Title: A Randomized, Double-Blind, Placebo-Controlled, Multicenter Study to Evaluate the Efficacy and Safety of Vedolizumab in the Prophylaxis of Intestinal Acute Graft-Versus-Host Disease in Subjects Undergoing Allogeneic Hematopoietic Stem Cell Transplantation

NCT Number: NCT03657160

SAP Approve Date: 22 October 2021

Certain information within this Statistical Analysis Plan has been redacted (ie, specific content is masked irreversibly from view with a black/blue bar) to protect either personally identifiable (PPD) information or company confidential information (CCI).

This may include, but is not limited to, redaction of the following:

- Named persons or organizations associated with the study.
- Proprietary information, such as scales or coding systems, which are considered confidential information under prior agreements with license holder.
- Other information as needed to protect confidentiality of Takeda or partners, personal information, or to otherwise protect the integrity of the clinical study.

STATISTICAL ANALYSIS PLAN

STUDY NUMBER: Vedolizumab-3035

Randomized, Double-Blind, Placebo-Controlled, Multicenter Study to Evaluate the Efficacy and Safety of Vedolizumab in the Prophylaxis of Intestinal Acute Graft-Versus-Host Disease in Subjects Undergoing Allogeneic Hematopoietic Stem Cell Transplantation

Vedolizumab in the prophylaxis of intestinal acute graft vs host disease in subjects undergoing allogeneic hematopoietic stem cell transplantation.

PHASE 3

Version**:** 3.0

Date**:** 22 October 2021

Prepared by:

PPD

Based on:

Protocol Version: Amendment 07

Protocol Date: 18 Sep 2019

1.1 Approval Signatures

Electronic signatures can be found on the last page of this document.

2.0 TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

3.0 LIST OF ABBREVIATIONS

Acronym	Definition
AE	adverse event
AESI	adverse event of special interest
aGvHD	acute graft-versus-host disease
allo-HSCT	allogeneic hematopoietic stem cell transplantation
ALT	alanine aminotransferase
APC	antigen-presenting cell
AST	aspartate aminotransferase
ATG	antithymocyte globulin
ATG-F	antithymocyte globulin-Fresenius
CCI	[redacted]
CD	Crohn's disease
CMV	cytomegalovirus
CNI	calcineurin inhibitor
CRO	contract research organization
CYS	cyclosporine
DLI	donor leukocyte infusion
DMC	Data Monitoring Committee
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data capture
EQ-5D	EuroQol-5 Dimension
ET	early termination
FACT-BMT	Functional Assessment of Cancer Therapy-Bone Marrow Transplant Scale
FAS	full analysis set
FDA	Food and Drug Administration
FSH	follicle-stimulating hormone
GALT	gut-associated lymphoid tissue
GCP	Good Clinical Practice
GI	gastrointestinal
GvHD	graft-versus-host disease
HIV	human immunodeficiency virus
HRQOL	health-related quality of life
HSC	hematopoietic stem cell
HSCT	hematopoietic stem cell transplantation
IAC	Independent Adjudication Committee
IB	Investigator's Brochure
IBD	inflammatory bowel disease

Acronym	Definition
IBMTR	International Bone Marrow Transplant Registry Database
ICF	informed consent form
ICH	International Council for Harmonisation
ID	identification
IEC	independent ethics committee
INR	international normalized ratio
IRB	institutional review board
IRR	infusion-related reaction
IV	intravenous
IRT	interactive response technology system
LFT	liver function test
LTFU	long-term follow-up
mAb	monoclonal antibody
MAdCAM-1	mucosal addressin cell adhesion molecule-1
MAGIC	Mount Sinai Acute GVHD International Consortium
MedDRA	Medical Dictionary for Regulatory Activities
MHRA	Medicines and Healthcare products Regulatory Agency
MMF	mycophenolate mofetil
MTX	methotrexate
MUD	matched unrelated donor
NRM	nonrelapse mortality
OS	overall survival
CCI	[redacted]
PMDA	Pharmaceuticals and Medical Devices Agency of Japan
PML	progressive multifocal leukoencephalopathy
PTE	pretreatment event
Q4W	every 4 weeks
Q8W	every 8 weeks
QOL	quality of life
RAMP	Risk Minimization Action Plan for PML
SAE	serious adverse event
SAP	statistical analysis plan
SOE	schedule of events
TAC	tacrolimus
TB	tuberculosis
UC	ulcerative colitis
ULN	upper limit of normal

4.0 OBJECTIVES

4.1 Primary Objectives

The primary objective is:

- To evaluate the efficacy of vedolizumab when added to background aGvHD prophylaxis regimen compared to placebo and background aGvHD prophylaxis regimen on intestinal aGvHD-free survival by Day +180 in subjects who receive allo-HSCT as treatment for a hematologic malignancy or myeloproliferative disorder.

4.2 Safety Objective

The safety objective is to evaluate the safety of vedolizumab when added to background aGvHD prophylaxis regimen compared to placebo and background aGvHD prophylaxis regimen.

4.3 Secondary Objectives

The secondary objectives are:

- To evaluate the effect of vedolizumab compared to placebo on intestinal aGvHD-free, relapse-free (free of underlying malignancy) survival by Day +180.
- To evaluate the effect of vedolizumab compared to placebo on Grade C-D aGvHD-free (any organ involvement) survival by Day +180.
- To evaluate the effect of vedolizumab compared to placebo on nonrelapse mortality (NRM) in subjects by Day +180.
- To evaluate the effect of vedolizumab compared to placebo on OS by Day +180.
- To evaluate the effect of vedolizumab compared to placebo on Grade B-D aGvHD-free (any organ involvement) survival by Day +180.

