6	<2.0	<4.0	<8

3.4 Randomization and Blinding

Not applicable, as this is an open-label trial with one treatment arm.

4. DATA DEFINITIONS AND PRE-PROCESSING

4.1 Baseline Definitions

The baseline factor IX activity will be imputed based on the patient's historical hemophilia B severity that is documented on the CRF. If the patient has documented severe factor IX deficiency (factor IX plasma level <1%) their baseline factor IX activity level will be imputed as 1%. If the patient has documented moderately severe factor IX deficiency (factor IX plasma level >= 1% and <= 2%) their baseline factor IX activity level will be imputed as 2%.

The baseline ABR will be the number of bleeds in the previous year as assessed at screening.

Baseline age is the age in years at the time of the Screening Visit.

For CCI and CCI that are assessed at visits, the Baseline value is period-specific. For the lead-in period, the Baseline value is the latest value prior to the lead-in period that is not within 2 weeks of a bleed. For the post-treatment period, the Baseline value is the latest value prior to IMP that is not within 2 weeks of a bleed. Note that the Baseline value is not used as a covariate in statistical-modeling-based treatment-period-comparative analyses for this study, because each subject serves as one's own control.

For vital signs and safety laboratory values, Baseline is period-specific. The Baseline value for the lead-in period is the last non-missing central laboratory value or vital signs value on or prior to Visit L1. The Baseline value for the post-treatment period is the last central laboratory value or vital signs value prior to the first dose of AMT-061.

For height, weight, and BMI, the baseline value is the last value prior to the start of the lead-in period.

4.2 Data Handling Rules and Definitions, Including Handling of Missing Data

Missing data will be maintained as missing in the safety and efficacy datasets, unless specified otherwise.

Data for Adverse Events Summaries by Severity and Relationship to Study Drug

For the AE summaries by severity (mild, moderate, or severe), an AE with missing severity will be deemed as severe. For the AE summaries by relationship to study drug, an AE with a missing relationship to study drug will be deemed as related. Seriousness cannot be imputed as 'Yes' by default, since this would affect the reconciliation between trial database and registry of SAEs.

Data for Laboratory Summaries (Continuous Parameters)

Data from unscheduled visits will not be used for by-visit summaries (unless they have been assigned to a scheduled visit according to the Time Windows for Statistical Analysis). Data from both scheduled and unscheduled visits will be used for determining incidence of clinically significant values.

Data for All Laboratory Summaries

Definitions are provided in Appendix 1 Data Handling Rules.

Study Dates and Day of Assessment or Event

Study Day and Day of Assessment or Event definitions are provided in Appendix 1 Data Handling Rules.

Duration of Event

Definitions are provided in Appendix 1 Data Handling Rules.

Distance between Events

Definitions are provided in Appendix 1 Data Handling Rules.

The Use of Lead-In Period Month 6 Visit versus Lead-In Final Visit for Analyses of Factor IX Activity

For factor IX activity, for the Lead-In Month 6 assessment and the Lead-In Final assessment, some specific instructions are as follows. For factor IX activity (aPTT and chromogenic from the central laboratory) for analysis purposes if a Lead-In Month 6 value is available, then it will be used (for analysis); otherwise if a Lead-In Final value is available, then it will be used for the purpose of (i.e. as if it were) the Month 6 value. The rationale is that (1) Lead-In Month 6 is a planned assessment (for factor IX activity) and that in the presence of a Month 6 assessment, a Lead-In Final value is not essential for analysis and (2) the planned duration of the Lead-In period is approximately 6 months.

Missing Data

If causality is missing for a TEAE, the TEAE will be regarded as ‘Related’. If causality is missing for an AE with onset before administration of trial drug, the AE will be regarded as 'Not related'. If the intensity is missing, the intensity of the AE will be regarded as “Severe.” In the case where seriousness is missing, this should be queried. Seriousness cannot be imputed as 'Yes' by default, since this would affect the reconciliation between trial database and registry of SAEs.

Time Windows for Data Collection:

The below assessment time windows (in Table 4) are given in the protocol as "target" time windows for assessments to be carried out. However, these are not the time windows to be applied for statistical analysis, which are described farther below.

Table 4: Protocol-Recommended Study Time Intervals for Collection of Efficacy and Safety Evaluation and Laboratory Parameters

Nominal Time for Visits or Assessment	Protocol-Recommended Time Interval for the Visit or Assessment
Screening	Approximately -28 days prior to Visit L1
Lead-In	
Visit L1 (L-W0)	0 days
Visit L2 to LX (every 2 months, starting at L-W8)	±14 days
Visit L-Final	-28 days (± 7 days) from Visit D
IMP Dose	
Post-IMP 3 hours	±15 minutes
Post-treatment Follow-Up	
Week 1 to Week 12	±2 days
Month 4 to Month 11	±5 days
Month 12/Week 52	±5 days
Long-Term Follow-up	
Month 18	±2 weeks
Month 24	±2 weeks
Month 30	±2 weeks
Month 36	±2 weeks
Month 42	±2 weeks
Month 48	±2 weeks
Month 54	±2 weeks
Month 60	±2 weeks

Time Windows for Statistical Analysis:

Scheduling difficulties due to the coronavirus disease (COVID-19) pandemic may result in an increased number of missed, delayed, or unscheduled visits. Scheduling difficulties may also result in the performance of assessments at a scheduled visit where performance of the assessment was not originally planned. As an action to mitigate risk, analysis windows will utilize such unplanned assessments as follows. A schema for the assignment of such unplanned

assessments to scheduled time points for visit-based endpoint analysis and summary (for visit-based efficacy and safety endpoints) will be defined as follows:

- An unplanned assessment will be assigned to a scheduled visit only if that visit has a missing value for the relevant endpoint and the visit was a scheduled time point for the performance of the assessment per the study protocol.

- Analysis windows for the assignment of unplanned assessments will range from the previous visit at which the endpoint is planned to be collected to the next visit at which the endpoint is planned to be collected.

- The unplanned assessment closest in time within the analysis window (either before or after) will be used to replace a missing assessment for a scheduled visit. If two unplanned assessments are both the closest in time, with one being before and the other being after, the earlier assessment will be used.

- For efficacy endpoints, values obtained on a post-treatment actual study day < Day 21 will not be candidates to be used for imputing missing values for post-treatment visits at a nominal visit time subsequent to post-treatment Study Day 21.

- Only values from unplanned assessments (not planned assessments) will be used to replace a missing scheduled assessment.

- An unplanned assessment may be used more than once provided it lies within the analysis window for two consecutive scheduled time points at which the assessment was not performed as planned.

- Analysis windows will be defined separately for the lead-in and post-treatment periods. This means that values will not be carried from one treatment period to the other.

- For laboratory-based efficacy assessments, the above instructions (in the first 8 bullet points above) are applicable pertaining to all planned assessments; except, however, for the Lead-In Month 6 assessment and the Lead-In Final assessment, some specific instructions are as follows. For relevant laboratory-based efficacy endpoints – i.e. factor IX activity (aPTT and chromogenic) – and for the laboratory endpoints of Total (IgM and IgG) and Neutralizing Antibodies to AAV5 –, eligible unplanned assessment values will be used (as applicable) to replace a missing assessment for the Lead-In Month 6 Visit (but not used to provide a value for the Lead-In Final Visit). The rationale is that Lead-In Month 6 is a planned assessment and that – for such endpoints – in the presence of a Month 6 assessment, a Lead-In Final value is not essential for analysis. For these endpoints, the analysis window for the assignment of unplanned assessments to the Lead-In Month 6 Visit will range from the previous visit at which the endpoint is planned to be collected (i.e. the planned date of the Lead-In Month 4 Visit) to the actual Day before AMT-061 dosing. If both the Lead-In Month 6 Visit and the Lead-In Final Visit have no value (for the endpoint), then an unplanned-assessment value will be sought to be

assigned to the Lead-In Month 6 Visit; if either the Lead-In Month 6 Visit or the Lead-In Final Visit has a value (for the endpoint), then no such assignment (of an unplanned assessment to Lead-In Month 6) is needed. Unscheduled-visit values should never be assigned to the Lead-In Final Visit.

- For CCI efficacy endpoints, the above instructions (in the first 8 bullet points above) are applicable pertaining to all planned assessments; however, for certain CCI -endpoints some specific instructions for the imputation of the Lead-In final value are provided as follows. For CCI efficacy endpoints that are scheduled to be collected at the Lead-In Month 4 and Lead-In Final Visits – i.e. CCI, CCI, CCI, CCI, CCI, CCI, and PROBE – eligible unplanned assessment values will be used (as applicable) to replace a missing assessment for the Lead-In Final Visit. For these endpoints, the analysis window for the assignment of unplanned assessments to the L-Final Visit will range from the previous visit at which the endpoint is planned to be collected (i.e. the planned date of the Lead-In Month 4 Visit) to the actual Day before AMT-061 dosing. The latest available eligible unplanned assessment will be assigned to the Lead-In Final Visit. By the way, the Lead-In Month 6 Visit is not a planned assessment time and therefore is not to be receiving values from unplanned assessments.

- For CCI efficacy endpoints, the above instructions (in the first 8 bullet points above) are applicable pertaining to all planned assessments; however, for certain (other) CCI -endpoints some specific instructions for the imputation of the Lead-In final value are provided as follows. For visit-based efficacy and safety endpoints that are scheduled be collected for the first time (post-screening) at the Lead-In Final Visit – i.e. MSKUS (efficacy) and Abdominal Ultrasound (safety) – eligible unplanned assessment values will be used (as applicable) to replace a missing assessment for the Lead-In Final Visit. For these endpoints, the analysis window for the assignment of unplanned assessments to the L-Final Visit will range across the entire duration of the Lead-In Period. The latest available eligible unplanned assessment will be assigned to the Lead-In Final Visit. By the way, the Lead-In Month 6 Visit is not a planned assessment time and therefore is not to be receiving values from unplanned assessments.

Unscheduled or unplanned assessment values will not be assigned to scheduled (analysis) visits for vector DNA (genome) assessments.

Any other rules for missing data handling will be given in the endpoint-specific sections.

4.3 Bleed Counting Rules

Bleeds will be counted irrespective of assessments by the investigator as to the trueness or newness of the bleed (except for a small number of designated sensitivity analyses).

For designated supportive analyses, only exogenous-factor-IX-treated bleeds will be counted. If the field for whether the bleed was treated is missing, then (for conservativeness) it will be assumed that the bleed was treated with factor IX.

For a small number of designated sensitivity analyses, only bleeds that are assessed to be new and true bleeds will be counted. If the assessment field for newness of the bleed is missing, then (for conservativeness) it will be assumed that the bleed is new. If the assessment field for trueness of the bleed is missing, then (for conservativeness) it will be assumed that the bleed is true. The rationale for these sensitivity analyses is given in the following bullet points:

- The occurrence of bleeds during the conduct of the study are self-reported by enrolled subjects through the daily use of electronic and paper diaries, up to the 52 Week post-treatment visit and from the 52 Week visit until the five year visits, respectively. In some cases, due to previous chronic joint bleeds and damage, the patient may experience pain, mistake it for a new bleed, and self-infuse factor IX. Investigators will review and assess reported bleeds as a means of verifying that patient-reported events meet the clinical criteria required to be characterized as new, true bleeds. The steps in the bleed reporting and assessment process are presented below.

- The patient reports signs and symptoms in a daily e-diary up to the Week 52 visit, and in paper diaries through to the end of the study.

- The Principal Investigator or designee reviews the diary data and, if needed, requests further information from the patient prior to their evaluation of the signs and symptoms.

- The Principal Investigator or designee evaluates the signs and symptoms reported in the diary and/or during discussions with the patient and assesses whether the reported event was a true bleed and whether the reported event was a new bleed. For example, based on such sign and symptom evaluation, the investigator in some cases may need to distinguish whether there is a new bleed or whether the patient is experiencing pain (due e.g. to previous chronic joint bleeds and damage) that is not really a new bleed.

- When patients are next at the study site, the physician may elect to use a diagnostic scan (X-ray, ultrasound, MRI, CT scan, etc.) to confirm the presence of blood or signs of acute inflammation. Blood or signs of acute inflammation observed using one or more of these confirmatory methods coupled with the physician's assessment will serve as sufficient confirmation to identify an event as a true bleed.

5. DATA AND ANALYTICAL QUALITY ASSURANCE

The overall quality assurance procedures for the study data, statistical programming and analyses are described in Standard Operating Procedures (SOPs) of Everest Clinical Research. Detailed data management procedures are documented in the study Data Management Plan, Data Validation Check Specifications, and Integrated Safety Data Review Plan.

6. ANALYSIS POPULATIONS

6.1 Population Definitions

6.1.1 Screen Failures

The screen failure population will include all subjects who were screened but never entered the lead-in period.

6.1.2 Lead-in Discontinuers

The lead-in discontinuers population will include all subjects who entered the lead-in period but discontinued from the study prior to AMT-061 dosing.

6.1.3 Safety Population

The lead-in safety population will consist of all subjects who are enrolled into the lead-in period. The post-treatment safety population will consist of all subjects who receive AMT-061, irrespective of any protocol deviations. Period-specific safety tabulations will use the period-specific safety population for the "N" and denominator (for percentages). The safety population will consist of all subjects who are in either the lead-in safety population or the post-treatment safety population.

6.1.4 Full Analysis Set (FAS)

The FAS will include all subjects who are enrolled, entered the lead-in phase, were dosed with AMT-061, and provide at least one efficacy endpoint assessment for any efficacy endpoint subsequent to AMT-061 dosing. The FAS population will be the primary population for all efficacy statistical analyses.

6.1.5 Per-Protocol Population

The PP population will include all subjects from the FAS population who adhere to a stable and adequate prophylaxis use during the lead-in phase, who complete at least 18 months of efficacy assessments (52 weeks after achieving stable FIX expression) for the 18-month (data cut) analysis who complete at least a full year of efficacy assessments for the 12-month (data cut) analysis, or who complete at least 6 months of efficacy assessments for the 6-month (data cut) analysis, and who have no major protocol deviations that impact the interpretation of efficacy. The PP population will be used for sensitivity analyses. Protocol deviations that impact the interpretation of efficacy include the unwillingness to discontinue continuous prophylaxis use after receipt of AMT-061.

7. STATISTICAL ANALYSIS

7.1 Subject Disposition

A disposition table for CT-AMT-061-02 for all subjects will be provided. This tabulation will include the number of subjects who were screen failures, not treated with AMT-061 (i.e. lead-in discontinuers), who were prematurely discontinued from treatment (treated with only a partial dose of AMT-061), who received the full dose of study treatment, who withdrew early from the study post dose of AMT-061, and who completed the study. The number and percentage of subjects included in the FAS, PP, lead-in safety population, and post-treatment safety population will also be tabulated. The number and percentage of subjects in the PROBE sub-study and the MSKUS sub-study will also be tabulated. A subject is considered to be in the PROBE sub-study if the subject had at least one post-treatment assessment of PROBE. A subject is considered to be in the MSKUS sub-study if the subject had at least one post-treatment assessment of MSKUS.

The reason for exclusion from the FAS, PP, lead-in safety population, and post-treatment safety populations will be summarized. Reasons for premature discontinuation from study treatment will be summarized for the post-treatment safety population (this would include any partial dosing).

The data on subject disposition, missed visits, and protocol deviations (including those related to COVID-19) will be listed.

7.2 Demographic and Baseline Characteristics

Descriptive statistics of demographics and baseline characteristics will be presented for the FAS, PP, lead-in safety population, and post-treatment safety populations. For quantitative variables, all summaries will include the number of non-missing observations, mean, standard deviation (SD), first quartile (Q1), median, third quartile (Q3), minimum, and maximum. For the qualitative variables, the summaries will include the number and percentage of subjects in each category or level. All data will be included in listings.

7.2.1 Demography

Demographics collected at Screening include year of birth (i.e., age at Screening Visit), race, ethnic group, and gender according to local regulations. According to inclusion criterion 1, all patients are males.

7.2.2 Baseline Disease Characteristics

Baseline disease characteristics include duration of disease, endogenous factor IX activity level at time of diagnosis, severity of disease, indicator of family history of hemophilia B disease, number of bleeds in the year prior to screening (total, spontaneous, traumatic, joint, and unknown), and the type of factor IX therapy used. The baseline disease characteristics are tabulated according to the information collected on the electronic Case Report Form (eCRF).

Severity of hemophilia B will be categorized as severe (factor IX plasma level < 1%) or moderately severe (factor IX plasma level >= 1% and <= 2 %).

7.2.2.1 Hemophilia B History

All hemophilia B history data will be listed, and the listing will include the following: date first presented symptoms, date of initial diagnosis, duration of disease, endogenous factor IX activity level at diagnosis (if available), severity of hemophilia B at time of diagnosis, number of factor IX exposure days (an exposure day is defined as a day when the subject received at least one injection of factor IX treatment), and family members with a history of factor IX inhibitors.

7.2.2.2 Medical and Surgical History

All medical and surgical history will be listed, including the following information: surgical or medical history event, start date and end date or current status. Medical history will be coded using the most recent version of the MedDRA at the time of the database lock.

7.2.2.3 Target Joints at Screening

Target joints are defined as joints with three or more spontaneous bleeds into a single joint within a consecutive six-month period. Once there have been ≤ 2 bleeds into the joint within a consecutive 12-month period, the joint is no longer considered a target joint and the target joint is then considered to have resolved. Target joints at screening will be listed.

7.2.2.4 Baseline Antibody Parameters

The baseline antibody parameters include anti-FIX antibody titer levels, the presence of factor IX inhibitors, total (IgG and IgM) antibodies to AAV5, neutralizing antibody levels to AAV5, and AAV5 capsid-specific T-cells. These data will be listed.

Box and whisker plots over post-treatment study time will also be produced for these parameters, if applicable.

Please refer to Appendix 1 Data Handling Rules on how special laboratory values will be handled in quantitative analyses.

7.2.2.5 FIX Gene Mutation

Factor IX gene sequence analyses will be performed for all subjects who provide consent during the Screening Visit, even if they already have factor IX gene mutation information available. The data will be presented in a listing.

7.2.2.6 Prior and Concomitant Medications

Prior and concomitant medications will be collected and coded using the most recent World Health Organization (WHO) drug dictionary at the time of database lock. Prior and concomitant medications will be listed.

Prior medications are defined as those treatments with a start date before Visit L1 for the lead-in period. A medication/therapy will be identified as a "lead-in" concomitant medication/therapy if it is being continued by the subject at the date of the L1 Visit or is any new medication/therapy received during the lead-in period prior to the date of AMT-061 dosing. A medication/therapy will be identified as a "post-treatment concomitant" medication/therapy if it is being continued by the subject at the date of AMT-061 dosing or is any new medication/therapy received during the post-treatment period. A medication with end date that is the same as the AMT-061 dosing date will not be considered to be "post-treatment concomitant".

7.2.2.7 Prior Factor IX Therapy Use

Factor IX therapy use during the year prior to screening will be summarized and listed. Factor IX therapy use during the screening period will be listed.

7.3 Investigational Product Exposure

A listing for exposure to investigational product will be provided showing the date of exposure and dose received. The listing will also state whether the full dose was received.

Also, the time of subject's routine factor IX product/dose, incremental recovery, maximum concentration, and identity of subject's routine factor IX product/dose will be listed.

7.4 Blood Sample for Future Research

Data regarding the additional blood samples drawn at Screening (Visit S), Baseline (Visit D pre-IMP), Visit F12 (Week 12), and Visit F-Final (Month 12/Week 52), for the purpose of potential future research in the hemophilia B disease area will be listed with the corresponding informed consent date.

7.5 Efficacy Analyses

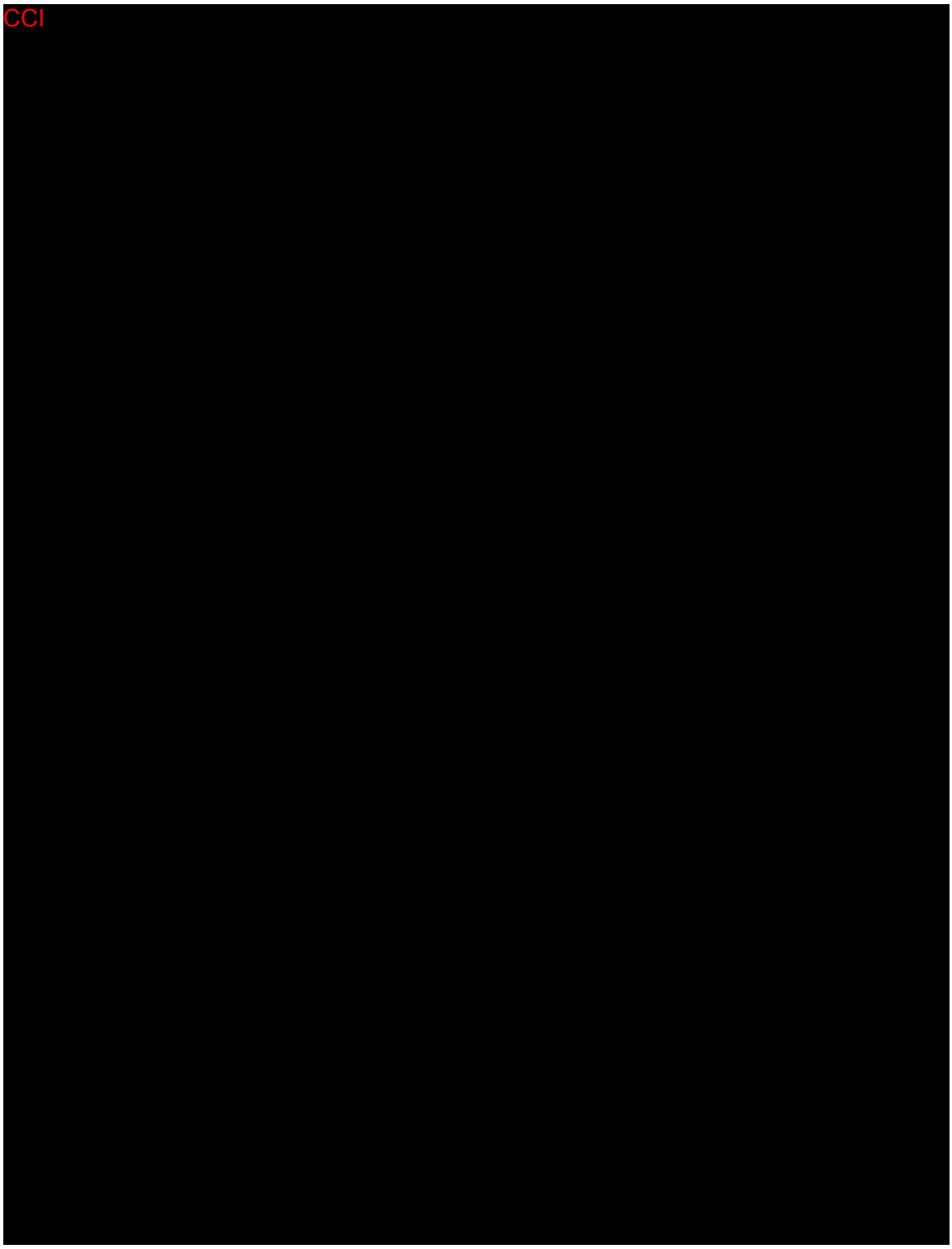
7.5.1 Primary Endpoint

The primary efficacy endpoint is as follows:

- ABR comparison between AMT-061 and prophylaxis for non-inferiority between the 52 weeks following stable FIX expression (6-18 months) post treatment (AMT-061) follow-up and the lead in phase

The analysis of the primary endpoint is discussed below.

CCI


CCI

CCI [redacted]

7.5.2.1 Sensitivity Analysis 1: PP Population

A sensitivity analysis using the PP population will be performed. With the exception of the study population, the analysis will be identical to the main analysis described in Section 7.5.4.

7.5.2.2 Sensitivity Analysis 2: Including (not excluding) Periods Subsequent to Exogenous Factor IX Use

A second sensitivity analysis, using the FAS population, will be conducted for ABR to evaluate the robustness of the analysis findings to inclusion (i.e. non-exclusion) of time intervals with exogenous factor IX use during the post-treatment period. In this analysis, person-time during the post-treatment period (that is) within 5 half-lives subsequent to exogenous factor IX use will not be excluded from (i.e. will be included in) the time at risk for a bleeding event. However, as with the primary ABR analysis, bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) (to Day 21 for pre-Month 18 data cuts) post-AMT will not be included in the analysis.

7.5.2.3 Sensitivity Analysis 3: Bleeds Treated with Exogenous Factor IX

A third sensitivity analysis will repeat the main analysis for ABR using the FAS population while considering only bleeds treated with exogenous factor IX. Bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) (to Day 21 for pre-Month 18 data cuts) post-AMT will not be included in the analysis.

7.5.2.4 Sensitivity Analysis 4: Cumulative Responder Analysis using Subject-Specific Bleeding Rates

To provide a descriptive summary of per-subject annualized bleeding rates in the respective treatment periods, a cumulative responder analysis (as described in Farrar 2006) will be performed characterizing ABR in the lead-in period and the post-treatment period. This analysis will use the FAS population. The observed ABR for the post-treatment period will be plotted on the *x*-axis, and the proportion of "responders" (subjects that have an ABR equal or greater than the level specified) will be plotted on the *y*-axis. The same graph will plot the observed ABR across the lead-in period (also) on the *x*-axis, with the proportion of "responders" (subjects that have an ABR equal or greater than the level specified) plotted on the *y*-axis. Thus, a cumulative distribution plot will be produced where the proportion of responders can be compared by treatment period across a continuous range of ABR values. The ABR for the post-treatment period will be calculated using all available data at the time of one year after the subject's dosing with AMT-061 (with AMT-061), the time of study completion, or the time of early withdrawal from the study, whichever is earlier. As with the main ABR analysis, person-time during time intervals "contaminated" with exogenous factor IX – according to the 5-half-life contamination rule – will be excluded from the denominator for the subject-specific bleeding rates, but bleeds during such time intervals will still be counted and included in the numerator for the rate. However, as with the primary ABR analysis, bleeds and person-time on or after Day 1 and prior

to stable FIX expression (Month 6) (to Day 21 for pre-Month 18 data cuts) post-AMT will not be included in the analysis.

7.5.2.5 Sensitivity Analysis 5: New and True Bleeds

A fifth sensitivity analysis will repeat the main analysis for ABR using the FAS population while considering only bleeds that are assessed to be new and true by the investigator. Bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) (to Day 21 for pre-Month 18 data cuts) post-AMT will not be included in the analysis.

7.5.2.6 Sensitivity Analysis 6: New and True Bleeds Treated with Exogenous Factor IX

A sixth sensitivity analysis will repeat the main analysis for ABR using the FAS population while considering only bleeds treated with exogenous factor IX that are assessed to be new and true by the investigator. Bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) (to Day 21 for pre-Month 18 data cuts) post-AMT will not be included in the analysis.

7.5.2.7 Sensitivity Analysis 7: Excluding Periods Contaminated by Systemic Corticosteroid Exposure

A seventh sensitivity analysis will repeat the main analysis for ABR using the FAS population and will be conducted to evaluate the robustness of the analysis findings to exclusion of time intervals with systemic corticosteroid use during the post-treatment period. In this analysis, person-time during the post-treatment period (that is) during systemic corticosteroid use or within 5 half-lives subsequent to the end of systemic corticosteroid use will be excluded from the time at risk for a bleeding event. However, as with the primary ABR analysis, bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) (to Day 21 for pre-Month 18 data cuts) post-AMT will not be included in the analysis.

7.5.2.8 Sensitivity Analysis 8: Optional Zero-Inflated Negative Binomial Regression

An optional sensitivity analysis may be conducted for ABR to evaluate the robustness of the analysis findings to account for the possibility of there being a very large number of subjects having 0 bleeds in either treatment period. This sensitivity analysis will be considered optional given that literature indicates that adding a zero-inflation component to a negative binomial regression is likely superfluous (https://statisticalhorizons.com/zero-inflated-models) given the inherent flexibility of the negative binomial model (to accommodate a substantial number of counts of zero). Some relatively recent literature has been written about methodology for repeated measures GEE zero-inflated negative binomial regression (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4303594/); however, the methodology may not yet have been incorporated into standard statistical analysis packages.

CCI

[redacted]

CCI

[redacted]

7.6.1 Secondary Efficacy Analysis

The secondary efficacy endpoints are as follows:

1. Endogenous factor IX activity at 6 months after AMT-061 dosing
2. Endogenous factor IX activity at 12 months after AMT-061 dosing
3. Endogenous factor IX activity at 18 months after AMT-061 dosing
4. Annualized consumption of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up, excluding factor IX replacement for invasive procedures compared to the lead-in phase
5. Annualized infusion rate of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up, excluding factor IX replacement for invasive procedures compared to the lead-in phase
6. Proportion of subjects remaining free of previous continuous routine prophylaxis during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up
7. Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 over the 52 weeks following stable FIX expression (6-18 months).
8. ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up
9. Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up compared to the lead-in phase
10. Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up compared to the lead-in phase
11. Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a “correlation” analysis) (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)

12. Correlation of factor IX activity levels during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)

13. Occurrence of (and resolution of) new target joints during the 52 weeks following stable FIX expression (6-18 months) following AMT-061 dosing and resolution of pre-existing target joints following AMT-061 dosing (these endpoints will not have hypothesis testing and therefore are not included in the Type I error control)

14. Proportion of subjects with zero bleeds in the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)

15. CCI [redacted] during the 12 months following AMT-061 dosing compared with the lead-in phase

16. CCI [redacted] during the 12 monthsfollowing AMT-061 dosing compared with the lead-in phase.

The analysis of each secondary endpoint is discussed in its own subsection below.

For the secondary efficacy endpoints, main analyses will be performed using the FAS. Analyses of superiority using the PP population will be considered to be supportive analyses.

All data will be listed.

7.6.2 Endogenous Factor IX Activity at 6 months After AMT-061 Dosing using the "Original" Contamination Rule

The main analysis of the first secondary endpoint "Uncontaminated Endogenous Central Laboratory One-Stage aPTT Factor IX activity at 6 months after AMT-061 dosing" actually took place using the 6-month data cut and was done according to the existing Statistical Analysis Plan text (at the time) (SAP text v2.0, 19 Oct 2020). The subsequent refinement to the definition of the "contamination period", mentioned in this version of the Statistical Analysis Plan text, and the associated plan for statistical analysis, is described below (in the "18 months" section) (in Section 7.6.4).

The update to the definition of the "contamination period" – will be applied to the 12-month second secondary efficacy endpoint analysis and the 18-month third secondary efficacy endpoint analysis. The effect of the refinement to the contamination rule on the analysis at the 6-month time point will be viewable within the 18-month data cut analysis, because the 6-month time point is one of the time points to be displayed in the main analysis table for the 18-month data cut analysis (wherein the main time point for analysis is the 18-month time point).

Given that the 6-month time point is a time point that is contained also within the 18-month data-cut analysis, any sensitivity analyses to be reported in the CSR to support the secondary factor-IX-activity endpoint can be done within the framework of the 18-month-data-cut analysis.

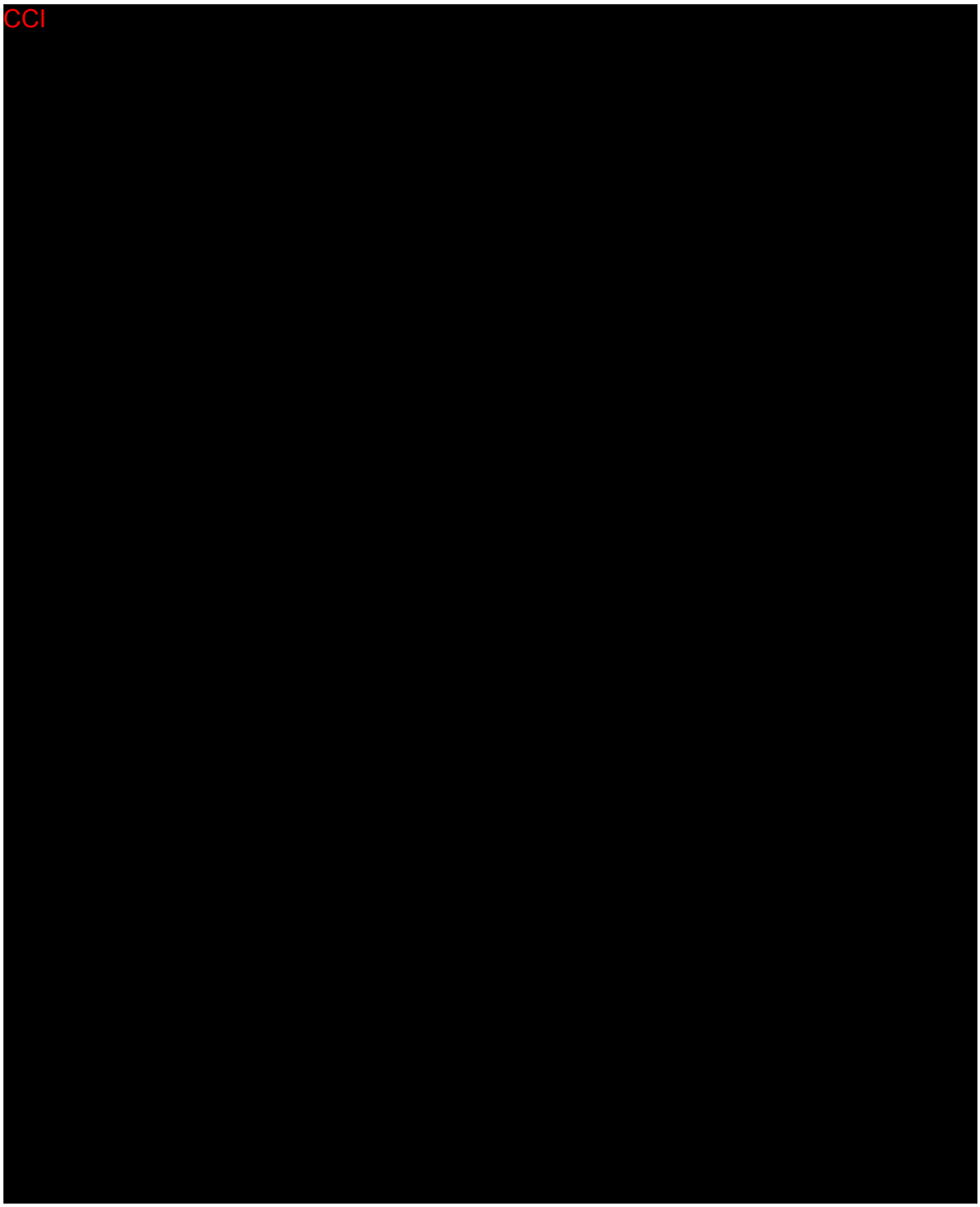
7.6.3 Endogenous Factor IX Activity at 12 months After AMT-061 Dosing using the "Refined" Contamination Rule

The main analysis of the second secondary endpoint "Endogenous Central Laboratory One-Stage aPTT Factor IX activity at 12 months after AMT-061 dosing" will take place using the 12-month data cut and will be done according to the methodology described below (in Section 7.6.4) for the third secondary efficacy endpoint.

The update to the definition of the "contamination period" (described in Section 7.6.4 for the 18-month third secondary efficacy endpoint analysis) will be applied also to the 12-month second secondary efficacy endpoint analysis.

Given that the 12-month time point is a time point that is contained also within the 18-month data-cut analysis, any sensitivity analyses to be reported in the CSR to support the secondary factor-IX-activity endpoint can be done within the framework of the 18-month-data-cut analysis.

CCI


CCI

7.6.4.1 Sensitivity Analysis 1: PP Population

A sensitivity analysis using the PP population on the secondary factor IX activity efficacy analysis will be performed.

7.6.4.2 Sensitivity Analysis 2: To Account for Missing Data

A second sensitivity analysis will be conducted for factor IX activity 18 months after a single AMT-061 treatment to evaluate the robustness of the main analysis findings to missing data. For this analysis, any missing values (that are still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits) will be imputed using the most recent previous post-treatment uncontaminated factor IX activity value that was observed for the subject.

7.6.4.3 Sensitivity Analysis 3: Cumulative Responder Analysis

A third sensitivity analysis, a (single-treatment) cumulative responder analysis, will be conducted for the change from baseline in uncontaminated central-laboratory factor IX activity 18 months after a single AMT-061 treatment. A cumulative distribution plot will also be produced. The full analysis set will be used. The observed change from baseline in factor IX activity at 18 months will be plotted on the *x*-axis and the proportion of responders (subjects that have an equal or greater level of change) will be plotted on the *y*-axis. If a subject still lacks the 18-month value (being still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value that is closest in time to (either before or after) 18 months post-AMT will be employed in the analysis. If two such assessments are both the closest in time, with one being before and the other being after, the earlier assessment will be used. If a subject has zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that is to be used in the analysis) will be imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity.

7.6.4.4 Sensitivity Analysis 4: Excluding Subjects with No Uncontaminated Central-Laboratory Values

The secondary factor IX activity efficacy analysis will be repeated (on the FAS) without the imputation of post-baseline values based on historical hemophilia B severity for subjects with zero uncontaminated central-laboratory values after administration of AMT-061.

7.6.4.5 Sensitivity Analysis 5: Excluding Visits Contaminated by Corticosteroid Exposure

The secondary factor IX activity efficacy analysis will be repeated (on the FAS) to evaluate the robustness of the findings to the impact of systemic corticosteroid therapy on factor IX activity levels. Visits post-AMT-061 that are in proximity to corticosteroid exposure will be characterized as contaminated and will be excluded from this analysis. The duration of contamination associated with corticosteroid usage will be dependent on the specific product (and its associated half-life), its route of administration, and the dosing frequency. A listing detailing (1) Medication Name, (2) Route of Administration, (3) Dosing Frequency, (4) Half-life, and (5) the Contamination Period (days) (subsequent to the medication stop date) for systemic corticosteroid use during either the lead-in or post-treatment period will accompany this analysis. An additional listing detailing subject-specific factor IX activity levels relative to instances of corticosteroid exposure will be provided.

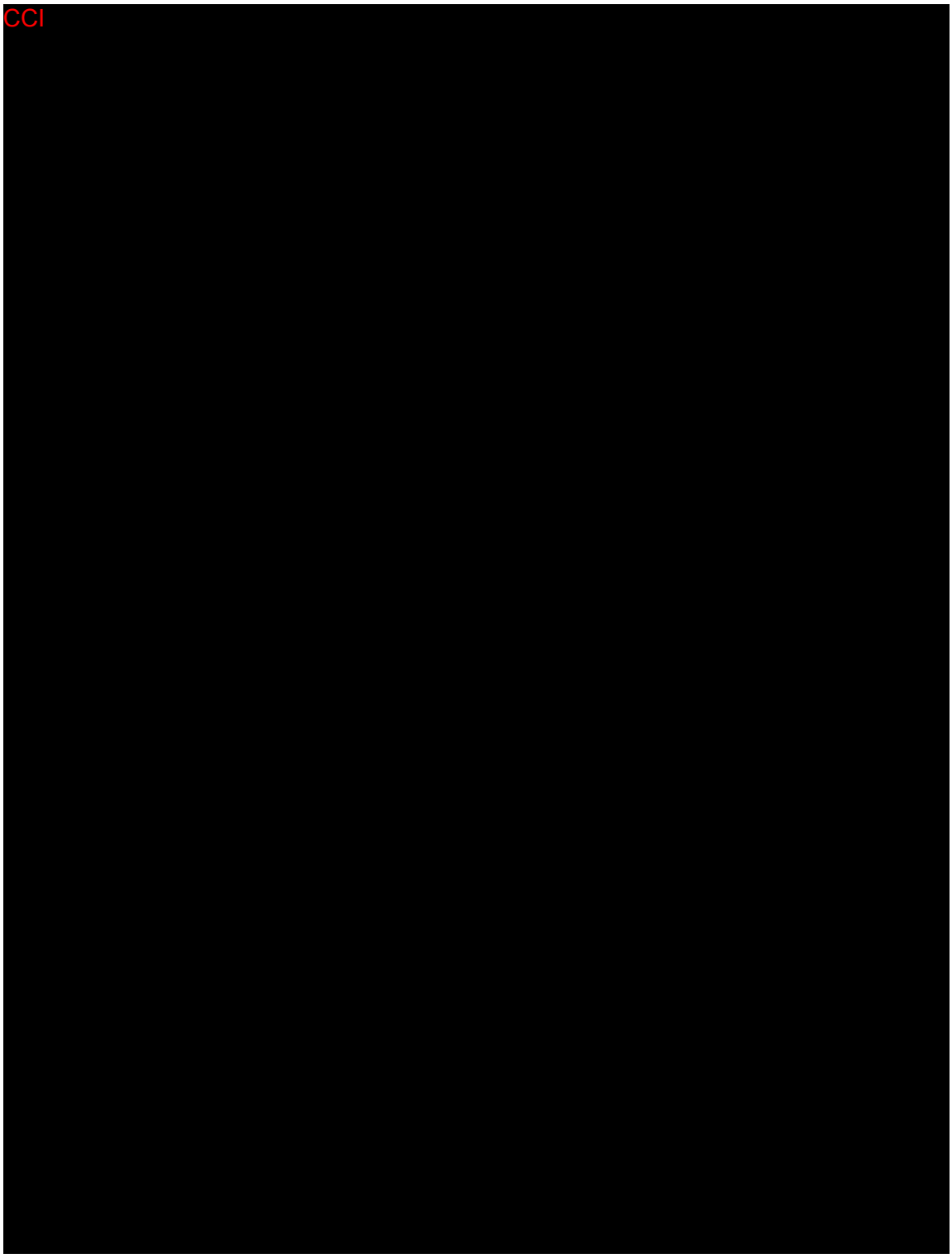
7.6.4.6 Sensitivity Analysis 6: Excluding Visits Contaminated by Exogenous Factor IX Using a 10-Day (Date/Time-Based) Contamination Rule

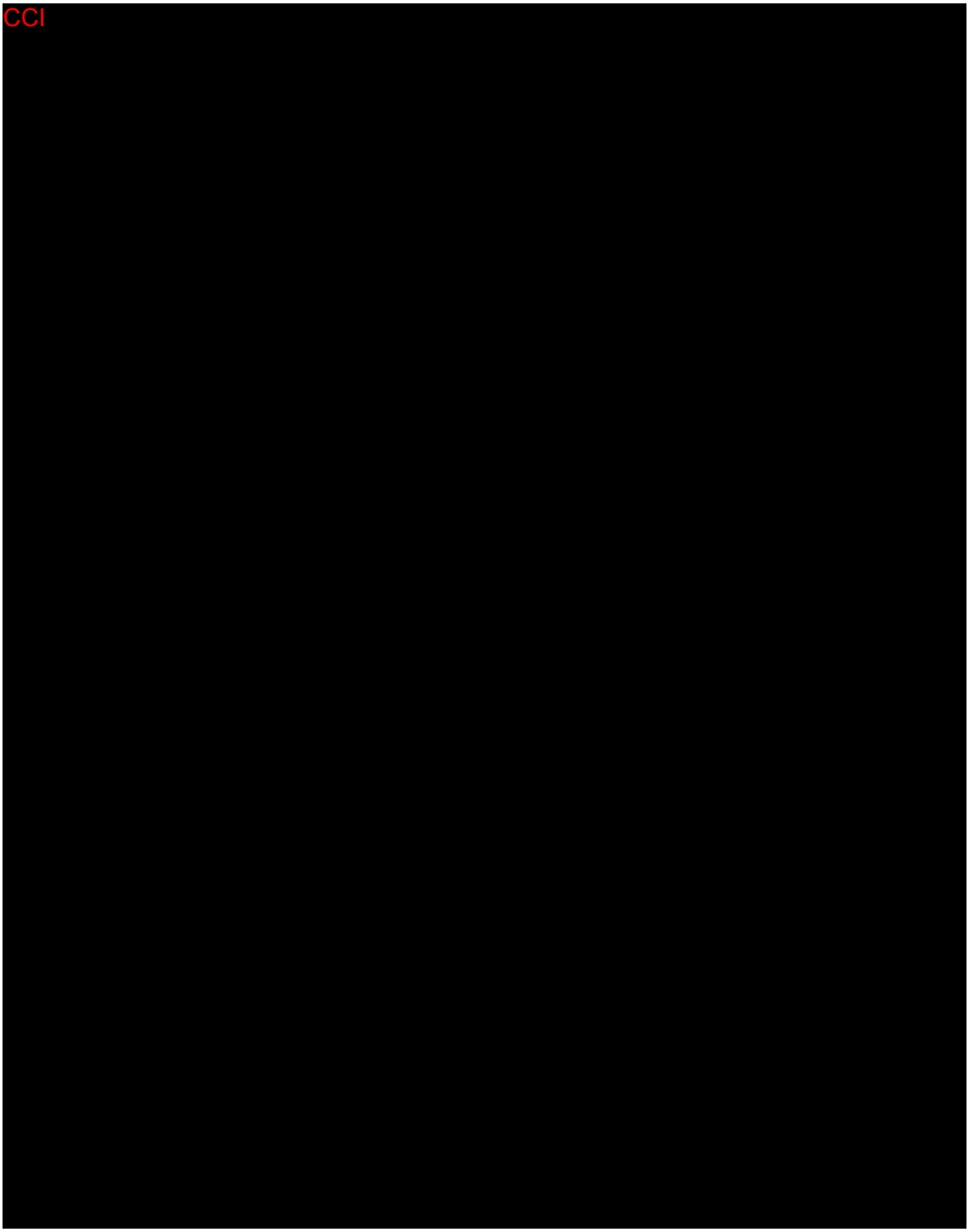
The secondary factor IX activity efficacy analysis will be repeated (on the FAS) to evaluate the robustness of the findings to the impact of using an alternative rule for determining contamination due to infused exogenous factor IX use. Visits post-AMT-061 that are within 10 days (240 hours) of exogenous factor IX use are considered contaminated and will be excluded from this analysis. Both the date and time of the exogenous factor IX injection start and the blood sampling for factor IX activity assessment will be taken into account to determine whether there was contamination. A blood sample drawn at a date/time prior to a given exogenous factor IX infusion start time is not considered to be contaminated by that infusion. See further detail in the Data Handling Rules Appendix (Appendix 1 Data Handling Rules) in this document [under the category of “Contamination due to exogenous factor IX (infusion) use”]. This sensitivity analysis will not use the 5 half-life contamination rule. Note that this sensitivity-analysis 10-day contamination rule – based on date-and-time – is distinct from the 10-discrete-calendar-day contamination rule that was used at the 26-week-data-cut interim efficacy analysis.

CCI

CCI

[redacted]


CCI


CCI

7.6.8.1 Sensitivity Analysis 1: Cumulative Responder Analysis Comparing the Lead-In Period with the 52 weeks following stable FIX expression (6-18 months) after Treatment with AMT-061

A sensitivity analysis, a (comparative between-treatment period) cumulative responder analysis (as described in Farrar 2006), will be conducted for the factor IX activity between the lead-in period and the mean across the scheduled visits between the six-month and twelve-month time points (i.e. Month 6, 7, 8, 9, 10, 11, and 12) in the post-treatment period. The full analysis set will be used. Only central-laboratory values will be used. For the post-treatment period, only uncontaminated values will be used. The uncontaminated factor IX activity mean across visits from Month 6 to Month 18 (to Month 12 for the 12-month data cut) for the post-treatment period will be plotted on the *x*-axis, and the proportion of "responders" (subjects that have an equal or greater level) will be plotted on the *y*-axis. The same graph will plot the mean of the observed factor IX activity across the Month 2, Month 4, and Month 6 visits for the lead-in period (also) on the *x*-axis, and the proportion of "responders" (subjects that have an equal or greater level) will be plotted on the *y*-axis. Thus, a cumulative distribution plot by treatment period will be produced. The observed factor IX activity is plotted on the *x*-axis and the proportion of

"responders" (subjects that have an equal or greater level) for each treatment period is plotted on the *y*-axis. If a subject has zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that is to be used in the analysis) will be imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity. The cumulative responder curves for each treatment period will then be compared using a two-sample Kolmogorov-Smirnov test.

If a subject still lacks having at least one uncontaminated factor IX activity value at least one of the following time points – post-treatment month 6, 7, 8, 9, 10, 11, 12, and 18 (6, 7, 8, 9, 10, 11, and 12 for the 12-month data cut) – value (that is still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the single (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value that is closest in time to (either before or after) 18 months (12 months for the 12-month data cut) will be employed in the analysis. If two such assessments are both the closest in time, with one being before and the other being after, the earlier assessment will be used.

7.6.8.2 Sensitivity Analysis 2: Excluding Contaminated Values Using a 10-day Contamination Rule

The analysis of "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 over 52 weeks following stable FIX expression (6-18 months)" will be repeated (on the FAS) to evaluate the robustness of the findings to the impact of using an alternative rule for determining contamination due to infused exogenous factor IX use. For this analysis, visits post-AMT-061 that are within 10 days of exogenous factor IX use are considered contaminated. Both the date and time of the exogenous factor IX injection start and the blood sampling for factor IX activity assessment will be taken into account to determine whether there was contamination. See further detail in the Data Handling Rules Appendix (Appendix 1 Data Handling Rules) in this document [under the category of "Contamination due to exogenous factor IX (infusion) use"]. Data from post-AMT-061 visits that are within 10 days of exogenous factor IX use will be excluded from the analysis. This sensitivity analysis will not use the 5 half-life contamination rule.

7.6.9 ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in phase and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up

ABR will be determined for the lead-in period and post-treatment period (52 weeks following stable FIX expression [6-18 months] [up to Year 1 post-treatment for the 12-month data cut]). Analysis of the number of reported bleeding events will be performed using a repeated measures GEE negative binomial regression model accounting for the paired design of the trial with an offset parameter to account for the differential collection periods. An unstructured covariance matrix will be employed. If the model fails to converge, then a compound symmetry covariance structure will be used. If convergence is still not attained, then initial parameter estimates will be provided. The model will include treatment (i.e. period) as a categorical variable.

The estimated rate ratio (between post-AMT-061 and the lead-in) will be tested using the following hypotheses:

H_o: rate ratio (post-treatment)/(lead in) = 1 (no effect of treatment)
H_1: rate ratio (post-treatment)/(lead in) < 1.

The hypothesis that *(post-treatment)/(lead in) = 1* (i.e. no difference between the two treatment periods) will be tested and a one-sided p-value <=0.025 will be regarded statistically significant. The one-sided p-value and two-sided 95% CI for the rate ratio will be obtained and presented in a table. The treatments will be compared for superiority. The main population will be the FAS. A sensitivity analysis will use the PP population.

Rules for computing time at risk for a bleeding event were already described along with the analysis methodology for the ABR endpoint (for the non-inferiority comparison). See Sections 7.5.2 and 4.3 for further details.

7.6.10 Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase

Similar to ABR, the number of spontaneous bleeding events and person-time at risk of (having) spontaneous bleeding events will be determined for the lead-in and post-treatment periods. Analysis of the annualized spontaneous bleeding rate will be performed using a repeated measures negative binomial regression model with the log of the time at risk of spontaneous bleeding (in the respective period) used as an offset parameter to account for the differential reporting periods. An unstructured covariance matrix will be employed; if that model fails to converge then a compound symmetry covariance structure will be used. If convergence is still not attained, initial parameter estimates will be provided. Treatment (i.e., period) will be a categorical covariate. To allow time for AMT-061 to become fully active and to allow the subject the opportunity to stop the lead-in prophylactic factor IX therapy, at stable FIX expression (Month 6) (at Day 21 for the 12-month data cut) (of the post-treatment period) will be used in the analysis.

The estimated rate ratio (between post-AMT-061 and the lead-in) will be tested using the following hypotheses:

H_o: rate ratio (post-treatment)/(lead in) = 1 (no effect of treatment)
H_1: rate ratio (post-treatment)/(lead in) < 1.

The hypothesis that *(post-treatment)/(lead in) = 1* (i.e. no difference between the two treatment periods) will be tested and a one-sided p-value <=0.025 will be regarded statistically significant. The one-sided p-value and two-sided 95% CI for the rate ratio will be obtained and presented in a table. The treatments will be compared for superiority. The main population will be the FAS. A sensitivity analysis will use the PP population.

Rules for computing time at risk for a bleeding event were already described along with the analysis methodology for the ABR endpoint (for the non-inferiority comparison).

The number of spontaneous bleeding events and the time-at-risk of spontaneous bleeding events will be listed by period for each subject.