CCI

[redacted]

CCI

[redacted]

4.5 Study Design

This is a phase 3, randomized, double-blind, placebo-controlled, multicenter study to evaluate the efficacy and safety of vedolizumab when added to a background aGvHD prophylaxis regimen as prophylaxis for intestinal aGvHD in subjects undergoing allo-HSCT. The subject population will consist of subjects with hematologic malignancies or myeloproliferative disorders for whom allo-HSCT from an unrelated donor is planned using either peripheral blood or bone marrow as the stem cell source.

Eligibility will be determined during the screening period, which may last for up to 30 days before Day -1 (designation of the day of the first IV infusion of study drug). Subjects who meet all eligibility criteria and provide written informed consent will be randomized into this study within 2 days of the first dose of study drug on Day -1. Approximately 558 subjects will be randomized in a 1:1 fashion to 2 treatment arms (vedolizumab IV or placebo IV). Randomization

will be stratified by age (≥18 years or adolescents aged 12 to <18 years), HLA match or mismatch (8/8 or 7/8), conditioning regimen (myeloablative or reduced intensity conditioning), and treatment with or without ATG (ATG-F or thymoglobulin). Subjects randomized to the vedolizumab IV arm will receive 7 doses of vedolizumab IV, beginning on Day -1 before allo-HSCT and then on Days +13, +41, +69, +97, +125, and +153 after allo-HSCT. Subjects randomized to the placebo IV arm will receive 7 doses of placebo IV, beginning on Day -1 before allo-HSCT and then on Days +13, +41, +69, +97, +125, and +153 after allo-HSCT.

5.0 ANALYSIS ENDPOINTS

5.1 Primary Endpoint

The primary endpoint is intestinal aGvHD-free survival by Day +180 after allo-HSCT. Intestinal aGvHD is defined as Stage 1-4 intestinal involvement per Acute Graft-versus-Host Disease Clinical Stage criteria.

5.2 Secondary Endpoints

The secondary endpoints are:

- Intestinal aGvHD-free and relapse-free (of the underlying malignancy) survival by Day +180.
- Grade C-D aGvHD-free (any organ involvement) survival by Day +180.
- Non-relapse Mortality (NRM) by Day +180.
- Overall Survival (OS) by Day +180.
- Grade B-D aGvHD-free (any organ involvement) survival by Day +180.

5.3 Safety Endpoints

- Safety as assessed by adverse events (AEs), adverse events of special interest (AESIs), SAEs, vital signs, results of standard laboratory test and procedures (eg, clinical chemistry, hematology, coagulation).

CCI

[redacted]

CCI

6.0 DETERMINATION OF SAMPLE SIZE

Assuming the event rate for the intestinal aGvHD-free survival by Day +180 after allo-HSCT is 34.1% for the placebo group and 21.8% for the vedolizumab group and the rate of loss-to-follow-up is 10% for both groups, a sample size of 279 subjects per group (558 subjects total) is expected to generate 148 events for the intestinal aGvHD-free survival and hence provide 90% power at the at alpha of 0.05 for a 2-sided hypothesis based on log-rank test. Assuming the event rate for the intestinal aGvHD-free and relapse-free survival is 39.5% for the placebo group and 27.1% for the vedolizumab group, this sample size is expected to generate 177 events for the intestinal aGvHD-free and relapse-free survival and hence provide approximately 86% power at alpha of 0.05 for a 2-sided hypothesis.

7.0 METHODS OF ANALYSIS AND PRESENTATION

The End of Treatment Analysis (ie, primary analysis) for efficacy and safety data will be performed at the end of treatment (EOT), when all subjects have completed Day +180 or withdrawn from the study, or when the planned number of primary endpoint events as determined by the sample size adaptation rule in the interim analysis are accrued. The End of

Study Analysis (ie, final analysis) will be performed after the final database lock at the end of the study (EOS). Please refer to Section 7.1.3 for details.

7.1 General Principles

All statistical analyses will be conducted using SAS® Version 9.4, or higher.

All statistical testing will be 2-sided performed at an alpha level of 0.05. All confidence intervals (CI) reported will be 2-sided 95%, unless stated otherwise. P-values will be rounded to 4 decimal places prior to assessment of statistical significance.

Means and medians will be presented to 1 more decimal place than the recorded data. The standard deviations (SDs) will be presented to 2 more decimal places than the recorded data. Confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate.

A visit window convention will be used to determine the analysis value for a given study visit for observed data analyses, unless otherwise stated.

Where appropriate, variables will be summarized descriptively by analysis visit. For the categorical variables, counts and percentages of each possible value will be tabulated. The denominator for the percentage will be based on the number of subjects who provided non-missing responses to the categorical variable. For continuous variables, the number of subjects with non-missing values, mean, median, SD, minimum, and maximum values will be presented.