7.6.11 Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase

Similar to ABR, the number of joint bleeding events and person-time at risk of joint bleeding events will be determined for the lead-in and post-treatment periods. Analysis of the reported number of joint bleeding events will be performed using a repeated measures negative binomial regression model with the log of the time at risk of joint bleeding (in the respective period) used as an offset parameter to account for the differential reporting periods. An unstructured covariance matrix will be employed; if that model fails to converge then a compound symmetry covariance structure will be used. If convergence is still not attained, initial parameter estimates will be provided. Treatment (i.e., period) will be a categorical covariate. To allow time for AMT-061 to become active and to allow the subject the opportunity to stop the lead-in prophylactic factor IX therapy, bleed counts beginning at stable FIX expression (Month 6) (at Day 21 for the 12-month data cut) (of the post-treatment period) will be used in the analysis.

The estimated rate ratio (between post-AMT-061 and the lead-in) will be tested using the following hypotheses:

H_o: rate ratio (post-treatment)/(lead in) = 1 (no effect of treatment)
H_1: rate ratio (post-treatment)/(lead in) < 1.

The hypothesis that *(post-treatment)/(lead in) = 1* (i.e. no difference between the two treatment periods) will be tested and a one-sided p-value <=0.025 will be regarded statistically significant. The one-sided p-value and two-sided 95% CI for the rate ratio will be obtained and presented in a table. The treatments will be compared for superiority. The main population will be the FAS. A sensitivity analysis will use the PP population.

Rules for computing time at risk for a bleeding event were already described along with the analysis methodology for the ABR endpoint (for the non-inferiority comparison).

7.6.12 Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a “correlation” analysis)

To examine the relationship (i.e. “correlation”) between ABR and baseline anti-AAV5 neutralizing antibodies (NAB), the following analysis will be carried out. A nonparametric, generalized additive model (GAM) will be implemented to graph the relationship of ABR to the natural logarithm of baseline anti-AAV5 neutralizing antibodies with a negative binomial model. Values of “< LOD” will be set to LOD/2 for the purpose of this analysis. Information for each subject across the 52 weeks following stable FIX expression (6-18 months) post-treatment will be employed. This endpoint will not have hypothesis testing and therefore is not included in the type I error control.

7.6.13 Correlation of factor IX activity levels during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay

The arithmetic mean of uncontaminated central laboratory one-stage aPTT factor IX activity level across all visits during the post-treatment period from the stable FIX expression (Month 6 Visit) through the 52 weeks following stable FIX expression (Month 18 Visit) (through the Month 12 Visit for the 12-month data cut) will be computed for each subject. The Pearson and Spearman correlation between this mean factor IX activity and the pre-IMP anti-AAV5 antibody titers (using the luciferase based NAB assay) will be tabulated as well as their 95% confidence intervals. The Pearson product-moment correlation coefficient (rp) will provide a measure of the strength of a linear association between factor IX activity levels and pre-IMP Anti-AAV5 antibody titers and the Spearman correlation coefficient (rs) will provide a measure of the strength of a monotone association between factor IX activity levels and pre-IMP Anti-AAV5 antibody titers. A scatter plot will be produced, with an overlaid linear regression line. This endpoint will not have hypothesis testing and therefore is not included in the type I error control.

If a subject has zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that is to be used in the analysis) will be imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity.

If a subject still lacks the Month 18 (Month 12 for the 12-month data cut) value (being still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value that is closest in time to (either before or after) 18 months (12 months for the 12-month data cut) post-AMT will be employed in the analysis. If two such assessments are both the closest in time, with one being before and the other being after, the earlier assessment will be used.

The main population will be the FAS. A sensitivity analysis will use the PP population.

7.6.14 Occurrence of (and resolution of) new target joints during the 6-18 month post-treatment follow-up

The rate of occurrence of new target joints per person-time of follow-up between stable FIX expression (post-treatment Month 6) (Day 21 for the 12-month data cut) and the time that is 52 weeks following stable FIX expression (18 months) (12 months for the 12-month data cut) after the subject's dosing (with AMT-061) will be summarized descriptively across subjects. Target joints that are being counted are ones that did not exist prior to stable FIX expression (post-treatment Month 6) (prior to Day 21 for the 12-month data cut). The percentage resolution of such new target joints will also be tabulated.

This endpoint will not have hypothesis testing and therefore is not included in the type I error control.

The main population will be the FAS. A sensitivity analysis will use the PP population.

7.6.15 Time to resolution of pre-existing target joints during the post-treatment follow-up

The time to resolution of pre-existing target joints (existing immediately prior to AMT dosing) will be summarized using a Kaplan-Meier curve. Time to resolution will be presented using the date of AMT dosing as the reference date. Each target joint will be handled as the experimental unit for this analysis, irrespective of subject. This analysis is censored at the subject's data cut-off date for the data cut.

This endpoint will not have hypothesis testing and therefore is not included in the type I error control.

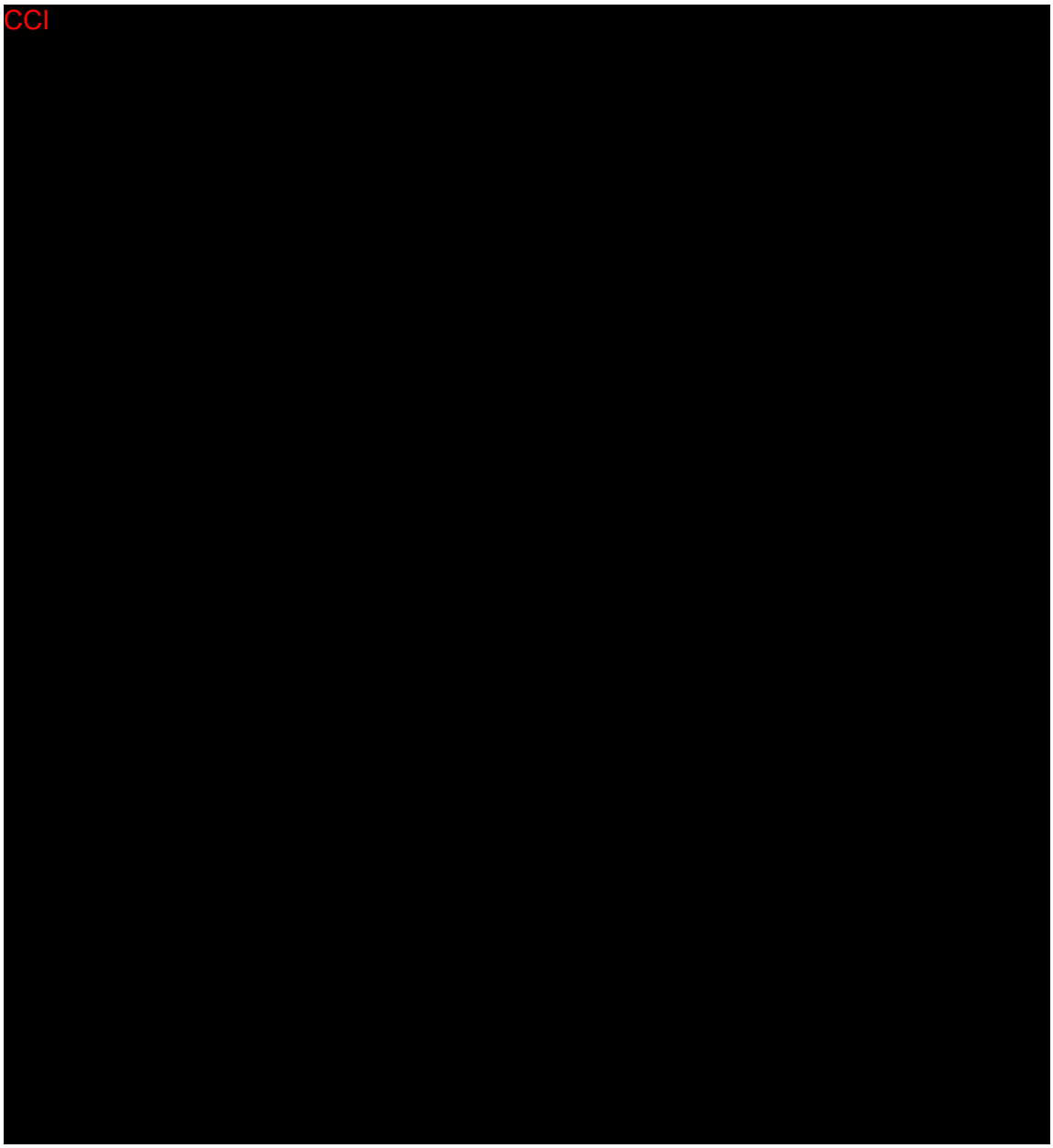
The main population will be the FAS. A sensitivity analysis will use the PP population.

7.6.16 Proportion of subjects with zero bleeds in the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up

The number and percentage of subjects with zero bleeds during the post-treatment period from stable FIX expression (Month 6) (from Day 21 for the 12-month data cut) to the time that is 52 weeks following stable FIX expression (18 months) (that is 12 months for the 12 month data cut) after the subject's dosing with AMT-061 will be tabulated with descriptive statistics and presented in a table. This endpoint will not have hypothesis testing and therefore is not included in the type I error control.

CCI

CCI



7.7 Type I Error Control and Simultaneous Confidence Intervals

Formal statistical testing of the efficacy endpoints will be performed using the closed testing principle (for Type I error control for multiple testing). Due to the closed testing principle, no correction for multiplicity is necessary. Among the endpoints being formally tested for statistical significance, all will be tested for superiority at a one-sided alpha level of 0.025 (except as otherwise noted). Superiority testing and non-inferiority testing will be accomplished using the FAS population.

Fixed sequential testing will be performed using a hierarchical approach and will be continued until a non-significant result is obtained (except as otherwise noted). The order of fixed sequential tests is specified below:

1. ABR comparison between AMT-061 and prophylaxis for non-inferiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up (primary efficacy endpoint)
2. Endogenous factor IX activity at 6 months after AMT-061 dosing (first secondary efficacy endpoint)
3. Endogenous factor IX activity at 12 months after AMT-061 dosing (second secondary efficacy endpoint)
4. Endogenous factor IX activity at 18 months after AMT-061 dosing (third secondary efficacy endpoint)
5. Annualized consumption of factor IX replacement therapy during the week52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase (secondary efficacy endpoint)
6. Annualized infusion rate of factor IX replacement therapy during the week52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase (secondary efficacy endpoint)
7. Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 52 weeks following stable FIX expression (6-18 months) (secondary efficacy endpoint)
8. ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up (secondary efficacy endpoint)
9. Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to lead-in phase (secondary efficacy endpoint)
10. Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase (secondary efficacy endpoint)
11. CCI [redacted] during the 12 months following AMT 061 dosing compared with the lead-in phase (secondary efficacy endpoint)
12. CCI [redacted] during the 12 months following AMT 061 dosing compared with the lead-in phase (secondary efficacy endpoint).

Simultaneous one-sided 97.5% CIs based on a graphical approach to multiple testing (Bretz et al. 2015; Guilbaud 2008; Strassburger and Bretz 2008) will be provided for the type I error controlled efficacy endpoints as a supportive analysis. For endpoints for which an increase is

favorable, the lower one-sided 97.5% confidence bound will be provided; for endpoints for which an increase is unfavorable, the upper one-sided 97.5% confidence bound will be provided.

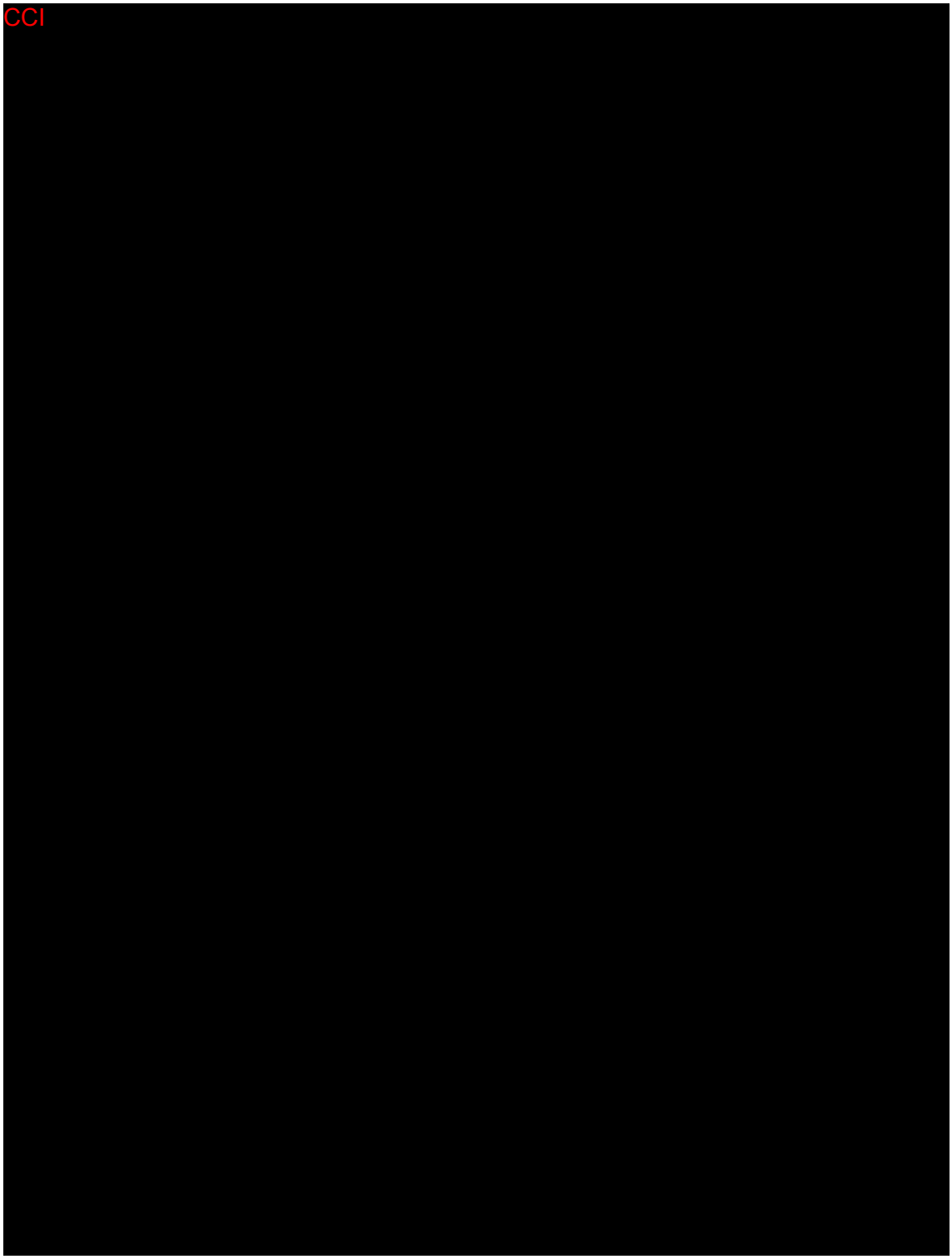
For any data cuts (i.e. analysis times) that are not the main data cut for a given endpoint, the p-values and CIs will be considered to be descriptive rather than inferential.

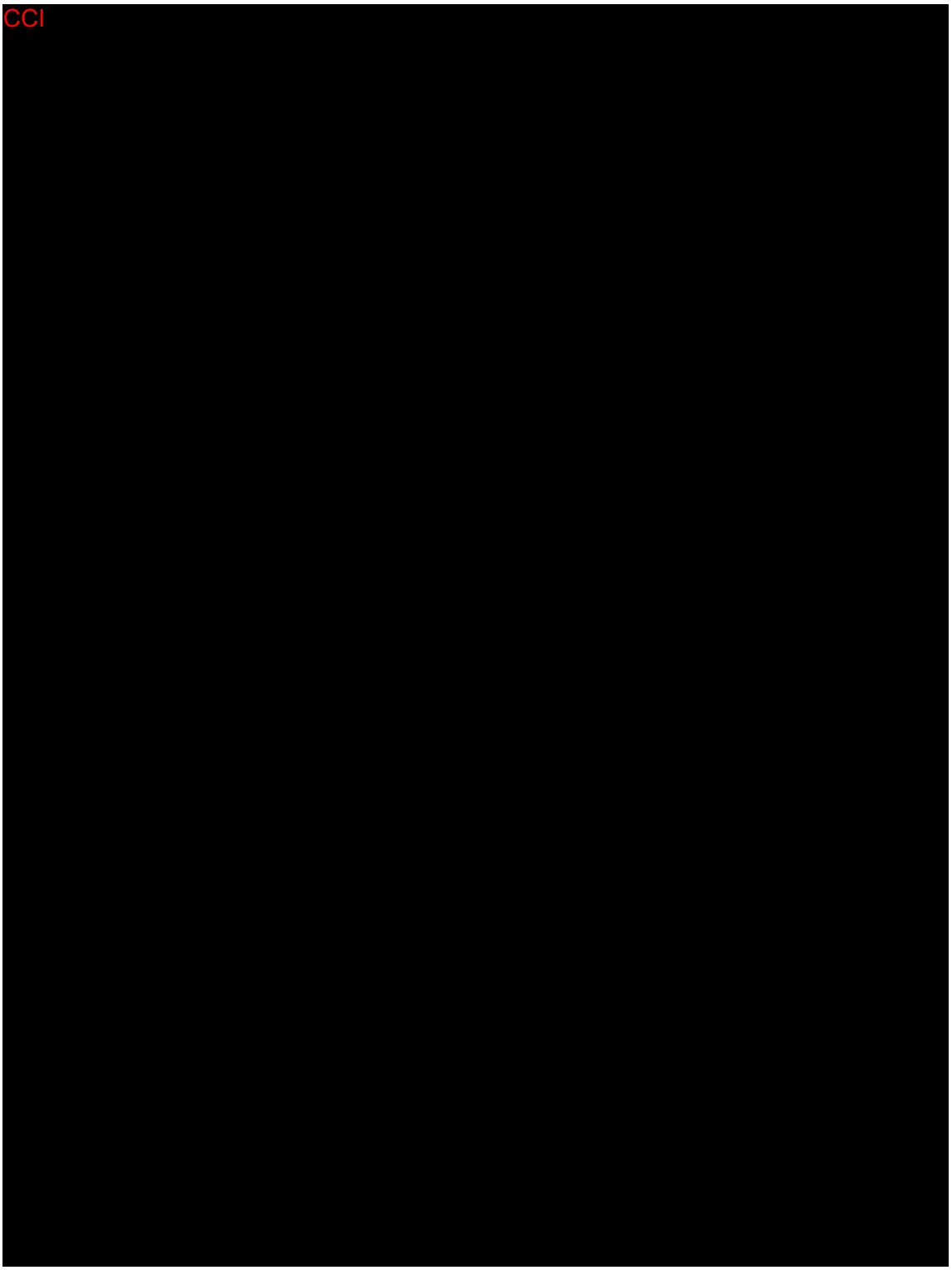
CCI

[redacted]

7.8.2 Factor IX protein levels during the 52 weeks following stable FIX expression (6-18 months) following AMT-061 dosing

Factor IX protein levels will be summarized descriptively by visit

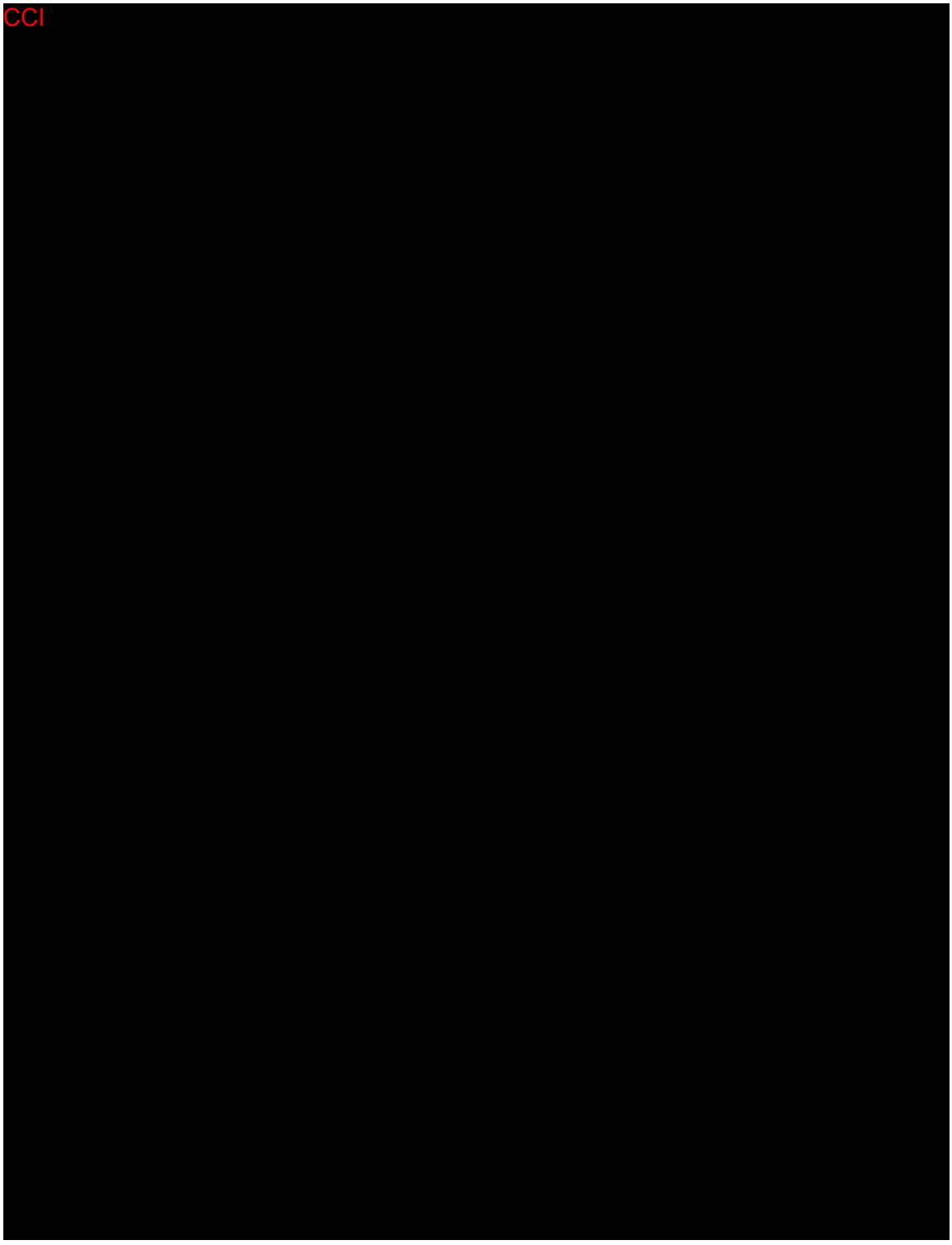

CCI


CCI

7.8.4 Impacted Responders Analysis: Correlation of factor IX activity levels at Month 18 with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay

An impacted-response curve will be developed – as an exploratory analysis – to examine the association between these two variables (factor IX activity levels at Month 18 and pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay). An impacted response (for the purpose of NAB effect determination) is defined as a subject's having an uncontaminated one-stage aPTT assay for factor IX activity (%) to be < 5% of normal at Month 18 (Month 12 for the 12-month data cut). The percentage of subjects having impacted response for the group of subjects with NAB titer >= x will be plotted as a function of x. The number of subjects with NAB titer >=x will also be indicated as a function of x on the graph. If, based on this graph, there exists a value "x" of NAB titer above which > 25% of subjects have impacted response for a group (and if 12 or more subjects have NAB titer above that value "x"), then that NAB titer value "x" is considered to be a potential candidate for being a meaningful NAB cutoff; otherwise, no such candidate NAB cutoff titer will have been identified. If a subject has zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that is to be used in the analysis will be imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity. If a subject lacks a month 18 (month 12 for the 12-month data cut) value (that is still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the uncontaminated central-laboratory post-AMT value that is closest in time to 18 months (12 months for the 12-month data cut) – either before or after – will be employed in the analysis. If two such assessments are both the closest in time, with one being before and the other being after, the earlier assessment will be used.

Specific levels of pre-treatment (i.e. baseline) NAB titer may be identified and used as the basis for subgroups in evaluating their potential relationship with factor IX activity levels. The pre-treatment NAB titer taken pre-dose on the day of dosing will be used. If this result is not available, the value closest in time pre-day of dosing (while being prior to the dose) will be used.


CCI

CCI

CCI

7.8.7 Subgroup Analyses

- Subgroup analyses will be carried out for the following endpoints (the subgroups are mentioned a bit farther below in this SAP):
    - Endogenous factor IX activity at Month 18

- Annualized consumption of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding replacement for invasive procedures, compared to the lead-in phase

- Annualized infusion rate of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding replacement for invasive procedures, compared to the lead-in phase

- ABR comparison between AMT-061 (during the 52 weeks following stable FIX expression [6-18 months] post-treatment follow-up) and factor IX prophylaxis (during the lead-in period)

- Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 over the 52 weeks following stable FIX expression (6-18 months)

- Proportion of subjects remaining free of previous prescribed continuous routine prophylaxis during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up.

The subgroup analyses for the aforementioned endpoints will be carried out for the following subgroups:

- Age categories: <40 years, 40 to <60 years, >= 60 years
- Race and/or Ethnicity subgroups (with categories to be specified later because the racial/ethnic frequencies are not well known in advance)
- Zero bleeds versus >=1 bleed in lead-in
    - Because this subgrouping is defined using information from the lead-in period, the analysis will provide descriptive statistics only and will provide those descriptive statistics for only the post-treatment period.
- Presence or absence of target joints at Screening
- Baseline NAB titer categories: positive titer (>= LOD) versus negative titer (<LOD), where LOD denotes limit of detection.
- HIV-negative vs. controlled HIV positive (CD4+ count >200 /μL) at Baseline
- History of Hepatitis B or C at Baseline
- Baseline liver pathology, according to Baseline FibroScan™ or equivalent SWE (shear wave elastography), MRE (magnetic resonance elastography) result:

- Degree of fibrosis [≥9Kpa versus <9Kpa]
- Degree of steatosis [Controlled Attenuation Parameter (CAP) score ≥S2 (≥ 260 dB/m) versus <S2 (<260 dB/m)] versus Missing.

All efficacy endpoints will be analyzed using data over the 2 to 5 year follow-up. Please see Section 10 for additional details.

7.8.8 Optional Sub-Study Efficacy and CCI Endpoint Analyses

Optional sub-study endpoints consist of:

- PROBE questionnaire sub-study summary scores
- Musculoskeletal ultrasound sub-study ultrasound results.

Analysis will be based on the FAS population for the set of subjects participating in the respective sub-study. Any subject with at least one assessment of the sub-study endpoint will be considered to be participating in the respective sub-study.

7.8.8.1 PROBE Questionnaire Summary scores

PROBE Questionnaire summary scores and individual item responses will be summarized descriptively by treatment and visit. The following statistics will be displayed: n, mean (SE), SD, Q1, median, Q3, minimum, and maximum. Summary scores and individual item responses also will be listed. A higher score indicates better health.

7.8.8.2 Musculoskeletal Ultrasound results

There will be a separate SAP document for the statistical analysis of the Musculoskeletal Ultrasound results.

7.9 Safety Analyses

All safety analyses will be based on the safety population.

The safety endpoints to be analyzed are:

- TEAEs
- Changes in abdominal ultrasound
- Anti-AAV5 antibodies (total [IgM and IgG], neutralizing antibodies)
- AAV5 capsid-specific T cells
- Anti-FIX antibodies

- Factor IX inhibitors and recovery
- Hematology and serum chemistry parameters
- ALT and AST levels and corticosteroid use for ALT and AST increases
- Vector DNA in blood and semen
- Inflammatory markers: IL-1β, IL-2, IL-6, IFNγ, MCP-1
- AFP.

7.9.1 Adverse Events

An adverse event is considered to be treatment-emergent for the AMT-061 treatment (i.e. a TEAE) if the event occurs after the administration of the IMP, or if the AE worsened during the study after the dose of study drug (intensity and/or severity changed to a worsened grade). An adverse event that begins on the same date as the IMP administration is treatment-emergent if the AE begins after the time of dose or if the time of AE onset is unknown. Additionally, if an AE has an onset date during post-treatment period and has an outcome of death, that death will be considered to be treatment-emergent. Furthermore, if the AE could possibly be treatment-emergent, based on the missing or incomplete date, then the AE will be regarded as treatment-emergent. A treatment-emergent adverse event can be described as having had "incidence" during the treatment period.

An adverse event will be counted as having had "incidence" during the lead-in period if it occurs during the lead-in period, or if the AE worsened during the lead-in period (intensity and/or severity changed to a worsened grade). Additionally, if an AE has an onset date during the lead-in period and has an outcome of death, that death will be counted as having incidence during the lead-in period. Furthermore, if the AE could have had incidence during the lead-in period, based on the missing or incomplete date, then the AE will be regarded as having incidence during the lead-in period.

An adverse event incidence table for the safety populations will be created displaying the number of subjects (and percentage) experiencing an incident event and the number of incident events for: any AEs, AEs of special notification, serious AEs, related AEs, serious and related AEs, AEs leading to early treatment discontinuation (i.e. to a partial dose), mild/moderate/severe AEs, and deaths.

The following AE incidence summary tables will be presented by decreasing frequency of occurrence based on SOC and Preferred Term:

1. AEs for the lead-in and post-treatment safety populations
2. Serious AEs for the lead-in and post-treatment safety populations

3. Related TEAEs for the post-treatment safety population
4. Related serious TEAEs for the post-treatment safety population
5. TEAEs leading to treatment discontinuation for the post-treatment safety population (treatment discontinuation means receiving only a partial dose)
6. Serious TEAEs leading to treatment discontinuation for the post-treatment safety population (treatment discontinuation means receiving only a partial dose)
7. Fatal AEs
8. TEAEs by highest severity for the post-treatment safety population
9. Related TEAEs by highest severity for the post-treatment safety population
10. Incidence of TEAEs for Special Notification for the post-treatment safety population
11. Incidence of non-serious TEAEs occurring in at least 5% of subjects in the post-treatment period
12. The incidence of TEAEs occurring in at least 10% of subjects in the post-treatment period.

All incident AEs will be tabulated by SOC and preferred terms within each SOC according to the Medical Dictionary for Regulatory Activities (MedDRA) terminology list. The version of the MedDRA that is current at the time of database lock will be used to code verbatim terms for AEs for final analysis of the data. A glossary of MedDRA preferred terms used for adverse events reported in the study along with the associated Investigator's verbatim term will be provided. No hypothesis tests will be performed.

The summary tables will be accompanied by individual subject listings of all AEs, including pre-treatment AEs and information on actual AE description, date/time of start and end of AE, preferred term (MedDRA), SOC (MedDRA), severity, relationship/causality, type of AE, action taken, seriousness and outcome. Pre-existing AEs will be flagged. Pre-existing AEs are not considered to be treatment emergent, except in case of worsening during/after trial treatment (to be collected as a separate AE). Separate listings will be created for AEs for special notification, deaths, and SAEs. All adverse events, whether treatment-emergent or not, will be included in the listings. A listing of any reported deaths during lead-in and post-treatment periods will be provided and will include the number of days since IMP administration.

The following will be done for events with irregular onset dates**.** All AEs will be included in the data listings regardless of the completeness of the onset dates. Any partial dates will be used in order to determine whether an AE is lead-in-incident or treatment-emergent using the rules in Appendix 1; however, imputed dates will not be provided in the data listings.

7.9.1.1 Adverse Events of Special Notification

Table 5 contains (S)AEs that qualify for special notification as they are seen as safety issues of particular concern for Advanced Therapy Medicinal Product (ATMP) (ENTR/F/2/SF/dn D (2009) 35810. Brussels, 03/12/2009) and gene therapy medicinal products (EMA/CHMP/GTWP/60436/2007):

Table 5: Adverse Events of Special Notification

AEs related to the IMP administration procedure
Suspected or confirmed cases of opportunistic or serious infections that in the investigator's opinion might be related to the IMP
Unexpected reactions (e.g., hypersensitivity, immunological, toxic or other as consequence of a change in the construction or function of the viral vector [e.g., generation of replication competent virus])
AEs related to product failure (including lack of efficacy)
AEs related to mandatory concomitant medication (e.g., immunosuppression)
AEs related to medical devices which form part of the product or are used for application of the product
Development of any new/recurrent cancer.

These AEs should be reported and followed in the same manner as SAEs. Note that the AEs may be serious or non-serious by definition (please see the protocol for more details). AEs of special notification are designated as such on the eCRF and therefore do not need to be derived.

7.9.1.2 Severity of Adverse Event

If an AE changes severity over time, the severity of maximum severity (i.e. intensity) will be reported.

7.9.1.3 Relationship Between IMP and Adverse Event

Please refer to the protocol for the definitions of related to IMP, probably related to IMP, possibly related to IMP, and not related to IMP.

7.9.2 Changes in Abdominal Ultrasound

To monitor subjects for liver fibrosis and potential occurrences of liver malignancies, abdominal ultrasounds will be performed. These ultrasounds will occur at the final Lead-In visit at the latest (to establish baseline status), at post-treatment Month 12, and then annually thereafter.

A shift table will be used to summarize normal and abnormal results at Month 12 and the subsequent follow-up visits relative to the results obtained at baseline. Baseline abdominal ultrasound is the most recent assessment prior to the dose of study medication.

All abdominal ultrasound data will be listed.

7.9.3 Anti-AAV5 Antibodies, Anti-Factor IX Antibodies, and Factor IX Inhibitors

Total IgG and IgM antibodies against the vector capsid is evaluated by the enzyme-linked immunosorbent assay (ELISA), anti-AAV5 neutralizing antibodies are assessed with the luciferase assay. Antibodies against FIX will be evaluated by ELISA and will be reported as IgG, IgM. The results from the total IgG and IgM antibodies against the vector capsid, neutralizing antibodies against the vector capsid and non-inhibitory factor IX antibodies will be tabulated by visit using descriptive statistics (for the titer) of n, mean (SE), SD, Q1, median, Q3, and max titer at each visit. The titer of factor IX inhibitors will be reported in Bethesda Units and the sub-class of immunoglobulin of the inhibitor will be displayed as IgG, IgM or others. These results will be displayed at each visit. All data will be listed.

A subject is said to suffer from factor IX inhibitors if the subject tested positive for factor IX inhibitors at two consecutive tests, performed preferably within two weeks. Occurrences of "suffering from factor IX inhibitors" will be flagged in the factor IX inhibitor listing.

Measurement of factor IX recovery (maximum concentration [C_{max}]) and incremental recovery measured as increase in activity per unit infused (IU/ml per U/kg) at 30 min after infusion of a dose of factor IX will be performed at baseline Visit L-Final. Additionally, measurement of factor IX recovery and incremental recovery should be done at suspicion of factor IX inhibitor as judged by the investigator.

All data will be listed.

7.9.4 AAV5 capsid-specific T cells

The AAV5 capsid-specific T cells testing (ELISPOT) will be summarized by visit using descriptive statistics of n, mean (SE), SD, Q1, median, Q3, minimum, and maximum.

The data will also be listed.

7.9.5 Clinical Laboratory Measurements

The lab parameters collected include the following:

Table 6: Safety Lab Parameters

Hematology	
Hemoglobin	White blood cells with differential count
Hematocrit	CD4+ count
Platelet count	
Red blood cells	
Serum Chemistry	
Sodium serum electrolytes	Alkaline phosphatase (ALP)
Potassium serum electrolytes	C-Reactive Protein
Creatinine	Albumin
Gamma-glutamyltransferase (GGT)	Total Bilirubin
AST	Glucose (non-fasting)
ALT	
Coagulation	
aPTT	
PT (or International Normalized Ratio [INR])	
Serology	
HIV 1/2 antibody differentiation	Hepatitis B extracellular antigen (HBeAG) *
HIV 1/2 screen	Hepatitis B virus (HBV) DNA
Hepatitis B surface antigen (HBsAG)	Hepatitis C virus (HCV) RNA
Alpha-fetoprotein	
AFP	
Local Laboratory	
AST	
ALT	

* This parameter was removed with Protocol Amendment 3.

A Clinically Significant Laboratory Abnormality as identified by the investigator after the study drug is administered will be recorded as an Adverse Event and tabulated as an AE in the AE analysis. Abnormalities occurring prior to the IMP administration will be noted in medical history and presented in a data listing.

All laboratory data will be stored in the database with the units in which they were originally reported. Laboratory data not reported in International System of Units (SI units; Système International d'Unités) will be converted to SI units before data analysis.

Individual clinical laboratory variables for hematology, serum chemistry, coagulation, serology, and local laboratory will be provided in listings. Comments for laboratory testing will be listed. For listings, laboratory values will be flagged as low or high based on the reference ranges provided by the central laboratory.

If there are multiple laboratory values for the same parameter for a visit, the last value will be chosen for analysis.

Summary statistics (n, mean, Q1, median, Q3, standard deviation, minimum, and maximum) for the baseline assessment and change from baseline at each post-baseline visit for scheduled lab assessments of continuous laboratory variables will be tabulated for post-treatment safety population. Data from unscheduled visits will not be used for the by-visit summaries (unless they have been assigned to a scheduled visit according to the Time Windows for Statistical Analysis). Data from both scheduled and unscheduled visits will be listed.

Shift tables will be produced using the categories defined by the Common Terminology Criteria for Adverse Events (CTCAE) grades for the post-treatment safety population for hematology and serum chemistry. For these shift tables, the subject's pre-IMP grade will be cross-tabulated by the subject's maximum post-treatment follow-up; also, the subject's maximum post-IMP grade during post-treatment follow-up will be tabulated for all baseline grades combined. Percentages of subjects in each maximum post-IMP grade will be calculated for each pre-dose grade for the treatment and also for all baseline grades combined. Laboratory abnormal values on-treatment will be flagged as High or Low values based on laboratory reference ranges provided by LabCorp Laboratories (found in Appendix 3). These flags along with the reference ranges will be provided in the laboratory data listings.

Potentially Clinically Significant Laboratory Values Above/Below a Clinically Relevant Threshold on-treatment, based on CTCAE and other criteria, will be identified based on the thresholds in the table below.

Table 7: Potentially Clinically Significant (PCS) Laboratory Parameter Criteria

Central Laboratory	Post-Baseline Criteria
Serum Chemistry	
Sodium serum electrolytes	NA
Potassium serum electrolytes	<3.0 mmol/L >6.0 mmol/L
Creatinine	>2 x ULN
Gamma-glutamyltransferase	NA
AST	>2 x Baseline
ALT	>2 x Baseline
ALP	>2 x ULN
CRP	NA
Albumin	NA
Total bilirubin	>2 x ULN
Glucose (non-fasting)	NA
Hematology	
Hemoglobin	<8.0 g/dL (<80 g/L) Increase of >40 g/L to a value above the ULN
Hematocrit	NA
Platelet count	<50 x 10^9/L >999 x 10^9/L
Red blood cells	NA
White blood cells with differential count	<2 x 10^9/L >35 x 10^9/L
CD4+ count	≤200/µL
Coagulation	
aPTT	NA
PT (or INR)	NA
Serology	
HIV viral load	>200 copies/mL
HBsAg	NA
HBeAG	NA
Hepatitis B Virus DNA (HBV DNA)	NA
Hepatitis C Virus RNA (HCV RNA)	NA
Alpha-fetoprotein	
AFP	NA
Local Laboratory	
AST	>3 x ULN
ALT	>3 x ULN

NA: Not Applicable

Clinically significant laboratory values will be tabulated for the lead-in safety population and the post-treatment safety population. For all laboratory data for the parameter identified as potentially clinically significant for a subject will be listed. Low platelet counts are counted as being clinically significant only if they occur >= 4 weeks after IMP administration.

On the listings, the reference range and flag indicating if the measurement in question is outside the reference range will be provided.

7.9.6 ALT Levels, AST Levels and Corticosteroid Use for ALT and AST Increases

Summary statistics (n, mean, Q1, median, Q3, standard deviation, minimum, and maximum) for the baseline assessment and change from baseline at each post-baseline visit for ALT levels, AST levels and corticosteroid use will be tabulated and listed. Data from unscheduled visits will not be used for the by-visit summaries (unless they have been assigned to a scheduled visit according to the Time Windows for Statistical Analysis). Data from both scheduled and unscheduled visits will be listed.

Plots of individual subject profiles of ALT and AST levels over time will also be displayed. Corticosteroid use will be indicated on the plots.

7.9.7 Vector Genome Detection

The number of days until vector DNA can no longer be detected in semen and blood will be tabulated. The number of days is calculated using the date of collection of the third consecutive negative sample for each matrix.

All data will be listed.

Time to first shedding negative will be defined for each type of matrix and each patient as the post-treatment time point where a negative result is measured for the first time in a consecutive order of 3 or more time points with a negative result. A negative result is defined as a result of either ‘0’ or ‘LOD’ (limit of detection). The time to first shedding negative will be flagged on the above-mentioned listings.

The time to first shedding negative for the post-treatment period will also be summarized using a Kaplan-Meier curve. The censoring time will be truncated at the data cut-off date, the time of completion of the study, or time of early withdrawal from the study, whichever is earlier. For the 5-year analysis (and 5-year CSR), there will be no data cut-off date.

7.9.8 Inflammatory Markers

Blood samples will be taken to assess IL-1β, IL-2, IL-6, IFNγ and MCP-1 (monocyte chemotactic protein-1) using ELISA.

The test results will be summarized by visit using descriptive statistics and will include a change from baseline calculation for each post-baseline measurement. All data will also be listed.

Inflammatory markers will not be summarized or presented following the interim six-month data cut as the full set of inflammatory marker data will not be available as part of the six-month database lock.

7.9.9 Alpha-fetoprotein

Alpha-fetoprotein results will be summarized by period and visit. They will also be listed.

7.9.10 Physical Examination (Including Height and Weight)

A physical examination will be performed at Screening (Visit S), L-Final, Visit D (pre-IMP), during the post-treatment follow-up at visits F1, F2, F4, F6, F12, F13, F15, F17, F19, and F-Final, and during the long-term follow-up at visits LTF1, LTF2, LTF3, LTF4, LTF6, and LTF8. Height will be measured only at screening and weight will be measured only at screening and Visit L-Final.

Height (without shoes) will be measured and recorded, rounded to the nearest centimeter. Body weight (without overcoat and shoes) will be measured and recorded, rounded to the nearest kilogram.

The physical examination will include general appearance and bedside examination of the following body systems: Lymph nodes, eyes and ears, mouth and throat, lungs, abdomen, extremities, musculoskeletal system, neurological system, cardiovascular system and skin.

Abnormal physical examination findings will be reported as adverse events.

Abnormalities (e.g., scar at the left side at knee following total knee replacement, or arthropathy of left ankle due to hemophilia B) identified at screening will be documented in the subject's source documents and on the medical history electronic case report form (eCRF). Changes after the Screening Visit will be captured as AEs on the AE eCRF page, as deemed clinically significant in the opinion of the investigator. These abnormalities are to be followed until they reached "final outcome" (please refer to the protocol).

7.9.11 Vital Signs

Blood pressure, pulse, and body temperature will be measured at Screening (Visit S), Visit L-Final, at pre-IMP and post-IMP (3 hours) on Visit D and at all visits during the post-treatment phase. Before measurement of blood pressure and pulse, the subject should rest for at least 5 minutes. For the individual subject, all measurements should be performed while the subject is in the same position (i.e., sitting or lying) throughout the trial.

A summary of baseline weight, height, and BMI will be presented by treatment period for the FAS, PP, and safety populations in the demographics table.

Vital signs values will be listed.

8. INTERIM 6 MONTH ANALYSIS

A partial database lock and data extraction will be performed once the last subject has achieved 6 months after AMT-061 therapy.

The first secondary efficacy endpoint, endogenous factor IX activity, will be analyzed. This endpoint/analysis will be included in (added to) the 18-month-data-cut CSR.

Factor IX activity will be summarized (and listed) by visit, overall and by patient, over the 6-month period since administration of AMT-061. By-subject plots of factor IX activity over time will be overlaid with plots of exogenous factor IX consumption (time of administration) and with the times of occurrence of bleeding events over the first 6 months subsequent to AMT-061 administration.

The ratio of factor IX activity (%) to factor IX protein (%) will be tabulated. A table will also be provided to summarize the factor IX activity (%) by patients with or without pre-existing neutralizing antibodies to factor IX. A scatter plot of factor IX activity (%) by baseline titer of neutralizing antibodies to AAV5 will also be presented to show the correlation of baseline titer of neutralizing antibodies to AAV5 and factor IX activity at Month 6.

Bleeding episodes will be tabulated by the following bleed types: all bleeds, spontaneous, traumatic, unknown, and medical/dental/other. The estimated ABR will be tabulated descriptively by treatment.

The factor IX replacement during the post-treatment period and the actual exogenous factor IX use will be tabulated overall and will be listed by subject.

The incidence of AEs, SAEs, AEs in descending frequency, related AEs, and related SAEs will be tabulated.

Additional figures showing ALT and AST levels (U/L) and corticosteroid use for ALT and AST elevations over time, T-cell (AAV5-capsid) ELISPOT ((SFC)/million PBMCs) over time will be produced. Subject disposition and demographic data will be listed for all subjects screened and all subjects treated. Baseline characteristics will be listed and will include the following: hemophilia B history, joint status at screening, bleeding history in the year prior to screening, history of previous factor IX replacement therapy use, prior medication/therapy, medical and surgical history, and factor IX gene sequencing.

Additional efficacy endpoints will be listed and include the following: factor IX protein concentration (%).

Safety listings will include the following: AEs, SAEs, AEs of special notification, PCS laboratory values, vital signs, other laboratory values, and vector shedding.

9. INTERIM 12 MONTH ANALYSIS

A partial database lock and data extraction will be performed once the last subject has achieved 12 months after AMT-061 therapy.

The second secondary efficacy endpoint, endogenous factor IX activity, will be analyzed. This endpoint/analysis will be included in (added to) the 18-month-data-cut CSR.

Factor IX activity will be summarized (and listed) by visit, overall and by patient, over the 12 month period since administration of AMT-061. By-subject plots of factor IX activity over time will be overlaid with plots of exogenous factor IX consumption (time of administration) and with the times of occurrence of bleeding events over the first 12 months subsequent to AMT-061 administration.

The ratio of factor IX activity (%) to factor IX protein (%) will be tabulated. A table will also be provided to summarize the factor IX activity (%) by patients with or without pre-existing neutralizing antibodies to factor IX. A scatter plot of factor IX activity (%) by baseline titer of neutralizing antibodies to AAV5 will also be presented to show the correlation of baseline titer of neutralizing antibodies to AAV5 and factor IX activity over months 6-12.

Bleeding episodes will be tabulated by the following bleed types: all bleeds, spontaneous, traumatic, unknown, and medical/dental/other. The estimated ABR will be tabulated descriptively by treatment.

The factor IX replacement during the post-treatment period and the actual exogenous factor IX use will be tabulated overall and will be listed by subject.

The incidence of AEs, SAEs, AEs in descending frequency, related AEs, and related SAEs will be tabulated.

Additional figures showing ALT and AST levels (U/L) and corticosteroid use for ALT and AST elevations over time, T-cell (AAV5-capsid) ELISPOT ((SFC)/million PBMCs) over time will be produced. Subject disposition and demographic data will be listed for all subjects screened and all subjects treated. Baseline characteristics will be listed and will include the following: hemophilia B history, joint status at screening, bleeding history in the year prior to screening, history of previous factor IX replacement therapy use, prior medication/therapy, medical and surgical history, and factor IX gene sequencing.

Additional efficacy endpoints will be listed and include the following: factor IX protein concentration (%).

Safety listings will include the following: AEs, SAEs, AEs of special notification, PCS laboratory values, vital signs, other laboratory values, and vector shedding.

10. CSR

After 52 weeks following stable FIX expression (18 months post-treatment), the database will be locked, and all available efficacy and safety data collected between screening and 52 weeks following stable FIX expression (18 months post-treatment follow-up time) will be analyzed and reported in a full CSR. All of the efficacy endpoints will be analyzed.

Data up to each analysis time point will be considered locked and will not be changed (with the exception of ending dates for continuing events and treatments).

Table, listing, and figure shells for the Final CSR will be provided in a separate document.

Factor IX activity will be summarized (and listed) by visit, overall and by patient, over the 18-month period since administration of AMT-061. By-subject plots of factor IX activity over time will be overlaid with plots of exogenous factor IX consumption (time of administration) and with the times of occurrence of bleeding events over the 52 weeks following stable FIX expression (6-18 months) months subsequent to AMT-061 administration.

Descriptive statistics will be provided for the estimated unadjusted ABR during time periods subsequent to stable FIX expression (6 months) after the dose of AMT-061 and during the Lead-in period. The unadjusted ABR is the number of bleeds divided by the person-time at risk during a given time period. For the 52 weeks following stable FIX expression (6-18 months) post-AMT, bleeds and person time on or after Day 1 and prior to stable FIX expression (Month 6) post-AMT will not be included in the calculation.

Bleeding events will be listed by subject. The number of bleeding events and the time-at-risk of bleeding events will be listed by period for each subject.

The ratio of factor IX activity (%) to factor IX protein (%) will be tabulated. A table will also be provided to summarize the factor IX activity (%) by patients with or without pre-existing neutralizing antibodies to factor IX. A scatter plot of factor IX activity (%) by baseline titer of neutralizing antibodies to AAV5 will also be presented to show the correlation of baseline titer of neutralizing antibodies to AAV5 and factor IX activity over the 52 weeks following stable FIX expression (6-18 months).

CCI

12. CHANGES FROM METHODS PLANNED IN THE PROTOCOL

- The full analysis set has become the primary population for the ABR non-inferiority analysis, while the per-protocol population has been relegated to a sensitivity analysis. The reason is that the FDA (Food and Drug Administration) statistical team requested this.

- The contamination period due to exposure to exogenous factor IX has been changed to the 5 half-life rule (from a 10-day rule) to allow greater accuracy.

- ABR has now become the sole primary endpoint. The reason is that the FDA (Food and Drug Administration) statistical and clinical teams requested this.

- Have changed the data cut for the main CSR to be at 18 months post-treatment. The reason is that the FDA asked for the efficacy analysis to pertain to the year after a stable factor IX activity level is reached.

13. STATISTICAL SOFTWARE

Data processing, statistical screening, descriptive reporting and analysis of the efficacy and safety data will be performed using SAS (Version 9.4 or higher).

14. REFERENCES

Bretz F, Maurer W, and Maca J “Graphical Approaches to Multiple Testing.” Clinical Trial Biostatistics and Biopharmaceutical Applications, edited by Young W and Chen D, CRC Press, 2015, 349-394.

EuroQol Group. EQ-5D. Accessed at www.euroqol.org/about-eq-5d/how-to-use-eq-5d.html (November 2014).

Farrar JT, Dworkin RH, Max MB. Use of the cumulative proportion of responders analysis graph to present pain data over a range of cut-off points: Making clinical trial data more understandable. J Pain Symptom Manage 2006; 31:369-377.

Guilbaud O. Simultaneous confidence regions corresponding to Holm's stepdown procedure and other closed testing procedures. Biometrical Journal, 2008; 50:678-692.

Paul J, Venugopal R, et al. Measurement of controlled attenuation parameter: A surrogate marker of hepatic steatosis in patients of nonalcoholic fatty liver disease on lifestyle modification – A prospective follow-up study. Arq. Gastroenterol. 55(1): São Paulo Mar./Apr. 2018.

Pickard S, Law E, Jiang R, Pullenayegum E, et al. United States Valuation of EQ-5D-5L Health States Using an International Protocol. Value in Health, 2019; 22(8):931–941.

SAS Institute Inc. 2008. SAS/STAT 9.2 User's Guide. Cary, NC: SAS Institute Inc.

Sasso M, Beaugrand M, de Ledinghen V, et al. Controlled attenuation parameter (CAP): a novel VCTE(TM) guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: Preliminary study and validation in a cohort of patients with chronic liver disease from various causes. Ultrasound Med Biol. 2010; 36:1825-35.

Simioni P, Tormene D, Tognin G, Gavasso S, Bulato C, Iacobelli NP, Finn JD, Spiezia L, Radu C, Arruda VR. X-Linked Thrombophilia with a Mutant Factor IX (Factor IX Padua). NEJM 2009.

Strassburger K and Bretz F. Compatible simultaneous lower confidence bounds for the holm procedure and other bonferroni based closed tests. Statistics in Medicine, 2008; 27: 4914-4927.

van Reenan M, Janssen B. EuroQol Group EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf (April 2015)

World Federation of Haemophilia. World Federation of Haemophilia report on the annual global survey 2017. dated October 2018. http://www1.wfh.org/publications/files/pdf-1714.pdf date accessed: 23 November 2018

APPENDIX 1: DATA HANDLING RULES

Programming of the tables, listings and figures will be performed using SAS Version 9.4 or a more recent version. The following table presents the algorithms to be used in SAS to calculate the derived variables, including rules for handling other missing data or partial dates, or irregular/unexpected data issues.

Category	Description	Data Handling Rule
1. Age (years)	Age (years)	Age = integer part of ([Screening Visit date – Birth date + 1]/365.25)
2. Medical History	Medical History Begin Date of Condition	Begin date of condition will be imputed for all subjects as the 1st of the month for the purpose of computing the onset day.
3. Surgical History	Surgical History Date of Surgery	Date of surgery will be imputed for all subjects as the 1st of the month for the purpose of computing the onset day.
4. Treatment Date	date/time of first study treatment	The date and time (24 hr. clock) of the dose of IMP (study treatment) will be taken from the Dosing eCRF. It is not necessary to define a first treatment date for the lead-in period since the lead-in treatment is not qualitatively different from pre-study therapies.
5. Last Visit Date	Date of Last Visit	Date of last visit according to the Visit eCRF.
6. Last Study Participation Date (STDM variable, typically named RFPENDTC)	Last Study Participation Date (STDM variable, RFPENDTC), where SDTM denotes Study Data Tabulation Model	Last study participation date is defined as last known date of contact, which would be the later of the following dates: last visit date, date of last contact if lost-to-follow-up, date of telephone follow-up, or death date.
7. Study Day Definitions	**Study Day** for assessment/event that occurs on or after the beginning of the period.	For the post-treatment period, Study Day = Date of assessment/event – date of IMP administration + 1. For the lead-in period (or for overall study day), Study Day is the Date of assessment/event – date of L1 Visit + 1.
	Study Day for assessments/events on days prior to the period	For the post-treatment period, Study Day = Date of assessment/event – date of IMP administration. For the lead-in period, Study

Category	Description	Data Handling Rule
		Day is the Date of assessment/event – date of L1 Visit.
	Dose Day	**Dose Day** in the study is defined as the study day of the trial drug administration (Study Day 1 for the post-treatment period).
	Last Study Day	**For subjects who did not receive the dose of trial drug**, Last Study Day is defined as (the later of the last visit date and the date of last contact for subjects lost-to-follow-up from the Study Completion/Early Discontinuation CRF) – Date of Screening Visit + 1. **For subjects who received the dose of trial drug,** Last Study Day is defined as (the later of the last visit date and the date of last contact for subjects lost-to-follow-up from the Study Completion/Early Discontinuation CRF) – date of IMP administration + 1.
	Days Since IMP drug administration for event (e.g., Death)	**Days Since IMP drug administration** is defined as date of event – date of IMP drug administration.
8. Duration of event	The duration of any event	The duration of any event is defined as (stop date – start date + 1).
9. Distance between Event	Distance between factor IX activity measurement and most recent factor IX replacement therapy administration	Date of factor IX activity measurement – Date Preceding factor IX Replacement Therapy Administration) + 1 The date and time of the factor IX activity measurement in question and the factor IX replacement therapy administrations respectively are used to find the latest factor IX replacement therapy administration preceding the factor IX activity measurement in question. In case the dates of the factor IX activity measurement in question and a factor IX replacement therapy administration are the same and no time is indicated, it is assumed that the factor IX replacement therapy administration precedes the factor IX activity

Category	Description	Data Handling Rule
		measurement in question, and the above defined distance therefore becomes equal to 1.
10. Multiple assessments for the same visit	Vital Sign and Laboratory assessments	● All data will be listed in data listings. ● The last of multiple valid assessments within a post-baseline study time window will be used for summaries. ● If there are multiple laboratory values for the same parameter at post-baseline pre-dose of a visit, the last value will be chosen for analysis. ●
11. Special Lab Value Handling for Safety Lab values	Lab values with a prefix such as '>', '<', '+' and 'Less than' etc.…	● '>': use the available original value +0.001 in the analyses. ● '<': use the available original value –0.001 in the analyses. ● '+': use the available original value without the prefix in the analyses. ● '>=': use the available original value in the analyses. ● '<=': use the available original value in the analyses.
12. Prior and concomitant medication	Prior, and lead-in concomitant, and post-treatment concomitant medication	**1.** Prior medication/treatment: is any medication/therapy (including herbal treatments, vitamins, non-pharmacological treatment such as psychotherapy as appropriate) received will be considered prior if the start date of the medication/therapy is missing or the medication/therapy start date is before Visit L1 for the lead-in period. **2.** A medication/therapy will be identified as a "post-treatment concomitant" medication/therapy if it is being continued by the subject at the date of AMT-061 dosing or is any new medication/therapy received during the post-treatment period. A medication with end date that is the same as the AMT-061 dosing date will not be considered to be "post-treatment concomitant". A medication/therapy will be identified as a "lead-in" concomitant

Category	Description	Data Handling Rule
		medication/therapy if it is being continued by the subject at the date of the L1 Visit or is any new medication/therapy received during the lead-in period prior to the date of AMT-061 dosing. The distinction will be made between lead-in concomitant medications and post-treatment concomitant medications. **3.** Any medication/therapy that cannot be identified as Prior, Lead-In Concomitant, or Post-Treatment Concomitant will be considered as being in each of the possible categories depending on available information. The designation of concomitant medication will be done in a manner that is specific to either the lead-in period or the post-treatment period. Given that the study treatment is permanent, there cannot be a medication category subsequent to "concomitant" with respect to the post-treatment period.
13. Adverse event	Missing severity	For the AE summary by severity, an AE with missing severity will be deemed as Severe.
	Missing relationship to study drug	For AE summary by relationship, an AE with a missing relationship to study drug will be deemed as related.
	Treatment-emergent adverse event	An adverse event is considered treatment-emergent for the post-treatment period if an event occurs (or if there was a worsening [intensity and/or severity changed to worsened grades]) on or after the date of dosing with AMT-061. A treatment-emergent adverse can be described as having incidence during the post-treatment period. An adverse event is considered to have had incidence during the lead-in period if an event occurs (or if there was a worsening [intensity and/or severity changed to worsened grades]) on or after the Visit L1 date and before the date of dosing with AMT-061.

Category	Description	Data Handling Rule
		Prior to the Visit L1 date, adverse events are considered to be a part of the medical history. A death is considered to be treatment-emergent for the post-treatment period if any of the adverse events that led to the death occurred on or after the date of administration of the IMP. A death is considered to have had incidence during the lead-in period if any of the adverse events that led to the death occurred on or after the Visit L1 date and before the date of administration of IMP.
		If the AE start date is partial/missing, then ● If AE start date is completely missing, then the AE is considered as both treatment-emergent during the post-treatment period and to have had incidence during the lead-in period. ● If both AE start month and day are missing and AE start year is the same or after the IMP dosing year, then the AE is considered as treatment-emergent for the post-treatment period. If both AE start month and day are missing and AE start year is the same or after the L1 Visit year and on or before the IMP dosing year, then the AE is considered as having had incidence during the lead-in period. ● If AE start day is missing and AE start year and month are the same or after the IMP dosing year and month, then the AE is considered as treatment-emergent for the post-treatment period. If AE start day is missing and AE start "year and month" are the same or after the L1 Visit "year and month" and on or before the IMP dosing "year and month", then the AE is

Category	Description	Data Handling Rule
		considered as having had incidence during the lead-in period. Missing/incomplete (partial) AE start and end dates will not be imputed for data listings.
14. Hard coding	Hard coding for data analysis	Hard Coding is not allowed during data analysis unless agreed to in writing by uniQure.
15. CCI data	Bleed event	Assessments within two weeks of a bleed event will not be included in any analysis.
16. Listing outputs	Data excluded	All data not used for efficacy analysis will be flagged in listings.
17. Contamination due to exogenous factor IX (infusion) use	Contamination of factor IX activity or protein assessment (or in some cases bleeding assessment) due to exogenous factor IX (infusion) use	The date/time (where available) – rather than just date – for the time of the exogenous factor IX infusion and the time of the blood draw for factor IX activity (or protein) assessment will be used for the determination of contamination. The use of date/time (instead of just date) should be applied for the 5-half-life contamination rule and for the 10-day sensitivity-analysis contamination rule. If only the date – but not the time – of the exogenous factor IX infusion is known, then the contamination period will (conservatively) be the time period beginning on midnight at the beginning of that day and ending at the time which is 24 hours plus five half-lives later (but would be ending at 24 hours plus 240 hours for the alternative (sensitivity) 10-day contamination rule). If only the date but not the time of the factor IX activity assessment is known and if any point in time on that date overlaps with the contamination period, then the activity assessment will be deemed contaminated. As alluded to in the SAP text section about secondary efficacy for factor IX activity (Section 7.6.2), the 6-month-data-cut analysis actually used a less refined contamination rule – whereby the date of exogenous factor IX infusion and the subsequent 9 days (10 discrete calendar days in total) were considered to be days of contamination

Category	Description	Data Handling Rule
		with factor IX. As also alluded to in that section, the more refined "5 half-life" contamination period is actually being applied to the 12-month-data-cut analysis and the 18-month-data-cut analysis.

APPENDIX 2: ANALYSIS DATASET SPECIFICATIONS

Analysis datasets will be built to gain efficiency and ensure consistency in data analyses and presentation for this trial. The specifications for each analysis data set will be prepared separately and will not be a part of this SAP.

APPENDIX 3: CENTRAL LABORATORY REFERENCE RANGES FOR USE IN FLAGGING ABNORMAL VALUES

This appendix is provided as an attachment to this document.

CCI

This appendix is provided as an attachment to this document.

APPENDIX 5: SAS CODE FOR STATISTICAL ANALYSES

The prototype SAS Code will be in a separate document.

APPENDIX 6: STATISTICAL DETAILS

The Statistical Details will be in a separate document.

APPENDIX 3 CENTRAL LABORATORY REFERENCE RANGES FOR USE IN FLAGGING ABNORMAL VALUES

Protocol Number:	CT-AMT-061-02
Investigational Drug and Drug Number:	AMT-061 (AAV5-hFIXco-Padua); CCI
Indication:	Hemophilia B
Dosage Form/Dose:	• 2×10^{13} gc/kg AMT-061
Client:	• uniQure biopharma B.V.

Protocol Title: Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B

Date of Issue: 10 Jun 2021

Version: Version 4.0

TABLE OF CONTENTS

1. CTCAE V5.0 LABORATORY TEST CRITERIA

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

Investigations				
	Grade			
Laboratory Analyte	1	2	3	4
GGT increased	>ULN – 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 – 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 – 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Hemoglobin increased (a)	Increase in >0 – 2 gm/dL above ULN (a)	Increase in >2 – 4 gm/dL above ULN (a)	Increase in >4 gm/dL above ULN (a)	n/a
Anemia (hemoglobin decreased)	LLN- 10g/dL; <LLN – 6.2 mmol/L; <LLN – 100 g/L	<10.0 – 8.0 g/dL; <6.2 – 4.9 mmol/L; <100 – 80g/L	<8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated
Leukocytosis (White blood cell increased)			>100,000/mm^3	Clinical manifestations of leucostasis (sic); urgent intervention indicated Note that the spelling is often “leukostasis” in the literature.
White blood cell decreased	<LLN – 3000/mm^3; <LLN – 3.0 x 10^9/L	<3000 – 2000/mm^3; <3.0 – 2.0 x 10^9 /L	<2000 – 1000/mm^3; <2.0 – 1.0 x 10^9 /L	<1000/mm^3; <1.0 x 10^9 /L
Platelet count decreased	<LLN - 75,000/mm3; <LLN - 75.0 x 10^9/L	<75,000 - 50,000/mm3; <75.0 - 50.0 x 10^9/L	<50,000 - 25,000/mm3; <50.0 - 25.0 x 10^9/L	<25,000/mm3; <25.0 x 10^9/L

2. PROTOCOL REFERENCE RANGE DEFINITIONS

Test Name	Subject Characteristics	Unit	Reference Range	Notable Values	Critical Values	Methodology
Chemistry						
Albumin	Adult	g/dL	3.5-5.5	-	-	Photometry
Alkaline Phosphatase	Adult	U/L	37-116	> 348	-	Photometry
ALT/SGPT	Adult	U/L	6-41	> 123	> 164	Photometry
AST/SGOT	Adult	U/L	9-34	> 102	> 164	Photometry
Bilirubin (Total)	Adult	mg/dL	0.10-1.10	> 2.00	-	Photometry
Creatinine	Adult	mg/dL	0.50-1.40	> 2.00	> 3.00	Photometry
Gamma Glutamyl Transferase (GGT)	Adult Male	U/L	11-52	> 156	-	Photometry
Glucose	Adult	mg/dL	60-115	< 50 or > 180	< 40 or > 450	Photometry
Potassium	Adult	Mmol/L	3.5-5.1	< 3.0 or > 6.0	< 2.8 or > 6.2	Ion Selective Electrode
Sodium	Adult	Mmol/L	134-144	-	< 120 or > 160	Ion Selective Electrode
Coagulation						
Activated Partial Thromboplastin Time (APTT)	All	Sec	23.9-40.0	-	> 70.0	Clotting
International Normalized Ratio (INR)	All	(None)	0.8-1.2	-	> 2.5	Calculation
EIA						

Test Name	Subject Characteristics	Unit	Reference Range	Notable Values	Critical Values	Methodology
Interleukin-1 beta	Adult	pg/mL	< 0.61	-	-	ECLIA
Interleukin-6	Adult	pg/mL	< 8.60	-	-	ECLIA
MCP-1	All	pg/mL	200.0-722.0	-	-	ELISA
Hematology						
Basophil %	All	%	0.0-4.0	-	> 7.0	Volume, Conductivity, Scatter
Basophil (Absolute)	All	10^3/µL	0.0-0.3	-	-	Volume, Conductivity, Scatter
Eosinophil %	All	%	0.0-10.0	-	> 20.0	Volume, Conductivity, Scatter
Eosinophil (Absolute)	All	10^3/µL	0.0-0.8	-	-	Volume, Conductivity, Scatter
Hematocrit	Adult Male	%	40-52	-	< 20 or > 60	Calculation
Hemoglobin	Adult Male	g/dL	13.6-18.0	-	< 7.0 or > 20.0	Photometry
Lymphocyte %	All	%	15.0-45.0	-	> 75.0	Volume, Conductivity, Scatter
Lymphocyte (Absolute)	All	10^3/µL	1.0-5.0	-	-	Volume, Conductivity, Scatter
Monocyte %	All	%	0.0-12.0	-	> 25.0	Volume, Conductivity, Scatter
Monocyte (Absolute)	All	10^3/µL	0.0-1.0	-	-	Volume, Conductivity, Scatter
Neutrophil %	All	%	40.0-80.0	-	-	Volume, Conductivity, Scatter
Neutrophil (Absolute)	All	10^3/µL	1.0-8.0	-	-	Volume, Conductivity, Scatter
Platelet	Adult	10^3/µL	140-400	-	< 40 or > 999	Impedence

Test Name	Subject Characteristics	Unit	Reference Range	Notable Values	Critical Values	Methodology
Red Blood Cells	Adult Male	10^6/µL	4.30-6.00	-	-	Impedence
White Blood Cells	Adult	10^3/µL	3.5-11.0	-	< 2.0 or > 30.0	Impedence
Immunology						
HBsAg	All	(None)	Non-reactive	-	-	Electrochemiluminescence Immunoassay
Hepatitis C Antibody	All	(None)	Non-reactive	-	-	Electrochemiluminescence Immunoassay
HIV 1 and 2 Screen	All	(None)	Non-reactive	-	-	
Molecular						
HCV RNA, Qualitative (by RT-PCR)	All	(None)	Not Detected	-	-	RT-PCR
Nephelometry						
hs-C-Reactive Protein	All	mg/L	0.0-3.0	-	-	Nephelometry
Special Chemistry (PPD)						
HIV-1 RNA	All	(None)	Not Detected	-	> Detected	Transcription Mediated Amplification

APPENDIX 4 CCI [redacted]

Protocol Number:	CT-AMT-061-02
Investigational Drug and Drug Number:	AMT-061 (AAV5-hFIXco-Padua); CCI [redacted]
Indication:	Hemophilia B
Dosage Form/Dose:	• 2×10^{13} gc/kg AMT-061
Client:	• uniQure biopharma B.V.

Protocol Title: Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B

Date of Issue: 10 Jun 2021

Version: Version 4.0


CCI


CCI

APPENDIX 5 SAS CODE FOR STATISTICAL ANALYSES

Protocol Number:	CT-AMT-061-02
Investigational Drug and Drug Number:	AMT-061 (AAV5-hFIXco-Padua); CCI [redacted]
Indication:	Hemophilia B
Dosage Form/Dose:	2×10^{13} gc/kg AMT-061
Client:	uniQure biopharma B.V.

Protocol Title: Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B

Date of Issue: 10 Jun 2021

Version: 4.0

Descriptive Statistics

Test	SAS Codes for Modeling

Summary statistics	`proc sort data = InputDataset;`
	`  by GroupVar1 GroupVar2;`
	`run;`
	
	`proc means data = InputDataset noprint;`
	`  by GroupVar1 GroupVar2;`
	`  var AVAL;`
	`  output out = Dat1`
	`  n=s1 mean=s2 stddev=s3`
	`  median=s4 min=s5 max= s6;`
	`run;`

Some General Rules

For visit-based modeling analyses, please do not include time points in the input dataset if the n for that time point (i.e. visit) is < 10.

One-Sided P-values

One-sided p-values	Please note that the TLG output requires the production of one-sided p-values in many instances. When submitting the code provided in this appendix, SAS produces output presenting two-sided p-values. Please see the SAP text, the TLG mock footnotes and the notes to the programmer to identify the instances where a one-sided p-value is required. To convert the SAS-provided two-sided p-value to a one-sided p-value, please employ the following approach: In instances when the (directionality of the) point estimate for the parameter suggests the alternative hypothesis is true (i.e. when the directionality of the point estimate favors the study treatment): One-sided p-value = (Two-sided p-value)/2 In instances when the (directionality of the) point estimate does not suggest the alternative hypothesis to be true (i.e. when the directionality of the point estimate does not favor the study treatment): One-sided p-value = 1 – ((two-sided p-value)/2) To find the alternative hypothesis for a given analysis, please check the SAP text and/or the TLG mock footnotes. An example of the specification from a footnote for the primary efficacy analysis using the repeated measures linear mixed model is as follows: `[c] One-sided P-value <= 0.025 for Post-treatment > baseline is regarded as statistically significant.`

Primary Efficacy Analyses

Test	SAS Codes for Modeling

Repeated Measures Linear Mixed Model – e.g. for Factor IX Activity	`proc mixed data = InputDataset;`
	`  class avisitn subjid;`
	`  model chg = avisitn;`
	`  repeated avisitn / type = TOEP`
	`  subject = subjid;`
	`lsmeans avisitn / cl;`
	`  ods output LSMeans = lsm;`
	`run;`
	
	where AVISIT is the study visit and CHG is change from baseline in FIX activity.
	
	If the Toeplitz model fails to converge, then a first-order autoregressive [AR(1)] structure will be used instead. In the AR(1) model, subject will be included as a random effect.
	
	`proc mixed data = InputDataset;`
	`  class avisitn subjid;`
	`  model chg = avisitn;`
	`  repeated avisitn / type = AR(1)`
	`  subject= subjid;`
	`  random subjid;`
	`  lsmeans avisitn / cl;`
	`  ods output LSMeans = lsm;`
	`run;`
	
	Please do not include time points in the input dataset if the n for that time point is < 10.

Repeated Measures Generalized Estimating Equations (GEE) Negative Binomial Regression Model – e.g. for bleeding count endpoints	`proc genmod data = InputDataset;`
	`  title 'Negative Binomial with Offset';`
	`  class aphase (ref = 'Lead-in') subjid;`
	`  model NumBleed = aphase`
	`  / dist = negbin link = log`
	`  offset = logCollTime;`
	`  repeated subject = subjid/ type = UN`
	`  within = aphase printmle corrw;`
	`  lsmeans aphase / ilink cl pdiff exp;`
	`  ods output ParameterEstimates = est`
	`  geewcorr = wcorr`
	`  LSMeans = lsm`
	`  Diffs = lsmdiff_ ;`
	`run;`

Test	SAS Codes for Modeling

	where NumBleed is the collected number of bleeds, Aphase is pre or post treatment, and logCollTime is the logarithm of the time frame over which the number of bleeds is collected.
	
	If the model fails to converge, then a compound symmetry covariance structure will be used.
	…
	`repeated subject = SubjectID type = CS`
	…
	To obtain the Adjusted Rate and corresponding 95% CI:
	
	`data adjrate;`
	`set lsm;`
	`adjrate = expestimate;`
	`arul = upperexp;`
	`arll = lowerexp;`
	`run;`
	
	To obtain the Rate Ratio, corresponding 95% CI, and two-sided p-value:
	`data raterat;`
	`set lsmdiff;`
	`rr = expestimate;`
	`rrll = lowerexp;`
	`rrul = upperexp;`
	`pval = probz;`
	`run;`
	
	To obtain the dispersion parameter:
	
	`data dispparm;`
	`set est;`
	`if strip(parameter) ne 'Dispersion' then delete;`
	`nbdp = estimate;`
	`u_nbdp = UpperWaldCL;`
	`l_nbdp = LowerWaldCL;`
	`run;`

Test	SAS Codes for Modeling

Correlation Between Treatment Periods – for Subject-Specific Unadjusted ABR	`proc corr data = <InputDataset> fisher(biasadj=no) pearson spearman;`
	`  var <Lead-in Unadj. ABR> <Post-Tx Unadj. ABR>;`
	`  ods output fisherpearsoncorr = pearson`
	`  fisherspearmancorr = spearman;`
	`run;`

Period-Specific Negative Binomial Regression Model – e.g. for bleeding count endpoints	`proc genmod data = InputDataset;`
	`  title 'Period-Specific Negative Binomial with Offset';`
	`  class aphase;`
	`  where aphase = 'Specify Period';`
	`  model NumBleed = period`
	`  / dist = negbin link = log`
	`  offset = logCollTime;`
	`  lsmeans aphase / ilink cL;`
	`  ods output ParameterEstimates = est_`
	`  LSMeans = lsm_ ;`
	`run;`
	
	where NumBleed is the collected number of bleeds, aphasen is the treatment period, logCollTime is the logarithm of the time frame over which the number of bleeds is collected. Separate models should be fit for the Lead-in and Post-Treatment periods.
	
	To obtain the period-specific model-based rate and the associated 95% confidence interval:
	
	`data psmbr;`
	`  set lsm;`
	`  psmbr = expestimate;`
	`  llpsmbr = lowerexp;`
	`  ulpsmbr = upperexp;`
	`run;`
	
	To obtain the dispersion parameter and the associated 95% confidence interval:
	
	`data dispparm;`
	`  set est;`
	`  if strip(parameter) ne 'Dispersion' then delete;`
	`  nbdp = estimate;`
	`  u_nbdp = UpperWaldCL;`
	`  l_nbdp = LowerWaldCL;`
	`run;`

Secondary Efficacy Analysis

Test	SAS Codes for Modeling

Paired t-test – for Annualized Exogenous Factor IX Consumption	`ods output ttests = ttests statistics = adjmean;`
	`proc ttest data = InputDataset;`
	`  title 't-test';`
	`  paired FIXALI*FIXAPT;`
	`run;`
	
	where FIXALI is the annualized FIX replacement therapy from the lead-in period and FIXAPT is the annualized FIX replacement therapy from the post-treatment period.

Repeated Measures Generalized Linear Mixed Model Logistic Regression – for % of Subjects with < 12% Factor IX Activity	**18-Month (data cut) Analysis**
	
	`proc genmod data = InputDataset;`
	`  title 'Logistic regression for correlated data`
	`  from the same subject across periods';`
	`  class subjid avisit;`
	`  model NumFIX(event = "1") = avisit`
	`  / dist = bin link = logit;`
	`  repeated subject = subjid/`
	`  type = UN`
	`  within = avisit`
	`  printmle;`
	
	`  lsmeans avisit/ ilink exp diff cl;`
	`  estimate '-"TrtComp" avisit -8 -8 -8 0 0 0 3 3 3 3 3 3 3 3/divisor = 24 exp;`
	`  ods output ParameterEstimates = est_`
	`  Diffs = lsmdiff_`
	`  Estimates = Estimate;`
	`run;`
	
	where NumFIX is the binary response of the status of trough FIX activity < 12% of normal. In the above example code NumFIX = 1 if aval < 12. Please note it is recommended to relabel the value of avisit for "Post-treatment Week 12" to "Post-treatment Month 3" in the input dataset so the model parameterizes the visits in sequential order.
	
	First, include all lead-in visits, and include all post treatment visits from Month 3 to Month 18.
	
	Note: the above and below SAS code is provided for the 18-month analysis.
	
	If the model using the unstructured covariance matrix fails to converge then the compound symmetry covariance structure will be employed across visits for the same subject and the same period:
	
	…
	
	`  repeated subject = Subjid/`
	`  type = CS`
	`  within = avisit;`
	…

	If the above model does not converge, then include all lead-in visits and include only the Post Treatment Month 12 and Month 18 visits in the input dataset. Again, first try unstructured covariance (`type = UN`); if convergence fails then use compound symmetry covariance (`type = CS`). The updated estimate statement would be:
	
	`estimate 'TrtComp' avisit -2 -2 -2 3 3/divisor = 6 exp;`
	
	If the above model does not converge, then include only the Month 6 visit from the Lead-In Period and include only the Post Treatment Month 12 and Month 18 visits in the input dataset. Again, first try unstructured covariance (`type = UN`); if convergence fails then use compound symmetry covariance (`type = CS`). The updated estimate statement would be:
	
	`estimate 'TrtComp' avisit -2 1 1/divisor = 2 exp;`
	
	If the above model does not converge, then include only the Month 6 visit from the Lead-In Period and include only the Post Treatment Month 18 visit in the input dataset. Again, first try unstructured covariance (`type = UN`); if convergence fails then use compound symmetry covariance (`type = CS`). The updated estimate statement would be:
	
	`estimate 'TrtComp' avisit -1 1/divisor = 1 exp;`
	
	pvalue is attained from PROBCHISQ field in the 'estimate' dataset where label = 'TrtComp'. If the OR (odds ratio) estimate favors the post-treatment period (< 1) then report the chi-square p-value/2. If the OR estimate favors the lead-in period (> 1) then report 1-(chi-square p-value)/2.
	OR estimate is obtained from LBETAESTIMATE field in the 'estimate' dataset where label = 'Exp(TrtComp)'
	95% CIs are obtained from LBETALOWERCL, LBETAUPPERCL fields from the 'estimate' dataset where label = 'Exp(TrtComp)'.

12-Month (data cut) Analysis

```
proc genmod data = InputDataset;
  title 'Logistic regression for correlated data
         from the same subject across periods';
  class subjid avisit;
  model NumFIX(event = "1") = avisit
    / dist = bin link = logit;
  repeated subject = subjid/
                     type = UN
                     within = avisit
                     printmle;

  lsmeans avisit/ ilink exp diff cl;
  estimate '-"TrtComp" avisit -7 -7 -7 0 0 0 3 3 3 3 3 3
3/divisor = 21 exp;
  ods output ParameterEstimates = est_
              Diffs = lsmdiff_
              Estimates = Estimate;
run;
```

where NumFIX is the binary response of the status of trough FIX activity < 12% of normal. In the above example code NumFIX = 1 if aval < 12. Please note it is recommended to relabel the value of avisit for "Post-treatment Week 12" to "Post-treatment Month 3" in the input dataset so the model parameterizes the visits in sequential order.

First, include all lead-in visits, and include all post treatment visits from Month 3 to Month 12.

Note: the above and below SAS code is provided for the 12-month analysis.

If the model using the unstructured covariance matrix fails to converge then the compound symmetry covariance structure will be employed across visits for the same subject and the same period:

…

```
  repeated subject = Subjid/
                     type = CS
                     within = avisit;
…
```

	If the above model does not converge, then include all lead-in visits and include only the Post Treatment Month 6 and Month 12 visits in the input dataset. Again, first try unstructured covariance (`type = UN`); if convergence fails then use compound symmetry covariance (`type = CS`). The updated estimate statement would be:
	
	`estimate 'TrtComp' avisit -2 -2 -2 3 3/divisor = 6 exp;`
	
	If the above model does not converge, then include only the Month 6 visit from the Lead-In Period and include only the Post Treatment Month 6 and Month 12 visits in the input dataset. Again, first try unstructured covariance (`type = UN`); if convergence fails then use compound symmetry covariance (`type = CS`). The updated estimate statement would be:
	
	`estimate 'TrtComp' avisit -2 1 1/divisor = 2 exp;`
	
	If the above model does not converge, then include only the Month 6 visit from the Lead-In Period and include only the Post Treatment Month 12 visit in the input dataset. Again, first try unstructured covariance (`type = UN`); if convergence fails then use compound symmetry covariance (`type = CS`). The updated estimate statement would be:
	
	`estimate 'TrtComp' avisit -1 1/divisor = 1 exp;`
	
	pvalue is attained from PROBCHISQ field in the 'estimate' dataset where label = 'TrtComp'. If the OR (odds ratio) estimate favors the post-treatment period (< 1) then report the chi-square p-value/2. If the OR estimate favors the lead-in period (> 1) then report 1-(chi-square p-value)/2.
	OR estimate is obtained from LBETAESTIMATE field in the 'estimate' dataset where label = 'Exp(TrtComp)'
	95% CIs are obtained from LBETALOWERCL, LBETAUPPERCL fields from the 'estimate' dataset where label = 'Exp(TrtComp)'.

Kolmogorov-Smirnov Test	`proc npar1way data = <InputDataset> edf;`
	`  class <treatment period>;`
	`  var <factor IX activity (%)>;`
	`run;`

Kaplan-Meier curves	`proc lifetest method = km plots = (s)`
	`  outsurv = out1;`
	`  time time*censor(0);`
	`  strata treatment / diff = all;`
	`run;`
	
	Get 95% CI from dataset 'out1'.
	** event=1; censored=0;
	
	See the SAP text regarding censoring rules.

Pearson and Spearman correlation	`proc corr data = InputDataset pearson spearman`
	`  fisher(biasadj=no);`
	`  var AVAL;`
	`  ods output`
	`  FisherPearsonCorr = p_corr`
	`  FisherSpearmanCorr = s_corr;`
	`run;`

Simple Linear Regression Line	`proc reg data = InputDataset;`
	`  model apfixact = l10_basenab;`
	`  ods output Parameter Estimates = PE FitStatistics=FS;`
	`run;`
	
	`Note: Use log base 10 of the NAb titer value when fitting the regression line.`

Repeated Measures Linear Mixed Model for CCI endpoints	**12-Month (data cut) and 18-Month Analysis**
	
	CCI
	
	`proc mixed data = InputDataset;`
	`  class visit period SubjectID;`
	`  model score = period visit period*visit;`
	`  random int / subject = SubjectID;`
	`  repeated visit / type = UN`
	`  subject = period*SubjectID;`
	`  lsmeans period / cl pdiff;`
	`  lsmeans period*visit / cl;  ods output LSMeans = lsm Diffs = lsmdiff_;`
	`run;`
	
	If the Unstructured model fails to converge, then a compound symmetric covariance structure will be used instead.
	
	`proc mixed data = InputDataset;`
	`  class visit period SubjectID;`
	`  model score = period visit period*visit;`
	`  random int / subject = SubjectID;`
	`  repeated visit / type = CS`
	`  subject = period*SubjectID;`
	`  lsmeans period / cl pdiff;`
	`  lsmeans period*visit / cl;`
	`  ods output LSMeans = lsm Diffs = lsmdiff_;`
	`run;`
	
	See the paragraph at the end of this section about what to do if convergence still is not attained from the above analyses. Also see the SAP text.
	
	6-Month (data cut) Analysis (and 12-Month CCI Analysis)
	
	CCI

CCI [redacted]

```
proc mixed data = InputDataset;
  class period SubjectID;
  model score = period;
  repeated period / type = UN
                    subject = SubjectID;
  lsmeans period / cl pdiff;
  ods output LSMeans = lsm Diffs = lsmdiff_;
run;
```

If the Unstructured model fails to converge, then a compound symmetric covariance structure will be used instead.

```
proc mixed data = InputDataset;
  class period SubjectID;
  model score = period;
  repeated period / type = CS
                    subject = SubjectID;
  lsmeans period / cl pdiff;
  ods output LSMeans = lsm Diffs = lsmdiff_;
run;
```

CCI [redacted]

Generalized Additive-Model Plot for ABR

```
proc gampl* data = InputDataset;
  model count = spline(FIX) /
    offset = log_time_at_risk dist = NegativeBinomial
    link = log;
  id count log_time_at_risk time_at_risk FIX usubjid;
  output out = OutDataset pred = p lower upper;
run;

data OutDataset
  set OutDataset;
  /* if daily rate, convert rate to yearly rate*/
  if cmiss(p, time_at_risk) = 0 then
    predrate = 365.25*(p/ time_at_risk);
  if cmiss(lower, time_at_risk) = 0 then
    predl = 365.25*(lower/ time_at_risk);
  if cmiss(upper, time_at_risk) = 0 then
    predu = 365.25*(upper/ time_at_risk);
run;

proc sgplot data = QCData noautolegend;
  series  y = predrate x = FIX;
    name='series';
  band x = FIX lower = predl upper = predu /
    transparency = 0.5;
  yaxis label = "Rate of Bleeding Events (per year)";
  xaxis type = linear
    label = "Factor IX Activity (%)";
run;
```

* PROC GAMPL allows for a negative binomial model; PROC GAM currently does not.
PROC GAMPL has spline fits but not LOESS fits.
FIX denotes aPTT central laboratory factor IX activity.
time_at_risk is time at risk in days.
log_time_at_risk is the natural logarithm of time_at_risk.

APPENDIX 6 STATISTICAL DETAILS

Protocol Number:	CT-AMT-061-02
Investigational Drug and Drug Number:	AMT-061 (AAV5-hFIXco-Padua); [redacted]
Indication:	Hemophilia B
Dosage Form/Dose:	• 2×10^{13} gc/kg AMT-061
Client:	• uniQure biopharma B.V.

Protocol Title: Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B

Date of Issue: 10 Jun 2021

Version: Version 4.0

TABLE OF CONTENTS

1. INTRODUCTION

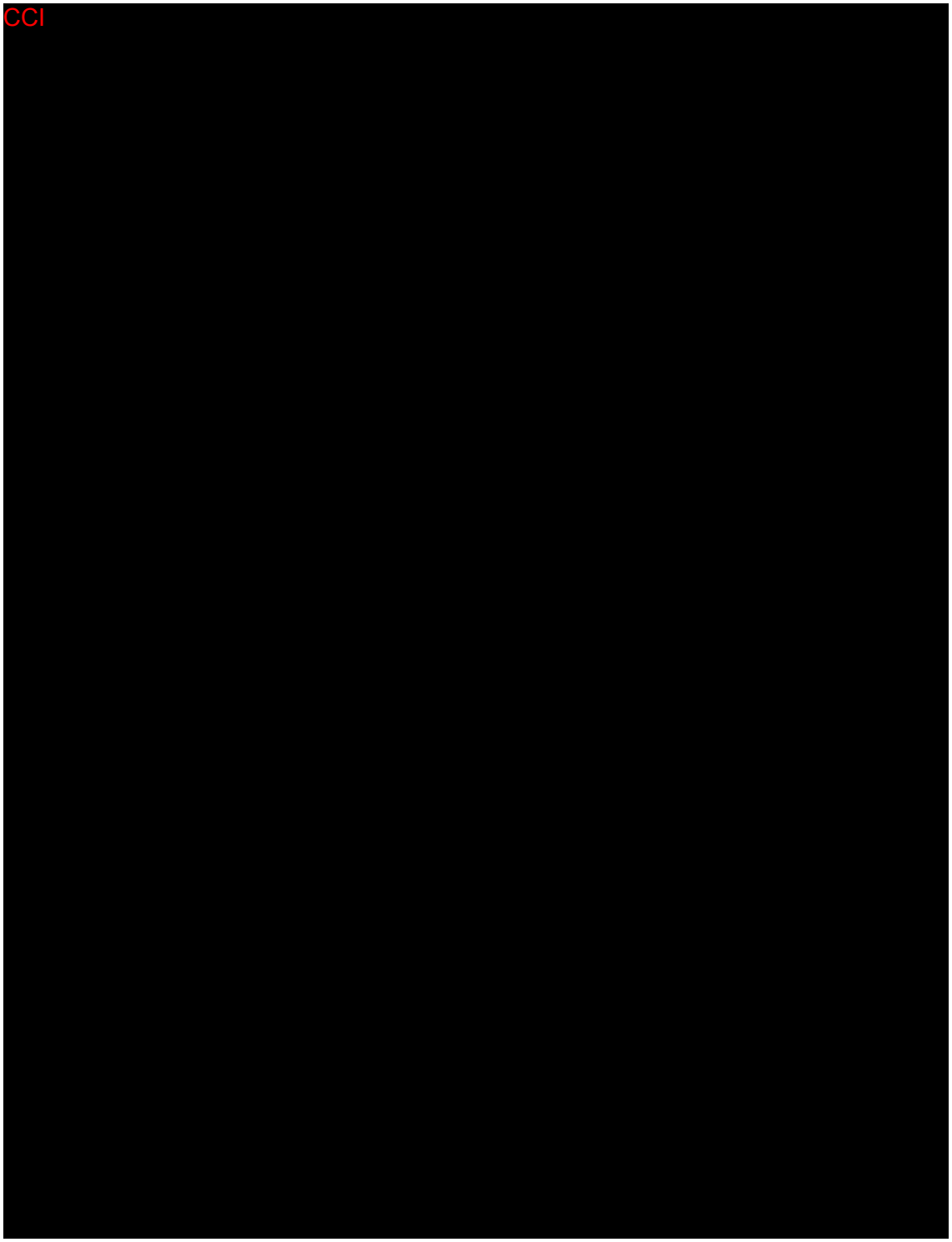
This SAP Details Appendix (1) states how modeling assumptions will be assessed and (2) shows how the simultaneous confidence intervals will be computed. The SAP text has already stated how robustness of results will be examined by sensitivity analyses; such sensitivity analyses do not need to be re-iterated here.

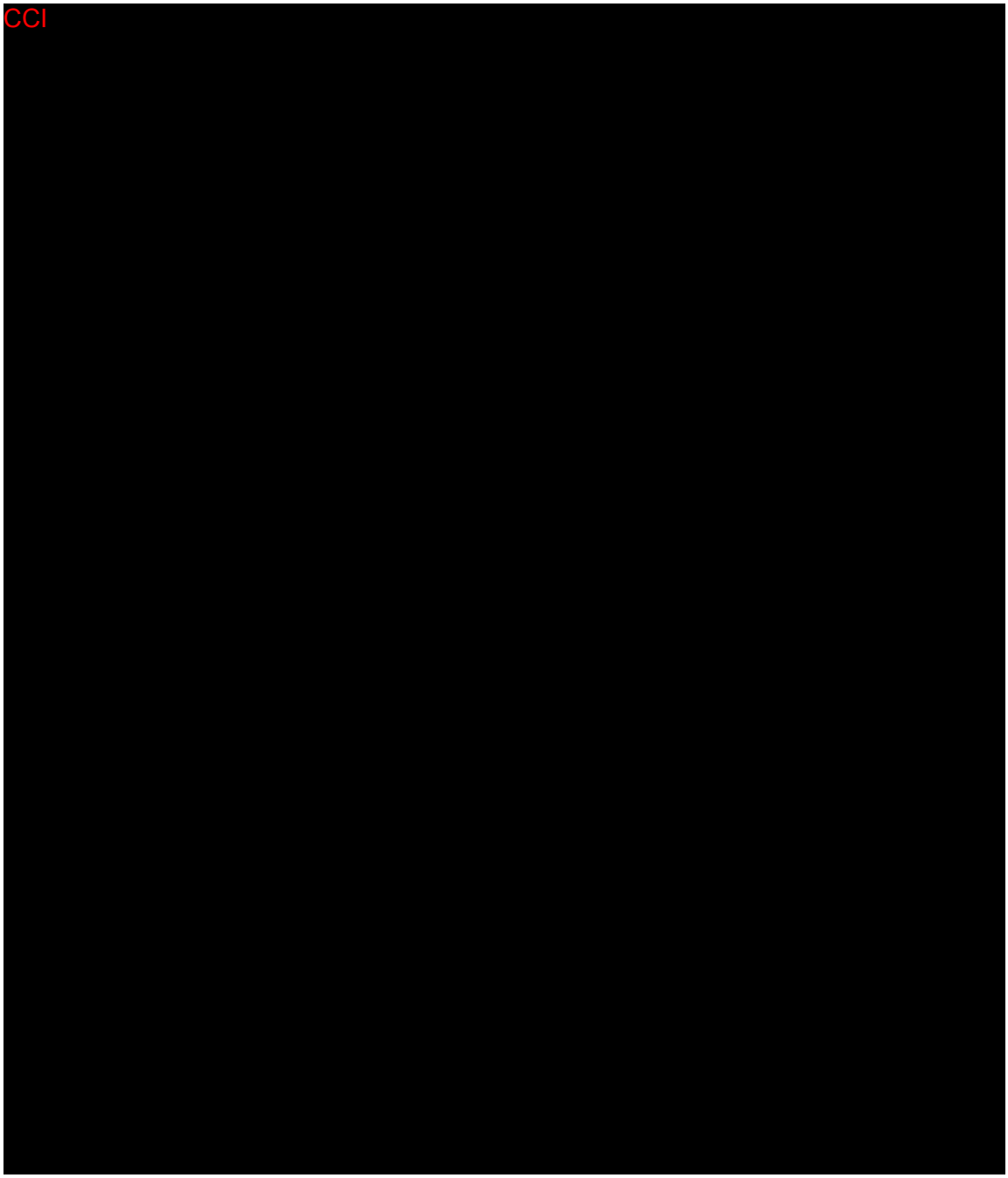
2. PRIMARY EFFICACY ENDPOINTS

The primary efficacy endpoint is as follows:

- ABR comparison between AMT-061 and prophylaxis for non-inferiority between the 52-week post treatment (AMT-061) follow-up and the lead in phase

CCI


CCI


CCI

5. TYPE I ERROR CONTROL AND CALCULATION OF SIMULTANEOUS CONFIDENCE INTERVALS

5.1 Type I Error Control

Formal statistical testing of the efficacy endpoints will be performed using the closed testing principle (for Type I error control for multiple testing). Due to the closed testing principle, no correction for multiplicity is necessary. Among the endpoints being formally tested for statistical significance, all will be tested for superiority at a one-sided alpha level of 0.025 (except as otherwise noted). Superiority testing and non-inferiority testing will be accomplished using the FAS population.

Fixed sequential testing will be performed using a hierarchical approach and will be continued until a non-significant result is obtained (except as otherwise noted). The order of fixed sequential tests is specified below:

1. ABR comparison between AMT-061 and prophylaxis for non-inferiority between the lead-in phase and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up
2. Endogenous factor IX activity at 6 months after AMT-061 dosing
3. Endogenous factor IX activity at 12 months after AMT-061 dosing
4. Endogenous factor IX activity at 18 months after AMT-061 dosing
5. Annualized consumption of factor IX replacement therapy during the 52-week post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase (secondary efficacy endpoint)
6. Annualized infusion rate of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase
7. Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 over the 52 weeks following stable FIX expression (6-18 months)
8. ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up
9. Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to lead-in phase
10. Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase
11. CCI [redacted] during the 12 months following AMT-061 dosing compared with the lead-in phase
12. CCI [redacted] during the 12 months following AMT-061 dosing compared with the lead-in phase

5.2 Simultaneous Confidence Intervals

Simultaneous one-sided 97.5% confidence intervals based on a graphical approach to multiple testing (Bretz et al. 2015; Guilbaud 2008; Strassburger and Bretz 2008) will be provided for the

type I error controlled efficacy endpoints as a supportive analysis. For endpoints for which an increase is favorable, the lower one-sided 97.5% confidence bound will be provided; for endpoints for which an increase is unfavorable, the upper one-sided 97.5% confidence bound will be provided.

For tractability of computing simultaneous confidence intervals (by a graphical approach) – it will be assumed that the three primary endpoints are tested in sequence (like the secondary endpoints). This does not change the requirement that for actual statistical significance for primary efficacy, individual statistical significance is required for each and every one of the three primary endpoints. As the reader will be able to see from the methodological details given below, having (in one scenario) individual statistical significance for each and every one of the three primary endpoints will place the simultaneous confidence intervals for these three endpoints (exactly) at the significance boundary; likewise having (in another scenario) individual statistical non-significance for at least one of the three primary endpoints will place the simultaneous confidence interval for at least one of these three endpoints beyond the significance boundary – in the direction of non-statistical-significance. Thus the inference inferred from the simultaneous confidence intervals will be the same as the actual inference that is being employed by the actual hypothesis testing approach.

For any data cuts (i.e. analysis times) that are not the main data cut for a given endpoint, the p-values and confidence intervals will be considered to be descriptive rather than inferential.

5.2.1 Steps to Compute Simultaneous Confidence Intervals for the Graphical Approach

In this section (Section 5.2.1), the simultaneous confidence interval computation is established for the general (weighted Bonferroni) graphical framework (for multiple testing). This is done in order to establish the background for the more specific calculations that will apply to this study's fixed-sequential hypothesis-testing framework (for multiple comparisons). Thus, for specific computations that fit this study's framework, see the next section (Section 5.2.2).

Consider a group of hypotheses (i.e. endpoints) to be tested by the graphical approach at overall family-wise level of significance α. Once each type I error-controlled statistical test has been carried out by the graphical testing, we will know whether statistical significance was attained for each test (i.e. endpoint).

The simultaneous confidence interval computation requires us to know only the following:

- whether the result was statistically significant (for a given endpoint),
- what the level of significance (α_i) was in the final graph (remaining) for a given endpoint, and
- whether all endpoints had statistical significance.

If not all tests (i.e. endpoints) had statistical significance, then do the following (for a given endpoint):

- If the result was statistically significant, then the simul-LCL (simultaneous lower confidence limit) is set to **(A)** zero* for a superiority-tested endpoint or to the non-inferiority margin for a non-inferiority-tested endpoint.

- If the result was not statistically significant, then set the simul-LCL equal to **(B)** the univariate LCL for the endpoint at the level of significance that the endpoint had in the final graph.

If all tests (i.e. endpoints) had statistical significance, then do the following (for a given endpoint):

- If the result was statistically significant, then the simul-LCL (simultaneous lower confidence limit) is set to the maximum of **(A)** and **(B)**.

* zero for a treatment difference; one for an odds ratio, hazard ratio, or rate ratio.

Note that for this exercise, all results must be first transformed so that the point estimate for the endpoint is on the scale of "larger-is-better" (i.e. with larger values being better than smaller values) as follows. If for a given endpoint "lower-is-better" (i.e. with lower values being better than larger values), then do the following:

- First convert the point estimate and its raw confidence intervals and the NI margin (if applicable) to those of a larger-is-better endpoint by doing the following:
  - For a continuous endpoint, multiply by -1.
  - For an odds ratio or hazard ratio or rate ratio (y), take the inverse: 1/y.
- Then, re-transform the computed endpoint-specific simultaneous LCL to a simultaneous UCL by doing the reverse transformation.
  - For a treatment difference, $-\infty$ would be reverse-transformed to ∞; for an odds ratio, hazard ratio, or rate ratio, zero would be reverse-transformed to ∞.

5.2.2 Steps to Compute Simultaneous Confidence Intervals for Fixed Sequential Testing

Consider a group of hypotheses (i.e. endpoints) to be tested by fixed sequential testing at overall family-wise level of significance α. Once each type I error-controlled statistical test has been carried out by the fixed sequential testing, we will know whether statistical significance was attained for each test (i.e. endpoint).

The simultaneous confidence interval computation requires us to know only the following:

- whether the result was statistically significant (for a given endpoint),
- the last endpoint – in the sequence of testing – to have statistical significance,
- whether all endpoints (in the sequence) had statistical significance.

If not all tests (i.e. endpoints) had statistical significance, then do the following (for a given endpoint):

- If the result was statistically significant, then the simul-LCL (simultaneous lower confidence limit) is set to zero* for a superiority-tested endpoint or to the non-inferiority margin for a non-inferiority-tested endpoint.
- If the result was not statistically significant, then set the simul-LCL equal to:
  - the univariate LCL for the endpoint at level α, if this is the first endpoint in the sequence or if the previous endpoint in the sequence had statistical significance

- -∞**, if the previous endpoint in the sequence did not have statistical significance.

If all tests (i.e. endpoints) had statistical significance, then do the following (for a given endpoint):

- If the result was statistically significant, then the simul-LCL (simultaneous lower confidence limit) is set to the maximum of **(A)** and **(B)**, where **(A)** = zero* for a superiority-tested endpoint or to the non-inferiority margin for a non-inferiority-tested endpoint, and **(B)** = the univariate LCL for the endpoint at level α.

---

* zero for a treatment difference; one for an odds ratio, hazard ratio, or rate ratio.
** -∞ for a treatment difference; zero for an odds ratio, hazard ratio, or rate ratio.

Note that for this exercise, all results must be first transformed so that the point estimate for the endpoint is on the scale of "larger-is-better" (i.e. with larger values being better than smaller values) as follows. If for a given endpoint "lower-is-better" (i.e. with lower values being better than larger values), then do the following:

- First convert the point estimate and its raw confidence intervals and the NI margin (if applicable) to those of a larger-is-better endpoint by doing the following:
  - For a continuous endpoint, multiply by -1.
  - For an odds ratio or hazard ratio or rate ratio (y), take the inverse: 1/y.
- Then, re-transform the computed endpoint-specific simultaneous LCL to a simultaneous UCL by doing the reverse transformation.
  - For a treatment difference, -∞ would be reverse-transformed to ∞; for an odds ratio, hazard ratio, or rate ratio, zero would be reverse-transformed to ∞.

6. REFERENCES

Bretz F, Maurer W, and Maca J "Graphical Approaches to Multiple Testing." Clinical Trial Biostatistics and Biopharmaceutical Applications, edited by Young W and Chen D, CRC Press, 2015, 349-394.

16.1.9.2 Statistical Methods Appendix for Study CT-AMT-061-02

STATISTICAL METHODS APPENDIX FOR STUDY CT-AMT-061-02

Protocol Number:	CT-AMT-061-02
Investigational Drug and Drug Number:	AMT-061 (AAV5-hFIXco-Padua); [redacted]
Indication:	Hemophilia B
Dosage Form/Dose:	2×10^{13} gc/kg AMT-061
Client:	uniQure biopharma B.V.

Protocol Title: Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B

Date of Issue: 18 Feb 2022

Version: Version 1.0

TABLE OF CONTENTS

List of Tables

List of Figures

GLOSSARY OF ABBREVIATIONS

Abbreviation	Term
AAV	adeno-associated viral
AAV5	adeno-associated viral vector serotype 5
AAV5-hFIXco	recombinant adeno-associated viral vector serotype 5 containing the wild type human FIX gene, codon-optimized for optimal expression in humans, under control of a liver-specific promoter (AMT-060)
AAV5-hFIXco-Padua	recombinant adeno-associated viral vector serotype 5 containing a codon-optimized Padua derivative of human coagulation FIX cDNA (AMT-061)
ABR	annualized bleeding rate
AE	adverse event
AFP	Alpha-fetoprotein
ALP	alkaline phosphatase
ALT	alanine aminotransferase
aPTT	activated partial thromboplastin time
AR	autoregressive
AST	aspartate aminotransferase
ATMP	advanced therapy medicinal product
CCI	[redacted]
cDNA	complementary deoxyribonucleic acid
CI	confidence interval
COVID-19	Coronavirus Disease (discovered in) 2019
CRP	c-reactive protein
CSR	Clinical Study Report
DMC	Data Monitoring Committee

DNA	deoxyribonucleic acid
eCRF	electronic case report form
ELISA	enzyme-linked immunosorbent assay
CCI	[redacted]
FAS	Full Analysis Set
FDA	Food and Drug Administration
FIX	coagulation factor IX
GAM	Generalized Additive Model
GGT	gamma-glutamyl transpeptidase
GEE	Generalized Estimating Equations
GTWP	Gene Therapy Working Party
CCI	[redacted]
HBeAg	hepatitis B extracellular antigen
HBsAg	hepatitis B surface antigen
HBV DNA	hepatitis B virus deoxyribonucleic acid
HCV RNA	hepatitis C virus ribonucleic acid
CCI	[redacted]
hFIX	human coagulation factor IX
CCI	[redacted]
IFNγ	interferon gamma
IgG	immunoglobulin G
IgM	immunoglobulin M
IL-1β	interleukin-1beta

IL-2	interleukin-2
IL-6	interleukin-6
IMP	investigational medicinal product
INR	International normalized ratio
CCI	[redacted]
IU	international unit
JADE	Joint Tissue Activity and Damage Exam
LOD	Limit of detection
MET	Metabolic equivalent of task
MCP-1	monocyte chemotactic protein-1
MedDRA	Medical Dictionary for Regulatory Activities
MRE	Magnetic resonance elastography
NAB	Neutralizing antibody
PCS	potentially clinically significant
PP	Per-protocol
CCI	[redacted]
PROBE	Patient Reported Outcomes, Burdens, and Experiences
Q1	First quartile
Q3	Third quartile
CCI	[redacted]
rAAV5	recombinant adeno-associated viral vector serotype 5
SAE	serious adverse event
SAP	Statistical Analysis Plan
SDTM	Study Data Tabulation Model

SOC	System Organ Class
SOP	Standard Operating Procedure
SWE	Shear wave elastography
TEAE	treatment emergent adverse event
US	United States
VAS	visual analogue scale
WFH	World Federation of Haemophilia
WHO	World Health Organization
CCI	[redacted]

Trademark Information

SAS

nQuery Advisor

1. INTRODUCTION

1.1 General Introduction

Congenital hemophilia B is an inherited bleeding disorder characterized by an increased bleeding tendency due to either a partial or complete deficiency of the essential blood coagulation factor IX (FIX). The deficiency is the result of mutations in the factor IX gene. Hemophilia B is an X-linked, recessive condition, since it occurs almost exclusively in males. Females typically are asymptomatic carriers. The number of people with hemophilia B worldwide is approximately 30,000 and in the United States (US) alone is approximately 4,000 (World Federation of Haemophilia [WFH], 2017). Approximately 1 in 20,000 - 50,000 live male newborns has hemophilia B.