7.1.1 Study Definitions

Endpoint	Definition
Intestinal aGvHD-free survival by Day +180 after allo-HSCT*	Time from the date of first study drug administration (Day-1, see 7.1.2) to intestinal aGvHD event/death, where an event is defined as death due to any cause or Stage 1-4 intestinal involvement per Acute Graft versus-Host Disease Clinical Stage criteria. It will be censored for subjects who have not had the intestinal aGvHD event or died or have had the event after pre-specified timing, eg, last contact or Day+180 after allo-HSCT whichever occurs first. If a subject had intestinal aGvHD event and died due to any cause including intestinal aGvHD, time to event will be derived as time to first qualifying event, ie, intestinal aGvHD event.(see Table 7.a)
Intestinal aGvHD-free survival by 12 months after allo-HSCT*	Time from the date of first study drug administration (Day-1, see 7.1.2) to intestinal aGvHD event/death, where an event is defined as death due to any cause or Stage 1-4 intestinal involvement per Acute Graft versus-Host Disease Clinical Stage criteria. It will be censored for subjects who have not had the intestinal aGvHD event or died or have had the event after pre-specified timing, eg, last contact or Day+365 after allo-HSCT whichever occurs first. If a subject had intestinal aGvHD event and died due to any cause including intestinal aGvHD, time to event will be derived as time to first qualifying event, ie, intestinal aGvHD event.(see Table 7.a)

Endpoint	Definition
Intestinal aGvHD-free and relapse-free survival by Day +180 after allo-HSCT*	Time from the date of first study drug administration (Day-1, see 7.1.2) to intestinal aGvHD event/death/relapse, where an event is defined as death or Stage 1-4 intestinal involvement per Acute Graft versus-Host Disease Clinical Stage criteria, or relapse. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+180 after allo-HSCT whichever occurs first.
Intestinal aGvHD-free and relapse-free survival by 12 months after allo-HSCT*	Time from the date of first study drug administration (Day-1, see 7.1.2) to intestinal aGvHD event/death/relapse, where an event is defined as death or Stage 1-4 intestinal involvement per Acute Graft versus-Host Disease Clinical Stage criteria, or relapse. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+365 after allo-HSCT whichever occurs first.
Grade C-D aGvHD-free (any organ involvement) survival by Day +180 after allo-HSCT	An event is defined as Grade C-D aGvHD any organ involvement per IBMTR Severity Index for aGvHD or death. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+180 after allo-HSCT whichever occurs first.
Grade C-D aGvHD-free (any organ involvement) survival by 12 months after allo-HSCT	An event is defined as Grade C-D aGvHD any organ involvement per IBMTR Severity Index for aGvHD or death. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or last contact or Day+365 after allo-HSCT whichever occurs first.
NRM by Day +180 after allo-HSCT	An event is defined as death without occurrence of a relapse. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+180 after allo-HSCT whichever occurs first.
NRM by 12 months after allo-HSCT	An event is defined as death without occurrence of a relapse. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+365 after allo-HSCT whichever occurs first.
OS by Day +180 after allo-HSCT	An event is defined as death from any cause. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+180 after allo-HSCT whichever occurs first.
OS by 12 months after allo-HSCT	An event is defined as death from any cause. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+365 after allo-HSCT whichever occurs first.
Grade B-D aGvHD-free (any organ involvement) survival by Day +180 after allo-HSCT	An event is defined as death or grade B-D any organ involvement per IBMTR Severity Index for aGvHD. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+180 after allo-HSCT whichever occurs first.
Grade B-D aGvHD-free (any organ involvement) survival by 12 months after allo-HSCT	An event is defined as death or grade B-D any organ involvement per IBMTR Severity Index for aGvHD. It will be censored for subjects who have not had the event or have had the event after pre-specified timing, eg, last contact or Day+365 after allo-HSCT whichever occurs first.

*For the determination of Intestinal aGvHD in these endpoints, if an investigator diagnoses a subject as having an Intestinal aGvHD event even though the corresponding Diarrhea Volume data is missing or not defined as Stage 1-4 on the eCRF, it will be considered as a valid primary efficacy event in the efficacy analysis.

Table 7.a Date of Event/Censor for primary endpoint by Day +180 after allo-HSCT

Reported Intestinal aGvHD Event?	Death?	Consider as Primary Efficacy Event?	Date of Primary Efficacy Event	Date of censoring
Yes	No	Yes	Date of reported intestinal aGvHD	
Yes	Yes	Yes	Date of first reorted event, ie, intestinal aGvHD	
No	Yes	Yes	Date of death	
No	No	No. Censor		Date of last observation, or last contact, EOS visit, Day+180, or interim datacut date whichever occurs first

Note: Event or Death that occurs after Day +180 will be censored at Day +180.
Note: In the case that liver or skin aGvHD has occurred but neither intestinal aGvHD nor death has occurred, such subjects will be censored subject.

7.1.2 Definition of Study Days

Per study protocol, Study Day 0 is defined as the date on which HSC infusion occurs (no sooner than 12 hours after completion of IV infusion of study drug on Day –1. Other study days are defined relative to the Study Day 0 with Day +1 being the day after, and Day -1 being the day prior to Study Day 0.

However, Day -1 per study protocol corresponds to Analysis Day 1 in ADaM datasets following the CDISC ADaM implementation guide. Similarly, other analysis study days in ADaM datasets are defined relative to the Analysis Day 1 (ie, Study Day -1 per protocol) with Analysis Day 2 being the day after (ie, Study Day 0 per protocol).

For time to event, the day until event/censor after allo-HSCT is following formula:

- The date of event/censor – the date of the first dose of study drug + 1.

7.1.3 Definition of Study Visit Windows

Baseline is defined as the last non-missing measurement prior to or on the date of the first dose of study drug (Study Day -1).