The severity of symptoms can vary and the severe forms become apparent early in life. About one-third of individuals with hemophilia B have a severe disorder characterized by factor IX levels that are less than 1% of normal (Kessler & Mariani, 2006). Moderate and mild hemophilia B, with 1 - 5% or 5 - < 40% of normal factor IX level, respectively, are each observed in about one-third of patients (Kessler & Mariani, 2006).

There is no cure for hemophilia B. The primary goals of hemophilia B therapy are the prevention of bleeding episodes, rapid and definitive treatment of bleeding episodes (breakthrough bleeds) that occur even while on a regular prophylactic regimen and provision of adequate hemostasis during surgery and emergencies. Currently, these goals are essentially met for hemophilia B subjects by intravenous (IV) injections of commercially available recombinant- or plasma-derived factor IX products, either at the time of a bleed (on-demand) or by regular infusions up to several times a week (prophylactically).

Somatic gene therapy offers the potential for a shift of the disease severity phenotype from severe to a moderate or mild hemophilia phenotype or complete amelioration through continuous production of stable factor IX levels after a single administration of vector, especially since a small rise in circulating factor IX to at least 1% of normal levels can substantially ameliorate the bleeding phenotype and potentially improve the CCI for patients.

AMT-061 has been developed for the treatment of hemophilia B. AMT-061 is a recombinant adeno-associated viral vector serotype 5 (rAAV5) containing the coding sequence for Padua variant of the human coagulation factor IX (hFIX Padua), codon-optimized, under control of a liver-specific promoter (also known as AAV5-hFIXco Padua).

AMT-061 is a derivative of AMT-060 which has been studied in a Phase I/II clinical trial in humans with severe and moderately severe hemophilia B. Both AMT-060 and AMT-061 have the same rAAV5 containing the codon-optimized wild-type human factor IX gene, but the latter incorporates two-nucleotide change in order to encode the naturally occurring Padua variant of human coagulation factor IX. The factor IX-Padua protein differs from the ‘wild-type’ human factor IX protein by a single amino acid and it is responsible for the observed increased factor IX activity per unit of dose achieved with AMT-061 as compared to its predecessor AMT-060 (Simioni et al, 2009).

This Phase III trial is to demonstrate the efficacy of AMT061 in terms of endogenous factor IX activity and annualized bleed rate (ABR), and to further describe its safety profile. The strong efficacy (protein expression) and safety results obtained during the Phase I/II trial with AMT-060 demonstrate 2×10^{13} gc/kg to be the selected dose for use in future trials. In addition, non-clinical data support the implementation of the prospectively defined product enhancement of AMT-060 (i.e. incorporation of the Padua mutation to form the new construct AMT-061) at a dose of 2×10^{13} gc/kg for this pivotal Phase III trial. The interim data analysis from the dose confirmation trial (CT-AMT-061-01) is now available and the dose for AMT-061 of 2×10^{13} gc/kg has been confirmed for the treatment phase of this Phase III trial.

1.2 This Statistical Methods Appendix (SMA)

This Statistical Methods Appendix (SMA) (for the Month 18 main CSR) outlines the statistical methods for the display, summary and analysis of data for the 26-week post-dose analysis, the 52-week post-dose analysis, and the 52-week analysis following the stable FIX expression (i.e. following 6 months post-treatment), and CSR addendum covering the long-term follow-up period. The SMA should be read in conjunction with the study protocol and the Statistical Analysis Plan (SAP. For the 26-week data cut, the subject-specific data cut-off date was the maximum of 31 Aug 2020 and the subject's month 6 visit end date. For the 52-week data cut, the subject-specific data cutoff date was the maximum of 28 Feb 2021 and the subject's month 12 visit end date.

2. STUDY OBJECTIVES AND ENDPOINTS

2.1 Study Objectives

2.1.1 Primary Objectives

The primary objectives were

- to demonstrate the non-inferiority of AMT-061 (2×10^{13} gc/kg) during the 52 weeks following establishment of stable FIX expression (months 6 to 18) post-treatment (AMT-061) follow-up compared to standard of care continuous routine factor IX prophylaxis during the lead-in phase, as measured by the ABR.

2.1.2 Secondary Objectives

The secondary objective was to demonstrate additional efficacy and safety aspects of systemic administration of AMT-061.

2.1.2.1 Secondary Efficacy Objectives

To investigate the effect of 2×10^{13} gc/kg AMT-061 on the following:

- Endogenous factor IX activity 6 months after a single AMT-061 treatment
- Endogenous factor IX activity 12 months after a single AMT-061 treatment

- Endogenous factor IX activity 18 months after a single AMT-061 treatment
- Annualized consumption of factor IX replacement therapy
- Annualized infusion rate of factor IX replacement therapy
- Discontinuation of previous continuous routine prophylaxis
- Trough factor IX activity
- Prevention of bleedings (comparison for superiority)
- Prevention of spontaneous bleeding
- Prevention of joint bleeding
- Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) – as a function of pre-investigational-medical-product (IMP) anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis)
- Correlation of pre-investigational medicinal product (IMP) anti-AAV5 antibody titers using the luciferase based neutralizing antibody (NAB) assay on factor IX activity levels after AMT-061 dosing
- Occurrence and resolution of target joints
- Proportion of subjects with zero bleeds during the 52 weeks following stable FIX expression (6-18 months) after AMT-061 dosing
- CCI [redacted]
- CCI [redacted]

CCI [redacted]

2.1.3 Safety Objectives

The safety objectives include evaluating the following:

- Adverse Events (AE)
- Changes in abdominal ultrasound
- Formation of anti-AAV5 antibodies (total immunoglobulin M and immunoglobulin G [IgM and IgG], neutralizing antibodies)
- AAV5 capsid-specific T cell response
- Formation of anti-FIX antibodies
- Formation of factor IX inhibitors and recovery

- Hematology and serum chemistry parameters
- AST and alanine aminotransferase (ALT) level increases and use of corticosteroids
- Shedding of vector deoxyribonucleic acid (DNA) in blood and semen
- Inflammatory markers: interleukin-1beta (IL-1β), interleukin-2 (IL-2), interleukin-6 (IL-6), interferon gamma (IFNγ), and monocyte chemotactic protein-1 (MCP-1)
- Alpha-fetoprotein (AFP).

2.2 Study Endpoints

2.2.1 Efficacy Endpoints

2.2.1.1 Primary Efficacy Endpoint

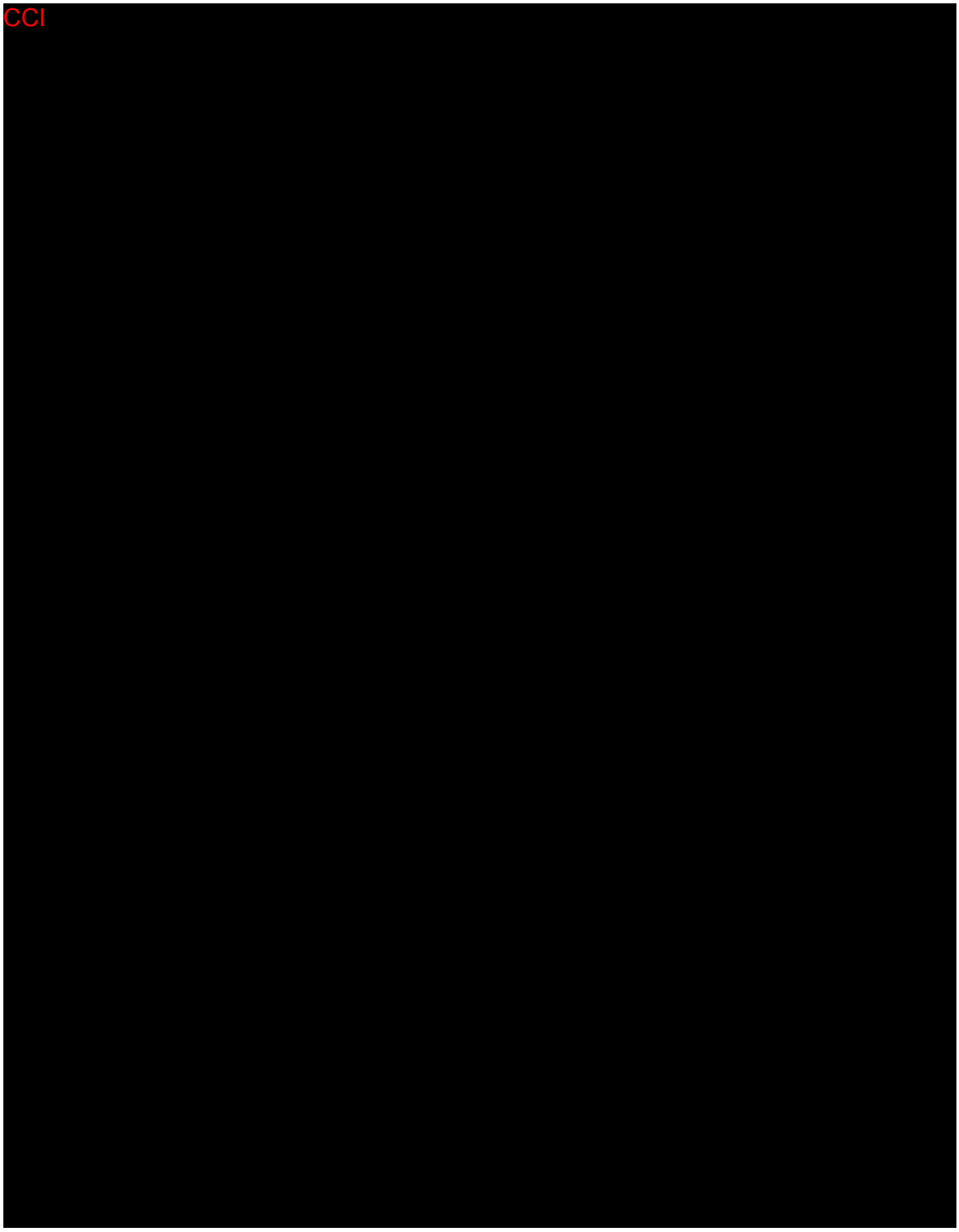
- ABR comparison between AMT-061 and prophylaxis for non-inferiority between the lead-in phase and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up.

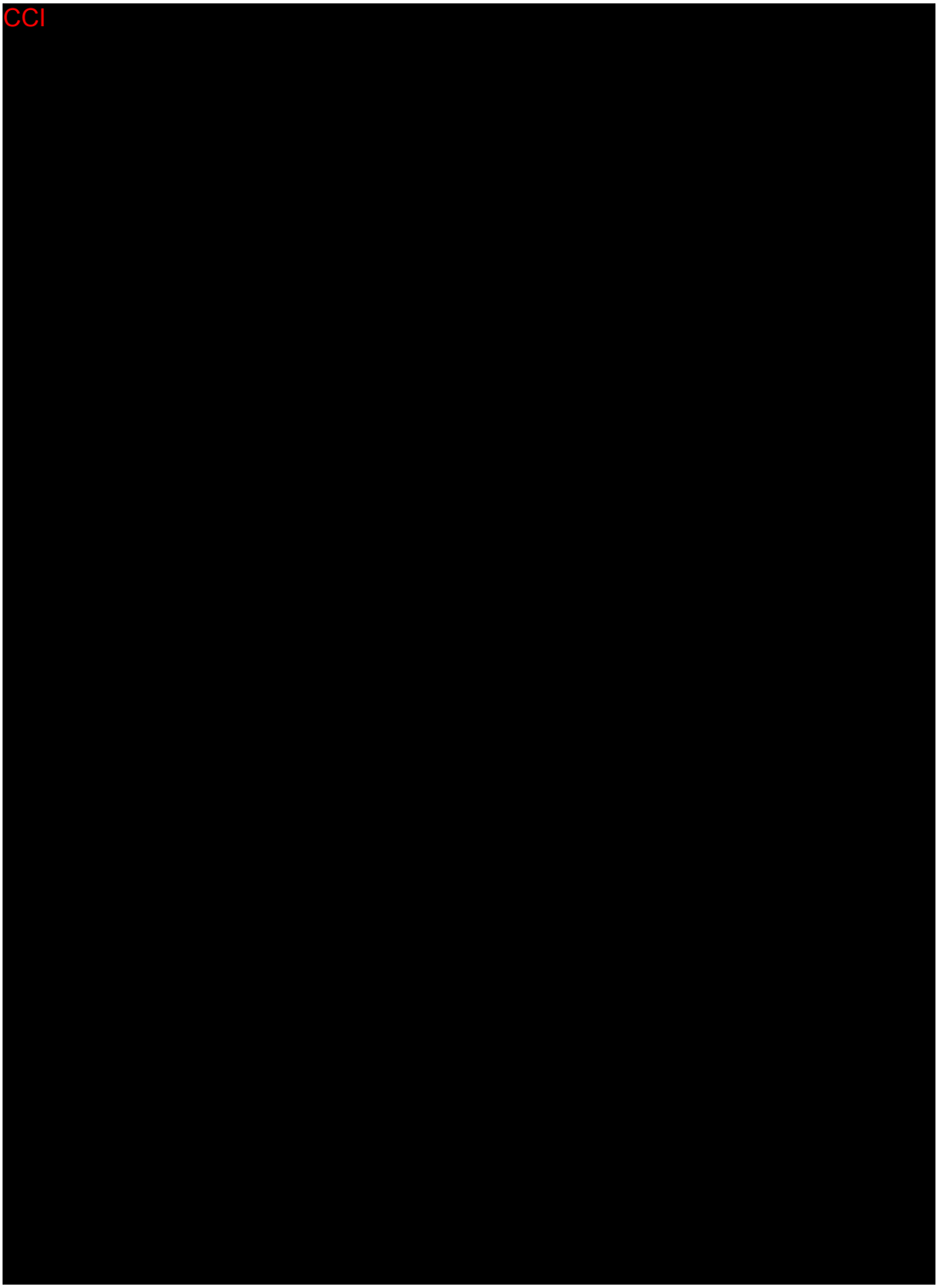
2.2.1.2 Secondary Efficacy Endpoints

- Endogenous factor IX activity at 6 months after AMT-061 dosing
- Endogenous factor IX activity at 12 months after AMT-061 dosing
- Endogenous factor IX activity at 18 months after AMT-061 dosing
- Annualized consumption of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase
- Annualized infusion rate of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase
- Proportion of subjects remaining free of previous continuous routine prophylaxis during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up
- Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 over the 52 weeks following stable FIX expression (6-18 months)
- ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up

- Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to lead-in phase
- Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase
- Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis)
- Correlation of factor IX activity levels during the 6-18 months post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay
- Occurrence of (and resolution of) new target joints during the 52 weeks following stable FIX expression (6-18 months) following AMT-061 dosing and resolution of pre-existing target joints following AMT-061 dosing
- Proportion of subjects with zero bleeds during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up
- CCI questionnaire scores from the CCI during the 12 months following AMT-061 dosing compared with the lead-in phase
- CCI questionnaire scores from CCI score during the 12 months following AMT-061 dosing compared with the lead-in phase.

CCI


CCI


CCI

CCI [redacted]

2.2.1.4 Sub-study Endpoints

- Patient Reported Outcomes, Burdens, and Experiences (PROBE) summary scores and individual item responses.
- Musculoskeletal Ultrasound Sub-Study Endpoints
    - Joint Tissue Activity and Damage Exam (JADE) scores.

2.2.2 Safety Endpoints

All adverse event (AE) data were collected from signing of the informed consent form until the end of the five-year follow-up. An AE, adverse drug reaction (ADR), and SAE were defined according to the ICH Guidelines E2A.

Safety analyses were based on the safety population and described as descriptive analyses. Safety set is defined as any study subjects receiving at least some amount of study treatment even when the full dose was not administered (including partial dose).

Safety data were analysed per study phase i.e., lead in phase, treatment emergent adverse events (TEAEs) during 18 months after study treatment start and during the follow up phase. The overall safety profile of AMT-061 was assessed given the below safety and tolerability criteria.

Secondary safety endpoints include the following:

- AEs
- Changes in abdominal ultrasound
- Anti-AAV5 antibodies (total [IgM and IgG], neutralizing antibodies)
- AAV5 capsid-specific T cells
- Anti-FIX antibodies
- Factor IX inhibitors and recovery
- Hematology and serum chemistry parameters
- ALT and AST levels, and corticosteroid use for ALT and AST increases
- Vector DNA in blood and semen
- Inflammatory markers:

- IL-1β
- IL-2
- IL-6
- IFNγ
- MCP-1

- AFP.

Other laboratory evaluations include coagulation and serology parameters. In addition, the following (S)AEs qualify for special notification as they are seen as safety issues of particular concern for Advanced Therapy (ATMP) (ENTR/F/2/SF/dn D(2009) 35810. Brussels, 03/12/2009) and gene therapy medicinal products (EMA / CHMP / Gene Therapy Working Party (GTWP) / 60436/2007):

- AEs related to the IMP administration procedure
- Suspected or confirmed cases of opportunistic or serious infections that in the investigator's opinion might be related to the IMP
- Unexpected reactions (e.g., hypersensitivity, immunological, toxic or other as consequence of a change in the construction or function of the viral vector [e.g., generation of replication competent virus])
- AEs related to product failure (including lack of efficacy), mandatory concomitant medication (e.g., immunosuppression), and medical devices which form part of the product or are used for application of the product
- Development of any new/recurrent cancer.

All TEAEs were tabulated displaying the number of subjects (and percentage) experiencing an event and the number of events by SOC and preferred terms within each SOC according to the Medical Dictionary for Regulatory Activities (MedDRA) terminology.;, TEAEs were also tabulated by severity and by relationship to trial medication, using frequency counts (number of subjects with event and number of events) and percentages. Similar tables were created for TEAEs leading to premature discontinuation or interruption, AESIs, deaths, seriousness, infusion related and hypersensitivity reactions if applicable. These summary tables are presented by decreasing frequency of occurrence based on SOC and Preferred Term.

The summary tables are accompanied by individual subject listings of all AEs, including information on AE number, actual AE description, date/time of start and end of AE, preferred term (MedDRA), SOC (MedDRA), severity, relationship/causality, type of AE, seriousness, and

outcome. Pre-existing AEs are flagged. Pre-existing AEs were not considered to be treatment-emergent, except in case of worsening during/after trial treatment (to be collected as separate AE). Separate listings were created for AEs for special notification, deaths, and SAEs, if applicable. Other safety data are presented using graphical displays, as applicable, descriptive statistics (including change from baseline, if applicable), and/or individual data listing.
The number of days until vector DNA could no longer be detected in semen and blood is tabulated. The number of days was calculated using the date of collection of the first of three consecutive negative samples for each matrix (The protocol should have said first of three instead of third of three).

3. STUDY DESIGN AND ANALYTICAL CONSIDERATIONS

3.1 Study Design

3.1.1 Overall Study Design and Plan

CT-AMT-061-02 is an open-label, single-dose, multi-center, multi-national trial, with a screening period, a lead-in phase, a treatment + post-treatment follow-up phase, and a long-term follow-up phase. Overviews of the trial and its design are presented in Figure 1 and Figure 2.

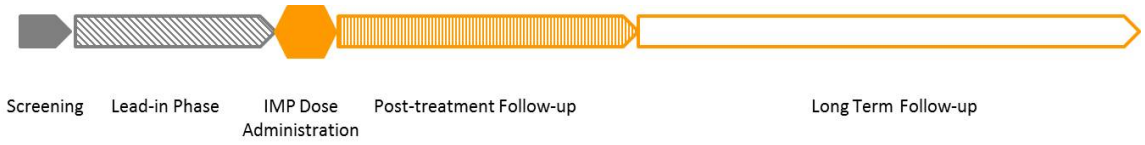


Figure 1: Study Overview

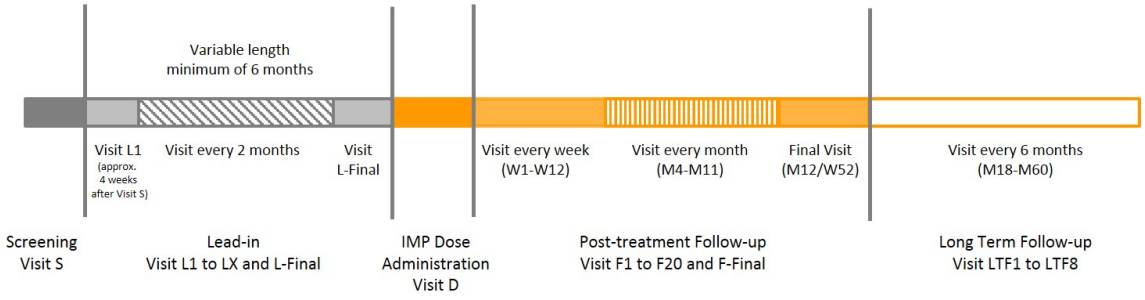


Figure 2: Study Design

Refer to the Protocol for the schedule of events for efficacy and safety evaluation for screening, lead-in, treatment and post-treatment, and follow-up visits. Refer to the Protocol for the schedule of events for laboratory parameters for screening, lead-in, treatment and post-treatment, and follow-up visits. Refer to the Protocol for the schedule of events for efficacy and safety

evaluation for long-term follow-up visits. Refer to the Protocol for the schedule of events for laboratory parameters for long-term follow-up visits.

After screening, eligible subjects entered a lead-in phase prior to the start of AMT-061 treatment. Visits occurred every 2 months during the lead-in phase with the final visit occurring a month prior to dosing. During the post-treatment follow-up, visits occurred weekly up to Week 12 and then monthly up to Month 12, after which subjects will enter a long-term follow-up with visits every 6 months.

Six months after IMP administration, the first secondary endpoint, endogenous factor IX activity at 6 months after IMP administration, was analyzed and reported via an interim analysis once the last subject had achieved 6 months after AMT-061 treatment. This assessment was based on clean data and a partially locked database.

Twelve months after IMP administration, the second secondary endpoint, endogenous factor IX activity at 12 months after IMP administration, was analyzed and reported via another interim analysis once the last subject had achieved 12 months after AMT-061 treatment. This assessment was based on clean data and a partially locked database.

After 52 weeks following stable FIX expression (18 months post-dose), all available efficacy and safety data collected between screening and 18 months post-treatment were analyzed and reported in a full CSR, including (but not limited to) the primary ABR endpoint and the third secondary endpoint. The first and second secondary efficacy endpoints (which will have been analyzed in their respective data cuts) are also added to the full (18-month) CSR. Data up to each analysis time point is considered locked and will not be changed (with the exception of ending dates and outcomes for continuing events and treatments) without explicit authorization. The subjects will be followed for an additional 3.5 years for evaluation of efficacy parameters and safety. At the end of that 3.5-year period, all safety and efficacy data will be reported in a CSR addendum covering the entire study duration, including the later 3.5-year period.

The overall trial participation will be approximately five and a half years.

3.2 Interim Analysis

There was an interim efficacy analysis of the first secondary endpoint, 6-month endogenous factor IX activity levels, after all subjects completed the 6-month assessment and the database is (partially) locked.

There was an interim efficacy analysis of the second secondary endpoint, 12-month endogenous factor IX activity levels, after all subjects completed the 12-month assessment and the database is (partially) locked.

3.3 Sample Size

The study sample size is constrained by the non-inferiority analysis of the primary endpoint, ABR.

Based on a literature search of trials in a similar clinical setting and the same underlying disease, as well as the previous AMT-060 Phase I/II trial, a non-inferiority margin of 1.8 is assessed for the rate ratio of ABR between AMT-061 (post-treatment) and factor IX prophylaxis (lead-in). For establishing the non-inferiority margin, an ABR of 2.4 between factor IX prophylaxis and placebo treatment has been assumed. Via simulation of ABR under a negative binomial distribution with a yearly rate of 2.4 events for lead-in and 1.9 for post-treatment, with a Pearson correlation of 0.05 for the number of events between the two periods, and with a common negative binomial dispersion parameter of 1.5, a sample size of N=50 will demonstrate non-inferiority with a non-inferiority margin of 1.8 and a power of 82.0%. Therefore, the study should consist of at least 50 analyzable subjects.

Given the sample size needed for ABR, this will produce a power >95% for the secondary statistical analysis of endogenous factor IX activity. For the secondary statistical analyses of factor IX activity at 6, 12, and 18 months, assuming a mean of 30.6 percent of normal (as observed at 6 weeks in study CT-AMT-061-01) and assuming a standard deviation of 6.97 (as observed at 6 weeks in study CT-AMT-061-01), assuming conservatively that the baseline factor IX activity is 2%, and assuming that the sample size is 50 subjects, for a one-sample t-test at the 0.025 one-sided level of significance to test whether the change from baseline is > 0, the statistical power is > 99%. Alternatively assuming that the standard deviation is 6.95, which is half of the range of factor IX activity values (23.9 to 37.8) observed at 6 weeks in study CT-AMT-061-01, the statistical power is still > 99%. The nQuery Advisor software was employed for this power calculation.

3.3.1 Non-inferiority Margin for Rate Ratio Analysis

Recent studies in similar Hemophilia B populations have shown the results – presented in the table below – relative to on-demand treatment. In the Idelvion publication, Group 2 (n=19) was followed for 6 months using on-demand treatment and then for 6 months using 7-day prophylaxis treatment. The ABR rates were estimated as 365.25 x (number of bleeding episodes)/ (number of days in the observed treatment period of interest). The rate reduction, 0.11, was estimated from a Poisson distribution, with a corresponding two-sided, 95% confidence interval (CI) for the rate reduction of (0.051, 0.238).

Table 1: Recent Results of Prophylaxis Compared to On-demand

Publication	ABR On-demand	ABR Prophylaxis	Rate Reduction
Alprolix (Powell, 2013)	18.67 (N=27)	3.12 (N=61)	0.17
Idelvion (Santagostino, 2016)	20.09 (N=19)	2.22 (N=19)	0.11
Nonacog (Collins, 2014)	15.58 (N=15)	40 IU: 2.51 (N=29)	0.16

Using a rate ratio analysis, as opposed to a difference in rates, results in an evaluation that is relatively independent of the magnitude of the baseline ABR. Thus, a rate reduction of 0.50 for a subject with 20 events during lead-in has the same meaning as for a subject with 4 events during lead-in. However, the difference in rates is quite different between such subjects (10 events and 2 events, respectively).

As currently proposed, the null and alternative hypotheses can be written as:

$$H_0: \frac{\lambda_{AMT-061}}{\lambda_{Prophy}} \geq M \; vs \; H_1: \frac{\lambda_{AMT-061}}{\lambda_{Prophy}} < M$$

where M represents the non-inferiority margin and λ_{XXX} represents the rate of events in the XXX group (AMT-061 representing the post-treatment period, Prophy representing the lead-in period). This can be equivalently rewritten as the difference between the rates on the natural logarithmic (base e) scale:

$$H_0: \log \lambda_{AMT-061} - \log \lambda_{Prophy} \geq \log M \;\; vs \; H_1: \log \lambda_{AMT-061} - \log \lambda_{Prophy} < \log M$$

If M1 represents the entire effect of prophylactic treatment compared to on-demand treatment on the natural logarithmic scale, then M1 = log(0.238) = -1.4354846, the upper limit of the CI from the Idelvion publication. If M2 represents the percentage of treatment effect relative to on-demand to be preserved, then M2 can be stated as p*M1 = p*log(0.238). Then, the hypotheses of interest can be restated as:

$$H_0: \frac{\lambda_{AMT-061}/\lambda_{on-d}}{\lambda_{Prophy}/\lambda_{on-d}} \geq M \; vs \;\; H_1: \frac{\lambda_{AMT-061}/\lambda_{on-d}}{\lambda_{Prophy}/\lambda_{on-d}} < M$$

or equivalently as differences on the natural logarithmic scale, where “on-d” represents on-demand treatment. Substitution of M1 in the denominator and M2 in the numerator gives the following equation:

$$(p-1)M1 \geq \log M$$

$$(1-p)x1.4354846 \geq \log M$$

This can then be solved for M or for p. This approach will be called Approach (1).

Alternatively, if M1 represents the entire effect of prophylactic treatment on the efficacy scale, then M1 = 1 – 0.238 = 0.762, using the relationship that efficacy = 1 – rate ratio, and M2 = pM1. Substituting 1-M2 into the numerator and 1-M1 into the denominator (RR = 1-efficacy), yields the following equation:

$$\frac{1 - p * 0.762}{1 - 0.762} < M$$

This approach will be called Approach (2).

The two approaches provide slightly different interpretations.

Table 2: Non-inferiority Margin (for the ABR rate ratio) for Retention of p%

	Percentage Retention				
	60%	**70%**	**75%**	**80%**	**90%**
Approach (1)	2.28	1.96	1.80	1.64	1.32
Approach (2)	1.776[a]	1.538	1.43	1.33	1.15
a. A margin of 1.8 using difference in log values corresponds to 59% retention of treatment effect.					

Hypothetical scenarios illustrating the upper limit of the non-inferiority margin under observations of varying pre-treatment annualized bleed rates are shown in the table below.

Table 3: Relation of pre-treatment ABR and NI margin using a 1.8 rate ratio

ABR observed during lead-in phase	2.0	2.5	5.0	10
Maximum ABR post-treatment to maintain NI (bleeds)	<3.6	<4.5	<9.0	<18
Permissible increase in number of bleeds	<1.6	<2.0	<4.0	<8

3.4 Randomization and Blinding

Not applicable, as this is an open-label trial with one treatment arm.

4. DATA DEFINITIONS AND PRE-PROCESSING

4.1 Baseline Definitions

The baseline factor IX activity was imputed based on the patient's historical hemophilia B severity that is documented on the CRF. If the patient has documented severe factor IX deficiency (factor IX plasma level <1%) their baseline factor IX activity level was imputed as 1%. If the patient has documented moderately severe factor IX deficiency (factor IX plasma level >= 1% and <= 2%) their baseline factor IX activity level was imputed as 2%.

The baseline ABR was the number of bleeds in the previous year as assessed at screening.

Baseline age is the age in years at the time of the Screening Visit.

For CCI [redacted] and CCI that are assessed at visits, the Baseline value was period-specific. For the lead-in period, the Baseline value was the latest value prior to the lead-in period that is not within 2 weeks of a bleed. For the post-treatment period, the Baseline value was the latest value prior to IMP that is not within 2 weeks of a bleed. Note that the Baseline value was not used as a covariate in statistical-modeling-based treatment-period-comparative analyses for this study, because each subject serves as one's own control.

For vital signs and safety laboratory values, Baseline was period-specific. The Baseline value for the lead-in period was the last non-missing central laboratory value or vital signs value on or prior to Visit L1. The Baseline value for the post-treatment period was the last central laboratory value or vital signs value prior to the first dose of AMT-061.

For height, weight, and BMI, the baseline value was the last value prior to the start of the lead-in period.

4.2 Data Handling Rules and Definitions, Including Handling of Missing Data

Missing data were maintained as missing in the safety and efficacy datasets, unless specified otherwise.

Data for Laboratory Summaries (Continuous Parameters)

Data from unscheduled visits was not used for by-visit summaries (unless they have been assigned to a scheduled visit according to the Time Windows for Statistical Analysis). Data from both scheduled and unscheduled visits were used for determining incidence of clinically significant values.

Data for All Laboratory Summaries

Definitions are provided in Appendix 1 Data Handling Rules.

Study Dates and Day of Assessment or Event

Study Day and Day of Assessment or Event definitions are provided in Appendix 1 Data Handling Rules.

Duration of Event

Definitions are provided in Appendix 1 Data Handling Rules.

Distance between Events

Definitions are provided in Appendix 1 Data Handling Rules.

The Use of Lead-In Period Month 6 Visit versus Lead-In Final Visit for Analyses of Factor IX Activity

For factor IX activity, for the Lead-In Month 6 assessment and the Lead-In Final assessment, some specific instructions are as follows. For factor IX activity (aPTT and chromogenic from the central laboratory) for analysis purposes if a Lead-In Month 6 value is available, then it was used (for analysis); otherwise if a Lead-In Final value was available, then it was used for the purpose of (i.e. as if it were) the Month 6 value. The rationale is that (1) Lead-In Month 6 is a planned assessment (for factor IX activity) and that in the presence of a Month 6 assessment, a Lead-In Final value is not essential for analysis and (2) the planned duration of the Lead-In period is approximately 6 months.

Missing Data

If causality is missing for a TEAE, the TEAE was regarded as 'Related'. If causality is missing for an AE with onset before administration of trial drug, the AE was regarded as 'Not related'. If the intensity is missing, the intensity of the AE was regarded as "Severe." In the case where seriousness is missing, this was queried. Seriousness cannot be imputed as 'Yes' by default, since this would affect the reconciliation between trial database and registry of SAEs.

Time Windows for Data Collection:

The below assessment time windows (in Table 4) are given in the protocol as "target" time windows for assessments to be carried out. However, these were not the time windows to be applied for statistical analysis, which are described farther below.

Table 4: Protocol-Recommended Study Time Intervals for Collection of Efficacy and Safety Evaluation and Laboratory Parameters

Nominal Time for Visits or Assessment	Protocol-Recommended Time Interval for the Visit or Assessment
Screening	Approximately -28 days prior to Visit L1
Lead-In	
Visit L1 (L-W0)	0 days
Visit L2 to LX (every 2 months, starting at L-W8)	±14 days
Visit L-Final	-28 days (± 7 days) from Visit D
IMP Dose	
Post-IMP 3 hours	±15 minutes
Post-treatment Follow-Up	
Week 1 to Week 12	±2 days
Month 4 to Month 11	±5 days
Month 12/Week 52	±5 days
Long-Term Follow-up	
Month 18	±2 weeks
Month 24	±2 weeks
Month 30	±2 weeks
Month 36	±2 weeks
Month 42	±2 weeks
Month 48	±2 weeks
Month 54	±2 weeks
Month 60	±2 weeks

Time Windows for Statistical Analysis:

Scheduling difficulties due to the coronavirus disease (COVID-19) pandemic resulted in an increased number of missed, delayed, or unscheduled visits. Scheduling difficulties also result in the performance of assessments at a scheduled visit where performance of the assessment was not originally planned. As an action to mitigate risk, analysis windows utilized such unplanned assessments as follows. A schema for the assignment of such unplanned assessments to scheduled time points for visit-based endpoint analysis and summary (for visit-based efficacy and safety endpoints) was defined as follows:

- An unplanned assessment was assigned to a scheduled visit only if that visit had a missing value for the relevant endpoint and the visit was a scheduled time point for the performance of the assessment per the study protocol.

- Analysis windows for the assignment of unplanned assessments ranged from the previous visit at which the endpoint was planned to be collected to the next visit at which the endpoint was planned to be collected.

- The unplanned assessment closest in time within the analysis window (either before or after) was used to replace a missing assessment for a scheduled visit. If two unplanned assessments were both the closest in time, with one being before and the other being after, the earlier assessment was used.

- For efficacy endpoints, values obtained on a post-treatment actual study day < Day 21 were not candidates to be used for imputing missing values for post-treatment visits at a nominal visit time subsequent to post-treatment Study Day 21.

- Only values from unplanned assessments (not planned assessments) were used to replace a missing scheduled assessment.

- An unplanned assessment may be used more than once provided it lay within the analysis window for two consecutive scheduled time points at which the assessment was not performed as planned.

- Analysis windows were defined separately for the lead-in and post-treatment periods. This means that values were not carried from one treatment period to the other.

- For laboratory-based efficacy assessments, the above instructions (in the first 8 bullet points above) are applicable pertaining to all planned assessments; except, however, for the Lead-In Month 6 assessment and the Lead-In Final assessment, some specific instructions were as follows. For relevant laboratory-based efficacy endpoints – i.e. factor IX activity (aPTT and chromogenic) – and for the laboratory endpoints of Total (IgM and IgG) and Neutralizing Antibodies to AAV5 –, eligible unplanned assessment values were used (as applicable) to replace a missing assessment for the Lead-In Month 6 Visit (but

not used to provide a value for the Lead-In Final Visit). The rationale is that Lead-In Month 6 is a planned assessment and that – for such endpoints – in the presence of a Month 6 assessment, a Lead-In Final value is not essential for analysis. For these endpoints, the analysis window for the assignment of unplanned assessments to the Lead-In Month 6 Visit ranged from the previous visit at which the endpoint was planned to be collected (i.e. the planned date of the Lead-In Month 4 Visit) to the actual Day before AMT-061 dosing. If both the Lead-In Month 6 Visit and the Lead-In Final Visit had no value (for the endpoint), then an unplanned-assessment value was sought to be assigned to the Lead-In Month 6 Visit; if either the Lead-In Month 6 Visit or the Lead-In Final Visit has a value (for the endpoint), then no such assignment (of an unplanned assessment to Lead-In Month 6) was needed. Unscheduled-visit values should never be assigned to the Lead-In Final Visit.

- For CCI efficacy endpoints, the above instructions (in the first 8 bullet points above) are applicable pertaining to all planned assessments; however, for certain CCI some specific instructions for the imputation of the Lead-In final value are provided as follows. For CCI efficacy endpoints that were scheduled to be collected at the Lead-In Month 4 and Lead-In Final Visits – i.e. CCI, CCI, CCI, CCI, CCI, CCI, and PROBE – eligible unplanned assessment values were used (as applicable) to replace a missing assessment for the Lead-In Final Visit. For these endpoints, the analysis window for the assignment of unplanned assessments to the L-Final Visit ranged from the previous visit at which the endpoint was planned to be collected (i.e. the planned date of the Lead-In Month 4 Visit) to the actual Day before AMT-061 dosing. The latest available eligible unplanned assessment will be assigned to the Lead-In Final Visit. By the way, the Lead-In Month 6 Visit was not a planned assessment time and therefore was not to be receiving values from unplanned assessments.

- For CCI efficacy endpoints, the above instructions (in the first 8 bullet points above) are applicable pertaining to all planned assessments; however, for certain (other) CCI some specific instructions for the imputation of the Lead-In final value are provided as follows. For visit-based efficacy and safety endpoints that were scheduled to be collected for the first time (post-screening) at the Lead-In Final Visit – i.e. MSKUS (efficacy) and Abdominal Ultrasound (safety) – eligible unplanned assessment values were used (as applicable) to replace a missing assessment for the Lead-In Final Visit. For these endpoints, the analysis window for the assignment of unplanned assessments to the L-Final Visit ranged across the entire duration of the Lead-In Period. The latest available eligible unplanned assessment was assigned to the Lead-In Final Visit. By the way, the Lead-In Month 6 Visit was not a planned assessment time and therefore was not to be receiving values from unplanned assessments.

Unscheduled or unplanned assessment values were not -- per the SAP text -- required to be assigned to scheduled (analysis) visits for vector DNA (genome) assessments. However, after it

was decided to add some informative by-visit displays of vector DNA results, analysis time windows were additionally applied for vector DNA results.

Any other rules for missing data handling are given in the endpoint-specific sections.

4.3 Bleed Counting Rules

Bleeds were counted irrespective of assessments by the investigator as to the trueness or newness of the bleed (except for a small number of designated sensitivity analyses).

For designated supportive analyses, only exogenous-factor-IX-treated bleeds were counted. If the field for whether the bleed was treated was missing, then (for conservativeness) it was assumed that the bleed was treated with factor IX.

For a small number of designated sensitivity analyses, only bleeds that were assessed to be new and true bleeds were counted. If the assessment field for newness of the bleed was missing, then (for conservativeness) it was assumed that the bleed was new. If the assessment field for trueness of the bleed is missing, then (for conservativeness) it was assumed that the bleed was true. The rationale for these sensitivity analyses is given in the following bullet points:

- The occurrence of bleeds during the conduct of the study were self-reported by enrolled subjects through the daily use of electronic and paper diaries, up to the 52 Week post-treatment visit and from the 52 Week visit until the five year visits, respectively. In some cases, due to previous chronic joint bleeds and damage, the patient may have experienced pain, mistaken it for a new bleed, and self-infused factor IX. Investigators reviewed and assessed reported bleeds as a means of verifying that patient-reported events met the clinical criteria required to be characterized as new, true bleeds. The steps in the bleed reporting and assessment process are presented below.

- The patient reported signs and symptoms in a daily e-diary up to the Week 52 visit, and in paper diaries through to the end of the study.

- The Principal Investigator or designee reviewed the diary data and, if needed, requested further information from the patient prior to their evaluation of the signs and symptoms.

- The Principal Investigator or designee evaluated the signs and symptoms reported in the diary and/or during discussions with the patient and assessed whether the reported event was a true bleed and whether the reported event was a new bleed. For example, based on such sign and symptom evaluation, the investigator in some cases may have needed to distinguish whether there was a new bleed or whether the patient was experiencing pain (due e.g. to previous chronic joint bleeds and damage) that was not really a new bleed.

- When patients were next at the study site, the physician may have elected to use a diagnostic scan (X-ray, ultrasound, MRI, CT scan, etc.) to confirm the presence of blood or signs of acute inflammation. Blood or signs of acute inflammation observed using one

or more of these confirmatory methods coupled with the physician's assessment served as sufficient confirmation to identify an event as a true bleed.

5. DATA AND ANALYTICAL QUALITY ASSURANCE

The overall quality assurance procedures for the study data, statistical programming and analyses are described in Standard Operating Procedures (SOPs) of Everest Clinical Research. Detailed data management procedures are documented in the study Data Management Plan, Data Validation Check Specifications, and Integrated Safety Data Review Plan.

6. ANALYSIS POPULATIONS

6.1 Population Definitions

6.1.1 Screen Failures

The screen failure population included all subjects who were screened but never entered the lead-in period.

6.1.2 Lead-in Discontinuers

The lead-in discontinuers population included all subjects who entered the lead-in period but discontinued from the study prior to AMT-061 dosing.

6.1.3 Safety Population

The lead-in safety population consisted of all subjects enrolled into the lead-in period. The post-treatment safety population consisted of all subjects who receive AMT-061, irrespective of any protocol deviations. Period-specific safety tabulations use the period-specific safety population for the "N" and denominator (for percentages). The safety population consisted of all subjects who were in either the lead-in safety population or the post-treatment safety population.

6.1.4 Full Analysis Set (FAS)

The FAS included all subjects who enrolled, entered the lead-in phase, were dosed with AMT-061, and provided at least one efficacy endpoint assessment for any efficacy endpoint subsequent to AMT-061 dosing. The FAS population was the primary population for all statistical analyses.

6.1.5 Per-Protocol Population

The PP population included all subjects from the FAS population who adhered to a stable and adequate prophylaxis use during the lead-in phase, who completed at least 18 months of efficacy assessments (52 weeks after achieving stable FIX expression) for the 18-month (data cut) analysis, who completed at least a full year of efficacy assessments for the 12-month (data cut) analysis, or who completed at least 6 months of efficacy assessments for the 6-month (data cut) analysis, and who had no major protocol deviations that impact the interpretation of efficacy. The PP population was used for sensitivity analyses. Protocol deviations that impacted the

interpretation of efficacy include the unwillingness to discontinue continuous prophylaxis use after receipt of AMT-061. At the interdisciplinary evaluability/data review meeting prior to database lock, it was determined that the only actual situation or protocol violation meriting exclusion from the PP Population was a subject's having received only a partial dose of AMT-061.

6.1.6 Subjects with Baseline NAb Titer < 3000

Subject PPD had very high NAb titers between Screening and Visit D, ranging from 2020 to 3291 (the next such highest NAb titer for a subject was 891). As such, additional analyses were run in some cases, excluding subject PPD. These are indicated for some of the endpoints described below.

7. STATISTICAL ANALYSIS

7.1 Subject Disposition

A disposition table for CT-AMT-061-02 for all subjects is provided. This tabulation includes the number of subjects who were screen failures, not treated with AMT-061 (i.e. lead-in discontinuers), who were prematurely discontinued from treatment (treated with only a partial dose of AMT-061), who received the full dose of study treatment, who withdrew early from the study post dose of AMT-061, and who completed the study. The number and percentage of subjects included in the FAS, PP, lead-in safety population, and post-treatment safety population are also tabulated. The number and percentage of subjects in the PROBE sub-study and the MSKUS sub-study are tabulated. A subject was considered to be in the PROBE sub-study if the subject had at least one post-treatment assessment of PROBE. A subject was to be considered to be in the MSKUS sub-study if the subject had at least one post-treatment assessment of MSKUS. Because MSKUS expert-reader interpretations were not available in time for the database lock, MSKUS data were left out of the CSR post-text tables, listings, and graphs for this Month 18 CSR.

The reason for exclusion from the FAS, PP, lead-in safety population, and post-treatment safety populations was summarized. Reasons for premature discontinuation from study treatment were summarized for the post-treatment safety population (this would include any partial dosing).

The data on subject disposition, missed visits, and protocol deviations (including those related to COVID-19) are listed.

7.2 Demographic and Baseline Characteristics

Descriptive statistics of demographics and baseline characteristics are presented for the FAS, PP, lead-in safety population, and post-treatment safety populations. For quantitative variables, all summaries include the number of non-missing observations, mean, standard deviation (SD), first quartile (Q1), median, third quartile (Q3), minimum, and maximum. For the qualitative variables, the summaries include the number and percentage of subjects in each category or level. All data are included in listings.

7.2.1 Demography

Demographics collected at Screening include year of birth (i.e., age at Screening Visit), race, ethnic group, and gender according to local regulations. According to inclusion criterion 1, all patients are males.

7.2.2 Baseline Disease Characteristics

Baseline disease characteristics included duration of disease, endogenous factor IX activity level at time of diagnosis, severity of disease, indicator of family history of hemophilia B disease, number of bleeds in the year prior to screening (total, spontaneous, traumatic, joint, and unknown), and the type of factor IX therapy used. The baseline disease characteristics were tabulated according to the information collected on the electronic Case Report Form (eCRF).

Severity of hemophilia B was categorized as severe (factor IX plasma level < 1%) or moderately severe (factor IX plasma level >= 1% and <= 2%).

7.2.2.1 Hemophilia B History

All hemophilia B history data are listed, and the listing includes the following: date first presented symptoms, date of initial diagnosis, duration of disease, endogenous factor IX activity level at diagnosis (if available), severity of hemophilia B at time of diagnosis, number of factor IX exposure days (an exposure day is defined as a day when the subject received at least one injection of factor IX treatment), and family members with a history of factor IX inhibitors.

7.2.2.2 Medical and Surgical History

All medical and surgical history are listed, including the following information: surgical or medical history event, start date and end date or current status. Medical history was coded using the most recent version of the MedDRA at the time of the database lock.

7.2.2.3 Target Joints at Screening

Target joints were defined as joints with three or more spontaneous bleeds into a single joint within a consecutive six-month period. Once there have been ≤ 2 bleeds into the joint within a consecutive 12-month period, the joint was no longer considered a target joint and the target joint was then considered to have resolved. Target joints at screening are listed.

7.2.2.4 Baseline Antibody Parameters

The baseline antibody parameters include anti-FIX antibody titer levels, the presence of factor IX inhibitors, total (IgG and IgM) antibodies to AAV5, neutralizing antibody levels to AAV5, and AAV5 capsid-specific T-cells. These data are listed.

Box and whisker plots over post-treatment study time are also produced for these parameters, if applicable.

Please refer to Appendix 1 Data Handling Rules on how special laboratory values were handled in quantitative analyses.

7.2.2.5 FIX Gene Mutation

Factor IX gene sequence analyses were performed for all subjects who provided consent during the Screening Visit, even if they already had factor IX gene mutation information available. The data were presented in a listing.

7.2.2.6 Prior and Concomitant Medications

Prior and concomitant medications were collected and coded using the most recent World Health Organization (WHO) drug dictionary at the time of database lock. Prior and concomitant medications are listed.

Prior medications were defined as those treatments with a start date before Visit L1 for the lead-in period. A medication/therapy was identified as a "lead-in" concomitant medication/therapy if it was being continued by the subject at the date of the L1 Visit or was any new medication/therapy received during the lead-in period prior to the date of AMT-061 dosing. A medication/therapy was identified as a "post-treatment concomitant" medication/therapy if it was being continued by the subject at the date of AMT-061 dosing or was any new medication/therapy received during the post-treatment period. A medication with end date that was the same as the AMT-061 dosing date was not considered to be "post-treatment concomitant".

7.2.2.7 Prior Factor IX Therapy Use

Factor IX therapy use during the year prior to screening was summarized and listed. Factor IX therapy use during the screening period is listed.

7.3 Investigational Product Exposure

A listing for exposure to investigational product is provided showing the date of exposure and dose received. The listing also states whether the full dose was received.

Also, the time of subject's routine factor IX product/dose, incremental recovery, maximum concentration, and identity of subject's routine factor IX product/dose is listed.

7.4 Blood Sample for Future Research

Data regarding the additional blood samples drawn at Screening (Visit S), Baseline (Visit D pre-IMP), Visit F12 (Week 12), and Visit F-Final (Month 12/Week 52), for the purpose of potential future research in the hemophilia B disease area are listed with the corresponding informed consent date.

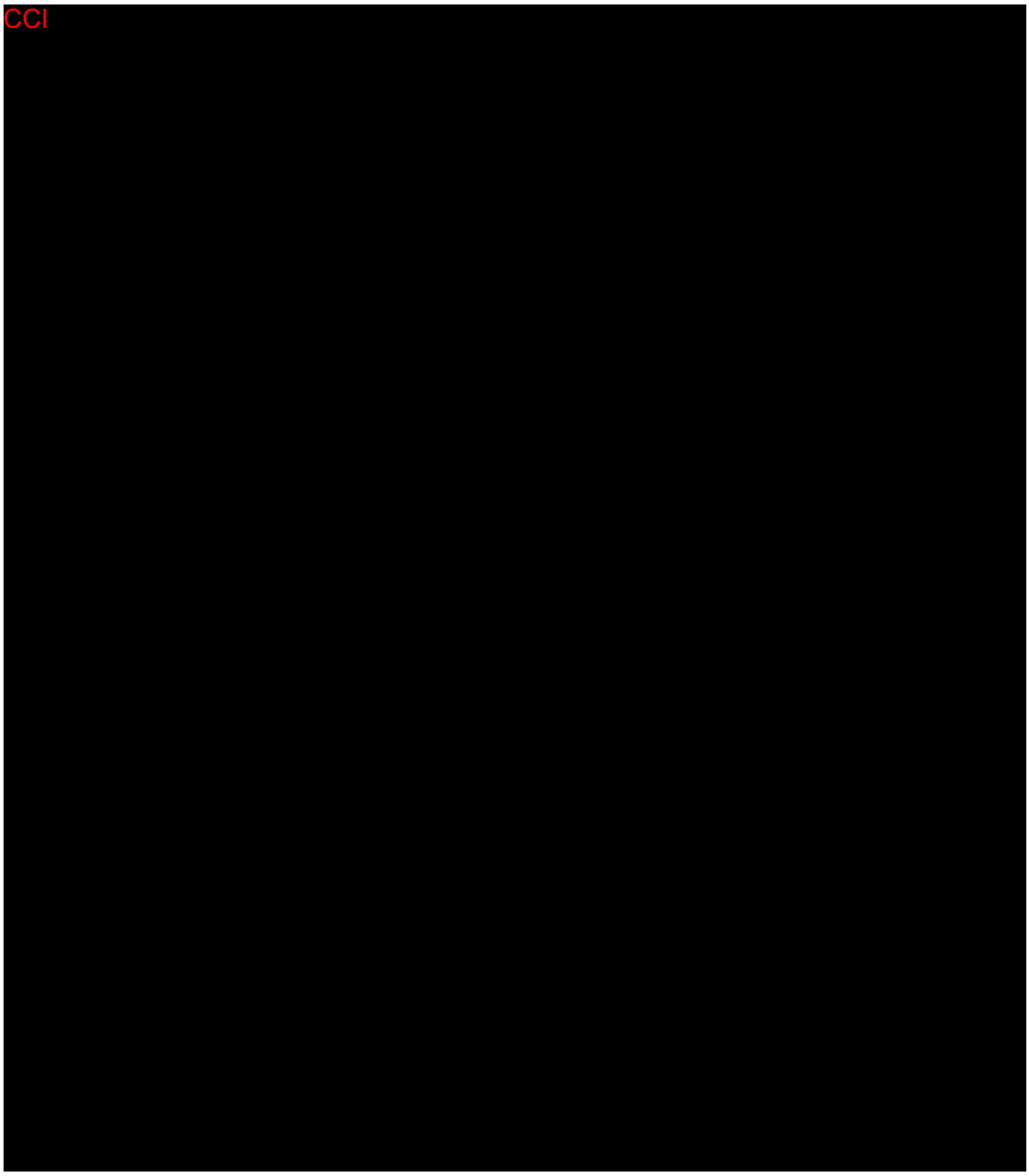
7.5 Efficacy Analyses

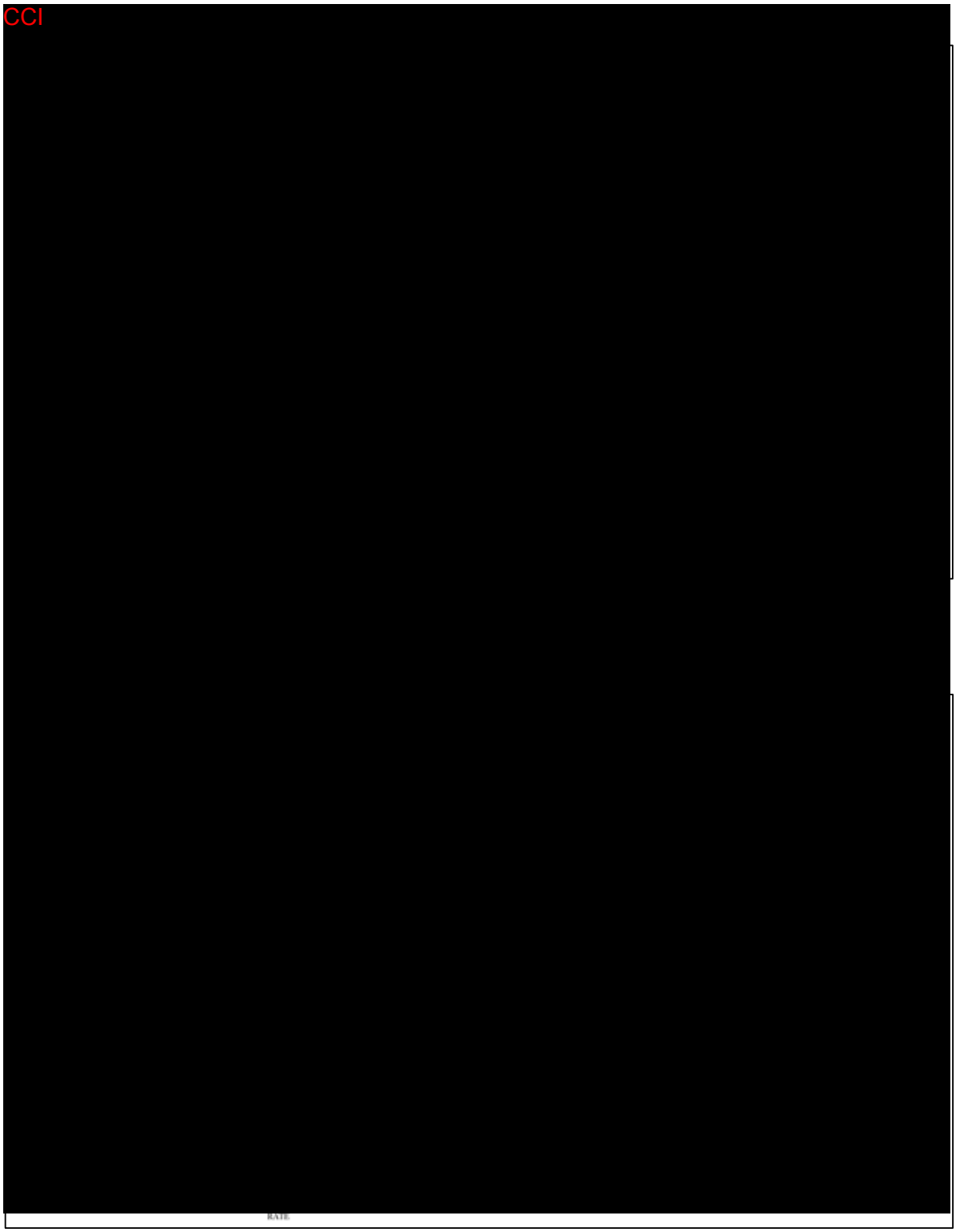
7.5.1 Primary Endpoint

The primary efficacy endpoint is as follows:

- ABR comparison between AMT-061 and prophylaxis for non-inferiority between the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up and the lead-in phase.

The analysis of the primary endpoint is discussed below.


CCI


CCI

CCI

7.5.2.1 Sensitivity Analysis 1: PP Population

A sensitivity analysis using the PP population was performed. With the exception of the study population, the analysis undertaken was identical to the main analysis described in Section 7.5.2. The results *(Table 2.2.1.3)* were consistent with those for the FAS population *(Table 2.2.1.1)*.

7.5.2.2 Sensitivity Analysis 2: Including (not excluding) Periods Subsequent to Exogenous Factor IX Use

A second sensitivity analysis, using the FAS, was conducted for ABR to evaluate the robustness of the analysis findings to the inclusion (i.e. non-exclusion) of time intervals with exogenous factor IX use during the post-treatment period *(Table 2.2.1.5)*. In this analysis, person-time during the post-treatment period that was within 5 half-lives subsequent to exogenous factor IX use was not excluded from (i.e. was included in) the time at risk for a bleeding event. However, as with the primary ABR analysis, bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) (and prior to Day 21 for pre-Month 18 data cuts) post-AMT were not included in the analysis.*)*.

7.5.2.3 Sensitivity Analysis 3: Bleeds Treated with Exogenous Factor IX

A third sensitivity analysis repeated the main analysis for ABR using the FAS while considering only bleeds treated with exogenous factor IX. Bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) (and prior to Day 21 for pre-Month 18 data cuts) post-AMT were not included in the analysis. (*Tables 2.2.1.6.1, 2.2.1.6.2*, for FAS and PP populations respectively)

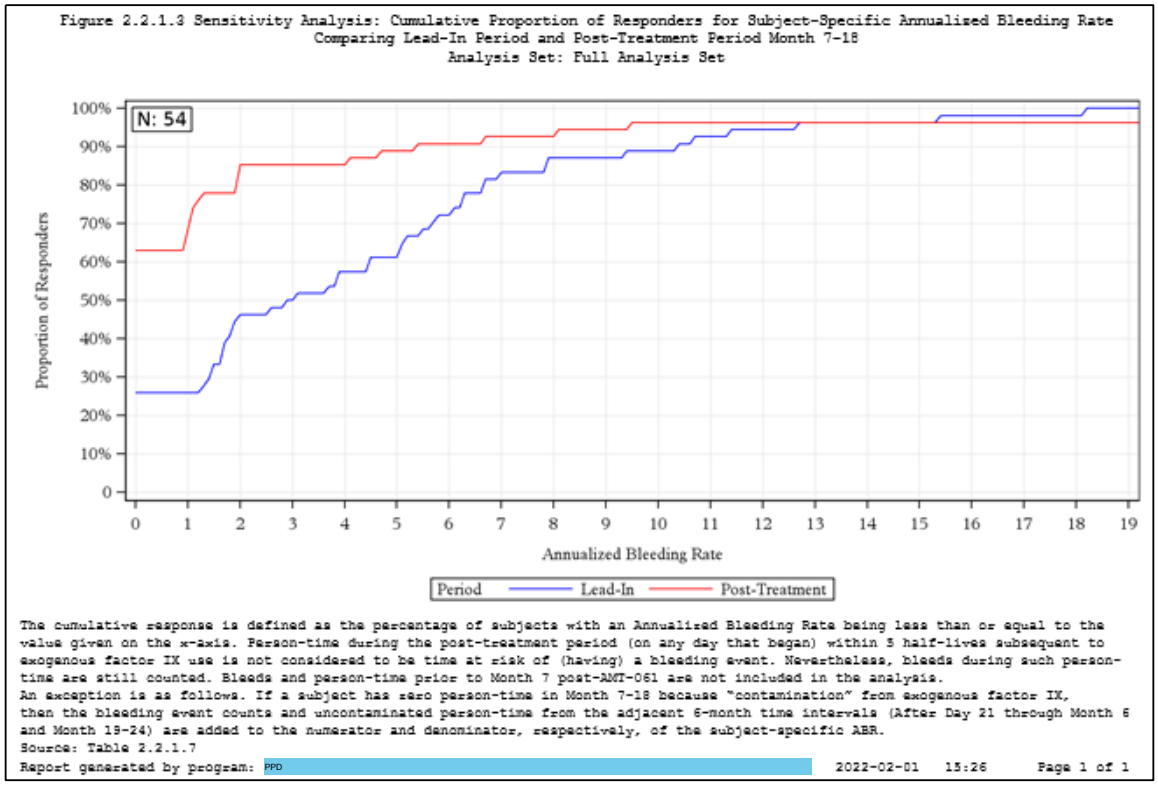
7.5.2.4 Sensitivity Analysis 4: Cumulative Responder Analysis using Subject-Specific Bleeding Rates

To provide a descriptive summary of per-subject annualized bleeding rates in the respective treatment periods, a cumulative responder analysis (as described in Farrar 2006) was performed characterizing ABR in the lead-in period and the post-treatment period *(Table 2.2.1.7)*. This analysis used the FAS. The observed ABR for the post-treatment period was plotted on the *x*-axis, and the proportion of "responders" (subjects that have an ABR equal to or less than the level specified) was plotted on the *y*-axis *(Figure 2.2.1.3)*. The same graph plotted the observed ABR across the lead-in period (also) on the *x*-axis, with the proportion of "responders" (subjects that have an ABR equal to or less than the level specified) plotted on the *y*-axis. Thus, a cumulative distribution plot was produced where the proportion of responders could be compared by treatment period across a continuous range of ABR values.

The ABR for the post-treatment period was calculated for the 52 weeks following stable FIX expression [Month 7-18]. As with the main ABR analysis, person-time during time intervals "contaminated" with exogenous factor IX – according to the 5-half-life contamination rule – were excluded from the denominator for the subject-specific bleeding rates, but bleeds during

such time intervals were still counted and included in the numerator for the rate. However, as with the primary ABR analysis, bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) post-AMT were not included in the analysis.

Figure 5: Annualized Bleeding Rate Cumulative Proportion of Responders – Lead-In and Post-Treatment Periods



7.5.2.5 Sensitivity Analysis 5: New and True Bleeds

A fifth sensitivity analysis repeated the main analysis for ABR using the FAS population while considering only bleeds assessed to be new and true by the investigator *(Table 2.2.1.10)*. Bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6 post-AMT) were not included in the analysis.

7.5.2.6 Sensitivity Analysis 6: New and True Bleeds Treated with Exogenous Factor IX

A sixth sensitivity analysis repeated the main analysis for ABR using the FAS population while considering only bleeds treated with exogenous factor IX that are assessed to be new and true by the investigator *(Table 2.2.1.11)*. Bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6 post-AMT) were not included in the analysis.

7.5.2.7 Sensitivity Analysis 7: Excluding Periods Contaminated by Systemic Corticosteroid Exposure

A seventh sensitivity analysis repeated the main analysis for ABR using the FAS population and was conducted to evaluate the robustness of the analysis findings to exclusion of time intervals with systemic corticosteroid use during the post-treatment period *(Table 2.2.1.12)*. In this analysis, person-time during the post-treatment period (that is) during systemic corticosteroid use or within 5 half-lives subsequent to exogenous factor IX use or systemic corticosteroid use was excluded from the time at risk for a bleeding event. However, as with the primary ABR analysis, bleeds and person-time on or after Day 1 and prior to stable FIX expression (Month 6) post-AMT was not included in the analysis.

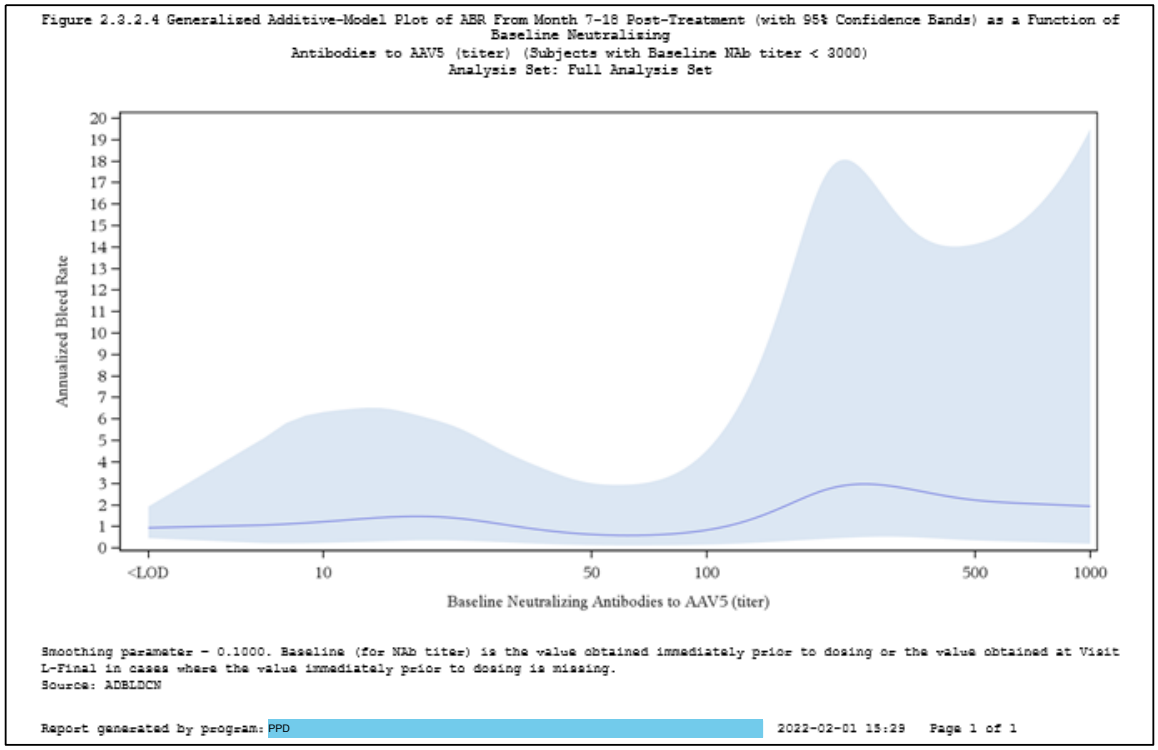
7.5.2.8 Sensitivity Analysis 8: Subjects with Baseline NAb titer < 3000

An eighth sensitivity analysis repeated the main analysis for ABR using the FAS population, and was conducted to evaluate the robustness of the analysis to exclusion of subjects with unusually large baseline NAb (neutralizing antibody) titers at post-treatment baseline *(Table 2.2.1.13)*.

One subject, #PPD, had a baseline titer > 3000 (titer of 3212) *(Listing 3.4.1)*, while the next highest titer in a subject was 678. Most subjects had a baseline titer < LOD (lower limit of detection). Subject #PPD had no bleeds during the Lead-In period and 5 bleeds during Months 7-18 post-treatment (days 233, 274, 366, 381, and 417), all assessed as "new" and "true" by the investigator *(Listing 2.2.1.3)* (incidentally, 3/54 subjects had more than 5 bleeds during the post-treatment period). For Months 7-12 there was only 1 day at risk of a bleeding episode, for Months 7-18, also only 1 day at risk, resulting in a subject-specific annualized bleeding rate of 1673.97 for Months 7-18, the highest among the 54 subjects in the FAS *(Listing 2.2.1.5.1)*.

Results for the adjusted ABR excluding this subject *(Table 2.2.1.13)* produced a slightly smaller rate ratio (Post-Treatment/Lead-In), and remained statistically significant (p<0.0001). This subject did have an effect on the generalized additive-model plots of ABR for Months 7-18 as a function of baseline NAb titer. In the analysis including this subject, the estimated ABR is consistently low up to a NAb titer of roughly 300, and increases substantially thereafter *(Figure 2.3.2.2)*. The dramatic rise in model-estimated ABR can be attributed to a single subject as follows. If subject #PPD is excluded, the estimated ABR is consistently low for NAb titers up to 1000 *(Figure 2.3.2.4; see also Figure 6 below)*. Scatterplots are also provided (*Figures 2.3.2.6, 2.3.2.7*). Figures as a function of Lead-In-Final-Visit NAb are also provided *(Figures 2.3.2.3, 2.3.2.5, 2.3.2.8)*.

Figure 6: GAM Plot of ABR from Month 7-18 post-treatment vs baseline NAb to AAV5 (Subjects with Baseline NAb < 3000)



The smoothing parameter value of 0.1 was chosen to attain a good balance between smoothness and granularity. The solid line in the graph shows the estimated ABR. The height of the grey area represents the degree of uncertainty in the estimate.

Additional tables for ABR for NAb titer < 3000 are available for the PP Population *(Table 2.2.1.14)*, the cumulative proportion of responders *(Table 2.2.1.15)*, excluding person-time contaminated with corticosteroids *(Table 2.2.16),* and for bleeding sub-types for the FAS and PP populations *(Tables 2.2.1.17, 2.2.1.18)*. As well, tables for ABR for NAb titer < 3000 are available by baseline NAb titer subgroup (Negative, Positive), for the FAS and PP Population (*Tables 2.2.1.21, 2.2.1.22*).

Period-specific dispersion parameter

In a supplemental analysis (*Table 2.2.1.2*), period-specific dispersion parameters were estimated using separate, period-specific negative binomial regression models (with an offset parameter to account for the length of the respective collection periods). The lead-in period analysis uses the entire lead-in period, whereas the post-treatment period analysis excludes information prior to stable FIX expression (Month 6). The same table shows the estimated correlation between periods for the subject-specific ABR.

Table 2.2.1.2.1 produces the same table but for subjects with baseline NAb titer < 3000.

7.5.2.9 Sensitivity Analysis 9: Optional Zero-Inflated Negative Binomial Regression

An optional sensitivity analysis was potentially to be conducted for ABR to evaluate the robustness of the analysis findings to account for the possibility of there being a very large number of subjects having 0 bleeds in either treatment period. This sensitivity analysis was considered optional given that literature indicates that adding a zero-inflation component to a negative binomial regression is likely superfluous (https://statisticalhorizons.com/zero-inflated-models) given the inherent flexibility of the negative binomial model (to accommodate a substantial number of counts of zero). Some relatively recent literature has been written about methodology for repeated measures GEE zero-inflated negative binomial regression (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4303594/); however, the methodology may not yet have been incorporated into standard statistical analysis packages. The FDA thus agreed that it is fine not to be doing this optional analysis, which was therefore not performed.

7.6 Secondary Efficacy and CCI Endpoints

Secondary endpoints of the trial focus on investigating the effect of 2×10^{13} gc/kg AMT-061 on endogenous factor IX activity, annualized consumption (and infusion rate) of factor IX replacement therapy, remaining free of previous continuous routine prophylaxis, assessment of trough factor IX activity, bleeding events, estimated ABR as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis), correlation of factor IX activity levels and observed anti-AAV5 antibody titers using the luciferase based NAb assay after AMT-061 dosing, occurrence and resolution of target joints, CCI , and CCI .

7.6.1 Secondary Efficacy Analysis

The secondary efficacy endpoints are as follows:

1. Endogenous factor IX activity at 6 months after AMT-061 dosing
2. Endogenous factor IX activity at 12 months after AMT-061 dosing
3. Endogenous factor IX activity at 18 months after AMT-061 dosing
4. Annualized consumption of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up, excluding factor IX replacement for invasive procedures compared to the lead-in phase
5. Annualized infusion rate of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up, excluding factor IX replacement for invasive procedures compared to the lead-in phase
6. Proportion of subjects remaining free of previous continuous routine prophylaxis during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up

7. Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 over the 52 weeks following stable FIX expression (6-18 months).

8. ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up

9. Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up compared to the lead-in phase

10. Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up compared to the lead-in phase

11. Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis) (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)

12. Correlation of factor IX activity levels during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)

13. Occurrence of (and resolution of) new target joints during the 52 weeks following stable FIX expression (6-18 months) following AMT-061 dosing and resolution of pre-existing target joints following AMT-061 dosing (these endpoints will not have hypothesis testing and therefore are not included in the Type I error control)

14. Proportion of subjects with zero bleeds in the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up (this endpoint will not have hypothesis testing and therefore is not included in the Type I error control)

15. CCI [redacted] questionnaire scores from the CCI [redacted] during the 12 months following AMT-061 dosing compared with the lead-in phase

16. CCI [redacted] questionnaire scores from the CCI [redacted] during the 12 monthsfollowing AMT-061 dosing compared with the lead-in phase.

The analysis of each secondary endpoint is discussed in its own subsection below.

For the secondary efficacy endpoints, main analyses were performed using the FAS. Analyses of superiority using the PP population were considered to be supportive analyses.

All data are listed.

7.6.2 Endogenous Factor IX Activity at 6 months After AMT-061 Dosing using the Original "10 calendar-days" Contamination Rule

The main analysis of the first secondary endpoint "Uncontaminated Endogenous Central Laboratory One-Stage aPTT Factor IX activity at 6 months after AMT-061 dosing" actually took place using the 6-month data cut and was done according to the existing Statistical Analysis Plan text (at the time) (SAP text v2.0, 19 Oct 2020). Robustness output for that analysis is presented in this section. The 6-month-data cut analysis used a less refined contamination rule – whereby the date of exogenous factor IX infusion and the subsequent 9 days (10 discrete calendar days in total) were considered to be days of contamination with factor IX. The subsequent refinement to the definition of the "contamination period", mentioned in the latest version of the Statistical Analysis Plan text, and the associated plan for statistical analysis, is described below (in the "18 months" section) (in Section 7.6.4).

The update to the definition of the "contamination period" – was applied to the 12-month second secondary efficacy endpoint analysis and the 18-month third secondary efficacy endpoint analysis. The effect of the refinement to the contamination rule on the analysis at the 6-month time point is viewable within the 18-month data cut analysis, because the 6-month time point is one of the time points to be displayed in the main analysis table for the 18-month data cut analysis (wherein the main time point for analysis is the 18-month time point).

Given that the 6-month time point is a time point that is contained also within the 18-month data-cut analysis, any sensitivity analyses reported in the CSR to support the secondary factor-IX-activity endpoint can be done within the framework of the 18-month-data-cut analysis.

For this analysis, the change in endogenous factor IX activity levels (by the one-stage aPTT assay) following a single treatment with AMT-061 were assessed after the last subject achieved 26 weeks post AMT-061 treatment.

To allow time for AMT-061 to become active and to allow the subject the opportunity to stop the lead-in prophylactic factor IX therapy, factor IX levels beginning with the Week 3 assessment were used in the analysis. For the original 26-Week analysis (26-Week data-cut), the contamination rule was "within 10 days of exogenous factor IX use", and it is these results that are presented below for the 26-Week analysis. The efficacy analysis was completed using the FAS population.

The change from baseline in factor IX activity (FIX_{DIFF}) tested:

$$H_o: FIX_{DIFF} = 0 \text{ (no effect of treatment)}$$

$$H_1: FIX_{DIFF} > 0.$$

The hypothesis that $FIX_{DIFF} = 0$ (i.e. that the change from baseline is zero) was tested and a one-sided p-value <=0.025 was regarded statistically significant.

The change from baseline in factor IX activity (percent of normal) at Week 26 was analyzed using a repeated measures linear mixed model. The baseline factor IX activity was imputed as described in Section 4.1. If a subject had zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at Week 26 (and at any other post-AMT planned assessment time point used in the analysis) was imputed based on the historical hemophilia B severity as documented on the CRF, in a manner identical to that used for baseline factor IX activity. Change from baseline at Week 26 for subjects with no uncontaminated central-laboratory values post-treatment was thus assigned to zero for this analysis. However, there were no subjects with zero uncontaminated central-laboratory values post-AMT-061 in the Month 18 data cut.

The model included visit as a categorical covariate. A Toeplitz covariance matrix was used to model correlation within a subject, and model convergence was attained. A contrast was used to carry out the comparison at Week 26. The change from baseline at Week 26, the two-sided, 95% confidence interval for the mean change, and the corresponding p-value for the comparison to zero was obtained from the model and provided in a table. The estimated mean change and CIs for each visit are displayed graphically.

For both the one-stage aPTT and the chromogenic assay, means, medians, and LS Means based on uncontaminated central laboratory data are presented for each visit and it's change from baseline, for which 95% confidence intervals and one-sided p-values are also provided (*Table 2.1.1.1, dated 2021-01-07*).

7.6.3 Endogenous Factor IX Activity at 12 Months After AMT-061 Dosing using the "5 half-life" Contamination Rule

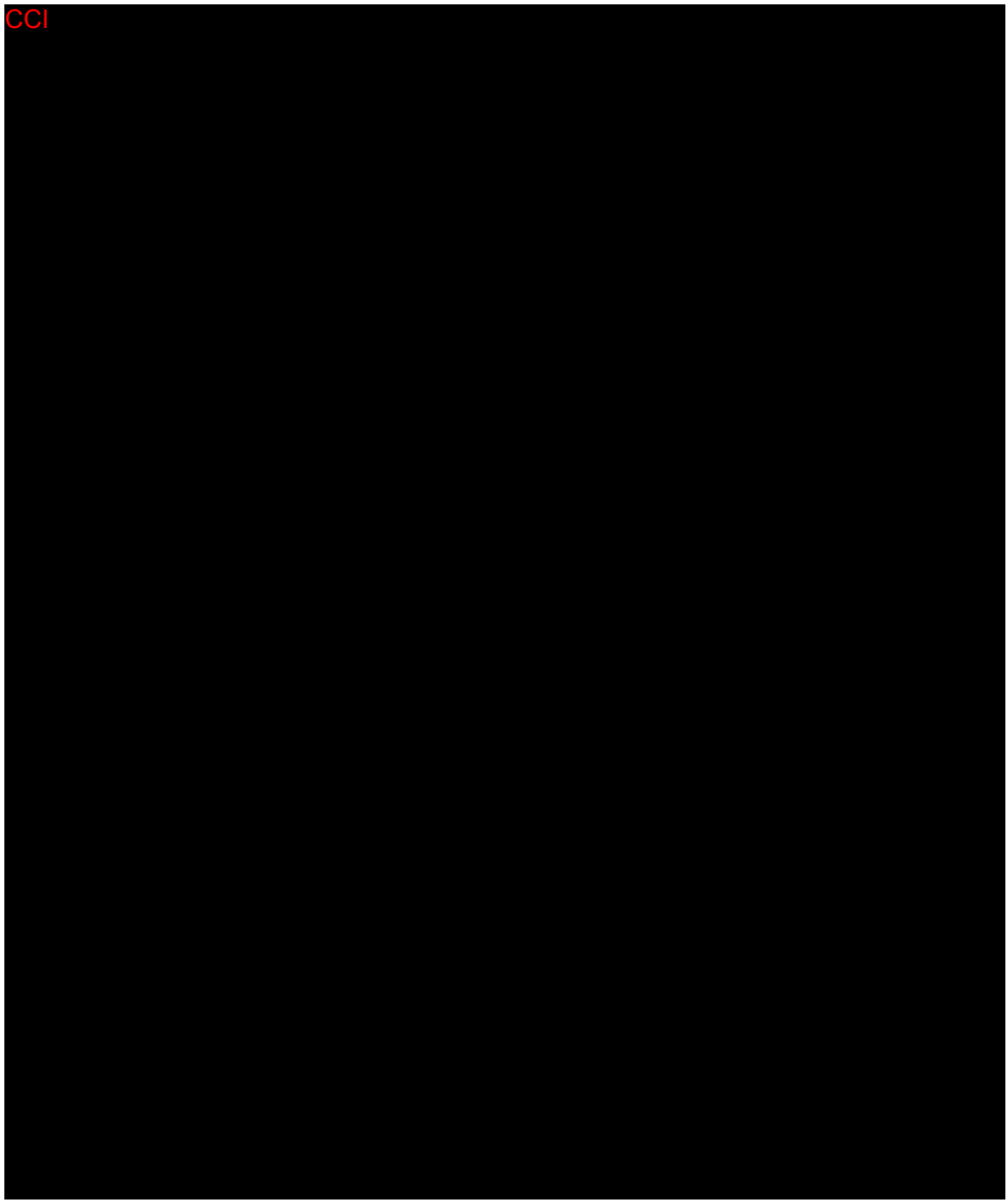
The main analysis of the second secondary endpoint "Endogenous Central Laboratory One-Stage aPTT Factor IX activity at 12 months after AMT-061 dosing" uses the 12-month data cut and was done according to the methodology described below (in Section 7.6.4) for the third secondary efficacy endpoint.

The update to the definition of the "contamination period" (described in Section 7.6.4 for the 18-month third secondary efficacy endpoint analysis) was applied also to the 12-month second secondary efficacy endpoint analysis.

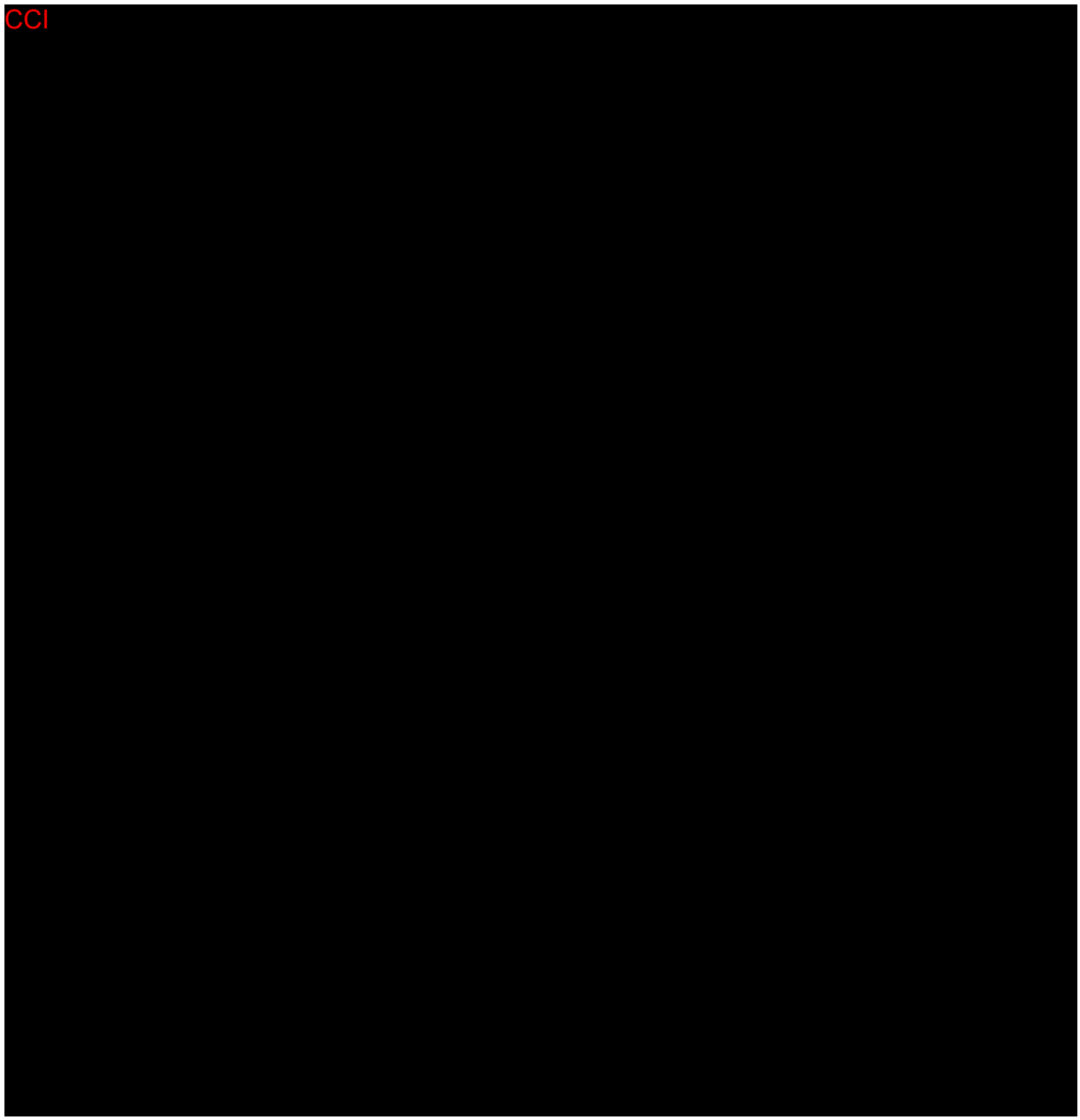
Given that the 12-month time point is a time point that is contained also within the 18-month data-cut analysis, any sensitivity analyses to be reported in the CSR to support the secondary factor-IX-activity endpoint can be done within the framework of the 18-month-data-cut analysis.

Endogenous factor IX activity at 52 weeks after AMT-061 dosing was analyzed in the same manner as the 6-month endpoint analysis, but this analysis employed all scheduled visits over the 52-week post-treatment period (except those early visits that were excluded also from the 26-week analysis). The change from baseline at Week 52 was estimated within the framework of the repeated measures model. The imputation of the baseline value was described earlier (Section 4.1).

For both the one-stage aPTT and the chromogenic assay, means, medians, and LS Means based on uncontaminated central laboratory data are summarized for each visit and it's change from baseline, for which 95% confidence intervals and one-sided p-values are also provided (*Table 2.1.1.1 dated at the time of the 52-week data cut)*.




CCI


CCI


CCI


CCI

7.6.4.1 Sensitivity Analysis 1: PP Population

A sensitivity analysis using the PP population on the secondary factor IX activity efficacy analysis was performed *(Table 2.1.1.2)*.

7.6.4.2 Sensitivity Analysis 2: To Account for Missing Data

A second sensitivity analysis was conducted for factor IX activity 18 months after a single AMT-061 treatment to evaluate the robustness of the main analysis findings to missing data. For this analysis, any missing values (that were still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits) were imputed using the most

recent previous post-treatment uncontaminated factor IX activity value that was observed for the subject *(Table 2.1.1.3)*. Such imputation was done for four subjects *(Tables 2.1.1.1 and 2.1.1.3)*.

7.6.4.3 Sensitivity Analysis 3: Cumulative Responder Analysis at Month 18

A third sensitivity analysis, a (single-treatment) cumulative responder analysis, was conducted for the change from baseline in uncontaminated central-laboratory factor IX activity 18 months after a single AMT-061 treatment *(Table 2.1.1.4)*. A cumulative distribution plot was also produced *(Figure 2.2.2)*. The full analysis set was used. The observed change from baseline in factor IX activity at 18 months is plotted on the *x*-axis and the proportion of responders (subjects that have an equal or greater level of change) is plotted on the *y*-axis. If a subject still lacked the 18-month value (being still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value closest in time to (either before or after) 18 months post-AMT was employed in the analysis. If two such assessments were both the closest in time, with one being before and the other being after, the earlier assessment was used.

7.6.4.4 Sensitivity Analysis 4: Excluding Subjects with No Uncontaminated Central-Laboratory Values

The secondary factor IX activity efficacy analysis was repeated (on the FAS) without the imputation of post-baseline values based on historical hemophilia B severity for subjects with zero uncontaminated central-laboratory values after administration of AMT-061 *(Table 2.1.1.5)*. Note however, there were no subjects with zero uncontaminated central-laboratory values post-AMT-061 in the Month 18 data cut.

7.6.4.5 Sensitivity Analysis 5: Excluding Visits Contaminated by Corticosteroid Exposure

The secondary factor IX activity efficacy analysis was repeated (on the FAS) to evaluate the robustness of the findings to the impact of systemic corticosteroid therapy on factor IX activity levels *(Table 2.1.1.6)*. Visits post-AMT-061 in proximity to corticosteroid exposure were characterized as contaminated and excluded from this analysis. Factor IX data from visits that occurred within 5 half-lives of systemic corticosteroid exposure (or during such exposure) were excluded from the analysis. As well, only values from blood sampling that did not occur within 5 half-lives of exogenous factor IX use are included in the analysis. The duration of contamination associated with corticosteroid usage was dependent on the specific product (and its associated half-life), its route of administration, and the dosing frequency. A listing detailing (1) Medication Name, (2) Route of Administration, (3) Dosing Frequency, (4) Half-life, and (5) the Contamination Period (days) (subsequent to the medication stop date) for systemic corticosteroid use during either the lead-in or post-treatment period accompanies this analysis. An additional listing detailing subject-specific factor IX activity levels relative to instances of corticosteroid exposure is provided *(Listing 2.1.5)*. However, as can be seen in *Table 2.1.1.6*, the results were nearly identical to those for the main analysis in *Table 2.1.1.1*.

7.6.4.6 Sensitivity Analysis 6: Excluding Visits Contaminated by Exogenous Factor IX Using a 10-Day (Date/Time-Based) Contamination Rule

The secondary factor IX activity efficacy analysis was repeated (on the FAS) to evaluate the robustness of the findings to the impact of using an alternative rule for determining contamination due to infused exogenous factor IX use *(Table 2.1.1.7)* . Visits post-AMT-061 within 10 days (240 hours) of exogenous factor IX use were considered contaminated and excluded from this analysis. Both the date and time of the exogenous factor IX injection start and the blood sampling for factor IX activity assessment were taken into account to determine whether there was contamination. A blood sample drawn at a date/time prior to a given exogenous factor IX infusion start time was not considered to be contaminated by that infusion. See further detail in the Data Handling Rules Appendix (Appendix 1 Data Handling Rules) in this document [under the category of "Contamination due to exogenous factor IX (infusion) use"]. This sensitivity analysis did not use the 5 half-life contamination rule. Note that this sensitivity-analysis 10-day contamination rule – based on date-and-time – is distinct from the 10-discrete-calendar-day contamination rule that was used at the 26-week-data-cut interim efficacy analysis. This 10-day contamination rule (now using date and time) is being provided (for transparency) as a sensitivity analysis because a 10-day rule (albeit based on discrete days) was the main contamination rule for an earlier version of the SAP. The use of a date-and-time-based 10-day window (240 hours) was employed for the sensitivity analysis because the use of discrete days was considered to be too crude (in retrospect) even for a sensitivity analysis. The FDA agreed to the use of the five-half-life contamination rule (in the amended SAP) as the main analysis.

7.6.4.7 NAb titer Subcategories: Patients with/without Pre-Existing Neutralizing Antibodies to AAV5.

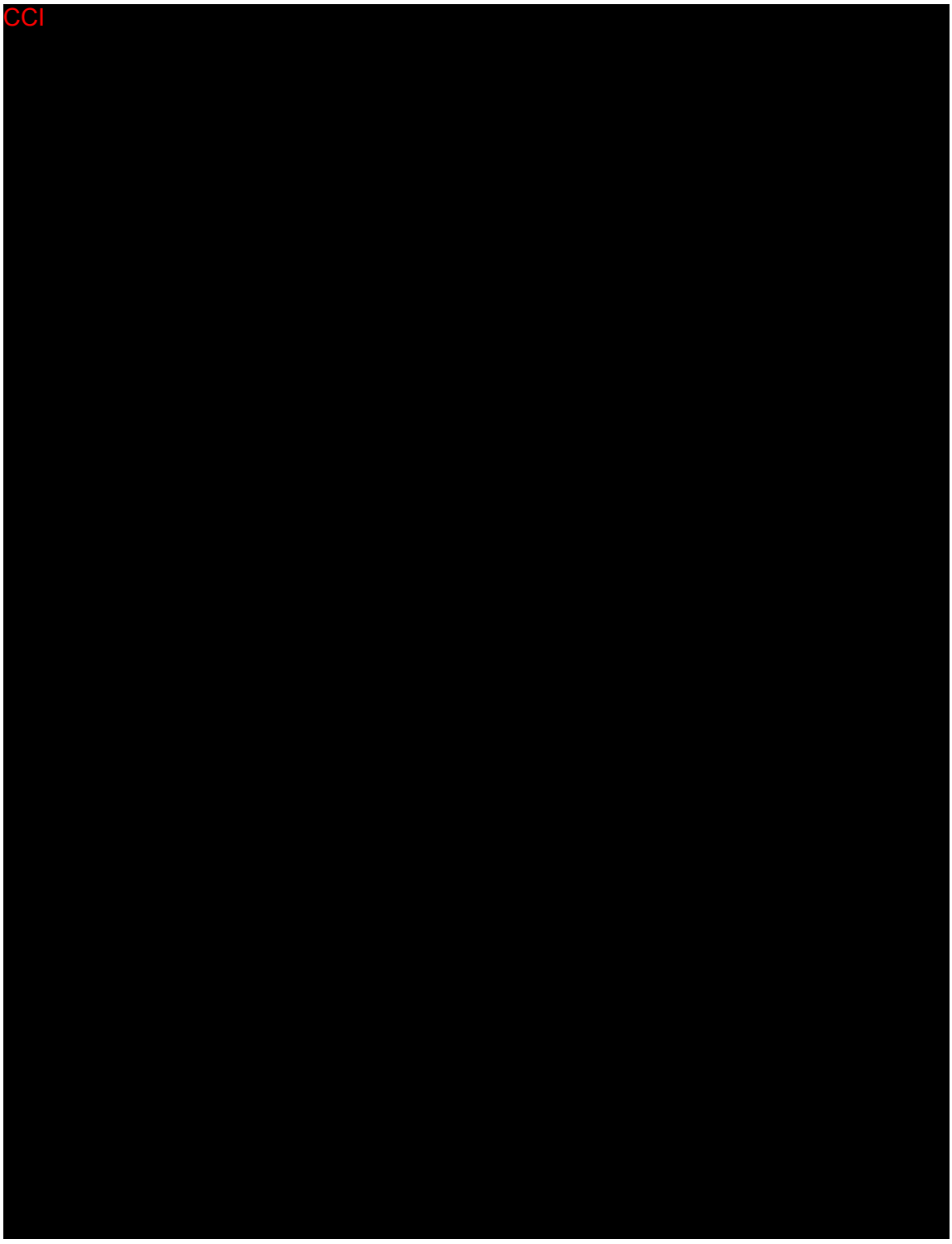
The secondary factor IX activity efficacy analysis was repeated, for the FAS and PP population, for two subcategories: patients with and without pre-existing neutralizing antibodies to AAV5 *(Table 2.1.2.1, 2.1.2.2)*. These analyses were both repeated for subjects with NAb titer < 3000 *(Table 2.1.2.3, 2.1.2.4)*.

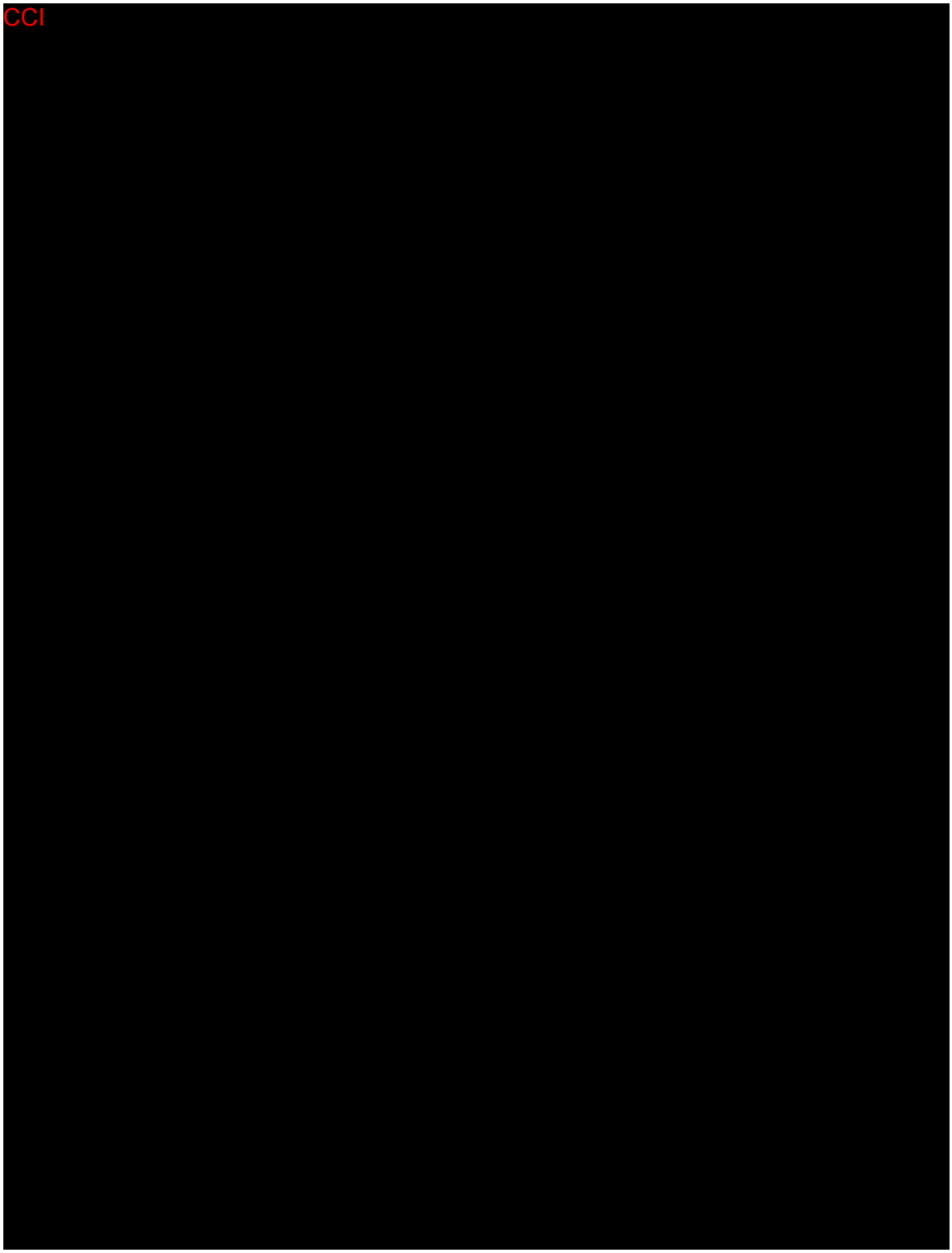
CCI

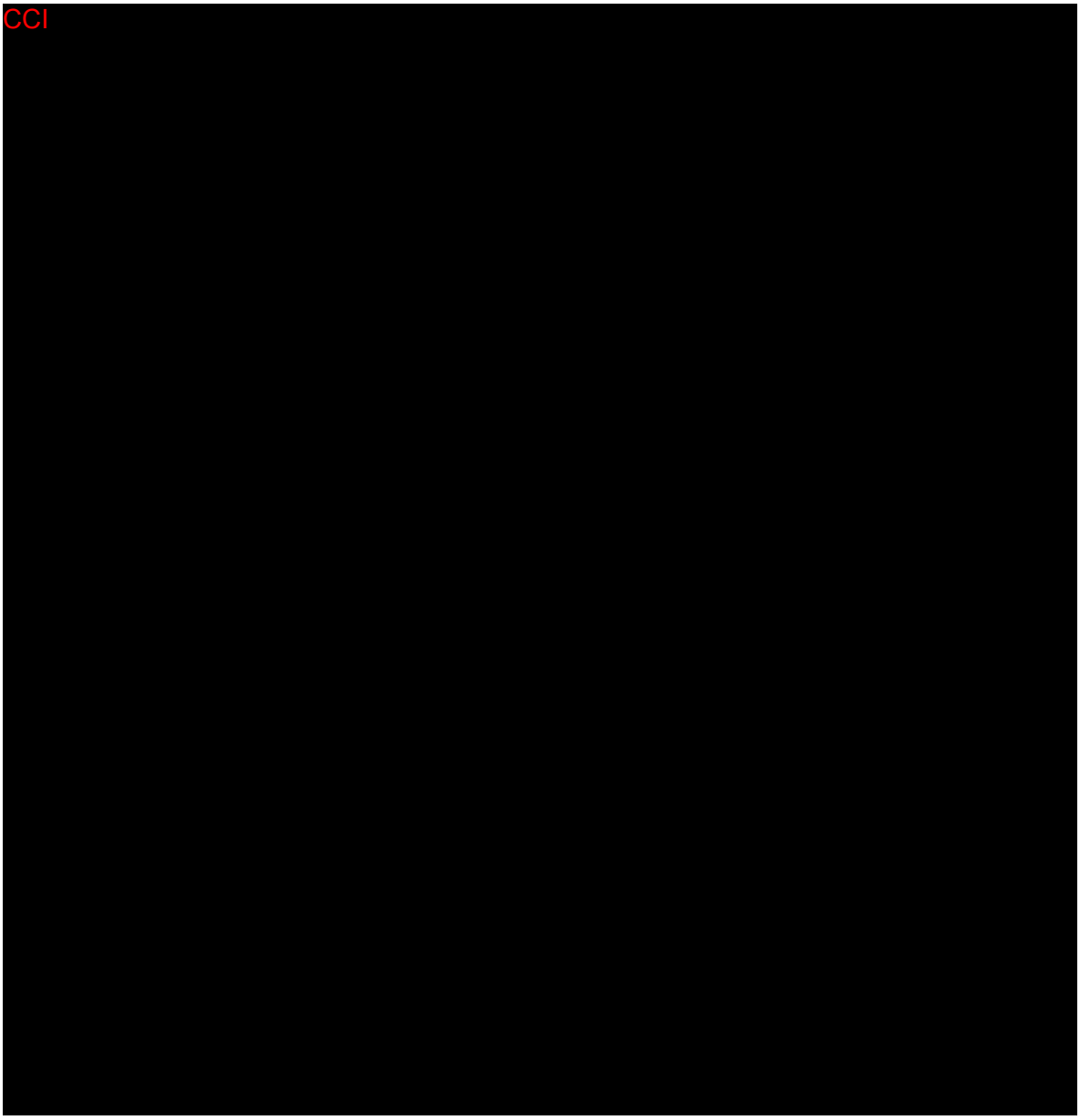

CCI

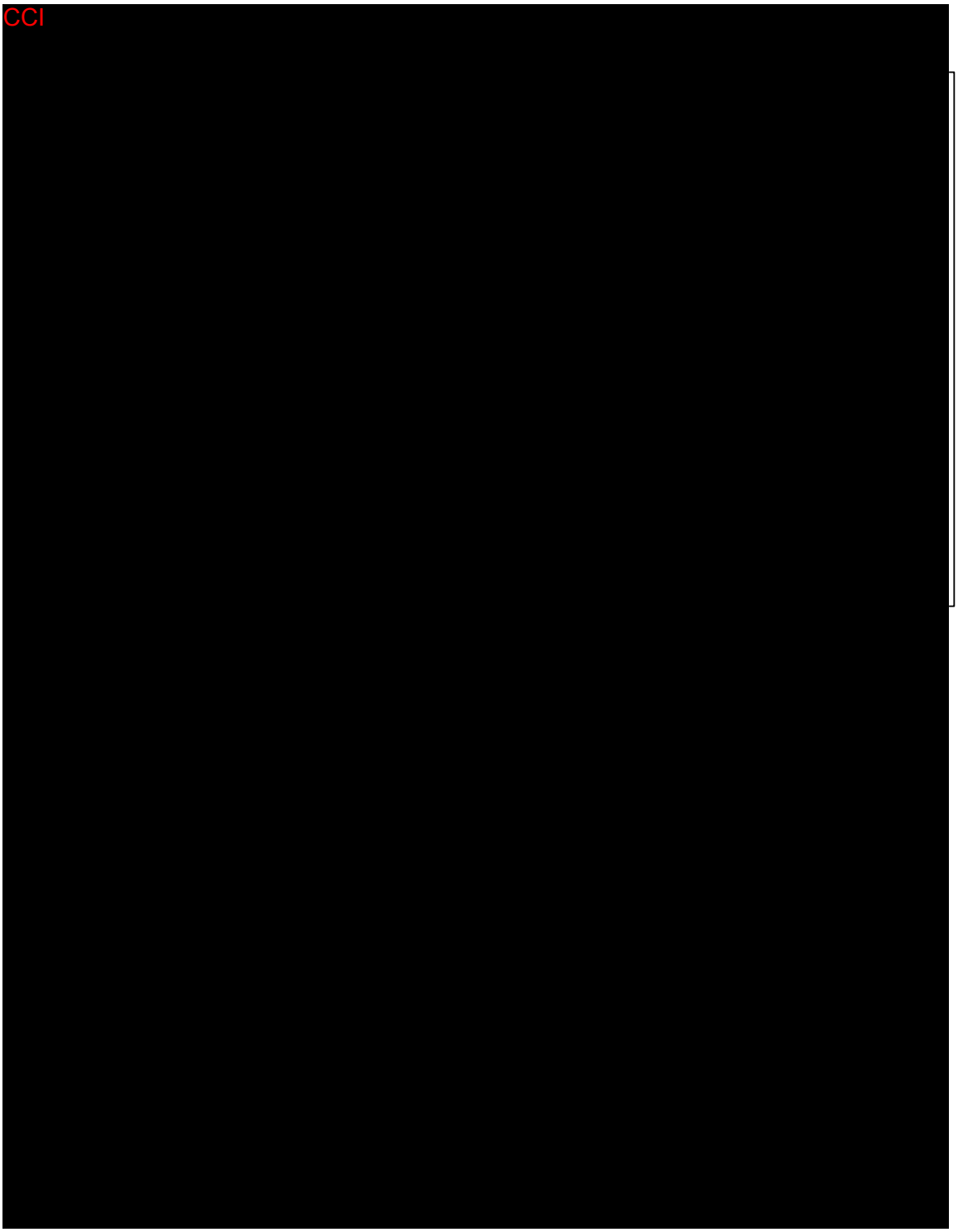

CCI

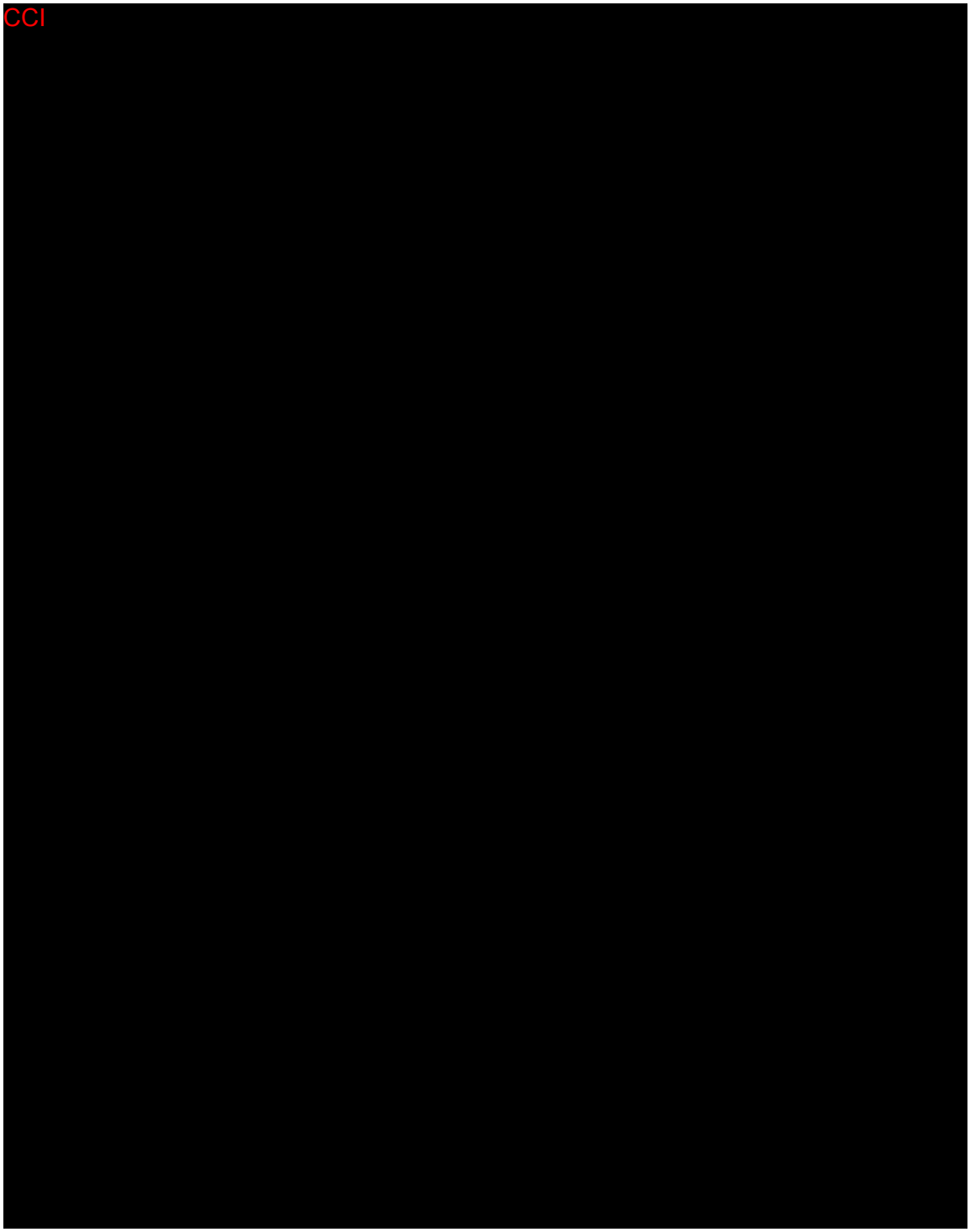

CCI


CCI


CCI


CCI


CCI


CCI

7.6.8.1 Sensitivity Analysis 1: Cumulative Responder Analysis Comparing the Lead-In Period with the 52 weeks following stable FIX expression (6-18 months) after Treatment with AMT-061

A cumulative responder sensitivity analysis (as described in Farrar 2006) compared the factor IX activity between the lead-in period (mean of Months 2, 4, and 6) with the mean across the scheduled visits between the six-month and 18-month (the SAP text accidentally said "12-month") time points (i.e. Month 6, 7, 8, 9, 10, 11, 12, and 18) in the post-treatment period. For each period, "response" was defined as having a change from baseline (mean minus baseline) greater than or equal to the specified value. The percent response (percentage of subjects with a response) was tabulated, cumulatively, for values starting with 0.5, 1.0, 2.0, and continuing on until the response fell to zero *(Table 2.2.8.3)*. The full analysis set was used. Only central-laboratory values were used. For the post-treatment period, only uncontaminated values were used.