Subjects do not always adhere strictly to the visit timing in the protocol. Therefore, the designation of visits throughout the study will be based on the actual day of evaluation relative to the start date of the study rather than the nominal visit recorded in the CRF. The visit windows for the post-baseline visits are defined by the middle point of the consecutive scheduled days, allowing equally distanced windows (See Table 7.b) between the scheduled days when applicable. The visit windows should be used to summarize the endpoints that are defined to target the scheduled visit days with the exception of time to event endpoints. If a subject has

more than 1 measurement in the same visit window, the measurement closest to the scheduled visit will be used. If 2 measurements in the same window are of equal distance to the scheduled visit, the measurement that occurs after the scheduled visit will be used. If 2 or more measurements occur on the same day, the mean value will be used.

Table 7.b Visit Windows to be used for the summary of the endpoints

Visit	Scheduled Visit Day	Assessment of GvHD (as binary endpoint)	Vital signs	Laboratory Assessment	FACT-BMT, EQ-5D	Immunogenicity
Baseline	-1*	≤-1	≤-1	≤-1	≤-1	≤-1
Day 6	+6	+0 — +10	+0 — +10	+0 — +10		
Day 13	+13	+11 — +16	+11 — +16[(a)]	+11 — +16		
Day 20	+20	+17 — +23	+17 — +23	+17 — +23		
Day 27	+27	+24 — +30	+24 — +30	+24 — +30		
Day 34	+34	+31 — +37	+31 — +37	+31 — +37		
Day 41	+41	+38 — +55	+38 — +55[(a)]	+38 — +55		+0 — +69
Day 69	+69	+56 — +83	+56 — +83[(a)]	+56 — +83		
Day 97	+97	+84 — +111	+84 — +111[(a)]	+84 — +111	+0 — +139	+70 — +138
Day 125	+125	+112 — +139	+112 — +139[(a)]	+112 — +139		
Day 153	+153	+140 — +167	+140 — +167[(a)]	+140 — +167		
Day 180	+180	+168 — +187	+168 — +230	+168 — +230	+140 — +270	+139 — +230
Day 280	+280	+188 — +320	≥+231	≥+231		≥+231
Month 12	+365	≥+321			≥+271	

(a) Vital signs will be obtained before and within 60 minutes of completion of IV infusion of vedolizumab.
* Scheduled visit Day -1 per protocol corresponds to Analysis Day 1 in ADaM.

7.1.4 Conventions for Missing Adverse Event Dates

Every effort will be made to determine the actual onset date for the event or to obtain a reliable estimate for the onset date from the investigator.

For AEs or SAEs, a missing or incomplete onset date will be imputed according to the following conventions:

1. If an onset date is missing, the derived onset date will be calculated as the first non-missing valid date from the following list (in order of precedence):
   - First study medication date.
   - Consent date (for SAEs only).

2. If an onset date is incomplete, the derived onset date will be calculated following:
   - Missing day, but month and year present: the day will be imputed as the 15th of the month. If the first study medication dose occurs after the imputed date in the same month, the derived onset date will be set equal to the first study medication date.
   - Missing day and month, but year present: the day and month will be imputed as the 30th June of the year. If the first study medication dose occurs after the imputed date, the derived onset date will be set equal to the first study medication date.

For AEs or SAEs, a missing or incomplete end date will be imputed according to the following conventions:

1. If an end date is missing, the derived end date will be imputed the last assessment date.
2. If an end date is incomplete, the derived end date will be calculated following:
   - Missing day, but month and year present: the day will be imputed as the last date (for example February 2009 will be imputed as 28 February 2009) of the month.
   - Missing day and month, but year present: the day and month will be imputed as the 31st December of the year.

7.1.5 Conventions for Missing Concomitant Medication Dates

Start and stop dates for all concomitant medications are collected on the CRF. However, in case of missing or partial information in these dates, the following rules will be used:

If the start date is unknown or partial:

- If the day is missing, the start day will be the first day of the month.
- If the month is missing,
  - If the year is the same as the date of first dose of study drug, the start month will be the month of the dose first study drug.
  - If the year is not the same as the date of first dose of study drug, the start month will be January.
- If the year is missing, the start year will be the minimum of the year of the first clinic visit or the year of the informed consent date.
- If the entire start date is unknown, the start date will be the date of first study drug administration.

If the stop date is missing, partial or "ongoing:"

- If the day is missing, the stop day will be the last day of the month reported.
- If the month is missing,
  - If the year is the same as the date of last assessment, the stop month will be the month during which the last assessment occurred.

  - If the year is not the same as the year of the last assessment, then the end month will be December.

- If the year is missing", the stop year will be the year in which the last in which the last assessment occurred.
- If the entire stop date is unknown or if the medication is "ongoing", the stop date will be the date of last assessment.

7.1.6 Methods for Handling of Missing Efficacy Data

Through the end of the double-blind period, the missing dichotomous efficacy data (eg, intestinal aGvHD event as dichotomous endpoint) will be handled using the non-responder imputation method, ie, any subject with missing information for determination of endpoint status will be considered as having an undesirable outcome in the analysis.

Missing data for continuous endpoints will be imputed using last available post-baseline observation carried forward (LOCF) method. For subjects without any non-missing post-baseline measurement, the missing data will be imputed using baseline observation carried forward method. Other missing data imputation method (eg, multiple imputations or repeated measure mixed effects model) may be explored.

Missing data for time to event endpoints will not be imputed.

Other missing data handling method may be explored.