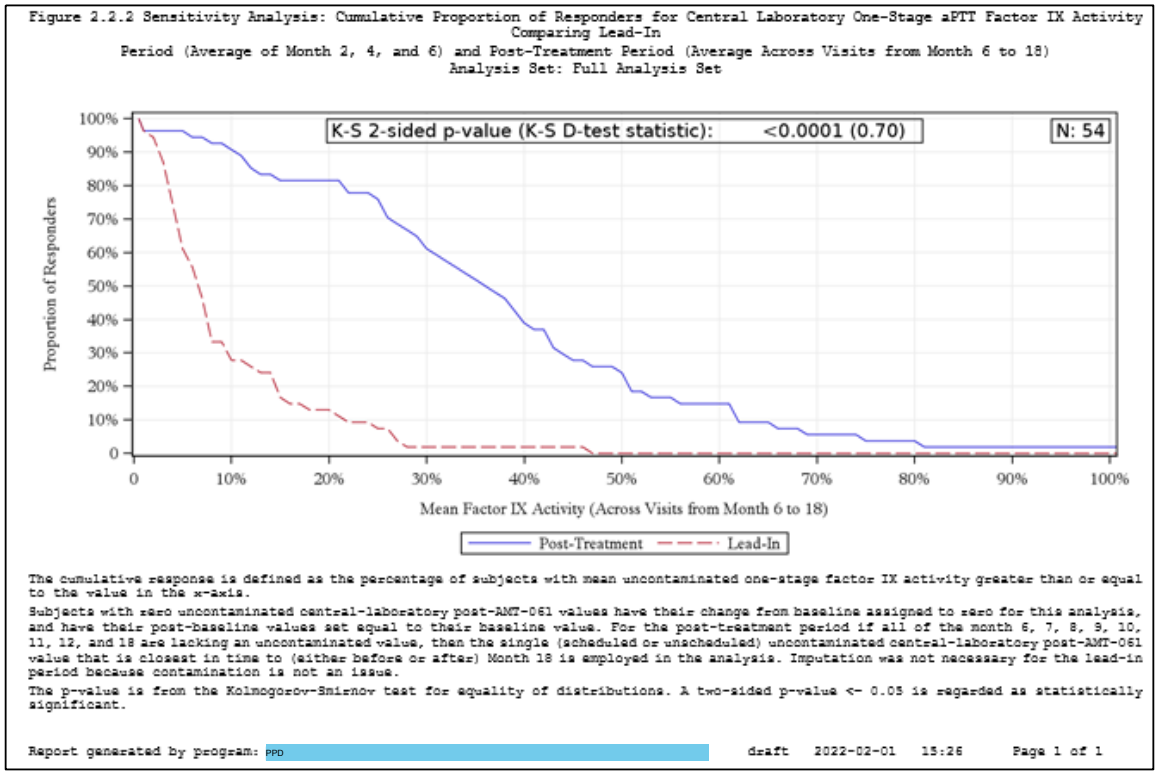
The uncontaminated factor IX activity mean across visits from Month 6 to Month 18 (to Month 12 for the 12-month data cut) for the post-treatment period was plotted on the x-axis, and the proportion of "responders" (subjects that have an equal or greater level) was plotted on the y-axis. The same graph plotted the mean of the observed factor IX activity across the Month 2, Month 4, and Month 6 visits for the lead-in period (also) on the x-axis, and the proportion of "responders" (subjects that have an equal or greater level) were plotted on the y-axis. Thus, a cumulative distribution plot by treatment period was produced *(Figure 2.2.2)*. The observed factor IX activity was plotted on the x-axis and the proportion of "responders" (subjects that have an equal or greater level) for each treatment period was plotted on the y-axis. If a subject

had zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that was to be used in the analysis) was to be imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity. Note however, there were no subjects with zero uncontaminated central-laboratory values post-AMT-061 in the Month 18 data cut. The cumulative responder curves for each treatment period were then compared using a two-sample Kolmogorov-Smirnov test.

If a subject still lacked having at least one uncontaminated factor IX activity value at least one of the following time points – post-treatment month 6, 7, 8, 9, 10, 11, 12, and 18 (6, 7, 8, 9, 10, 11, and 12 for the 12-month data cut) – value (missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the single (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value closest in time to (either before or after) 18 months (12 months for the 12-month data cut) was employed in the analysis. If two such assessments were both the closest in time, with one being before and the other being after, the earlier assessment was used.

Results are presented in Figure 16 below.

Figure 16: Cumulative Proportion of Responders for Factor IX Activity for Lead-In (Average of Months 2, 4, and 6) and Post-Treatment Period (Average Across Visits from Month 6 to 18)



7.6.8.2 Sensitivity Analysis 2: Excluding Contaminated Values Using a 10-day Contamination Rule

The analysis of "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 over 52 weeks following stable FIX expression (6-18 months)" was repeated (on the FAS) to evaluate the robustness of the findings to the impact of using an alternative rule for determining contamination due to infused exogenous factor IX use *(Table 2.2.8.4)*. For this analysis, visits post-AMT-061 within 10 days of exogenous factor IX use were considered contaminated. Both the date and time of the exogenous factor IX injection start and the blood sampling for factor IX activity assessment were taken into account to determine whether there was contamination. See further detail in the Appendix 1 Data Handling Rules in this document [under the category of "*Contamination due to exogenous factor IX (infusion) use*"]. Data from post-AMT-061 visits within 10 days of exogenous factor IX use were excluded from the analysis. This sensitivity analysis did not use the 5 half-life contamination rule. Results were similar to the main analysis.

7.6.8.3 Sensitivity Analysis 3: No Month 18 Imputation

The analysis of "Comparison of the percentage of subjects with factor IX activity < 12% of normal between the lead-in phase and after treatment with AMT-061 over 52 weeks following

stable FIX expression (6-18 months)" was repeated (on the FAS) to evaluate the robustness of the findings to the impact of Month 18 imputation *(Table 2.2.8.6)*. In the original analysis (*Table 2.2.8.1*), if the post-treatment Month 18 visit was lacking an uncontaminated value, then the single (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value that was closest in time to (either before or after) 18 months was employed in the analysis in place of the Month 18 value. In the sensitivity analysis conducted here for the FAS, no such imputation was made.

The same sensitivity analysis was also conducted for FAS subjects with a NAb titer < 3000 (*Table 2.2.8.7*).

7.6.10 Rate of spontaneous bleeding events during the 52-weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase

Similar to ABR, the number of spontaneous bleeding events and person-time at risk of (having) spontaneous bleeding events were determined for the lead-in and post-treatment periods. Analysis of the annualized spontaneous bleeding rate was performed using a repeated measures negative binomial regression model with the log of the time at risk of spontaneous bleeding (in the respective period) used as an offset parameter to account for the differential reporting

periods. An unstructured covariance matrix was employed, and the model converged. Treatment (i.e., period) was a categorical covariate.

The post-AMT-061 administration time at risk of (having) a bleeding event was the subject's (uncontaminated) time on the study between stable FIX expression (Month 6) and the time that was 52 weeks following stable FIX expression (18 months) (that was one year post-treatment for the 12-month data cut), the time of study completion, or the time of early withdrawal from the study, whichever was earlier. Any bleeds prior to stable FIX expression (Month 6) of the post-treatment period were not considered in the analysis. Events from the entire lead-in period were counted, and the entire lead-in period was considered to be time-at-risk.

The estimated rate ratio (between post-AMT-061 and the lead-in) was tested using the following hypotheses:

H_0: rate ratio (post-treatment)/(lead in) = 1 (no effect of treatment)
H_1: rate ratio (post-treatment)/(lead in) < 1.

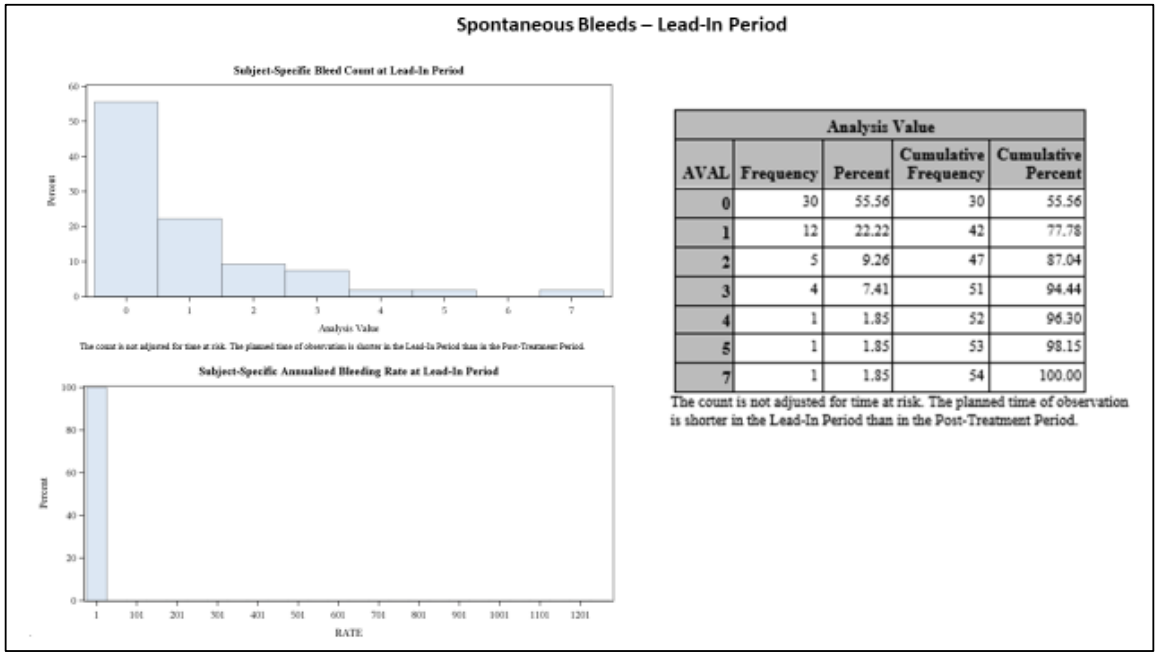
The hypothesis that *(post-treatment)/(lead in) = 1* (i.e. no difference between the two treatment periods) was tested and a one-sided p-value <=0.025 was regarded statistically significant. The one-sided p-value and two-sided 95% CI for the rate ratio are presented in a table. The treatments were compared for superiority. The main population was the FAS *(Table 2.2.4.1a)*. ABR for spontaneous bleeds treated with exogenous Factor IX were also analyzed *(Table 2.2.4.1b)*. A sensitivity analysis used the PP population *(Table 2.2.4.2)*.

Rules for computing time at risk for a bleeding event were already described along with the analysis methodology for the ABR endpoint (for the non-inferiority comparison).

The number of subjects with a spontaneous bleed, number of spontaneous bleeds, number of treated/not treated bleeds, bleeds per subject, and unadjusted annualized bleeding rates are provided in *Table 2.2.4.1a*. The adjusted annualized bleeding rates are also provided, along with 95% CIs, the rate ratio (Post-Treatment/Lead-In), the p-value, and estimates for the dispersion parameter.

The (frequency) distribution of the number of bleeds per subject was visually examined for each period with (1) a list of each unique bleed count along with its frequency (number of subjects having that count) and (2) a histogram. Also, the distribution of the subject-specific ABR (annualized bleeding rate) was visually examined for each period with a histogram. As the negative binomial distribution is flexible, there is no concern about the use of a negative binomial regression. There is a notable outlier at a subject-specific bleed rate of 1339 for Months 7-18 (for subject PPD) *(based on Listings 2.2.1.5.1 and 2.2.1.3)*. This is due to the conservative approach (requested by regulatory agencies) of removing time with contamination from exogenous factor IX from the time at risk.

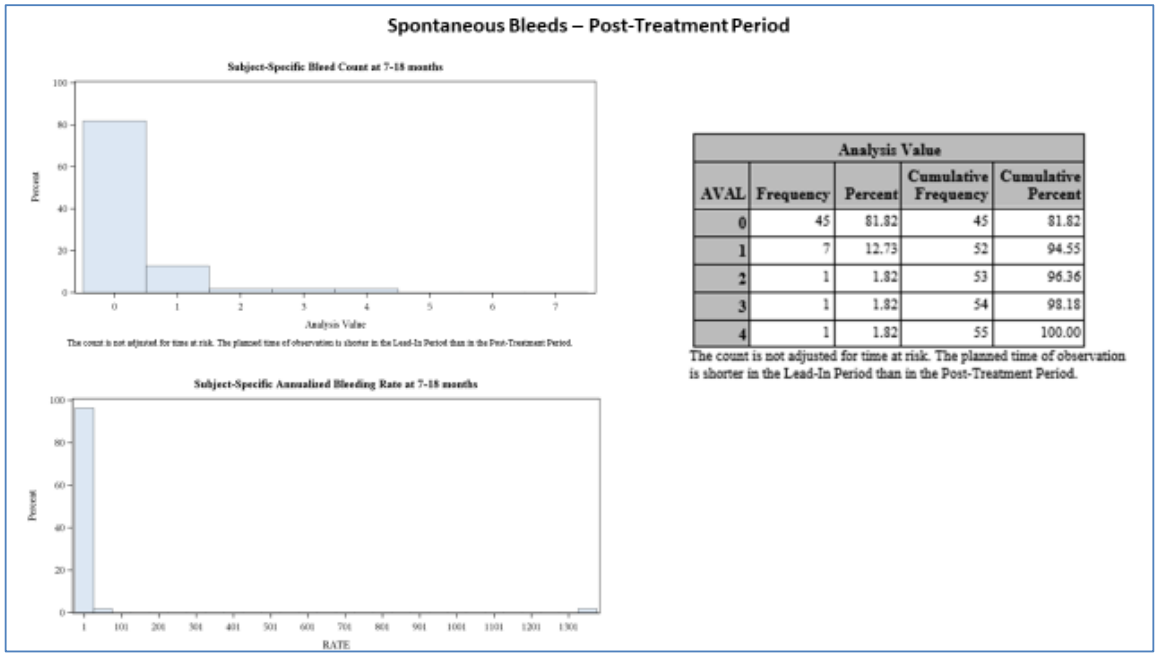
Figure 17: Spontaneous Bleed Rates – Lead-In Period



Analysis Value				
AVAL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	30	55.56	30	55.56
1	12	22.22	42	77.78
2	5	9.26	47	87.04
3	4	7.41	51	94.44
4	1	1.85	52	96.30
5	1	1.85	53	98.15
7	1	1.85	54	100.00

The count is not adjusted for time at risk. The planned time of observation is shorter in the Lead-In Period than in the Post-Treatment Period.

Figure 18: Spontaneous Bleed Rates – Post-Treatment Period



Analysis Value				
AVAL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	45	81.82	45	81.82
1	7	12.73	52	94.55
2	1	1.82	53	96.36
3	1	1.82	54	98.18
4	1	1.82	55	100.00

The count is not adjusted for time at risk. The planned time of observation is shorter in the Lead-In Period than in the Post-Treatment Period.

Bleeding events are listed by subject *(Listings 2.2.1.1 – 2.2.1.4)*. The number of bleeding events and the time-at-risk of bleeding events are listed by period for each subject *(Listing 2.2.1.5.1 and 2.2.1.5.2)*.

7.6.11 Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase

Similar to ABR, the number of joint bleeding events and person-time at risk of joint bleeding events were determined for the lead-in and post-treatment periods. Analysis of the reported number of joint bleeding events was performed using a repeated measures negative binomial regression model with the log of the time at risk of joint bleeding (in the respective period) used as an offset parameter to account for the differential reporting periods. An unstructured covariance matrix was employed, and the model converged. The post-AMT-061 administration time at risk of (having) a bleeding event was the subject's (uncontaminated) time on the study between stable FIX expression (Month 6) and the time that was 52 weeks following stable FIX expression (18 months) (that was one year post-treatment for the 12-month data cut), the time of study completion, or the time of early withdrawal from the study, whichever was earlier. Any bleeds prior to stable FIX expression (Month 6) of the post-treatment period were not considered in the analysis. Events from the entire lead-in period were counted, and the entire lead-in period was considered to be time-at-risk.

The estimated rate ratio (between post-AMT-061 and the lead-in) was tested using the following hypotheses:

H_0: *rate ratio (post-treatment)/(lead in) = 1 (no effect of treatment)*
H_1: *rate ratio (post-treatment)/(lead in) < 1*.

The hypothesis that *(post-treatment)/(lead in) = 1* (i.e. no difference between the two treatment periods) was tested and a one-sided p-value <=0.025 was regarded statistically significant. The one-sided p-value and two-sided 95% CI for the rate ratio is presented in a table. The treatments were compared for superiority. The main population was the FAS *(Table 2.2.5.1a*. ABR for joint bleeds treated with exogenous Factor IX were also analyzed *(Table 2.2.5.1b)*. A sensitivity analysis used the PP population *(Table 2.2.5.2)*.

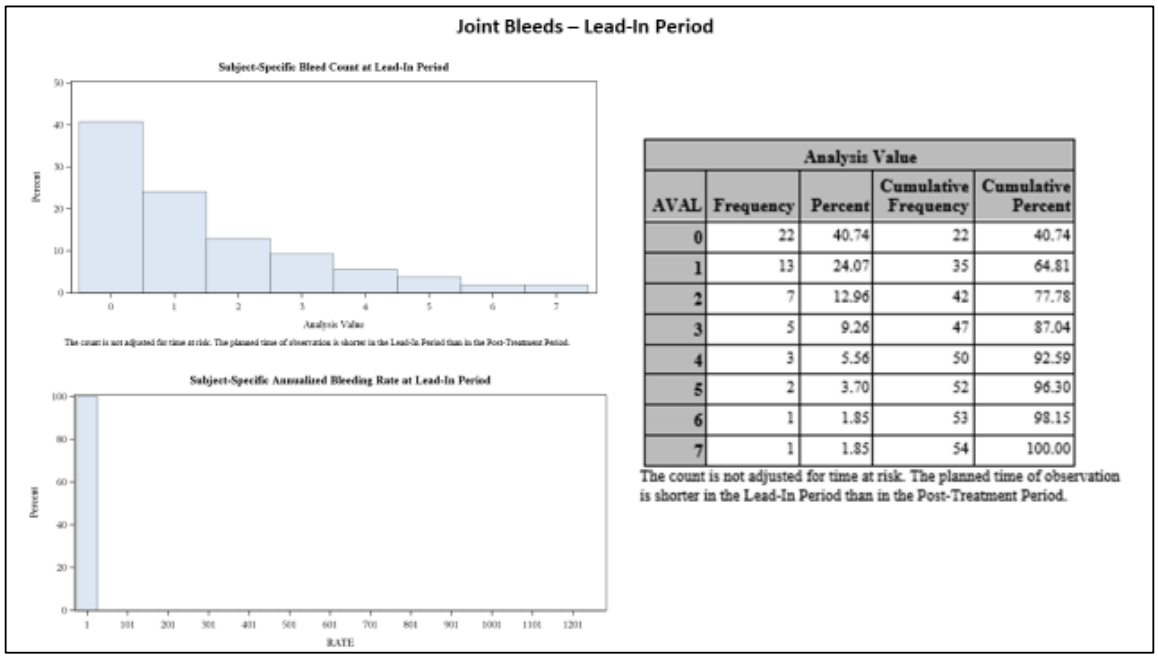
Rules for computing time at risk for a bleeding event were already described along with the analysis methodology for the ABR endpoint (for the non-inferiority comparison).

The number of subjects with a (joint) bleed, number of bleeds, number of treated/not treated bleeds, bleeds per subject, and unadjusted annualized bleeding rates are provided in *Table 2.2.5.1a*. The adjusted annualized bleeding rates are also provided, along with 95% CIs, the rate ratio (Post-Treatment/Lead-In), the p-value, and estimates for the dispersion parameter.

The (frequency) distribution of the number of bleeds per subject was visually examined for each period with (1) a list of each unique bleed count along with its frequency (number of subjects having that count) and (2) a histogram. Also, the distribution of the subject-specific ABR (annualized bleeding rate) was visually examined for each period with a histogram. As the

negative binomial distribution is flexible, there is no concern about the use of a negative binomial regression. There is a notable outlier at a subject-specific bleed rate of 1339 for Months 7-18 (for subject PPD) *(based on Listings 2.2.1.5.1 and 2.2.1.3)*. This is due to the conservative approach (requested by regulatory agencies) of removing time with contamination from exogenous factor IX from the time at risk, without a commensurate removal of bleeds during those periods.

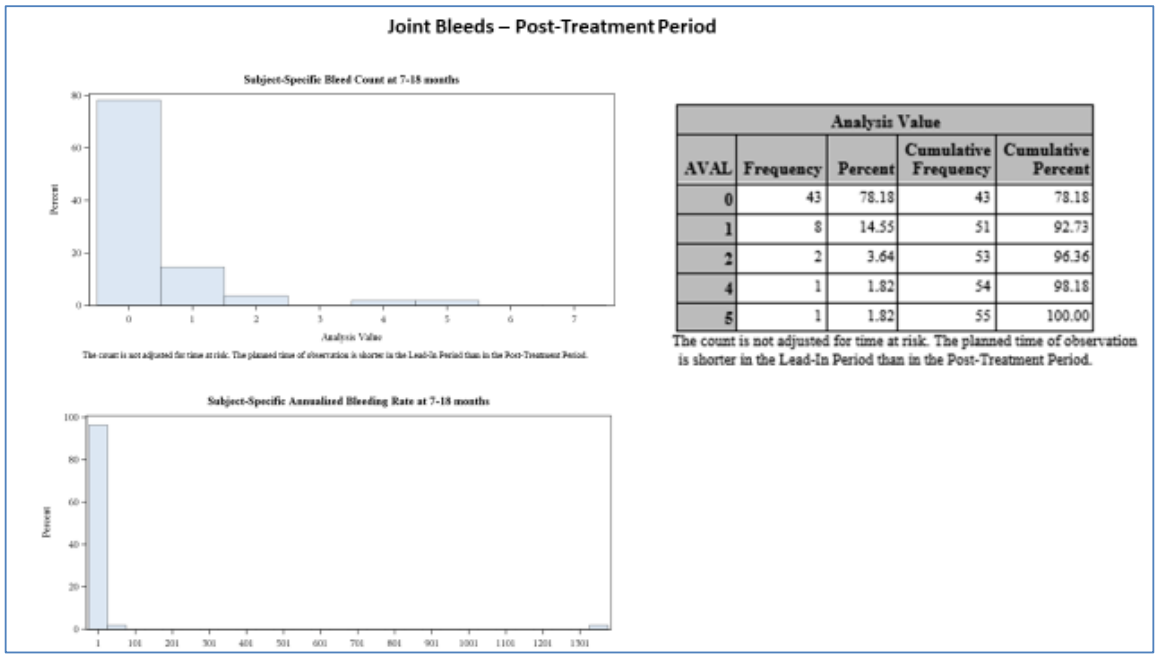
Figure 19: Joint Bleed Rates – Lead-In Period



Analysis Value				
AVAL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	22	40.74	22	40.74
1	13	24.07	35	64.81
2	7	12.96	42	77.78
3	5	9.26	47	87.04
4	3	5.56	50	92.59
5	2	3.70	52	96.30
6	1	1.85	53	98.15
7	1	1.85	54	100.00

The count is not adjusted for time at risk. The planned time of observation is shorter in the Lead-In Period than in the Post-Treatment Period.

Figure 20: Joint Bleed Rates – Post-Treatment Period



Analysis Value				
AVAL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	43	78.18	43	78.18
1	8	14.55	51	92.73
2	2	3.64	53	96.36
4	1	1.82	54	98.18
5	1	1.82	55	100.00

The count is not adjusted for time at risk. The planned time of observation is shorter in the Lead-In Period than in the Post-Treatment Period.

Bleeding events are listed by subject *(Listings 2.2.1.1 – 2.2.1.4)*.

7.6.12 Estimated ABR – during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up – as a function of pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay (as a "correlation" analysis)

To examine the relationship (i.e. "correlation") between ABR and baseline anti-AAV5 neutralizing antibodies (NAB), the following analysis was carried out. A nonparametric, generalized additive model (GAM) was implemented to graph the relationship of ABR to the natural logarithm of baseline anti-AAV5 neutralizing antibodies with a negative binomial model *(Figure 2.3.2.2)*. The relationship of ABR to the natural logarithm of the Lead-In Final Visit anti-AAV5 neutralizing antibodies was also examined *(Figure 2.3.2.3)*. Values of "< LOD" (7) were set to LOD/2 (3.5) for the purpose of this analysis. Information for each subject across the 52 weeks following stable FIX expression (6-18 months) post-treatment was employed. This endpoint does not have hypothesis testing and therefore is not included in the type I error control.

The two figures were also examined for subjects with Baseline NAb titer < 3000 *(Figures 2.3.2.4, 2.3.2.5)*, and can be seen to have been affected by the presence of the subject with a baseline NAb titer greater than 3000.

The generalized additive model graph of the relationship of ABR to pre-IMP anti-AAV5 antibody titers (with a negative binomial regression) was examined for robustness by varying the smoothing parameter. It is important not to over-smooth the curve, which could mask the relationship between variables. It is also important not to under-smooth the curve, which could make the graph too sensitive to variations at the individual-subject level. A reasonable compromise between these two extremes needs to be chosen. Graphs using a wide range of values of the smoothing parameter (e.g. varying by powers of the number 10) are presented below, to allow the effect of the choice of the smoothing parameter to be seen. For this analysis, a good balance was attained using a smoothing parameter value of 0.1, which is the value chosen for the formal figure.

Figure 21: Examination of Alternative Smoothing Parameters: 0.001, 0.01, 0.1, and 1

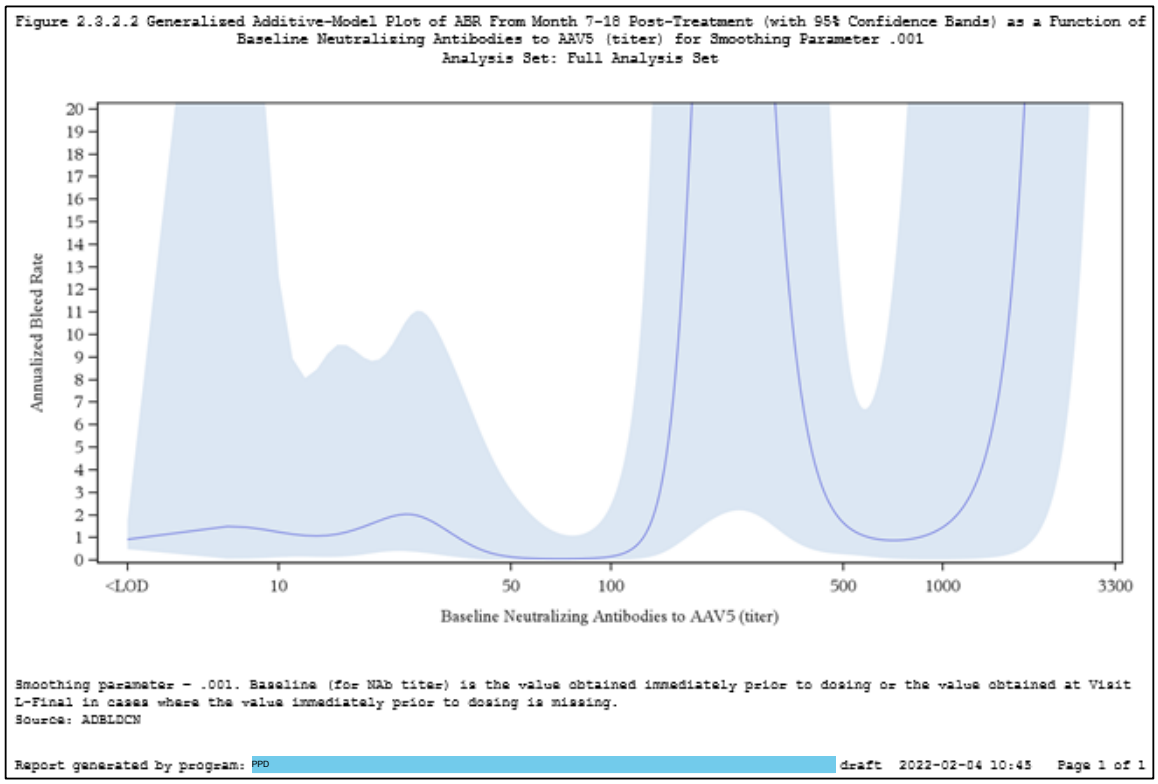
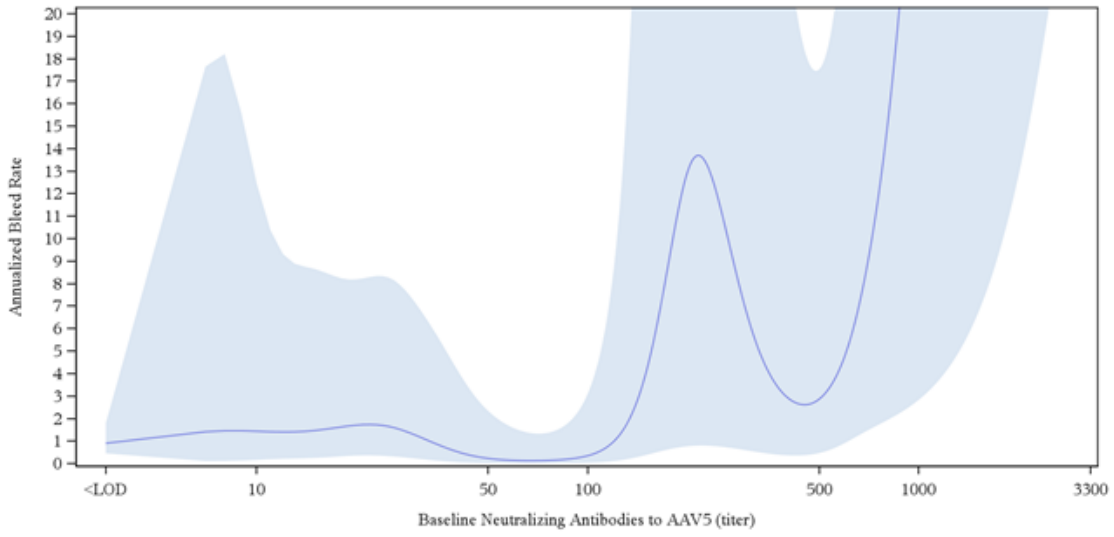


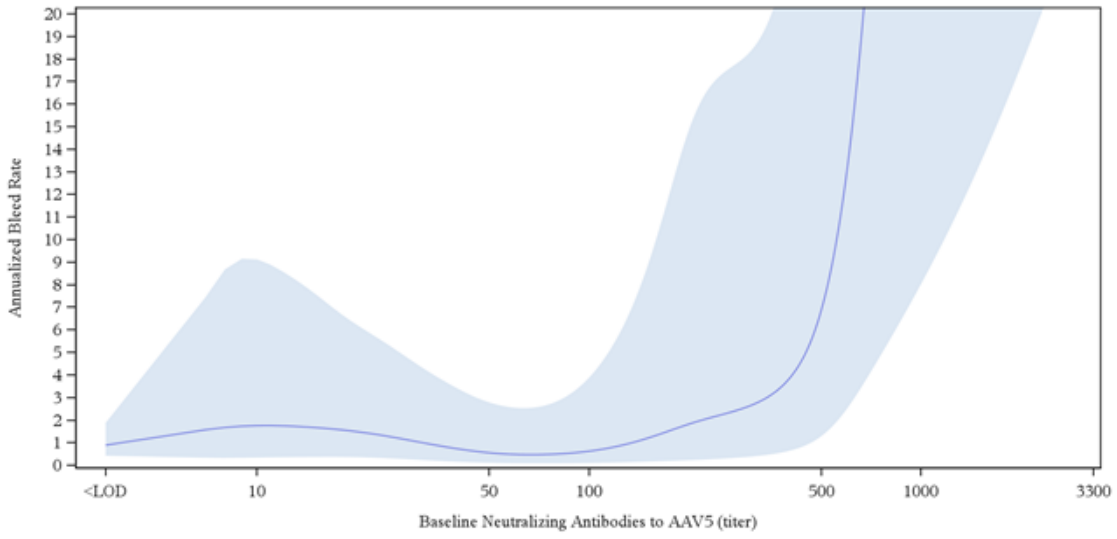
Figure 2.3.2.2 Generalized Additive-Model Plot of ABR From Month 7-18 Post-Treatment (with 95% Confidence Bands) as a Function of Baseline Neutralizing Antibodies to AAV5 (titer) for Smoothing Parameter .01
Analysis Set: Full Analysis Set



Smoothing parameter = .01. Baseline (for NAb titer) is the value obtained immediately prior to dosing or the value obtained at Visit L-Final in cases where the value immediately prior to dosing is missing.
Source: ADBLDCN

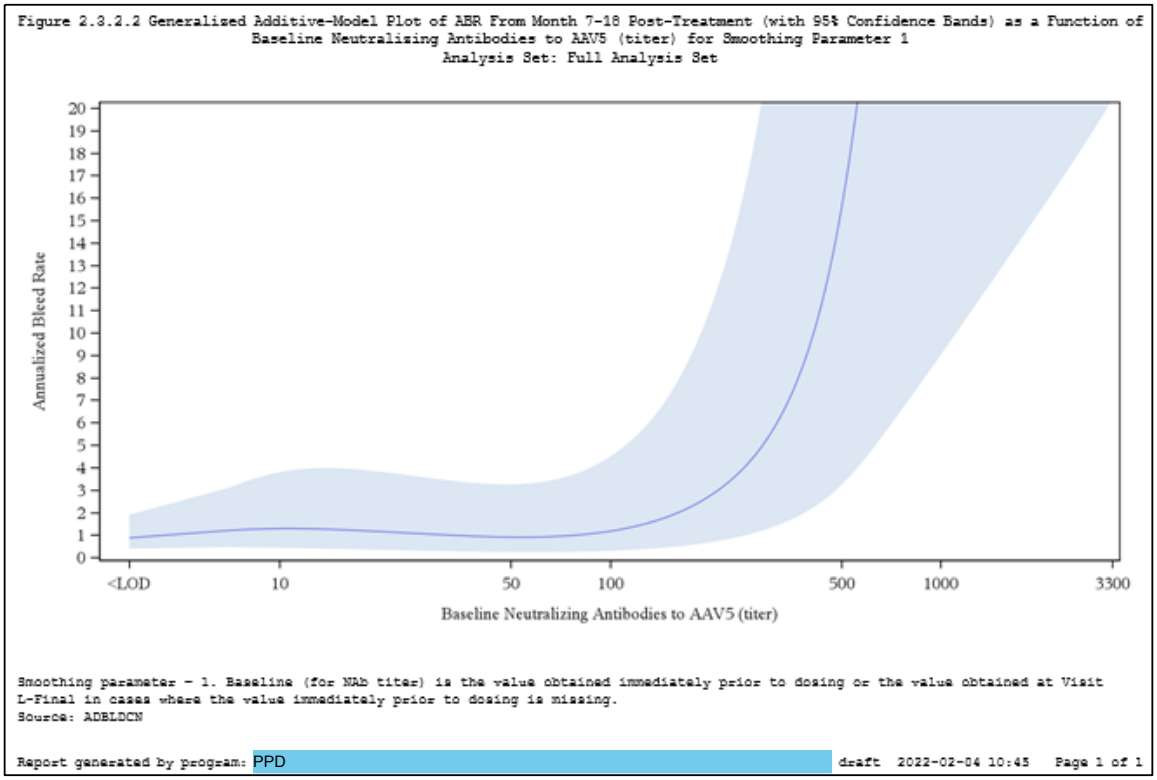
Report generated by program: PPD draft 2022-02-04 10:45 Page 1 of 1

Figure 2.3.2.2 Generalized Additive-Model Plot of ABR From Month 7-18 Post-Treatment (with 95% Confidence Bands) as a Function of Baseline Neutralizing Antibodies to AAV5 (titer) for Smoothing Parameter .1
Analysis Set: Full Analysis Set



Smoothing parameter = .1. Baseline (for NAb titer) is the value obtained immediately prior to dosing or the value obtained at Visit L-Final in cases where the value immediately prior to dosing is missing.
Source: ADBLDCM

Report generated by program: PPD 2022-02-04 10:45 Page 1 of 1



7.6.13 Correlation of factor IX activity levels during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay

The arithmetic mean of uncontaminated central laboratory one-stage aPTT factor IX activity level across all visits during the post-treatment period from the stable FIX expression (Month 6 Visit) through the 52 weeks following stable FIX expression (Month 18 Visit) was computed for each subject. The Pearson and Spearman correlation between this mean factor IX activity and the pre-IMP anti-AAV5 antibody titers (using the luciferase based NAb assay) were tabulated as well as their 95% confidence intervals *(Figure 2.3.1.1.3a)*. The Pearson product-moment correlation coefficient (rp) provides a measure of the strength of a linear association between factor IX activity levels and pre-IMP Anti-AAV5 antibody titers and the Spearman correlation coefficient (rs) provides a measure of the strength of a monotone association between factor IX activity levels and pre-IMP Anti-AAV5 antibody titers. A scatter plot was produced *(Figure 2.3.1.1.3b)*, with an overlaid linear regression line *(Figure 2.3.1.1.3)*. This endpoint does not have hypothesis testing and therefore is not included in the type I error control. The data showed a weak negative linear relationship, with a confidence interval for the correlation that barely excluded zero.

If a subject had zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that was to be used in the analysis was imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity. It turned out that there were no such subjects at the Month 18 data cut (or the 52-week data cut).

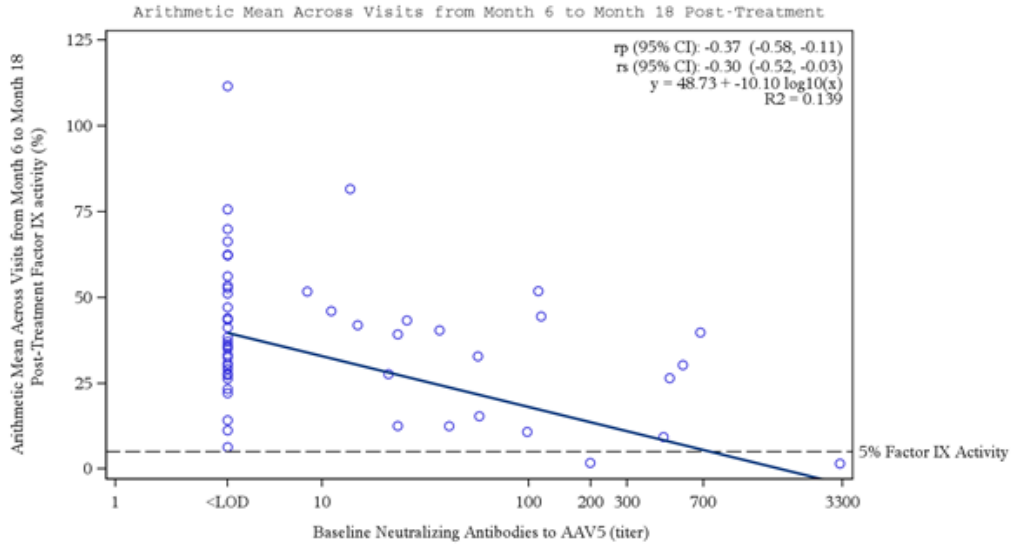
If a subject still lacked the Month 18 (Month 12 for the 12-month data cut) value (being still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value closest in time to (either before or after) 18 months (12 months for the 12-month data cut) post-AMT was employed in the analysis. If two such assessments were both the closest in time, with one being before and the other being after, the earlier assessment was used.

The main population is the FAS *(Figure 2.3.1.1.3)*. A sensitivity analysis used the PP population *(Figure 2.3.1.1.4)*. Analyses were also undertaken using the Lead-In Final Visit neutralizing antibodies rather than Baseline *(Figure 2.3.1.1.3c, 2.3.1.1.3d, 2.3.1.1.3e)*.

Finally, these analyses were repeated excluding subjects with NAb titer > 3000 *(Figures 2.3.1.1.5, 2.3.1.1.6, 2.3.1.1.7, 2.3.1.1.7a, 2.3.1.1.7b, 2.3.1.1.7c, 2.3.1.1.7d, 2.3.1.1.7e, and 2.3.1.1.8)*. A clear effect on the regression line and correlations was observed, and can be seen in the figures below.

Figure 22: Scatterplot of FIX Activity for Arithmetic Mean Across Month 6 to 18 by Baseline Neutralizing Antibodies to AAV5

Figure 2.3.1.1.3 Scatterplot of Uncontaminated Central Laboratory One-Stage aPTT Factor IX Activity (%) at Arithmetic Mean Across Visits from Month 6 to Month 18 Post-Treatment by Baseline Neutralizing Antibodies to AAV5 (titer) with Linear Regression
Analysis Set: Full Analysis Set



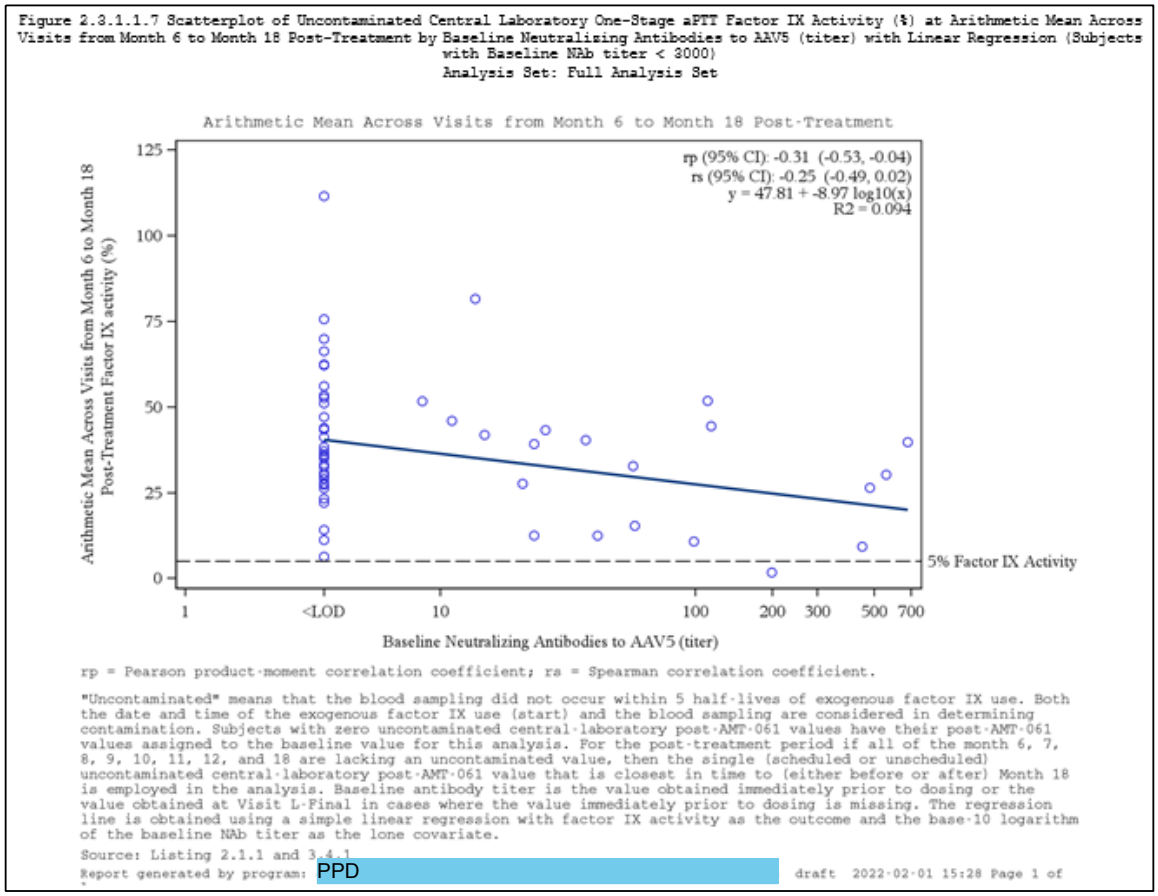
rp = Pearson product-moment correlation coefficient; rs = Spearman correlation coefficient.

"Uncontaminated" means that the blood sampling did not occur within 5 half-lives of exogenous factor IX use. Both the date and time of the exogenous factor IX use (start) and the blood sampling are considered in determining contamination. Subjects with zero uncontaminated central-laboratory post-AMT-061 values have their post-AMT-061 values assigned to the baseline value for this analysis. For the post-treatment period if all of the month 6, 7, 8, 9, 10, 11, 12, and 18 are lacking an uncontaminated value, then the single (scheduled or unscheduled) uncontaminated central-laboratory post-AMT-061 value that is closest in time to (either before or after) Month 18 is employed in the analysis. Baseline antibody titer is the value obtained immediately prior to dosing or the value obtained at Visit L-Final in cases where the value immediately prior to dosing is missing. The regression line is obtained using a simple linear regression with factor IX activity as the outcome and the base-10 logarithm of the baseline NAb titer as the lone covariate.

Source: Listing 2.1.1 and 3.4.1

Report generated by program: PPD draft 2022-02-01 15:26 Page 1 of 1

Figure 23: Scatterplot of FIX Activity for Arithmetic Mean Across Month 6 to 18 by Baseline Neutralizing Antibodies to AAV5 (Subjects with Baseline NAb titer < 3000)



7.6.14 Occurrence of (and resolution of) new target joints during the 6-18 month post-treatment follow-up

The rate of occurrence of new target joints per person-time of follow-up between stable FIX expression (post-treatment Month 6) and the time that is 52 weeks following stable FIX expression (18 months) after the subject's dosing (with AMT-061) was summarized descriptively across subjects *(Table 2.2.9.1)*. Target joints counted were ones that did not exist prior to stable FIX expression (post-treatment Month 6). The percentage resolution of such new target joints was also tabulated.

This endpoint did not have hypothesis testing and therefore was not included in the type I error control.

The main population was the FAS *(Tables 2.2.9.1)*. Sensitivity analyses used the PP population *(Tables 2.2.9.2)*, which were essentially identical to the main analyses.

As these analyses were descriptive and relatively simple in nature, no model diagnostics are presented.

7.6.15 Time to resolution of pre-existing target joints during the post-treatment follow-up

The time to resolution of pre-existing target joints (existing immediately prior to AMT dosing) was summarized *in Table 2.2.9.3*. Time to resolution was presented using the date of AMT dosing as the reference date. Each target joint was handled as the experimental unit for this analysis, irrespective of subject. There were 2 pre-existing target joints (subjects PPD , and PPD) that resolved during the study.

This endpoint did not have hypothesis testing and therefore was not included in the type I error control.

The main population was the FAS *(Table 2.2.9.3)*. A sensitivity analysis used the PP population *(Table 2.2.9.4)*. The results were identical.

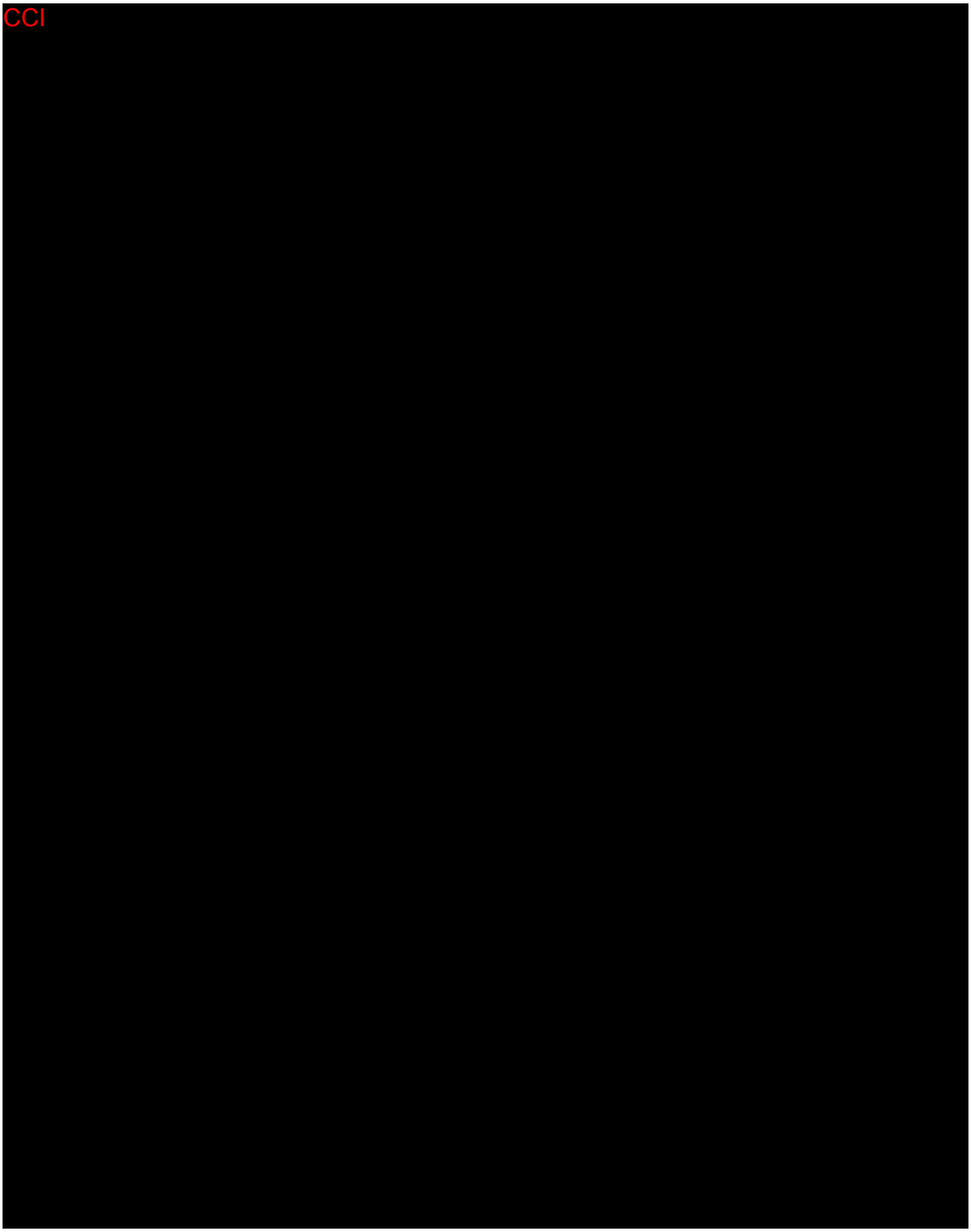
7.6.16 Proportion of subjects with zero bleeds in the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up

The number and percentage of subjects with zero bleeds during the post-treatment period from stable FIX expression (Month 6) to the time that is 52 weeks following stable FIX expression (18 months) after the subject's dosing with AMT-061 were summarized with descriptive statistics and presented in a table *(Table 2.2.10.1)*. This endpoint did not have hypothesis testing and therefore is not included in the type I error control.

The percentage of subjects with zero bleeds by baseline anti-AAV5 neutralizing antibodies subgroup (NAb titer negative/positive) were also presented *(Table 2.2.10.1a)*, as well the percentage of subjects with zero bleeds for subjects with a NAb titer < 3000 *(Table 2.2.10.2)*.

As this analysis is descriptive and simple in nature, no model diagnostics are presented.

CCI


CCI


CCI

CCI

CCI


CCI

7.7 Type I Error Control and Simultaneous Confidence Intervals

Formal statistical testing of the efficacy endpoints was performed using the closed testing principle (for Type I error control for multiple testing). Due to the closed testing principle, no correction for multiplicity is necessary. Among the endpoints being formally tested for statistical significance, all were tested for superiority at a one-sided alpha level of 0.025 (except as otherwise noted). Superiority testing and non-inferiority testing were accomplished using the FAS population.

Fixed sequential testing was performed using a hierarchical approach and was continued until a non-significant result was obtained (except as otherwise noted). The order of fixed sequential tests is specified below:

1. ABR comparison between AMT-061 and prophylaxis for non-inferiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up (primary efficacy endpoint)
2. Endogenous factor IX activity at 6 months after AMT-061 dosing (first secondary efficacy endpoint)
3. Endogenous factor IX activity at 12 months after AMT-061 dosing (second secondary efficacy endpoint)
4. Endogenous factor IX activity at 18 months after AMT-061 dosing (third secondary efficacy endpoint)

5. Annualized consumption of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase (secondary efficacy endpoint)
6. Annualized infusion rate of factor IX replacement therapy during the week52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding factor IX replacement for invasive procedures, compared to the lead-in phase (secondary efficacy endpoint)
7. Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 52 weeks following stable FIX expression (6-18 months) (secondary efficacy endpoint)
8. ABR comparison between AMT-061 and prophylaxis for superiority between the lead-in and the 52 weeks following stable FIX expression (6-18 months) post-treatment (AMT-061) follow-up (secondary efficacy endpoint)
9. Rate of spontaneous bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to lead-in phase (secondary efficacy endpoint)
10. Rate of joint bleeding events during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up compared to the lead-in phase (secondary efficacy endpoint)
11. CCI [redacted] questionnaire scores from the CCI [redacted] during the 12 months following AMT 061 dosing compared with the lead-in phase (secondary efficacy endpoint)
12. CCI [redacted] questionnaire scores from the CCI [redacted] during the 12 months following AMT 061 dosing compared with the lead-in phase (secondary efficacy endpoint).

Simultaneous one-sided 97.5% CIs based on a graphical approach to multiple testing (Bretz et al. 2015; Guilbaud 2008; Strassburger and Bretz 2008) were provided for the type I error controlled efficacy endpoints as a supportive analysis (Table 2.2.13). For endpoints for which an increase is favorable, the lower one-sided 97.5% confidence bound was provided; for endpoints for which an increase is unfavorable, the upper one-sided 97.5% confidence bound was provided

Figure 26: Type I Error-Controlled Endpoints

Table 2.2.13 Type I Error-Controlled Endpoints

Endpoint	Analysis set	Point Estimate	95% CI	One-sided P-value	Statistical Significance by Type I Error Control Framework [a]	Simultaneous One-Sided 97.5% Confidence Interval [b] LCL	UCL
Adjusted ABR Ratio (Month 7-18 Post-Treatment: Lead-In Period) for non-inferiority (Primary Efficacy)	FAS	0.36	0.20,0.64	NA	Y	NA	1.80
Change From Baseline One-Stage aPTT Factor IX Activity (%) at 6 Months Post-Treatment	FAS	36.00	31.47,40.54	<0.0001	Y	0.00	NA
Change From Baseline One-Stage aPTT (%) Factor IX Activity at Year 1 Post-Treatment	FAS	38.82	34.04,43.60	<0.0001	Y	0.00	NA
Change From Baseline One-Stage aPTT (%) Factor IX Activity at Month 18 Post-Treatment	FAS	34.31	29.55,39.08	<0.0001	Y	0.00	NA
Mean Difference in Annualized Consumption of Factor IX Replacement Therapy Use (IU/kg/yr) (Month 7-18 Post-Treatment – Lead-In Period)	FAS	-3056.8	-3642.8,-2470.8	<0.0001	Y	NA	0.0
Adjusted ratio for Annualized Infusion Rate of Factor IX Replacement Therapy (Month 7-18 Post-Treatment: Lead-In Period)	FAS	0.03	0.01,0.10	<0.0001	Y	NA	1.00
Odds Ratio One-Stage aPTT Factor IX Activity < 12% of normal (Month 6-18 Post-Treatment: Lead-In Period)	FAS	0.036	0.014,0.093	<0.0001	Y	NA	1.000
Adjusted ABR Ratio (Month 7-18 Post-Treatment: Lead-In Period) for Superiority	FAS	0.36	0.20,0.64	0.0002	Y	NA	1.00
Adjusted ABR Ratio (Month 7-18 Post-Treatment: Lead-In Period), Spontaneous Bleeds	FAS	0.29	0.12,0.71	0.0034	Y	NA	1.00
Adjusted ABR Ratio (Month 7-18 Post-Treatment: Lead-In Period), Joint bleeds	FAS	0.22	0.10,0.46	<0.0001	Y	NA	1.00

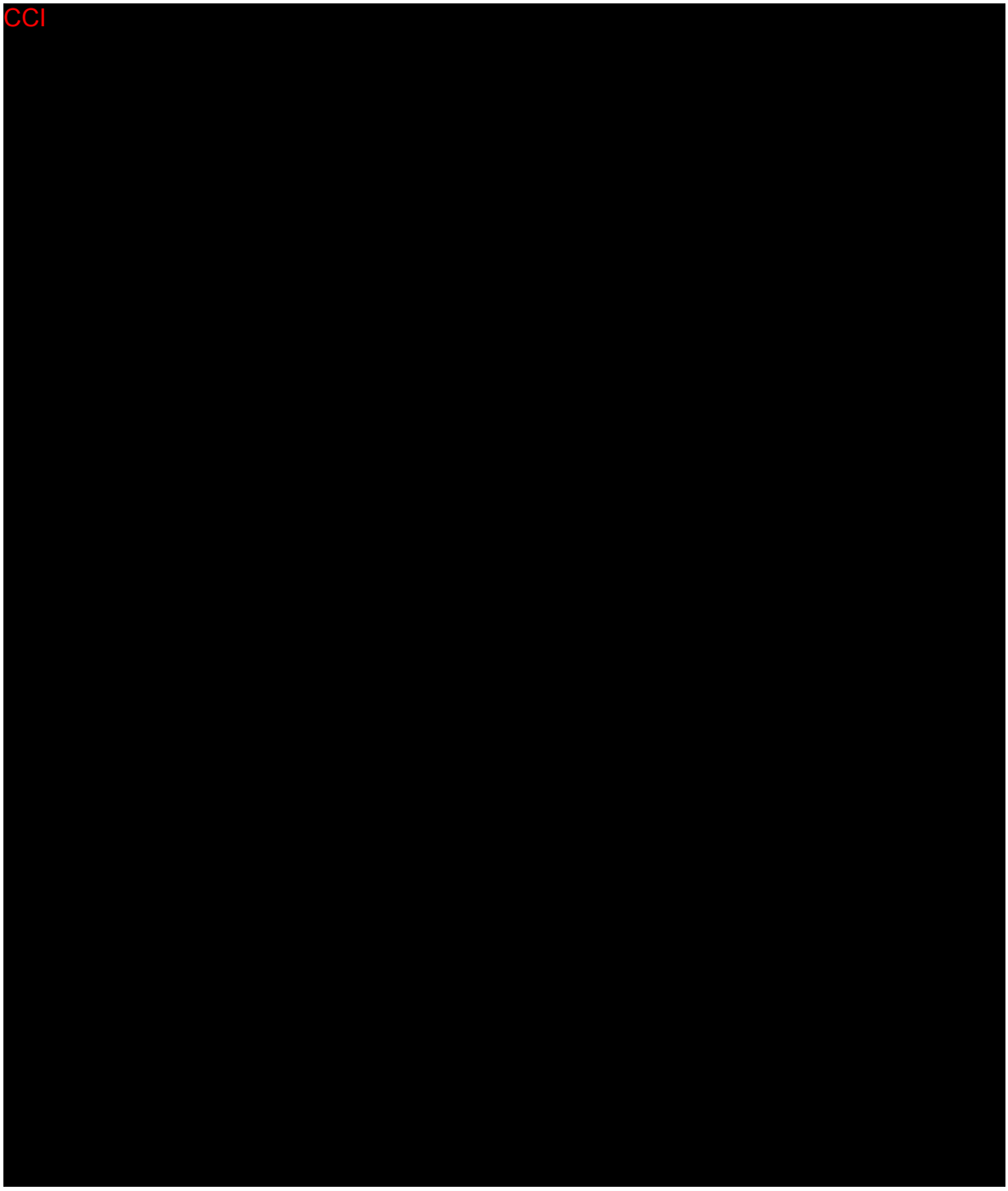
ence in CCI FAS -721.2 -1770.6,328.3 0.9121 N -1770.6 NA
(Post-T Lead-In
P

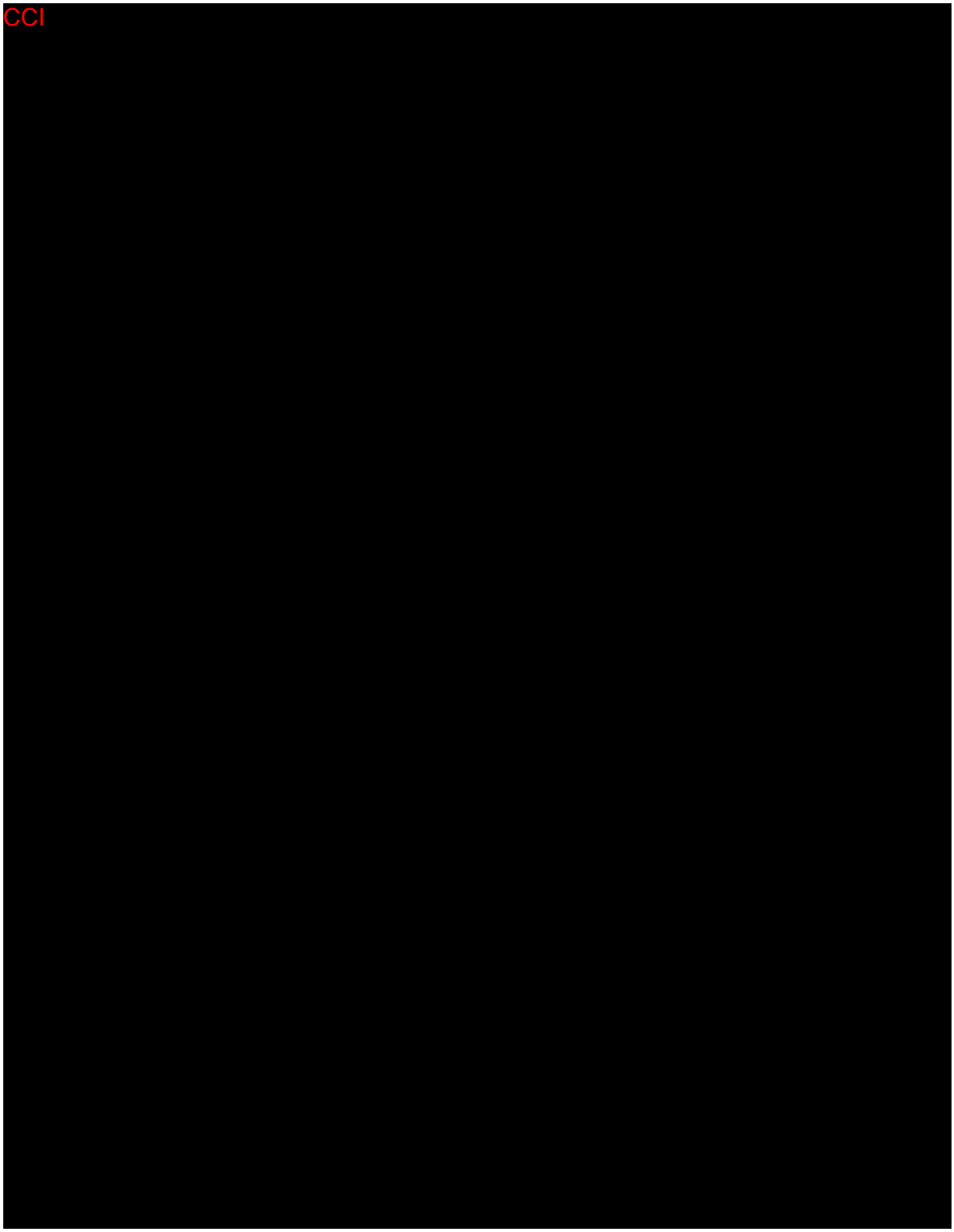
erence in CCI FAS 0.1 -3.5,3.8 0.4753 N -INF NA
(Post-Tre
P

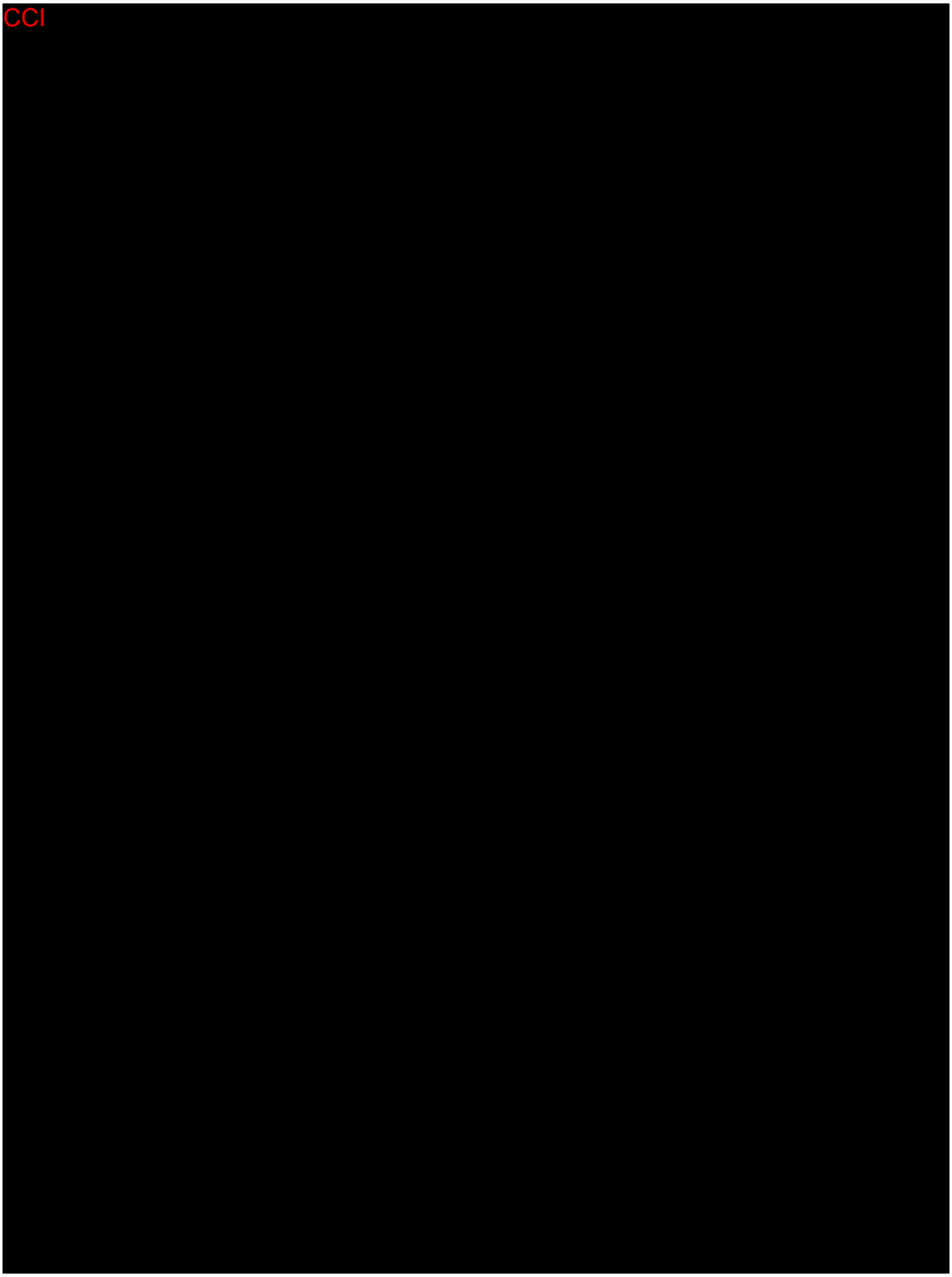
an; CCI ; CCI
; IN
esting and is not assessed for the endpoints listed after the first non-significant result.

[b] Supplemental Simultaneous One-Sided Confidence Limits are adjusted for the Type I Error Control Framework using the graphical approach to multiple testing that employs the sequential testing procedure.

Report generated by program: PPD draft 2021-11-22 14:46 (Page 2 of 2)


CCI


CCI


CCI

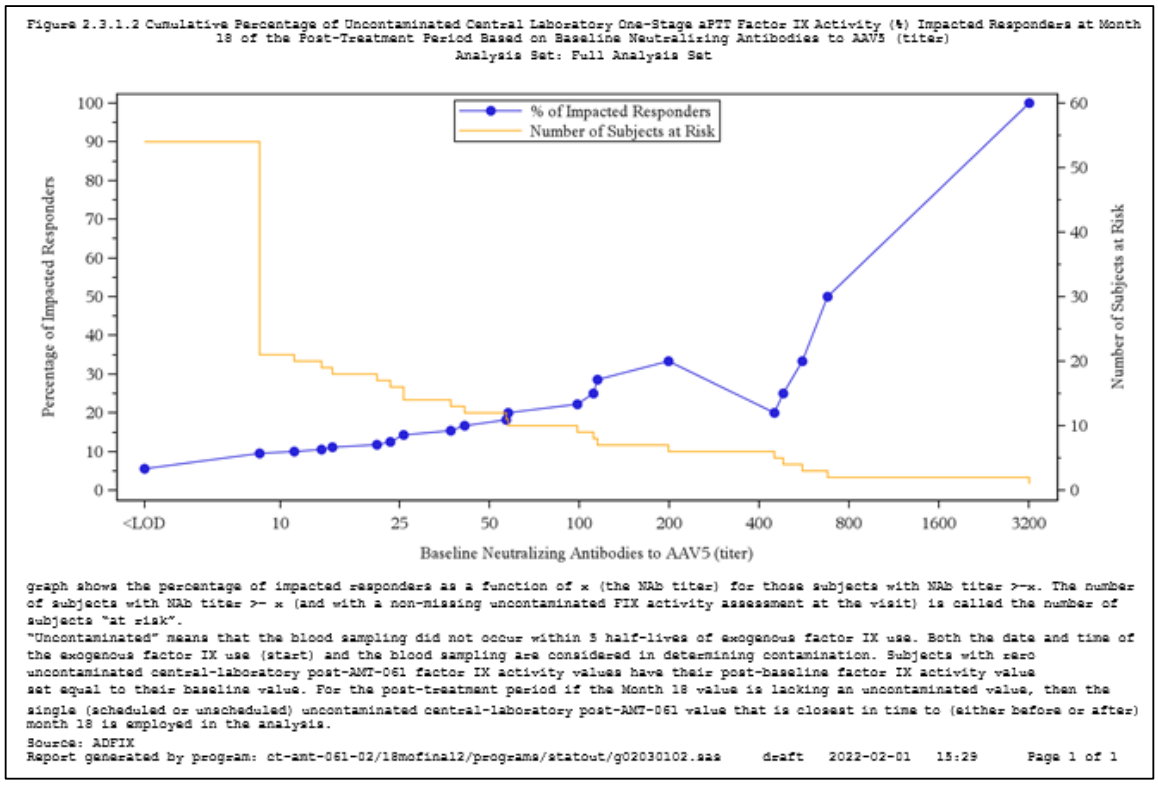
7.8.4 Impacted Responders Analysis: Correlation of factor IX activity levels at Month 18 with pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay

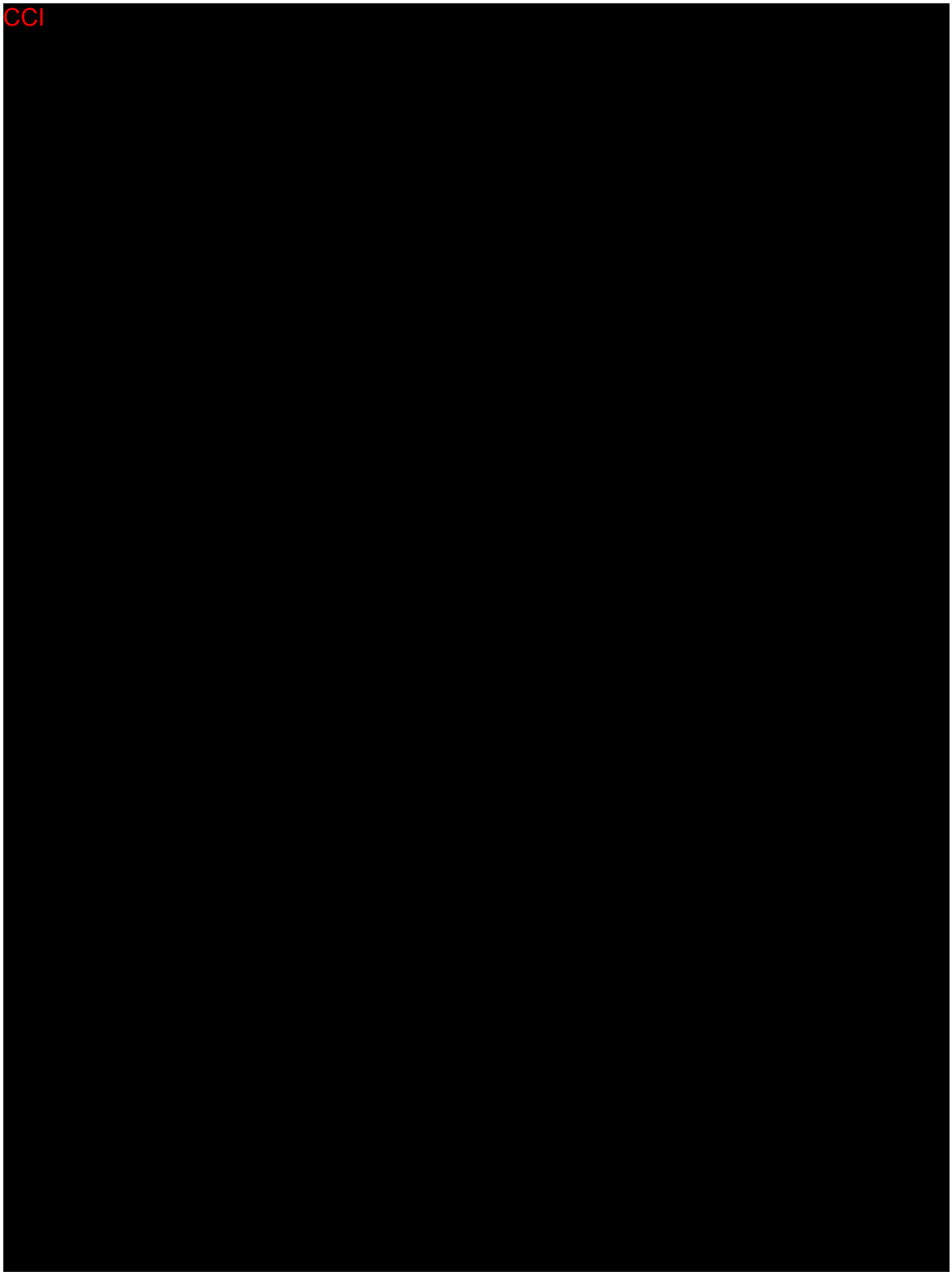
An impacted-response curve was created – as an exploratory analysis – to examine the association between these two variables (factor IX activity levels at Month 18 and pre-IMP anti-AAV5 antibody titers using the luciferase based NAB assay). An impacted response (for the purpose of NAB effect determination) was defined as a subject's having an uncontaminated one-stage aPTT assay for factor IX activity (%) to be < 5% of normal at Month 18. The percentage of subjects having impacted response for the group of subjects with NAB titer >= x was plotted as a function of x. The number of subjects with NAB titer >=x was also indicated as a function of x on the graph. If, based on this graph, there exists a value "x" of NAB titer above which > 25% of subjects have impacted response for a group (and if 12 or more subjects have NAB titer above that value "x"), then that NAB titer value "x" was considered to be a potential candidate for being a meaningful NAB cutoff; otherwise, no such candidate NAB cutoff titer was identified. If a subject had zero uncontaminated central-laboratory post-AMT-061 factor IX activity values, factor IX activity at any post-AMT planned assessment time point that was to be used in the

analysis was imputed based on the historical hemophilia B severity as documented on the CRF in a manner identical to that used for baseline factor IX activity. It turns out that there were no such subjects at the Month 18 data cut (and the Week 52 data cut). If a subject lacked a month 18 value (that was still missing even after the use of windowing to allow assignment of unplanned assessments to planned-assessment visits), then the uncontaminated central-laboratory post-AMT value closest in time to 18 months (12 months for the 12-month data cut) – either before or after – was employed in the analysis. If two such assessments were both the closest in time, with one being before and the other being after, the earlier assessment was used.

The pre-treatment NAb titer taken pre-dose on the day of dosing was used. If this result was not available, the value closest in time prior to dosing with AMT-061 (on or before the day of dosing) was used. Results are presented in Figure 27 below (*Figure 2.3.1.2 in CSR)*. Caution should be exerted when interpreting this figure because the sharp rise in the curve (of percentage of "impacted responders") beyond an NAb titer of 400 is due to a single impacted-response subject that is the sole impacted member of the increasingly smaller remaining "number of subjects at risk".

Figure 27: Impacted Responders Analysis




CCI

7.8.7 Subgroup Analyses

- Subgroup analyses were carried out for the following endpoints (the subgroups are mentioned a bit farther below in this document):
    - Endogenous factor IX activity at Month 18 *(Table 2.4.1)*
    - Annualized consumption of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up,

excluding replacement for invasive procedures, compared to the lead-in phase *(Table 2.4.2.1, 2.4.2.2)*

- Annualized infusion rate of factor IX replacement therapy during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up, excluding replacement for invasive procedures, compared to the lead-in phase *(Table 2.4.2.3)*
- ABR comparison between AMT-061 (during the 52 weeks following stable FIX expression [6-18 months] post-treatment follow-up) and factor IX prophylaxis (during the lead-in period) *(Table 2.4.3.1, 2.4.3.2, 2.4.3.3, 2.4.3.4, 2.4.3.5)*
- Comparison of the percentage of subjects with trough factor IX activity <12% of normal between the lead-in phase and after treatment with AMT-061 over the 52 weeks following stable FIX expression (6-18 months) *(Table 2.4.4)*
- Proportion of subjects remaining free of previous prescribed continuous routine prophylaxis during the 52 weeks following stable FIX expression (6-18 months) post-treatment follow-up *(Table 2.4.5)*.

The subgroup analyses for the aforementioned endpoints were carried out for the following subgroups:

- Age categories: <40 years, 40 to <60 years, >= 60 years
- Race and/or Ethnicity subgroups (with categories to be specified later because the racial/ethnic frequencies are not well known in advance)
- Zero bleeds versus >=1 bleed in lead-in
  - Because this subgrouping is defined using information from the lead-in period, the analysis will provide descriptive statistics only and will provide those descriptive statistics for only the post-treatment period.
- Presence or absence of target joints at Screening
- Baseline NAB titer categories: positive titer (>= LOD) versus negative titer (<LOD), where LOD denotes limit of detection.
- HIV-negative vs. controlled HIV positive (CD4+ count >200 /μL) at Baseline
- History of Hepatitis B or C at Baseline
- Baseline liver pathology, according to Baseline FibroScan™ or equivalent SWE (shear wave elastography), MRE (magnetic resonance elastography) result:
  - Degree of fibrosis [≥9Kpa versus <9Kpa]

- Degree of steatosis [Controlled Attenuation Parameter (CAP) score ≥S2 (≥ 260 dB/m) versus <S2 (<260 dB/m)] versus Missing.

All efficacy endpoints will be analyzed using data over the 2 to 5 year follow-up. Please see Section 11 for additional details.

7.8.8 Optional Sub-Study Efficacy and CCI Endpoint Analyses

Optional sub-study endpoints consist of:

- PROBE questionnaire sub-study summary scores
- Musculoskeletal ultrasound sub-study ultrasound results.

Analysis was to be based on the FAS population for the set of subjects participating in the respective sub-study. Any subject with at least one assessment of the sub-study endpoint was considered to be participating in the respective sub-study. Because MSKUS expert-reader interpretations were not available in time for the database lock, MSKUS data were left out of the CSR post-text tables, listings, and graphs for this Month 18 CSR.

7.8.8.1 PROBE Questionnaire Summary scores

PROBE Questionnaire summary scores and individual item responses are summarized descriptively by treatment and visit *(Table 2.5)*. The following statistics are displayed: n, mean (SE), SD, Q1, median, Q3, minimum, and maximum. Summary scores and individual item responses are also listed *(Listing 2.4.1.1.1, 2.4.1.1.2, 2.4.1.2.1, 2.4.1.2.2)*. A higher score indicates better health.

7.8.8.2 Musculoskeletal Ultrasound results

There will be a separate SAP document for the statistical analysis of the Musculoskeletal Ultrasound results. Because MSKUS expert-reader interpretations were not available in time for the database lock, MSKUS data were left out of the CSR post-text tables, listings, and graphs for this Month 18 CSR.

7.9 Safety Analyses

All safety analyses were based on the safety population.

The safety endpoints to be analyzed were:

- TEAEs
- Changes in abdominal ultrasound
- Anti-AAV5 antibodies (total [IgM and IgG], neutralizing antibodies)

- AAV5 capsid-specific T cells
- Anti-FIX antibodies
- Factor IX inhibitors and recovery
- Hematology and serum chemistry parameters
- ALT and AST levels and corticosteroid use for ALT and AST increases
- Vector DNA in blood and semen
- Inflammatory markers: IL-1 β, IL-2, IL-6, IFN γ, MCP-1
- AFP.

7.9.1 Adverse Events

An adverse event was considered to be treatment-emergent for the AMT-061 treatment (i.e. a TEAE) if the event occurred after the administration of the IMP, or if the AE worsened during the study after the dose of study drug (intensity and/or severity changed to a worsened grade). An adverse event that begins on the same date as the IMP administration was treatment-emergent if the AE began after the time of dose or if the time of AE onset was unknown. Additionally, if an AE had an onset date during post-treatment period and had an outcome of death, that death was considered to be treatment-emergent. Furthermore, if the AE could possibly be treatment-emergent, based on the missing or incomplete date, then the AE was regarded as treatment-emergent. A treatment-emergent adverse event can be described as having had "incidence" during the treatment period.

An adverse event was counted as having had "incidence" during the lead-in period if it occurred during the lead-in period, or if the AE worsened during the lead-in period (intensity and/or severity changed to a worsened grade). Additionally, if an AE had an onset date during the lead-in period and had an outcome of death, that death was counted as having incidence during the lead-in period. Furthermore, if the AE could have had incidence during the lead-in period, based on the missing or incomplete date, then the AE was regarded as having incidence during the lead-in period.

An adverse event incidence table for the safety populations was created displaying the number of subjects (and percentage) experiencing an incident event and the number of incident events for: any AEs, AEs of special notification, serious AEs, related AEs, serious and related AEs, AEs leading to early treatment discontinuation (i.e. to a partial dose), mild/moderate/severe AEs, and deaths *(Table 3.1.1)*.

The following AE incidence summary tables are presented by decreasing frequency of occurrence based on SOC and Preferred Term:

1. AEs for the lead-in and post-treatment safety populations *(Table 3.1.2.1, 3.1.3)*
2. Serious AEs for the lead-in and post-treatment safety populations *(Table 3.1.4)*
3. Related TEAEs for the post-treatment safety population *(Tables 3.1.5.1, 3.1.5.2)*
4. Related serious TEAEs for the post-treatment safety population *(Table 3.1.6)*
5. TEAEs leading to treatment discontinuation for the post-treatment safety population (treatment discontinuation means receiving only a partial dose) *(Table 3.1.7)*
6. Serious TEAEs leading to treatment discontinuation for the post-treatment safety population (treatment discontinuation means receiving only a partial dose) *(Table 3.1.8)*
7. Fatal AEs *(Table 3.1.9)*
8. TEAEs by highest severity for the post-treatment safety population *(Table 3.1.11)*
9. Related TEAEs by highest severity for the post-treatment safety population *(Table 3.1.12)*
10. Incidence of TEAEs for Special Notification for the post-treatment safety population *(Table 3.1.13)*
11. Incidence of non-serious TEAEs occurring in at least 5% of subjects in the post-treatment period *(Table 3.1.14)*
12. The incidence of TEAEs occurring in at least 10% of subjects in the post-treatment period *(Table 3.1.15)*.

Adverse event incidence summary tables are also provided for subjects with elevated transaminase values at dosing: AEs *(Table 3.1.2.2)*, related AEs *(Table 3.1.2.3)*. As well, tables are provided for subjects with baseline Nab titer < 3000: AEs *(Table 3.1.2.4),* serious AEs *(Table 3.1.17),* AEs of special notification *(Table 3.1.18),* AEs by highest severity *(Table 3.1.19, 3.1.20)*, and related AEs *(Table 3.1.21)*. Finally, tables are provided by baseline Anti-AAV5 neutralizing antibodies subgroup: *AEs (Table 3.1.22)*, related AEs *(Table 3.1.23),* serious AEs *(Table 3.1.24)*, and related serious AEs *(Table 3.1.25)*.

All incident AEs were tabulated by SOC and preferred terms within each SOC according to the Medical Dictionary for Regulatory Activities (MedDRA) terminology list. The version of the MedDRA current at the time of database lock was used to code verbatim terms for AEs for final analysis of the data (MedDRA 24.1). A glossary of MedDRA preferred terms used for adverse events reported in the study along with the associated Investigator's verbatim term was provided. No hypothesis tests were performed.

The summary tables are accompanied by individual subject listings of all AEs, including pre-treatment AEs and information on actual AE description, date/time of start and end of AE,

preferred term (MedDRA), SOC (MedDRA), severity, relationship/causality, type of AE, action taken, seriousness and outcome *(Listing 3.1.1)*. Pre-existing AEs are flagged. Pre-existing AEs were not considered to be treatment emergent, except in case of worsening during/after trial treatment (to be collected as a separate AE). Separate listings were created for AEs for special notification, deaths, and SAEs *(Listings 3.1.2, 3.1.3, 3.1.4)*. A listing of adverse events is also provided for subjects with elevated transaminase values at dosing *(Listing 3.1.5)*. All adverse events, whether treatment-emergent or not, are included in the listings. A listing of any reported deaths during lead-in and post-treatment periods is provided and includes the number of days since IMP administration.

The following was done for events with irregular onset dates**.** All AEs were included in the data listings regardless of the completeness of the onset dates. Any partial dates were used in order to determine whether an AE is lead-in-incident or treatment-emergent using the rules in Appendix 1; however, imputed dates are not provided in the data listings.

7.9.1.1 Adverse Events of Special Notification

Table 5 contains (S)AEs that qualify for special notification, as they are seen as safety issues of particular concern for Advanced Therapy Medicinal Product (ATMP) (ENTR/F/2/SF/dn D (2009) 35810. Brussels, 03/12/2009) and gene therapy medicinal products (EMA/CHMP/GTWP/60436/2007):

Table 5: Adverse Events of Special Notification

AEs related to the IMP administration procedure
Suspected or confirmed cases of opportunistic or serious infections that in the investigator's opinion might be related to the IMP
Unexpected reactions (e.g., hypersensitivity, immunological, toxic or other as consequence of a change in the construction or function of the viral vector [e.g., generation of replication competent virus])
AEs related to product failure (including lack of efficacy)
AEs related to mandatory concomitant medication (e.g., immunosuppression)
AEs related to medical devices which form part of the product or are used for application of the product
Development of any new/recurrent cancer.

These AEs were reported and followed in the same manner as SAEs *(Tables 3.1.13, 3.1.18, 3.1.26, Listing 3.1.2)*. Note that the AEs may be serious or non-serious by definition (please see the protocol for more details). AEs of special notification are designated as such on the eCRF and therefore did not need to be derived.

7.9.1.2 Severity of Adverse Event

If an AE changed severity over time, the severity of maximum severity (i.e. intensity) was reported.

7.9.1.3 Relationship Between IMP and Adverse Event

Please refer to the protocol for the definitions of related to IMP, probably related to IMP, possibly related to IMP, and not related to IMP.

7.9.2 Changes in Abdominal Ultrasound

To monitor subjects for liver fibrosis and potential occurrences of liver malignancies, abdominal ultrasound assessments were performed. These ultrasound assessments occurred at the final Lead-In visit at the latest (to establish baseline status), at post-treatment Month 12, and then annually thereafter.

A shift table was used to summarize normal and abnormal results at Month 12 and the subsequent follow-up visits relative to the results obtained at baseline *(Table 3.9)*.

All abdominal ultrasound data are listed.

7.9.3 Anti-AAV5 Antibodies, Anti-Factor IX Antibodies, and Factor IX Inhibitors

Total IgG and IgM antibodies against the vector capsid were evaluated by the enzyme-linked immunosorbent assay (ELISA), and anti-AAV5 neutralizing antibodies are assessed with the luciferase assay. Antibodies against FIX were evaluated by ELISA and reported as IgG, IgM. The results from the total IgG and IgM antibodies against the vector capsid, neutralizing antibodies against the vector capsid and non-inhibitory factor IX antibodies were tabulated by visit using descriptive statistics (for the titer) of n, mean (SD), SD, Q1, median, Q3, and max titer at each visit. The titer of factor IX inhibitors was reported in Bethesda Units and the sub-class of immunoglobulin of the inhibitor was displayed as IgG, IgM or others. These results are displayed at each visit *(Table 3.2.1, 3.2.2)*. All data is listed *(Listing 3.4.1, 3.4.2)*.

A subject was said to suffer from factor IX inhibitors if the subject tested positive for factor IX inhibitors at two consecutive tests, performed preferably within two weeks. Occurrences of "suffering from factor IX inhibitors" were flagged in the factor IX inhibitor listing *(Listing 3.5)*.

Measurement of factor IX recovery (maximum concentration [C_{max}]) and incremental recovery measured as increase in activity per unit infused (percent per U/kg) at 30 min after infusion of a dose of factor IX were performed at baseline Visit L-Final. Additionally, measurement of factor IX recovery and incremental recovery were done at suspicion of factor IX inhibitor as judged by the investigator *(Listing 3.9)*.

All data is listed *(Listings 3.4.1 [NAb to AAV5], 3.4.2 [Anti-AAV5 Antibodies], 3.5 [Anti-FIX Antibodies], 3.9 [Recovery])*.

7.9.4 AAV5 capsid-specific T cells

The AAV5 capsid-specific T cells testing (ELISPOT) was summarized by visit using descriptive statistics of n, mean (SD), SD, Q1, median, Q3, minimum, and maximum *(Table 3.3.1, 3.3.2)*.

The data is also listed *(Listing 3.6)*.

7.9.5 Clinical Laboratory Measurements

The lab parameters collected included the following:

Table 6: Safety Lab Parameters

Hematology	
Hemoglobin	White blood cells with differential count
Hematocrit	CD4+ count
Platelet count	
Red blood cells	
Serum Chemistry	
Sodium serum electrolytes	Alkaline phosphatase (ALP)
Potassium serum electrolytes	C-Reactive Protein
Creatinine	Albumin
Gamma-glutamyltransferase (GGT)	Total Bilirubin
AST	Glucose (non-fasting)
ALT	
Coagulation	
aPTT	
PT (or International Normalized Ratio [INR])	
Serology	
HIV viral load	Hepatitis B extracellular antigen (HBeAG) *
Hepatitis B surface antigen (HBsAG)	Hepatitis B virus (HBV) DNA
	Hepatitis C virus (HCV) RNA
Alpha-fetoprotein	
AFP	
Local Laboratory	
AST	
ALT	

* This parameter was removed with Protocol Amendment 3.

A Clinically Significant Laboratory Abnormality as identified by the investigator after the study drug was administered was recorded as an Adverse Event and tabulated as an AE in the AE analysis. Abnormalities occurring prior to the IMP administration were noted in medical history and presented in a data listing.

All laboratory data were stored in the database with the units in which they were originally reported. Laboratory data not reported in International System of Units (SI units; Système International d'Unités) were converted to SI units before data analysis.

Individual clinical laboratory variables for hematology, serum chemistry, coagulation, serology, and local laboratory are provided in listings. Comments for laboratory testing are listed. For listings, laboratory values are flagged as low or high based on the reference ranges provided by the central laboratory.

If there are multiple laboratory values for the same parameter for a visit, the last value was chosen for analysis.

Summary statistics (n, mean, Q1, median, Q3, standard deviation, minimum, and maximum) for the baseline assessment and change from baseline at each post-baseline visit for scheduled lab assessments of continuous laboratory variables were tabulated for post-treatment safety population *(Table 3.4.1, 3.4.3, 3.4.5, 3.4.6)*. Data from unscheduled visits was not used for the by-visit summaries (unless they had been assigned to a scheduled visit according to the Time Windows for Statistical Analysis). Data from both scheduled and unscheduled visits was listed.

Shift tables were produced using the categories defined by the Common Terminology Criteria for Adverse Events (CTCAE version 5.0) grades for the post-treatment safety population for hematology and serum chemistry *(Table 3.4.2, 3.4.4.1, 3.4.4.2)*. For these shift tables, the subject's pre-IMP grade was cross-tabulated by the subject's maximum post-treatment follow-up; also, the subject's maximum post-IMP grade during post-treatment follow-up was tabulated for all baseline grades combined. Percentages of subjects in each maximum post-IMP grade were calculated for each pre-dose grade for the treatment and also for all baseline grades combined. Laboratory abnormal values on-treatment were flagged as High or Low values based on laboratory reference ranges provided by LabCorp Laboratories (found in Appendix 3 of the SAP). These flags along with the reference ranges are provided in the laboratory data listings *(Listings 3.3.1, 3.3.2)*.

Potentially Clinically Significant Laboratory Values Above/Below a Clinically Relevant Threshold on-treatment, based on CTCAE and other criteria, will be identified based on the thresholds in the table below.

Table 7: Potentially Clinically Significant (PCS) Laboratory Parameter Criteria

Central Laboratory	Post-Baseline Criteria
Serum Chemistry	
Sodium serum electrolytes	NA
Potassium serum electrolytes	<3.0 mmol/L >6.0 mmol/L
Creatinine	>2 x ULN
Gamma-glutamyltransferase	NA
AST	>2 x Baseline
ALT	>2 x Baseline
ALP	>2 x ULN
CRP	NA
Albumin	NA
Total bilirubin	>2 x ULN
Glucose (non-fasting)	NA
Hematology	
Hemoglobin	<8.0 g/dL (<80 g/L) Increase of >40 g/L to a value above the ULN
Hematocrit	NA
Platelet count	<50 x 10^9/L >999 x 10^9/L
Red blood cells	NA
White blood cells with differential count	<2 x 10^9/L >35 x 10^9/L
CD4+ count	≤200/µL
Coagulation	
aPTT	NA
PT (or INR)	NA
Serology	
HIV viral load	>200 copies/mL
HBsAg	NA
HBeAG	NA
Hepatitis B Virus DNA (HBV DNA)	NA
Hepatitis C Virus RNA (HCV RNA)	NA
Alpha-fetoprotein	
AFP	NA
Local Laboratory	
AST	>3 x ULN
ALT	>3 x ULN

NA: Not Applicable

Clinically significant laboratory values were tabulated for the lead-in safety population and the post-treatment safety population *(Table 3.4.7)*. All laboratory data for the parameter identified as potentially clinically significant for a subject are listed. Low platelet counts were counted as being clinically significant only if they occur ≥ 4 weeks after IMP administration.

On the listings *(Listing 3.3.8)*, the reference range and flag indicating if the measurement in question is outside the reference range are provided.

7.9.6 ALT Levels, AST Levels and Corticosteroid Use for ALT and AST Increases

Summary statistics (n, mean, Q1, median, Q3, standard deviation, minimum, and maximum) for the baseline assessment and change from baseline at each post-baseline visit for ALT levels, AST levels and corticosteroid use were tabulated and listed *(Table 3.4.2, Listing 3.3.3, Listing 3.3.10)*. Data from unscheduled visits was not used for the by-visit summaries (unless they have been assigned to a scheduled visit according to the Time Windows for Statistical Analysis). Data from both scheduled and unscheduled visits are listed.

Plots of individual subject profiles of ALT and AST levels over time are also displayed *(Figure 3.1.1)*. Corticosteroid use is indicated on the plots.

7.9.7 Vector Genome Detection in Post Treatment Period

The number of days until vector DNA could no longer be detected in semen and blood was tabulated *(Tables 3.7, 3.8)*. The number of days was calculated using the date of collection of the third consecutive negative sample for each matrix.

All data is listed *(Listing 3.8.1, 3.8.2),* and individual plots of AAV5 Vector DNA over time are presented *(Figures 3.5.1 and 3.5.2)*.

Two subjects, PPD, and PPD, had two Visit D Pre-IMP blood values in the database, and in both cases one value was positive, while the other was negative. Subject PPD first had a positive result, followed by a negative result later that day. Subject PPD had two results with the same date/time stamp, one positive, and one negative.

For subject PPD, the recorded time of the positive result was considered incorrect. The raw data nominally showed “Post-IMP”, but we in ADaM programmed the “Pre-IMP” designation based on the date/time being ostensibly earlier than the dosing time. The raw data did show a nominal time slot of “Post-IMP”, but the ADaM programming has been relying on the date/time to make such designations. To resolve this dilemma, a programming rule was established whereby if based on the recorded date/time two ostensibly pre-dose values are available (for a subject) for analysis with one being nominally “Pre-IMP” and the other nominally “Post-IMP”, then the nominally “Pre-IMP” value is used for the Pre-Dose time point. This rule was applied in *Figure 3.5.2*.

For subject PPD, by the same reasoning, the recorded time of the positive result was considered incorrect. The raw data nominally showed “Post-IMP” but we in ADaM programmed

the "Pre-IMP" designation based on the date/time being ostensibly earlier than the dosing time. The raw data did show a nominal time slot of "Post-IMP", but the ADaM programming is generally relying on the date/time to make such designations. To reconcile this issue, if two values both have exactly the same date/time but one is nominally pre-IMP and the other is nominally post-IMP, then the nominally pre-IMP value is to be considered as pre-treatment baseline.

Time to first shedding-negative was defined for each type of matrix and each patient as the post-treatment time point where a negative result was measured for the first time in a consecutive order of 3 or more time points with a negative result. A negative result was defined as a result of either '0' or 'LOD' (limit of detection). The time to first shedding-negative is flagged on the above-mentioned listings.

The time to first shedding-negative for the post-treatment period was also summarized using a Kaplan-Meier curve *(Figures 3.5.3, 3.5.4)*. The censoring time was truncated at the data cut-off date, the time of completion of the study, or time of early withdrawal from the study, whichever was earlier. For the 5-year analysis (and 5-year CSR), there will be no data cut-off date.