7.2 Analysis Sets

The analysis sets used for analysis will include the following:

Randomized Set will be included all subjects who are randomized.

Full Analysis Set (FAS): Following the intent-to-treat principle, the FAS will include all subjects who are randomized, receive at least 1 dose of the treatment, and undergo allo HSCT. The analysis will be based on the treatment to which the subjects are randomized. This population will be used for efficacy analysis.

Safety Analysis Set (SAF) will include all subjects who receive at least 1 dose of study drug. Subjects in this set will be analyzed according to the treatment actually received. This population will be used for safety analysis.

CCI [redacted]

7.3 Disposition of Subjects

7.3.1 Study Information

General study information will be provided, including the date the first subject signed informed consent, the date of the last subject's last visit/contact, the date of the last subject's last visit for

collection of data for primary endpoint, MedDRA version, WHODrug version and SAS version used for creating the datasets.

7.3.2 Disposition of Subjects

Summary of the subject disposition will be summarized for each treatment group and overall, and will include the following:

- Summary of screen failures using all subjects who were not randomized.
- Summary of eligibility for randomization using all subjects who signed the informed consent form.
- Summary of study subjects randomized by site, geographic region (See Appendix D) using randomized set.
- Summary of patients for each pre-defined population using randomized set: full analysis set, per-protocol population, safety analysis set, CCI population.
- Summary of subjects completing study drug using randomized set.
- Summary of subjects completing study using randomized set.
- Disposition of subjects based on the reasons for discontinuation of treatment and for failing to complete the study using randomized set.
- Significant protocol deviations will be summarized descriptively using randomized set.

To assess the impact of COVID-19 on the safety of participating subjects, the following analyses of COVID-19 impact will be conducted if data permits. Additional tables and listings may be generated.

- Disposition of subjects based on the COVID-19 related discontinuation, including discontinuation due to adverse events in light of COVID-19 infection and discontinuation due to COVID-related reasons other than COVID-infection (eg, travel limitation, reduced site staff, etc) using randomized set.
- Significant protocol deviations due to COVID-19 pandemic will be summarized descriptively using randomized set.
- Summary of COVID-19 Impact including the summary statistics of Number of Missed Visits, Visits with Missed Assessments and Visits with Alternative method of Contact and the number and percentage of subjects with at least one missed visit and missed visits by study visit.

When calculating percentages for the reasons for not being treated, the total number of subjects not treated by the study drug will be used as the denominator. When calculating percentages for the reasons for discontinuation, the total number of subjects who prematurely discontinued will be used as the denominator.

7.4 Demographic and Other Baseline Characteristics

Demographic and baseline characteristics variables will be summarized using the FAS.

For continuous variables (age, weight, height, body mass index [BMI], Total CD34+, Stem Cell Dose: Total nucleated cell dose, and Stem Cell Dose: CD3+, underlying disease duration since diagnosis), summary statistics will be generated. BMI (in kg/m^2) will be calculated using the subject's height at screening and baseline weight measurements and summarized.

For categorical variables (age [<18, and ≥-18] [<40, 40 to 60, and >60], gender, ethnicity, race), the number and percentage of subjects in each category will be presented.

The following GvHD-related baseline characteristics will be summarized by treatment group for subjects:

Baseline Disease Characteristics	**Categories**
Geographic region	North America, South America, Western/ Northern Europe, Central Europe, Eastern Europe, Asia / Australia
Baseline ECOG Status	0, 1, 2, 3, 4
HLA Match	7/8, 8/8
Conditioning Regimen	Myeloablative, Reduced Intensity Conditioning
ATG	With, Without
CNI	Tacro vs Cyclosporine
GvHD Prophylaxis	Tacro + MTX, Tacro + MMF, Tacro + MTX + ATG, Tacro +MMF + ATG, Cyclo + MTX, Cyclo + MMF, Cyclo +MTX +ATG, Cyclo + MMF + ATG
Primary Disease	Acute Myeloid Leukaemia, Acute Lymphoid Leukaemia, Chronic Myeloid Leukaemia, Other Myeloproliferative Disorder, Myelodisplastic Syndrome, Chronic Lymphocytic Leukemia, Non-Hodgkin Lymphoma
Donor/Recipient CMV Match	D-/R+, D+/R+, D+/R-, D-/R-
Cytogenetic results (for AML/MDS)	AML • Favourable (for AML only) • Intermediate (for AML only) • Unfavourable (for AML only) • Unknown MDS • Low (for MDS only) • Intermediate1 (for MDS only) • Intermediate2 (for MDS only) • High (for MDS only) • Unknown

Baseline Disease Characteristics	Categories
Prior Exposure-Disease Status at time of HSCT (ALL and AML and CML)	ALL and AML • Complete Remission 1 (ALL and AML only) • Complete Remission >1 (ALL and AML only) • Other (ALL and AML only), CML • 1st Chronic phase (CML only) • Failing TKI (CML only) • Accelerated phase or >1st Chronic phase (CML only) • Blast crisis (CML only) • Progression (CML only)
Source of Stem Cells	Bone Marrow, Peripheral Blood Mononuclear Cell
Donor-recipient Gender Match	Female Subject-Female Donor, Female Subject-Male Donor, Male Subject-Female Donor, Male Subject-Male Donor
Locus of Mismatch	HLA-A Gene, HLA-B Gene, HLA-C Gene, HLA-DRB1 Gene, Other Gene

7.5 Medical History and Concurrent Medical Conditions

During the screening period, a complete medical history will be compiled for each subject. The history will emphasize the background and progress of the subject's malignancy and include a description of prior therapies for it. Medical and Surgical History will be summarized by system organ class and preferred term based on Safety analysis set.

7.6 Medication History and Concomitant Medications

Medications used by the subject and therapeutic procedures completed by the subject will be recorded in the eCRF from the first dose of study drug on Day -1 through 12 months after allo-HSCT.

Concomitant medications will be coded using WHO Drug. The number and percentage of patients taking concomitant medications will be tabulated by standardized medication name WHO drug generic term in the Safety analysis set, and categorized as follows:

- Concomitant medications that started and stopped prior to baseline.
- Concomitant medications that started prior to and were ongoing at baseline and those that started after baseline.
- Data listings of Concomitant medications of COVID-19 Vaccine will be presented separately from other Concomitant mediations.

7.7 Study Drug Exposure and Compliance

Vedolizumab will be administered or dispensed only to eligible patients under the supervision of the investigator or identified subinvestigator(s). The appropriate study personnel will maintain

records of vedolizumab receipt and dispensing. The number of doses will be summarized as both continuous and ordinal categorical variable using the safety analysis set.

The duration of exposure will be calculated as (date of last dose of study drug - date of first dose of study drug) + 127 days. If last dose date is missing, then 127 days will be imputed as treatment period.

Overall study drug compliance (%) will be determined as (total count of complete infusions taken / total number of infusions expected during study treatment, ie, 7 infusions expected for all subjects) × 100%. A subject must receive at least 75% of the infusion in order for it to be considered complete for each dose.

The number and percentage of subjects with overall study drug compliance will also be summarized in the following categories[<80%, 80 to 90%, ≥90%].

If data permits, additional data listing and/or summary table may be presented to evaluate the subject's adherence to study treatment as affected by COVID-19 pandemic, and subject's treatment accessibility as affected by COVID-19 pandemic.

7.8 Efficacy Analysis

Efficacy data will be analyzed based on the FAS. Missing data will be handled according to Section 7.1.6.

Efficacy endpoints and the associated supporting data will be descriptively summarized by age groups and overall respectively, and presented in supporting data listings. Unless stated otherwise, all statistical comparisons will be conducted in the overall population.

All statistical testing will be 2-sided performed at an alpha level of 0.05. To control the overall Type I error rate for the comparison between vedolizumab and placebo groups for the primary and key secondary efficacy endpoints in the overall population, a fixed-sequence testing approach will be used. Specifically, the statistical testing of the intestinal aGvHD-free and relapse-free survival will only be performed if the treatment difference for primary efficacy endpoint is statistically significant (ie, $p < 0.05$). The next secondary efficacy endpoint will only be tested if the treatment difference for the first secondary efficacy endpoint is significant (ie, $p < 0.05$), and so on for each subsequent secondary efficacy endpoint.

The order of the statistical testing for the primary and key secondary endpoints is as follows:

1. Intestinal aGvHD-free survival by Day +180 after allo-HSCT.
2. Intestinal aGvHD-free and relapse-free survival by Day +180.
3. Grade C-D aGvHD-free (any organ involvement) survival by Day +180.
4. NRM by Day +180.
5. OS by Day +180.
6. Grade B-D aGvHD-free (any organ involvement) survival by Day +180.

The testing of the additional efficacy endpoints will not be multiplicity-adjusted. Nominal p-values will be presented.

7.8.1 Primary Efficacy Endpoint(s)

The primary endpoint is intestinal aGvHD-free survival by Day +180 after allo-HSCT. Intestinal aGvHD is defined as Stage 1-4 intestinal involvement per Acute Graft-versus-Host Disease Clinical Stage criteria. The primary efficacy endpoint event is defined as death due to any cause or Stage 1-4 intestinal involvement per Acute Graft versus-Host Disease Clinical Stage criteria.

7.8.1.1 Primary Efficacy Analyses

The null and alternative hypotheses for the primary efficacy endpoint, intestinal aGvHD-free survival by Day +180 after allo-HSCT, are:

H_0: Intestinal aGvHD-free Survival by Day +180 $_{Vedolizumab}$ = Intestinal aGvHD-free Survival by Day +180 $_{Placebo}$

versus

H_A: Intestinal aGvHD-free Survival by Day +180 $_{Vedolizumab}$ ≠ Intestinal aGvHD-free Survival by Day +180 $_{Placebo}$

The primary endpoint will be analyzed using the log-rank test, with Kaplan-Meier estimates presented. Subjects without documented intestinal aGvHD events or death before reaching Day +180 after allo-HSCT will be censored at the date of last assessment/visit/contact or Day +180, whichever occurs first. The randomized subjects who did not receive any of their assigned study drug will be censored at Day -1. The statistical significance of treatment effect will be tested against 2-sided alpha level of 0.05.

The primary efficacy endpoint will also be analyzed using a Cox proportional hazards model with treatment group, stratified by randomization strata. Point estimate of hazard ratio and the corresponding 95% CI will be presented. The randomization strata are as follows:

- Age (≥18, ≥12 and <18).
- HLA Match (7/8, 8/8).
- Conditioning Regimen (Myeloablative, Reduced Intensity Conditioning).
- ATG (With, Without).