7.9.8 Inflammatory Markers

Blood samples were taken to assess IL-1β, IL-2, IL-6, IFNγ and MCP-1 (monocyte chemotactic protein-1) using ELISA.

The test results were summarized by visit using descriptive statistics *(Table 3.5)* and included a change from baseline calculation for each post-baseline measurement *(Listing 3.3.7)*. All data are listed *(Listing 3.3.7)*. Inflammatory markers were not summarized or presented following the interim six-month data cut as, the full set of inflammatory marker data was not available as part of the six-month database lock.

7.9.9 Alpha-fetoprotein

Alpha-fetoprotein results were summarized by period and visit *(Table 3.6)*. They are also listed *(Listing 3.3.6)*.

7.9.10 Physical Examination (Including Height and Weight)

A physical examination was performed at Screening (Visit S), L-Final, Visit D (pre-IMP), during the post-treatment follow-up at visits F1, F2, F4, F6, F12, F13, F15, F17, F19, and F-Final, and during the long-term follow-up at visits LTF1, LTF2, LTF3, LTF4, LTF6, and LTF8. Height was measured only at screening and weight was measured only at screening and Visit L-Final.

Height (without shoes) was measured and recorded, rounded to the nearest centimeter. Body weight (without overcoat and shoes) was measured and recorded, rounded to the nearest kilogram.

The physical examination included general appearance and bedside examination of the following body systems: lymph nodes, eyes and ears, mouth and throat, lungs, abdomen, extremities, musculoskeletal system, neurological system, cardiovascular system and skin.

Abnormal physical examination findings were reported as adverse events.

Abnormalities (e.g., scar at the left side at knee following total knee replacement, or arthropathy of left ankle due to hemophilia B) identified at screening were documented in the subject's source documents and on the medical history electronic case report form (eCRF). Changes after the Screening Visit were captured as AEs on the AE eCRF page, deemed clinically significant in the opinion of the investigator. These abnormalities were to be followed until they reached "final outcome" (please refer to the protocol).

7.9.11 Vital Signs

Blood pressure, pulse, and body temperature were measured at Screening (Visit S), Visit L-Final, at pre-IMP and post-IMP (3 hours) on Visit D and at all visits during the post-treatment phase. Before measurement of blood pressure and pulse, the subject was to rest for at least 5 minutes. For the individual subject, all measurements were to be performed while the subject was in the same position (i.e., sitting or lying) throughout the trial.

A summary of baseline weight, height, and BMI is presented by treatment period for the FAS, PP, and safety populations in the demographics table *(Table 1.2.1)*.

Vital signs values are listed *(Listing 3.7.1)*.

8. INTERIM 6 MONTH ANALYSIS

A partial database lock and data extraction were performed once the last subject achieved 6 months after AMT-061 therapy.

The first secondary efficacy endpoint, endogenous factor IX activity, has been analyzed. This endpoint/analysis is included in (added to) the 18-month-data-cut CSR.

9. INTERIM 12 MONTH ANALYSIS

A partial database lock and data extraction was performed once the last subject achieved 12 months after AMT-061 therapy.

The second secondary efficacy endpoint, endogenous factor IX activity, was analyzed. This endpoint/analysis is included in (added to) the 18-month-data-cut CSR.

10. CSR

After 52 weeks following stable FIX expression (18 months post-treatment), the database was locked, and all available efficacy and safety data collected between screening and 52 weeks following stable FIX expression (18 months post-treatment follow-up time) were analyzed and reported in a full CSR. All of the efficacy endpoints were analyzed.

Data up to each analysis time point was considered locked and was not to be changed (with the exception of ending dates for continuing events and treatments) without explicit authorization from the sponsor.

Table, listing, and figure shells for the Final CSR are provided in a separate document.

Factor IX activity is summarized (and listed) by visit, overall and by patient, over the 18-month period since administration of AMT-061. By-subject plots of factor IX activity over time have been overlaid with plots of exogenous factor IX consumption (time of administration) and with the times of occurrence of bleeding events over the 52 weeks following stable FIX expression (6-18 months) months subsequent to AMT-061 administration.

Descriptive statistics are provided for the estimated unadjusted ABR during time periods subsequent to stable FIX expression (6 months) after the dose of AMT-061 and during the Lead-in period. The unadjusted ABR is the number of bleeds divided by the person-time at risk during a given time period. For the 52 weeks following stable FIX expression (6-18 months) post-AMT, bleeds and person time on or after Day 1 and prior to stable FIX expression (Month 6) post-AMT are not included in the calculation.

Bleeding events are listed by subject. The number of bleeding events and the time-at-risk of bleeding events are listed by period for each subject.

The ratio of factor IX activity (%) to factor IX protein (%) has been tabulated. A table is also provided to summarize the factor IX activity (%) by patients with or without pre-existing neutralizing antibodies to factor IX. A scatter plot of factor IX activity (%) by baseline titer of neutralizing antibodies to AAV5 is also presented to show the correlation of baseline titer of neutralizing antibodies to AAV5 and factor IX activity over the 52 weeks following stable FIX expression (6-18 months).

11. CSR ADDENDUM

12. CHANGES FROM METHODS PLANNED IN THE PROTOCOL

- The full analysis set has become the primary population for the ABR non-inferiority analysis, while the per-protocol population has been relegated to a sensitivity analysis. The reason is that the FDA (Food and Drug Administration) statistical team requested this.
- The contamination period due to exposure to exogenous factor IX has been changed to the 5 half-life rule (from a 10-day rule) to allow greater accuracy.
- ABR has now become the sole primary endpoint. The reason is that the FDA (Food and Drug Administration) statistical and clinical teams requested this.
- Have changed the data cut for the main CSR to be at 18 months post-treatment. The reason is that the FDA asked for the efficacy analysis to pertain to the year after a stable factor IX activity level is reached.

13. STATISTICAL SOFTWARE

Data processing, statistical screening, descriptive reporting and analysis of the efficacy and safety data were performed using SAS (Version 9.4 or higher).

14. REFERENCES

Bretz F, Maurer W, and Maca J “Graphical Approaches to Multiple Testing.” Clinical Trial Biostatistics and Biopharmaceutical Applications, edited by Young W and Chen D, CRC Press, 2015, 349-394.

EuroQol Group. EQ-5D. Accessed at www.euroqol.org/about-eq-5d/how-to-use-eq-5d.html (November, 2014).

Farrar JT, Dworkin RH, Max MB. Use of the cumulative proportion of responders analysis graph to present pain data over a range of cut-off points: Making clinical trial data more understandable. J Pain Symptom Manage 2006; 31:369-377.

Guilbaud O. Simultaneous confidence regions corresponding to Holm's stepdown procedure and other closed testing procedures. Biometrical Journal, 2008; 50:678-692.

Paul J, Venugopal R, et al. Measurement of controlled attenuation parameter: A surrogate marker of hepatic steatosis in patients of nonalcoholic fatty liver disease on lifestyle modification – A prospective follow-up study. Arq. Gastroenterol. 55(1): São Paulo Mar./Apr. 2018.

Pickard S, Law E, Jiang R, Pullenayegum E, et al. United States Valuation of EQ-5D-5L Health States Using an International Protocol. Value in Health, 2019; 22(8):931–941.

SAS Institute Inc. 2008. SAS/STAT 9.2 User's Guide. Cary, NC: SAS Institute Inc.

Sasso M, Beaugrand M, de Ledinghen V, et al. Controlled attenuation parameter (CAP): a novel VCTE(TM) guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: Preliminary study and validation in a cohort of patients with chronic liver disease from various causes. Ultrasound Med Biol. 2010; 36:1825-35.

Simioni P, Tormene D, Tognin G, Gavasso S, Bulato C, Iacobelli NP, Finn JD, Spiezia L, Radu C, Arruda VR. X-Linked Thrombophilia with a Mutant Factor IX (Factor IX Padua). NEJM 2009.

Strassburger K and Bretz F. Compatible simultaneous lower confidence bounds for the holm procedure and other bonferroni based closed tests. Statistics in Medicine, 2008; 27: 4914-4927.

van Reenan M, Janssen B. EuroQol Group EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf (April 2015)

World Federation of Haemophilia. World Federation of Haemophilia report on the annual global survey 2017. dated October 2018. http://www1.wfh.org/publications/files/pdf-1714.pdf date accessed: 23 November 2018

APPENDIX 1: DATA HANDLING RULES

Programming of the tables, listings and figures was performed using SAS Version 9.4 or a more recent version. The following table presents the algorithms used in SAS to calculate the derived variables, including rules for handling other missing data or partial dates, or irregular/unexpected data issues.

Category	Description	Data Handling Rule
1. Age (years)	Age (years)	Age = integer part of ([Screening Visit date – Birth date + 1]/365.25)
2. Medical History	Medical History Begin Date of Condition	Begin date of condition will be imputed for all subjects as the 1st of the month for the purpose of computing the onset day.
3. Surgical History	Surgical History Date of Surgery	Date of surgery will be imputed for all subjects as the 1st of the month for the purpose of computing the onset day.
4. Treatment Date	date/time of first study treatment	The date and time (24 hr. clock) of the dose of IMP study treatment will be taken from the Dosing eCRF. It is not necessary to define a first treatment date for the lead-in period since the lead-in treatment is not qualitatively different from pre-study therapies.
5. Last Visit Date	Date of Last Visit	Date of last visit according to the Visit eCRF.
6. Last Study Participation Date (STDM variable, typically named RFPENDTC)	Last Study Participation Date (STDM variable, RFPENDTC), where SDTM denotes Study Data Tabulation Model	Last study participation date is defined as last known date of contact which would be the later of the following dates: last visit date, date of last contact if lost-to-follow-up, date of telephone follow-up, or death date.
7. Study Day Definitions	**Study Day** for assessment/event which occurs on or after the beginning of the period.	For the post-treatment period, Study Day = Date of assessment/event – date of IMP administration + 1. For the lead-in period (or for overall study day), Study Day is the Date of assessment/event – date of L1 Visit + 1.
	Study Day for assessments/events on days prior to the period	For the post-treatment period, Study Day = Date of assessment/event – date of IMP administration. For the lead-in period, Study

Category	Description	Data Handling Rule
		Day is the Date of assessment/event – date of L1 Visit.
	Dose Day	**Dose Day** in the study is defined as the study day of the trial drug administration (Study Day 1 for the post-treatment period).
	Last Study Day	**For subjects who did not receive the dose of trial drug**, Last Study Day is defined as (the later of the last visit date and the date of last contact for subjects lost-to-follow-up from the Study Completion/Early Discontinuation CRF) – Date of Screening Visit + 1. **For subjects who received the dose of trial drug,** Last Study Day is defined as (the later of the last visit date and the date of last contact for subjects lost-to-follow-up from the Study Completion/Early Discontinuation CRF) – date of IMP administration + 1.
	Days Since IMP drug administration for event (e.g., Death)	**Days Since IMP drug administration** is defined as date of event – date of IMP drug administration.
8. Duration of event	The duration of any event	The duration of any event is defined as (stop date – start date + 1).
9. Distance between Event	Distance between factor IX activity measurement and most recent factor IX replacement therapy administration	Date of factor IX activity measurement – Date Preceding factor IX Replacement Therapy Administration) + 1 The date and time of the factor IX activity measurement in question and the factor IX replacement therapy administrations respectively are used to find the latest factor IX replacement therapy administration preceding the factor IX activity measurement in question. In case the dates of the factor IX activity measurement in question and a factor IX replacement therapy administration are the same and no time is indicated, it is assumed that the factor IX replacement therapy administration precedes the factor IX activity

Category	Description	Data Handling Rule
		measurement in question, and the above defined distance therefore becomes equal to 1.
10. Multiple assessments for the same visit	Vital Sign and Laboratory assessments	● All data will be listed in data listings. ● The last of multiple valid assessments within a post-baseline study time window will be used for summaries. ● If there are multiple laboratory values for the same parameter at post-baseline pre-dose of a visit, the last value will be chosen for analysis.
11. Special Lab Value Handling for Safety Lab values	Lab values with a prefix such as: '>', '<', '+' and 'Less than' etc.…	● '>': use the available original value +0.001 in the analyses. ● '<': use the available original value –0.001 in the analyses. ● '+': use the available original value without the prefix in the analyses. ● '>=': use the available original value in the analyses. ● '<=': use the available original value in the analyses.
12. Prior and concomitant medication	Prior, and lead-in concomitant, and post-treatment concomitant medication	**1.** Prior medication/treatment: is any medication/therapy (including herbal treatments, vitamins, non-pharmacological treatment such as psychotherapy as appropriate) received will be considered prior if the start date of the medication/therapy is missing or the medication/therapy start date is before Visit L1 for the lead-in period. **2.** A medication/therapy will be identified as a "post-treatment concomitant" medication/therapy if it is being continued by the subject at the date of AMT-061 dosing or is any new medication/therapy received during the post-treatment period. A medication with end date that is the same as the AMT-061 dosing date will not be considered to be "post-treatment concomitant". A medication/therapy will be identified as a "lead-in" concomitant

Category	Description	Data Handling Rule
		medication/therapy if it is being continued by the subject at the date of the L1 Visit or is any new medication/therapy received during the lead-in period prior to the date of AMT-061 dosing. The distinction will be made between lead-in concomitant medications and post-treatment concomitant medications. **3.** Any medication/therapy which cannot be identified as Prior, Lead-In Concomitant, or Post-Treatment Concomitant will be considered as being in each of the possible categories depending on available information. The designation of concomitant medication will be done in a manner that is specific to either the lead-in period or the post-treatment period. Given that the study treatment is permanent, there cannot be a medication category subsequent to "concomitant" with respect to the post-treatment period.
13. Adverse event	Missing severity	For the AE summary by severity, an AE with missing severity will be deemed as Severe.
	Missing relationship to study drug	For AE summary by relationship, an AE with a missing relationship to study drug will be deemed as related.
	Treatment-emergent adverse event	An adverse event is considered treatment-emergent for the post-treatment period if an event occurs (or if there was a worsening [intensity and/or severity changed to worsened grades]) on or after the date of dosing with AMT-061. A treatment-emergent adverse can be described as having incidence during the post-treatment period. An adverse event is considered to have had incidence during the lead-in period if an event occurs (or if there was a worsening [intensity and/or severity changed to worsened grades]) on or after the Visit L1 date and before the date of dosing with AMT-061.

Category	Description	Data Handling Rule
		Prior to the Visit L1 date, adverse events are considered to be a part of the medical history. A death is considered to be treatment-emergent for the post-treatment period if any of the adverse events that led to the death occurred on or after the date of administration of the IMP. A death is considered to have had incidence during the lead-in period if any of the adverse events that led to the death occurred on or after the Visit L1 date and before the date of administration of IMP.
		If the AE start date is partial/missing, then • If AE start date is completely missing, then the AE is considered as both treatment-emergent during the post-treatment period and to have had incidence during the lead-in period. • If both AE start month and day are missing and AE start year is the same or after the IMP dosing year, then the AE is considered as treatment-emergent for the post-treatment period. If both AE start month and day are missing and AE start year is the same or after the L1 Visit year and on or before the IMP dosing year, then the AE is considered as having had incidence during the lead-in period. • If AE start day is missing and AE start year and month are the same or after the IMP dosing year and month, then the AE is considered as treatment-emergent for the post-treatment period. If AE start day is missing and AE start "year and month" are the same or after the L1 Visit "year and month" and on or before the IMP dosing "year and month", then the AE is

Category	Description	Data Handling Rule
		considered as having had incidence during the lead-in period. Missing/incomplete (partial) AE start and end dates will not be imputed for data listings.
14. Hard coding	Hard coding for data analysis	Hard Coding is not allowed during data analysis unless agreed to in writing by uniQure.
15. CCI data	Bleed event	Assessments within two weeks of a bleed event will not be included in any analysis.
16. Listing outputs	Data excluded	All data not used for efficacy analysis will be flagged in listings.
17. Contamination due to exogenous factor IX (infusion) use	Contamination of factor IX activity or protein assessment (or in some cases bleeding assessment) due to exogenous factor IX (infusion) use	The date/time (where available) – rather than just date – for the time of the exogenous factor IX infusion and the time of the blood draw for factor IX activity (or protein) assessment will be used for the determination of contamination. The use of date/time (instead of just date) should be applied for the 5-half-life contamination rule and for the 10-day sensitivity-analysis contamination rule. If only the date – but not the time – of the exogenous factor IX infusion is known, then the contamination period will (conservatively) be the time period beginning on midnight at the beginning of that day and ending at the time which is 24 hours plus five half-lives later (but would be ending at 24 hours plus 240 hours for the alternative (sensitivity) 10-day contamination rule). If only the date but not the time of the factor IX activity assessment is known and if any point in time on that date overlaps with the contamination period, then the activity assessment will be deemed contaminated. As alluded to in the SAP text section about secondary efficacy for factor IX activity (Section 7.6.2), the 6-month-data-cut analysis actually used a less refined contamination rule – whereby the date of exogenous factor IX infusion and the subsequent 9 days (10 discrete calendar days in total) were considered to be

Category	Description	Data Handling Rule
		days of contamination with factor IX. As also alluded to in that section, the more refined "5 half-life" contamination period is actually being applied to the 12-month-data-cut analysis and the 18-month-data-cut analysis.

APPENDIX 2: ANALYSIS DATASET SPECIFICATIONS

Analysis datasets were built to gain efficiency and ensure consistency in data analyses and presentation for this trial. The specifications for each analysis data set were prepared separately and are not part of this document.

APPENDIX 3: CENTRAL LABORATORY REFERENCE RANGES FOR USE IN FLAGGING ABNORMAL VALUES

This appendix is provided as an attachment to the SAP.

CCI [redacted]

This appendix is provided as an attachment to the SAP.

APPENDIX 5: SAS CODE FOR STATISTICAL ANALYSES

This appendix is provided as an attachment to the SAP.

APPENDIX 6: STATISTICAL DETAILS

This appendix is provided as an attachment to the SAP.

16.1.9.3 Data Monitoring Committee (DMC)

Data Monitoring Committee Charter

Trial IDs:	CT-AMT-061-01 and CT-AMT-061-02
Trial Titles:	CT-AMT-061-01: Phase IIb, open-label, single-dose, single-arm, multi-center trial to confirm the Factor IX activity level of the serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B CT-AMT-061-02: Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B
Trial Indication:	Haemophilia B
Investigational Medicinal Product (IMP):	AAV5-hFIXco-Padua (adeno-associated viral vector containing the naturally occurring Padua variant of human factor IX gene)
Date:	21 January 2021
Version:	3.0
Sponsor:	uniQure biopharma B.V. Paasheuvelweg 25a 1105 BP Amsterdam The Netherlands Phone: PPD

CT-AMT-061-01 AND CT-AMT-061-02 DATA MONITORING COMMITTEE CHARTER REVISION LOG

Version #	Revision Date	Section(s) Modified	Brief Description of Revision(s) or Reason(s) for Revision
Version 1.0	06 April 2018	N/A	Original issue.
Version 2.0	27 April 2020	Section 1 and 2 Section 5, 10, and 11 Section 5.1 Section 8.3 and 8.4 Section 8.7	Changes to required document signatory. Clarification regarding NAB review timeline requirements. Update to DMC Membership requirements from 5 to 4 members. Update to DMC member meeting quorum number and voting number requirements from 4 to 3 members. Clarification regarding timeline for DMC member meeting minutes comments.
Version 3.0	21 January 2021	Section 2 and Appendix 2	Sponsor contact updated.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND TERMS

Abbreviation	Definition
AAV5	Adeno-associated viral vector serotype 5
AAV5-hFIXco-Padua	AMT-061; Adeno-associated viral vector containing the naturally occurring Padua variant of human coagulation Factor IX gene
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
DMC	Data Monitoring Committee
FIX	Factor IX
gc	Gene copy
IMP	Investigational Medical Product
IRB	Institutional Review Board
NAB	Neutralizing antibody
PCS	Potentially clinically significant
rAAV5	Recombinant adeno-associated viral vector serotype 5
SAE	Serious adverse event
SC	Steering Committee
TFLs	Tables, figures, and listings

1. DMC MEMBERS' APPROVAL OF DATA MONITORING COMMITTEE CHARTER

DMC Chair PPD PPD

Signature Date

DMC Member PPD PPD

Signature Date

DMC Member PPD PPD

Signature Date

DMC Member PPD

Signature Date

DocUUID : 03bf62bc-55ae-44a4-b7cf-4169702b5682

1. DMC MEMBERS' APPROVAL OF DATA MONITORING COMMITTEE CHARTER

DMC Chair PPD

Signature _____ Date _____

DMC Member PPD

Signature _____ Date _____

DMC Member PPD

Signature _____ Date _____

DMC Member PPD

Signature PPD Date PPD

2. SPONSOR'S APPROVAL OF DATA MONITORING COMMITTEE CHARTER

PPD

PPD

PPD

Signature

Date

3. INTRODUCTION

The Data Monitoring Committee (DMC) is a committee established by uniQure, composed of individuals with appropriate and relevant experience, to review data from the Phase IIb study CT-AMT-061-01 and the pivotal Phase III study CT-AMT-061-02 in their AMT-061 clinical development program.

The DMC will monitor the safety of the subjects during both trials. The DMC will evaluate 6 week interim efficacy data from CT-AMT-061-01 and confirm if the dose of 2 x 10^{13} gene copies (gc)/kg AMT-061 should be used in CT-AMT-061-02. In addition, the DMC will assess whether there is an impact of pre-existing neutralizing antibodies (NABs) titers on clinical outcomes following single treatment with AMT-061.

This Charter defines the roles and responsibilities of the DMC members, Sponsor, and other DMC administrative personnel, and will detail the frequency and timing of meetings, communication methods between the DMC and Sponsor, Principal Investigator, and study representatives, the maintenance of confidentiality, and details on data to be provided to the DMC and statistical considerations. This Charter will serve as the standard operating procedure for the DMC. The operating procedures of the DMC are based on and are in compliance with the Food and Drug Administration's "Guidance for Clinical Trial Sponsors [on the] Establishment and Operation of Clinical Trial Data Monitoring Committees" (March 2006),[1] World Health Organization's "Operational Guidelines for the Establishment and Functioning of Data and Safety Monitoring Boards" (2005),[2] and the European Medicines Agency's "Guideline on Data Monitoring Committees" (2006).[3]

4. TRIAL DESIGNS

4.1 Study Treatment

uniQure is developing AMT-061 as a somatic gene therapy for hemophilia B.

AMT-061 (AAV5-hFIXco-Padua) is a recombinant adeno-associated viral vector serotype 5 (rAAV5) encoding the naturally occurring Padua variant of human coagulation Factor IX (FIX), under control of a liver-specific promoter. The FIX-Padua protein differs from the 'wild type' human FIX protein by a single amino acid and possesses an activity approximately 6-fold higher than wildtype FIX.

4.2 CT-AMT-061-01

CT-AMT-061-01 is a Phase IIb, open-label, single-dose, single-arm, multi-center trial to confirm the FIX activity level of the AAV5 vector containing the Padua variant of a codon-optimized human FIX gene (AAV5 hFIXco-Padua, AMT-061) administered to adult subjects with severe (FIX <1%) or moderately severe (FIX ≤2%) hemophilia B. In addition, the safety profile of AMT-061 will be assessed.

The trial will be conducted at multiple centers in the United States in 3 male (possibility of 6) subjects aged ≥18 years with congenital hemophilia B with either a known severe FIX deficiency, for which the subject is on continuous routine FIX prophylaxis or on-demand FIX

replacement therapy, or a known moderately severe FIX deficiency and a severe bleeding phenotype.

Subjects will receive a single infusion of 2 x 10^{13} gc/kg AMT-061 and will be followed for 1 year in a post-treatment follow-up phase, and then for an additional 4 years during a long-term follow-up phase. During the post-treatment follow-up phase, subjects will return to the clinic for scheduled safety and efficacy assessments. During the long-term follow-up, subjects will return to the clinic every half year for evaluation of efficacy parameters and safety. Occurrence of adverse events (AEs) will be continuously monitored, with at least quarterly contact moments between site staff and subject to discuss.

4.3 CT-AMT-061-02

CT-AMT-061-02 is a Phase III, open-label, single-dose, multi-center multinational trial investigating a serotype 5 adeno-associated viral vector containing the Padua variant of a codon-optimized human FIX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B. The purpose of this Phase III trial is to demonstrate the efficacy of AMT-061 in terms of endogenous FIX activity and annualized bleeding rate, and to further describe its safety profile.

The trial will be conducted at multiple centers in multiple countries in approximately 56 male subjects aged ≥18 years with congenital hemophilia B with a known severe or moderately severe FIX deficiency for which the subject is on continuous routine FIX prophylaxis.

Subjects will participate in a lead-in phase (minimum of 6 months) were they will record their use of prophylactic FIX replacement therapy and bleeding episodes in their e-diary.

Criteria for Investigational Medicinal Product (IMP) dose administration in this trial include the following: availability of the interim results of the CT-AMT-061-01 dose confirmation trial that confirm the dose for AMT-061 for this trial (as confirmed by this DMC; see **Section 9**); the subject has adequate record keeping and adequate compliance with study activities during the lead-in phase as determined by the investigator, and subjects must still meet all inclusion and exclusion criteria.

Subjects will receive a single infusion of AMT-061 and will be followed for 1 year in a post-treatment follow-up phase, and then for an additional 4 years during a long-term follow-up phase. During the post-treatment follow-up phase, subjects will continue to record their use of FIX replacement therapy and bleeding episodes in the e-diary, and will return to the site for scheduled safety and efficacy assessments. During the long-term follow-up, subjects will return to the clinic every half year for evaluation of efficacy parameters and safety. Occurrence of AEs will be continuously monitored, with at least quarterly contacts between site staff and subject.

5. RESPONSIBILITIES AND TASKS OF THE DMC

The DMC is created for these 2 studies to review accumulating safety data, to confirm if the dose in the Phase IIb study should be used in the Phase III study, and to assess the impact of pre-existing NABs. The DMC's activities will protect the rights, safety, and well-being of the

subjects in the studies, will assist with monitoring the overall conduct of the studies, and will protect the validity and credibility of the studies.

The primary responsibilities of the DMC are the following:

1. Complete a confidentiality document and a Disclosures of Conflict of Interest.
2. Review the draft DMC Charter, comment and recommend modifications, and, after it is finalized, follow the directives of the Charter for the duration of the study.
3. Review the protocol and other provided study material as necessary to perform their DMC activities.
4. Review monthly serious adverse event (SAE) reports.
5. Review data provided after every set of approximately 10 subjects enrolled (depending on enrollment) to assess the impact of NABs, as described in **Section 10** and **Section 11**.
6. Review the DMC Data Report provided prior to each DMC safety data review meeting (safety tables and listings as described in **Section 11**).
7. Attend the DMC scheduled meetings and if necessary, any ad hoc DMC meetings.
8. Following review of the DMC Data Report and discussion of data during closed sessions of each DMC safety data review meeting, prepare recommendations to the Sponsor as described in **Section 8.5**, indicating if the studies should continue, be temporarily stopped to investigate a safety issue, be modified, or stopped due to an urgent safety issue.
9. Following review of 6 week interim efficacy data from CT-AMT-061-01 at the first DMC safety data review meeting, prepare a recommendation to the Sponsor confirming if the dose of 2 x 10^{13} gc/kg AMT-061 used in the Phase IIb study CT-AMT-061-01 is suitable or not for administration in the Phase III study CT-AMT-061-02, as described in **Section 9**. The DMC may, in discussion with the Sponsor, Principal Investigator, or the Steering Committee (SC), decide if 1, 2, or 3 additional patients are required for further confirmation of efficacy and safety with the dose under investigation, or they may recommend that a second dose be evaluated.
10. Review and approve the DMC meeting minutes as prepared by the DMC Secretary, following each meeting (see **Section 8.7**).
11. The DMC Chair will be responsible for the following additional functions:
    - Act as the primary contact for the DMC.
    - Lead the data review discussions during the closed sessions of the DMC safety data review meetings (see **Section 8.2.2**).
    - Review, approve, and sign the DMC recommendation letter drafted by the DMC Secretary, who will then provide the final letter to the Sponsor Contact, Principal Investigator, and DMC Administrator (see **Section 8.6**).

5.1 DMC Membership

The DMC will consist of 4 non-sponsor, independent voting members who have been approved by the Sponsor and the SC. They will be physicians and clinicians with expertise in the area of hemophilia. Each DMC member will be considered and recognized as an expert in his/her fields of practice and will be experienced and knowledgeable with the conduct of clinical trials. DMC members may not participate in the study as Investigators and are not

allowed to knowingly administer medical or any other type of care to any subject in the study during the study conduct.

One member (1) will be designated as the Chair of the DMC. The Chair will be the primary contact for the DMC with the Sponsor, SC, and Principal Investigator, communicating through the DMC Administrator, and will also lead the discussions during the closed sessions of the DMC safety data review meetings (see **Section 8.2.2**).

It is understood that the DMC members will be available to fulfill the DMC obligations for the entire planned 5 year period of the two studies specified in this Charter, until the studies are completed or terminated, and the final required DMC meeting has occurred and safety data have been reviewed by the DMC. In the event that a DMC member has to stop his/her participation in the DMC, it will be determined by the Sponsor in consultation with the DMC Chair if a replacement will be required. If a replacement is deemed required, they will be selected by the Sponsor and the SC. No replacement can be made without the approval of the Sponsor.

The members of the DMC are identified in **Appendix 1**.

5.2 Conflict of Interest

Members of the DMC should remain independent of the study. Prior to their participation on the DMC, the members will be asked to sign a form disclosing any conflicts of interest, whether scientific, financial, or other, and will be required to notify the Sponsor or designee of any changes that occur during their participation on the DMC.

Members of the DMC, and their immediate families, will not buy, sell, or hold stock, options, derivatives or any other financial instrument related to the Sponsor's company for the following periods: from the date applicable to the member's first DMC meeting until the first to occur of the following: (a) the last meeting of the DMC; (b) the study results are made public; or (c) until a year after the member's active personal involvement in the DMC ends.

Certain other activities are not viewed as constituting conflicts of interest but must be reported annually to the Sponsor: the participation of a member in educational activities supported by the Sponsor and occasional scientific consulting to the Sponsor on issues not related to the product in the study.

5.3 Confidentiality

All members of the DMC will treat all reports, meeting discussions, minutes, and recommendations of the DMC as confidential.

5.4 Reimbursement

DMC members will receive an honorarium for their time and effort on the committee. DMC members will be reimbursed for reasonable expenses related to attending DMC meetings, such as transportation, lodging, and meals. No other payment or future consideration will be provided.

5.5 Indemnity

DMC members will be indemnified by the Sponsor from third party claims relating to harm arising to study participants or other parties as a result of such DMC members' actions and decisions taken in good faith in accordance with the terms of this Charter. The foregoing shall not apply in the case of harm caused by the negligence or wrongful conduct of a DMC member.

6. RESPONSIBILITIES OF THE SPONSOR

The Sponsor is responsible for supplying safety and efficacy information to the DMC, Regulatory Authorities, and the Investigators. The Sponsor's responsibilities include the following:

1. Assure the availability of resources to the DMC so that the DMC may fulfill its responsibilities as described in this Charter.
2. Review the DMC recommendations. The Sponsor has the ultimate responsibility for all final decisions concerning whether or not the DMC recommendations will be implemented.
3. The Sponsor will be the responsible party for the design and conduct of the study, and regulatory reporting of SAEs in accordance with the applicable regulations.

The Sponsor Contact, identified in **Appendix 2**, or their designee, will be responsible for the following:

1. Review DMC recommendations and arrange for Sponsor review and decision.
2. Serve as a resource to the DMC for requests of additional data or other information.
3. Be available to address safety issues that occurred in the study.
4. Attend the open sessions of all DMC meetings.
5. On a monthly basis, ensure that the DMC Secretary is provided with a monthly cumulative SAE listing with event narratives from the study for distribution to the DMC.
6. Ensure that the DMC Secretary is provided with a cumulative SAE listing with event narratives for inclusion with the DMC Data Report prior to any DMC safety data review meeting.

7. RESPONSIBILITES OF THE DMC ADMINISTRATOR AND DMC SECRETARY

The Sponsor will engage a DMC Administrator and DMC Secretary to aid in planning, organizing, and coordinating DMC activities, disseminating meeting reports, and communicating between the DMC and the Sponsor. The DMC Administrator and DMC Secretary will sign a confidentiality agreement for their participation in the DMC activities. The DMC Administrator and DMC Secretary for the study listed in **Appendix 2**.

The DMC Administrator is responsible for the following:

1. Develop and facilitate the review and signing of the DMC Charter.
2. Plan, organize, and coordinate DMC activities in accordance with the established DMC Charter, including timely delivery of study data (e.g., DMC Data Reports).

3. Provide administrative and logistic services to the DMC, serving as the primary contact point for day-to-day operations and between the DMC and the Sponsor.
4. Plan, organize, and host each of the DMC meeting open sessions.
5. Provide periodic progress status reports of the DMC activities to the Sponsor as needed.

The DMC Secretary will assist the DMC Administrator and is responsible for the following:

1. Forward the DMC Data Reports, monthly cumulative SAE listings and narratives, and any safety-related information to the DMC members. Materials will be forwarded at least 5 business days before any scheduled DMC meeting or teleconference, and as soon as possible before any ad hoc teleconference, when applicable.
2. Attend both the open and closed sessions of the DMC meetings. Produce, distribute for review, and finalize DMC meeting minutes for each meeting, separately for the open and closed sessions (see **Section 8.7**).
3. Facilitate drafting of the Recommendations Letter based on the meeting minutes (see **Section 8.6**).
4. Maintain a secure central file of all data reports, all minutes, and all DMC Recommendation Letters submitted to the Sponsor.
5. Provide an archived copy of all maintained files to the Sponsor after the study database is locked and the DMC activities have concluded.

8. DMC MEETINGS

The DMC meetings are planned to either be face-to-face or via teleconference (see **Appendix 3** for schedule). The DMC will participate in an initial (kick-off) meeting (**Section 8.1**) and at least 9 safety data review meetings (**Section 8.2**). At the first data review meeting, the DMC will review subject safety and will also review efficacy data from study CT-AMT-061-01 and confirm the dose of AMT-061 to be used in study CT-AMT-061-02 (**Section 9**). Subsequent data review meetings will focus on subject safety and assessment of the impact of NABs and will occur approximately every 6 months during subject enrollment and then every 12 months until both studies have completed or have been terminated. Meetings focused on the impact of NABs will be organized as needed depending on the rate of subject enrollment.

Ad hoc meetings may take place if requested by the DMC or Sponsor (**Section 8.8**).

8.1 Kick-off Meeting

The initial (kick-off) meeting of the DMC will be held prior to enrolment of the first subject in the trials and will be an open meeting to provide the DMC with an understanding of the study expectations and to establish the DMC procedures. At the meeting, the DMC members will meet with Sponsor representatives to discuss both of the studies, review and finalize the DMC Charter, and discuss the proposed content of the DMC Data Report (see **Section 11**). Additional topics which may be discussed include: the frequency of meetings, the format of future meetings, specific analyses approaches, key safety variables to be reviewed for decision making, and the potential need for additional meetings. The DMC Charter will be finalized and signed off shortly after the meeting.

The DMC Chair approves the DMC Charter on behalf of the DMC. The DMC Charter must be approved prior to start of screening of subjects. Once the approved DMC Charter is accepted by all DMC members by signature, the DMC will be considered activated.

8.2 DMC Safety Data Review Meetings

The DMC will participate in at least 9 safety data review meetings. DMC data review meetings will consist of two open and one closed session. The DMC Secretary will record both open and closed sessions, and prepare minutes of the meeting (see **Section 8.7**).

8.2.1 Open Session I

The first open session will be hosted by the DMC Administrator and will include the DMC members, DMC Secretary, Sponsor representatives (may include the Principal Investigator), and the Study Biostatistician. At the open session, the Sponsor's representative(s) and study personnel involved in trial management will present an overview of the study's current progress status to the DMC. The open session should focus on the conduct and progress of the trial with special attention to safety and efficacy data. Data presented in the open session may include data on enrolment data, individual adverse events, baseline characteristics, and subjects lost to follow-up.

8.2.2 Closed Session

The closed session will be led by the DMC Chair and will only include the members of the DMC and the DMC Secretary.

The Sponsor's representative(s) as well as study personnel involved in trial management will leave the meeting/teleconference prior to the start of the closed session. The DMC closed session will begin once the DMC Secretary confirms only the attendees for the closed session are on the teleconference web-link and telephone call-in line; connections to the session will be monitored during the session to ensure the closed nature of the discussion.

Prior to the meeting, the DMC Data Report and other safety information will have been provided to the DMC via the DMC Secretary for their review.

During the closed session, the DMC members will discuss the provided data summaries and formulate their recommendations to the Sponsor.

DMC members have the right to request detailed information regarding any of the reported individual clinical endpoint events if deemed necessary. The Sponsor Contact and Study Biostatistician will be "on call" during the DMC closed session, in case the DMC requests additional safety information and/or needs to discuss statistical considerations of the provided data; the Study Statistician will maintain confidentiality concerning these discussions.

8.2.3 Open Session II

At the conclusion of the closed session, the open session attendees will be invited to rejoin the meeting/teleconference for a second open session. This second open session will be hosted by the DMC Administrator and will include the DMC members, DMC Secretary, and the Sponsor Contact at a minimum. At this session, the DMC will share their

recommendations with the Sponsor Contact and other Sponsor representatives who are present. Concerns or questions from the DMC may be addressed at this time as well.

8.3 Quorum

For a meeting to proceed, a quorum of three (3) DMC members is required at each safety data review meeting.

At the beginning of the open sessions, the DMC Administrator will confirm the attendance of the quorum. If the quorum is not met, the Chair will confirm that the DMC meeting should not proceed and the DMC Administrator will reschedule the meeting.

The quorum at an ad hoc data review meeting will be three (3) DMC members.

8.4 Voting

All 4 DMC members have voting privileges. To vote, a DMC member must be a participant in the DMC meeting. Voting is by voice and will be recorded by the DMC Chair and in the closed portion of the meeting minutes.

DMC members are responsible for voting on the following:

- All recommendations that will be submitted to the Sponsor (see **Section 8.5**).

A unanimous vote by all members is required for all recommendations to implement a modification to the conduct of the study(ies), to temporarily halt enrollment, or to stop the study(ies).

Any recommendations to modify the study's conduct (e.g., early termination of the study due to alarming safety concerns) must be agreed upon by all members of the DMC and will be made on the basis of the provided study safety data.

A simple voting majority at a meeting will be required for a routine proposal, motion, or recommendation to be made to the Sponsor that does not involve modification to the conduct of the study or discontinuation of the study.

8.5 DMC Recommendations

Based upon their review and discussion of the safety data at each meeting, the DMC has the responsibility to recommend to the Sponsor whether the study should proceed as planned, be modified, or be stopped.

Anticipated recommendations are as follows:

- Continue the studies according to the protocols and current amendments.
- Temporarily stop enrollment in the CT-AMT-061-02 study to investigate a safety issue.
- Modify the study protocol(s) and/or the informed consent.
- Stop the studies due to an urgent situation (serious safety issue).

If the DMC detects a safety signal or has concerns about the provided data in the DMC Data Report or other safety information, they may request additional safety-related information from the Sponsor Contact or designee.

Modifications recommended by the DMC may include, but are not limited to, changes in the exclusion criteria due to NABs (see **Section 10**), other changes in the inclusion/exclusion criteria, frequency of safety monitoring, and instituting changes in the study procedures. While modifications are put into place, enrollment would continue in the CT-AMT-061-02 study unless the DMC also recommends for a temporary stop.

In the case that the DMC recommends to stop the studies or to temporarily halt enrollment in CT-AMT-061-02, the DMC Chair will notify the Sponsor Contact immediately. If the Sponsor needs additional clarification, the Sponsor Contact, or designee, and the DMC members will meet as quickly as possible to discuss the recommendations.

If the DMC recommends to stop the studies, the Sponsor's response would be to place study recruitment on hold pending final decision by the Sponsor. Investigators, Institutional Review Boards (IRBs), and other Regulatory Authorities will be notified by the Sponsor, as appropriate.

8.6 Recommendation Letter

Following each safety data review meeting, the DMC will submit a letter to the Sponsor outlining their recommendations for the study. The DMC Secretary will facilitate the drafting of the letter, based on the DMC recommendations discussed and agreed upon during the closed session. If there is a recommendation for action, the brief rationale for such recommendation will be included in the letter. The DMC Recommendation Letter will be sent to the DMC Chair for review, approval, and signature following the meeting (no later than 1 working day after the DMC meeting). The DMC Secretary will then provide a copy of the final signed Recommendation Letter to the Sponsor Contact, Principal Investigator, and the DMC Administrator on the DMC Chair's behalf via email. If the recommendation letter is not provided to the Sponsor on the same day as the meeting, the DMC Secretary will provide the Sponsor with the DMC members' recommendation decision via email.

The Sponsor will be responsible for communicating to the Investigators, IRBs, and/or other regulatory bodies, as appropriate, a summary of the DMC recommendation(s) and the Sponsor's proposed action.

8.7 DMC Meeting Minutes

Minutes of all meetings will be prepared by the DMC Secretary.

The DMC minutes for open and closed meeting sessions will be circulated to the DMC members for comments within 5 business days after a meeting. If comments/confirmation are not received within 5 business days, except for a pre-specified reason from a DMC member, it will be assumed there are no comments. At the same time, the minutes for the open sessions will be circulated to the Sponsor for comments. If the DMC has a recommendation for action, this will be outlined in the open session minutes with any rationale. The DMC Secretary will incorporate the comments and then route the final minutes to the DMC members for approval signatures.

Each DMC member will be provided with final copies of the complete summary meeting minutes by the DMC Secretary and the Sponsor will be provided with a final copy of the open session meeting minutes with DMC recommendations.

At the end of the study, the DMC Secretary will forward to the Sponsor Contact a complete set of summary meeting minutes (including both open and closed sessions).

8.8 Ad hoc Meetings

An emergency meeting of the DMC may be called at any time by the Chair should questions of subject safety arise. The request will be sent to the DMC Administrator, who will then schedule and organize the meeting and meeting materials. Further data review meetings in addition to those already planned may be required based on study enrollment extending beyond the planned duration of the 2 trials.

Ad hoc specialists can be invited to DMC meetings as recommended by the DMC or uniQure.

9. AMT-061 DOSE CONFIRMATION FOR PHASE III TRIAL

At their first safety data review meeting, the DMC will confirm if the dose of 2 x 10^{13} gc/kg AMT-061 used in the Phase IIb study CT-AMT-061-01 is suitable for administration in the Phase III study CT-AMT-061-02, in addition to reviewing the study safety.

In addition to safety data, the DMC will be provided with the 6 week interim efficacy data from CT-AMT-061-01 to review prior to the meeting. The DMC will evaluate response to treatment with a single dose of 2 x 10^{13} gc/kg of AMT-061 in terms of FIX activity levels and assess whether observed FIX activity levels are ≥5%.

Following their review, the DMC has the responsibility to provide a recommendation to Sponsor concerning the dose of 2 x 10^{13} gc/kg AMT-061, either confirming this dose is suitable for use in the Phase III study or to suggest further evaluation. If the DMC determines that the observed response is not within the expected range, or they do not observe enough consistency of effect to proceed to Phase 3 dosing, they can elect to recommend up to three more subjects be treated at the same dose or recommend a second dose be studied.

Anticipated recommendations concerning the dose of AMT-061 are as follows:

- Proceed with dosing in CT-AMT-061-02 with a single treatment of 2 x 10^{13} gc/kg AMT-061
- Treat 1, 2, or 3 additional subjects at the dose of 2 x 10^{13} gc/kg AMT-061 and reevaluate the data
- Evaluate a second dose of AMT-061

This recommendation will be in addition to a safety recommendation as per **Section 8.5**.

10. NEUTRALIZING ANTIBODIES

The DMC will assess whether there is an impact of pre-existing NAB titers on clinical outcome following treatment with AMT-061. Should the DMC determine that there is a recognizable impact of a certain titer and above, they can recommend institution of an exclusion criterion based on these titers for further enrollment in the study.

The DMC will review data after every set of approximately 10 subjects enrolled depending on enrollment rates (enrollment will not stop during this review). Review of this data will be

included in the safety data review meetings or in NAB-specific data review meetings, depending on the timing of the review due to the rate of enrollment. NAB review meetings may be combined and include multiple set of 10 subjects due to a high rate of enrollment. Ad hoc meetings may be called if there is a safety concern, impacted outcomes, or need for a protocol amendment. During their review, the DMC will provide one of the following assessments:

- No clear impact, continue with prospective analysis
- There are impacted outcomes (safety and/or efficacy), institution of a data-driven cut-point
- Protocol amendment needed

An enactment of a NAB titer cut-off mid-trial is recommended if all of the following criteria are met:

- Two patients with NAB titers and FIX <5%
- No patients with NAB titers equal or greater to NAB titers in patients with FIX ≥5%
- No patients without NAB titer and FIX <5%

11. DATA REPORTS FOR REVIEW BY THE DMC

For each data review meeting, the DMC Data Reports will be provided to each DMC member by the DMC Secretary via a suitable secure method for their prior review. The report will be delivered approximately 10 business days before each DMC meeting.

The DMC Data Report for the DMC includes the tables, figures, and listings (TFLs) to be reviewed by the DMC members at the scheduled and/or ad hoc meetings. Mock-ups of the report format and content (i.e., TFL shells) are provided for CT-AMT-061-01 in **Appendix 4** and for CT-AMT-061-02 in **Appendix 5**. The Sponsor and DMC members will review and approve the format and content of the mock-up DMC Data Report prior to the first DMC data review meeting.

The following descriptive summaries and listings will be included in the DMC Data Report based on the cumulative data in the study.

- Baseline demographic tables and results of the number of subjects screened, randomized or discontinued (with reasons for discontinuation) for the study
- List of frequent AEs (>5%) listed in order of frequency, from high to low and summarized of all SAEs and AEs leading to study discontinuation
- Summaries of all potentially clinically significant (PCS) laboratory results and vital signs
- Summaries of increases in alanine aminotransferase (ALT) and aspartate aminotransferase (AST)
- NAB titer
- Capsid-specific T-cells
- FIX activity levels

A subset of these outputs will be provided for DMC review after every set of approximately 10 subjects enrolled, for assessment of the impact of NABs. This subset of outputs will be identified in the DMC Data Report shells.

12. DMC CLOSURE

The responsibilities of the DMC will end when both clinical trials have closed. The DMC may be closed during a formal DMC meeting with all DMC members.

The DMC Secretary will provide an archived copy of all maintained files (recommendation letters, DMC data reports, open and closed meeting minutes) to the Sponsor after the study database is locked and the DMC activities have concluded.

13. REFERENCES

1. Food and Drug Administration Final Guidance: Guidance for clinical study sponsors on the establishment and operation of clinical study data monitoring committees. March, 2006.
2. World Health Organization. Operational Guidelines for the Establishment and Functioning of Data and Safety Monitoring Boards. 2005.
3. European Medicines Agency. Guideline on Data Monitoring Committees. January, 2006.

APPENDIX 1 DMC MEMBERS

Name, Degrees:	PPD **(DMC CHAIR)**
Role/Title:	PPD, University of Rochester School of Medicine & Dentistry PPD, Mary M. Gooley Hemophilia Treatment Center
Address:	Rochester, New York, USA
E-mail:	PPD Tel: PPD
Name, Degrees:	PPD
Role/Title:	PPD, National Hemophilia Center and the Institute of Thrombosis & Hemostasis PPD, Amalia Biron Research Institute of Thrombosis & Hemostasis, Sheba Medical Center
Address:	Tel Hashomer, Israel
E-mail:	PPD Tel: PPD
Name, Degrees:	PPD
Role/Title:	PPD , Emory University
Address:	Atlanta, Georgia, USA
E-mail:	PPD Tel: PPD
Name, Degrees:	PPD
Role/Title:	PPD, Nationwide Children's Hospital
Address:	Columbus, Ohio, USA
E-mail:	PPD Tel: PPD

APPENDIX 2 DMC SUPPORT PERSONNEL

Sponsor Contact[1]	Name, Degree(s):	PPD		
	Title:	PPD, Clinical Development		
	Organization:	uniQure		
	E-mail:	PPD	Tel:	PPD PPD (cell)
DMC Administrator	Name, Degree(s):	PPD		
	Title:	PPD, Medical Writing and Clinical Safety Monitoring, Clinical Safety Associate		
	Organization:	Everest Clinical Research Corporation		
	E-mail:	PPD	Tel:	PPD
DMC Secretary	Name, Degree(s):	PPD		
	Title:	PPD, Clinical Safety Monitoring		
	Organization:	Everest Clinical Research Corporation		
	E-mail:	PPD	Tel:	PPD

1 Primary Sponsor Contact for the DMC through the DMC Administrator or DMC Secretary (e.g., receiving the DMC Recommendations Letter from the DMC Chair).

APPENDIX 3 DMC MEETING PLAN

Meeting Chair: PPD		Minutes By: PPD	
Meeting		**Date/Time**	**Meeting Purposes and Objectives**
1	Kick-off Meeting	PPD	Initial meeting: • Review of the studies. • Review and approve DMC Charter.
2	Data Review Meeting: Dose Confirmation and Safety Review	Tentatively: PPD (once interim analysis data is available from CT-AMT-061-01)	Open Session I: • Overview of study progress. Closed Session: • Review of the DMC Data Report • Evaluation of the safety and tolerability of study treatment • Evaluation of the FIX activity data to confirm the AMT-061 dose for CT-AMT-061-02 • Determination of DMC Recommendations to the Sponsor Open Session II: • DMC shares recommendations
3-9	Data Review Meetings: Safety Review	During subject enrollment: approximately every 6 months. Once all subjects enrolled, every 12 months until study completion	Open Session I: • Overview of study progress. Closed Session: • Review of the DMC Data Report • Evaluation of the safety and tolerability of study treatment • Determination of DMC Recommendations to the Sponsor Open Session II: • DMC shares recommendations
	Data Review Meetings: NAB Review	During subject enrollment: approximately every 10 subjects dosed. May be combined with Safety Data Review Meeting	Open Session I: • Overview of study progress. Closed Session: • Review of the NAB DMC Data Report • Determination of DMC Recommendations to the Sponsor Open Session II: DMC shares recommendations
	Ad hoc Meetings (if required)	TBD	DMC may meet if required to discuss and communicate any safety concerns noted between scheduled safety data review meetings.

APPENDIX 4 DMC DATA REPORT SHELLS FOR CT-AMT-061-01

The DMC Data Report TFL shells for CT-AMT-061-01 will be provided in a separate document.

APPENDIX 5 DMC DATA REPORT SHELLS FOR CT-AMT-061-02

The DMC Data Report TFL shells for CT-AMT-061-02 will be provided in a separate document.

DMC Recommendations to Trial Sponsor

Trial Sponsor:	uniQure biopharma B.V.	**Protocol Number:**	CT-AMT-061-01
Protocol Title:	Phase IIb, open-label, single-dose, single-arm, multicenter trial to confirm the Factor IX activity level of the serotype 5 adeno-associated viral vector containing the Padua variant of a codonoptimized human factor IX gene (AAV5-hFIXco-Padua, AMT-061) administered to adult subjects with severe or moderately severe hemophilia B		
DMC Meeting:	Data Review Meeting #1	**DMC Meeting Date:**	PPD
DMC Members:	PPD (Chair) PPD PPD PPD PPD (absent)	**DMC Meeting Time:**	8:00 am PT

From: DMC Chair – PPD
To: PPD (Sponsor Contact)
Date: PPD

☑	Proceed with dosing in CT-AMT-061-02 with a single treatment of 2 x 10^{13} gc/kg AMT-061
☐	Treat 1, 2, or 3 additional subjects at the dose of 2 x 10^{13} gc/kg AMT-061 and reevaluate the data Additional subjects requested: _____
☐	Evaluate a second dose of AMT-061; dose: _____
√	Other: **Data review after the first 7 patients entered in the Phase III trial as opposed to the first 10 patients as well further review of longer term data of the first 3 subjects.**

Regards,

PPD

PPD

Name: PPD
Title: DMC Chair

Date (dd-Mmm-yyyy)

CONFIDENTIAL **Page 1 of 1**

Signature Page

CT-AMT-061-02 - Statistical Analysis Plan - statistical-methods

Signed By	Date (GMT)
PPD	PPD 17:12:15
Approved-Internal Approval	

--- Do Not Detach This Signature Page ---