In addition, summary statistics and Kaplan-Meier estimates by treatment group will be provided, including:

- Number and percentage of events,
- Number and percentage of censored observations,
- Median time to event and 95% CI,
- 25th percentile of the time to event and 95% CI,

- 75th percentile of the time to event and 95% CI,
- Range (minimum and maximum time to event),
- Kaplan-Meier estimates of the event rates (percentage and *n*) at 6 months,
- Kaplan-Meier estimates of the event rates (percentage and *n*) at 12 months,
- Plots of the Kaplan-Meier estimated survival curves.

The Kaplan-Meier mentioned above will be based on LIFETEST procedure in SAS®.

7.8.2 Secondary Efficacy Endpoint(s)

The secondary endpoints are:

- Intestinal aGvHD-free and relapse-free (of the underlying malignancy) survival by Day +180.
- Grade C-D aGvHD-free (any organ involvement) survival by Day +180.
- NRM by Day +180.
- OS by Day +180.
- Grade B-D aGvHD-free (any organ involvement) survival by Day +180.

All time-to-event endpoints will be analyzed in a similar way as primary efficacy endpoint, using log-rank test in the FAS. Cox proportional hazards model will be fitted with treatment group, stratified by the randomization strata. Point estimate of hazard ration and the corresponding 95% CI will be presented.

CCI

[redacted]

CCI

[redacted]

CCI

7.8.4 Sensitivity Analysis

To assess the robustness of the analysis for primary efficacy endpoint and key secondary endpoints, the following sensitivity analysis will be performed.

- To accommodate the events of interest that occur within the protocol-defined +/-7 day window, the primary and secondary efficacy endpoints will be analyzed by Day +187 after allo-HSCT. Subjects who have not had an event by Day +187 after allo HSCT will be censored at last contact or Day+187 after allo HSCT whichever occurs first.

- If any clinical site has detected or reported significant noncompliance with regulatory requirements during the course of study, sensitivity analysis may be conducted for the in the FAS excluding all subjects from that particular site.
- Primary and secondary efficacy endpoint will be analyzed using stratified log-rank test using the FAS, stratified by randomization stratification factors.
- Primary, secondary, CCI [redacted] efficacy endpoint may be explored using the competing risk analysis to accommodate the competing nature of multiple causes to the same event.

Additional sensitivity analysis may be performed as appropriate (eg, related to COVID-19 impact). If there is inconsistency in stratification information collected in IRT versus EDC, sensitivity analysis for inconsistency among them may also be performed.

7.8.5 Subgroup Analysis

This section applies to primary efficacy endpoint as well as the following selected efficacy endpoints:

1. Intestinal aGvHD-free survival by Day +180 after allo-HSCT
2. Intestinal aGvHD-free and relapse-free survival by Day +180.
3. Grade C-D aGvHD-free (any organ involvement) survival by Day +180.
4. NRM by Day +180.
5. OS by Day +180.
6. Grade B-D aGvHD-free (any organ involvement) survival by Day +180.

The subpopulations of interest are defined by the following baseline characteristics outlined Table 7.c.

Table 7.c List of subgroups of interest

Subgroup of Interest	Subgroup Categories
Gender	Male, Female
Race	White, Non-White
Geographic Region	North America, South America, Western/ Northern Europe, Central Europe, Eastern Europe, Asia / Australia
HLA Antigen Match	7/8, 8/8
Conditioning Regimen	Myeloablative, Reduced Intensity Conditioning
ATG	With, Without
Primary Disease	Acute Leukemia, Chronic Myelogenous Leukemia, Myelodysplasia, Chronic Lymphocytic Leukemia, Small Lymphocytic Lymphoma with chemosensitive disease, Other nonHodgkin or Hodgkin Lymphoma, Myelofibrosis and other Myeloproliferative disorders
GvHD Prophylaxis Therapies	CNI +MTX+ATG, CNI +MTX-ATG, CNI + MMF+ATG, and CNI + MMF-ATG
CNI	Tacrolimus vs Cyclosporine
Stem Cell Source	Bone Marrow, Peripheral Blood Mononuclear Cel

CCI [redacted]

If the value of the baseline grouping variable cannot be determined, the subject will be excluded from the corresponding subgroup analysis. If the number of subjects in any subgroup is less than 10, that subgroup will not be presented.

The results will be tabulated and the corresponding forest plots for the subgroup analyses will be presented as well.

To evaluate the impact of COVID-19 on the efficacy of the participating subjects, additional subgroup analysis may be performed (eg, COVID-19 infected subjects vs COVID-19 not-infected subjects).

CCI [redacted]

7.9.2 Pharmacodynamic Analysis

Not applicable

7.10 Other Outcomes

CCI

[redacted]

CCI

[redacted]

7.11 Safety Analysis

Safety analysis will be performed using the Safety Analysis Set with the actual treatment received. The analysis of safety endpoints will include AEs, clinical laboratory values, and vital signs. No statistical inference will be made for safety analyses.

7.11.1 Adverse Events

Adverse events (AEs) will be coded by MedDRA, and summarized by System Organ Class, High Level Term, and Preferred Term. AEs will also be summarized by severity, and by relationship to study investigational product respectively. All AEs and SAEs collected in the database (including those starting prior to first dose of study drug) will be listed. Any other information collected (eg, relatedness to study drug, action taken etc.) will be listed as appropriate.

A Treatment Emergent AE (TEAE) is defined as any AEs newly occurring or worsening from the first dose and 18 weeks after last dose of study treatment. A Serious Treatment Emergent AE (TEAE) is defined as SAE newly occurring or worsening from the first dose up to 18 weeks after last dose of study treatment, regardless of relationship to study drug. SAEs will also be collected from 18 weeks after last dose of study treatment to EOS. The number of percentage of subjects with TEAEs, serious TEAEs, and SAEs will be summarized. Key guidelines for determining the incidence of AEs are as follows:

- AEs with missing or unknown severity will be considered as severe.

- AEs with missing or unknown relationship to study drug will be counted as related.
- A subject with 2 or more AEs within the same level of the MedDRA term will be counted only once in that level.

The incidence rates of treatment-emergent AEs, as well as the frequency of occurrence of overall toxicity categorized by maximum toxicity grades (severity), will be described. Toxicity grade will be evaluated according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE version 5.0). In addition, treatment-emergent AEs will be summarized by causal relationship to study treatment (in the opinion of the investigator) and TEAE leading to drug discontinuation. Serious treatment-emergent AEs (Serious TEAEs) and SAEs, TEAE leading to death, both overall and by causal relationship to study treatment, will also be summarized. Exposure adjusted TEAE rates (eg, per 100 subject-years) will be summarized as well to accommodate variable subject follow-up time, where the exposure-adjusted incidence rate is defined as the number of subjects experiencing the AE divided by the total exposure-time among subjects, where each subject's exposure is calculated as ([(date of last dose of study drug + 126) – date of first dose of study drug] +1) / 365.25.

- TEAEs by system organ class (SOC), high level term (HLT), and preferred term (PT).
- Exposure-adjusted TEAEs by SOC, HLT, and PT.
- Drug related TEAEs by SOC, HLT and PT.
- Serious Pre-treatment Events by SOC, HLT, and PT.
- Serious TEAEs by SOC, HLT and PT.
- Serious AEs occurring beyond 18 weeks after last dose until the EOS by SOC, HLT and PT.
- Serious drug related TEAEs by SOC, HLT and PT.
- Serious drug related AEs occurring beyond 18 weeks after last dose until the EOS by SOC, HLT and PT.
- TEAEs leading to treatment discontinuation by SOC, HLT, and PT.
- TEAEs by PT by descending order of frequency.
- Grade 3 or higher TEAEs by SOC, HLT, and PT.
- Grade 3 or higher Drug-Related TEAEs by SOC, HLT, and PT.
- Most frequent AEs by PT (sorted by frequency, occurring in ≥5% of subjects in any treatment arm).
- COVID-19 related TEAEs (eg, PT containing "COVID-19", "SARS-CoV-2", "Coronavirus" Terms).
- Serious AEs in COVID-19 infected subjects.

- AEs of special interest by SOC, HLT, and PT.
  - Serious Infections.
  - PML.
  - Malignancy of AEs.
  - Liver injury.
  - Hypersensitivity Reactions including IRRs and Injection Site Reactions of AEs.
  - Leukopenia or Lymphopenia.
  - CMV colitis.
  - CMV Reactivation.
  - All AESI.

On-study deaths will be summarized based on Serious Adverse Event and End of Study CRF.

All adverse events will be listed by treatment, study center, subject number and onset date of the adverse event. The listing will contain: subject identifier, age, sex, body weight, race, adverse event (preferred term and reported term), SOC, onset date, end date or whether the event was ongoing, duration, frequency, intensity, action taken concerning study drug, causality to study drug, the outcome, whether the adverse event was an SAE and whether the event was an adverse event of special interest (AESI).

The incidence rates of CMV Reactivation will be described by donor CMV IgG status and recipient CMV IgG status.

All CMV Reactivation infections of AEs will also be listed by treatment, study center, subject number and onset date of the adverse event. The listing will contain the reported verbatim term in addition to the contents mentioned above.

All deaths and PML checklist data will be listed.

7.11.2 Clinical Laboratory Evaluations

Blood samples for analysis of the Hematology, Serum Chemistry, Coagulation, and Liver Functions parameters shown in following table.

Table 7.d Clinical Hematology, Coagulation, Chemistry, and LFTs

Hematology	**Serum Chemistry**	
Hematocrit	Albumin	Glucose
Hemoglobin	Bicarbonate (HCO_3)	Phosphate
Red blood cells (RBC)	Blood urea nitrogen (BUN)	Potassium
Leukocytes with differential including neutrophils, monocytes, and lymphocytes	Calcium	Sodium
	Creatinine	Urate
Platelets	Chloride	
Coagulation [a]	**Liver Functions**	
Prothrombin (PT)/INR	Alkaline phosphatase (ALP)	Bilirubin (total and indirect)
Activated partial thromboplastin time (aPTT)	ALT	Lactate dehydrogenase (LDH)
	AST	

The analysis of the laboratory data will include parameters tabulation for each assessment visit by treatment group, using the following:

- Number of subjects with non-missing values (n),
- Arithmetic mean,
- Median,
- Standard deviation (SD),
- Minimum and maximum observed values (Min, Max).

For the post-baseline assessments, the summary will also include the change from baseline values.

Mean laboratory values over time will be plotted for key laboratory parameters (eg, hematology, chemistry, and LFTs). Shift tables for laboratory parameters will be generated based on changes in the classification with low, normal and high values relative to the normal range from baseline to post-baseline values. Subjects with markedly abnormal values for laboratory tests will be tabulated (see Appendix E for details).

7.11.3 Vital Signs

Vital signs, including heart rate, respiratory rate, systolic and diastolic blood pressure, and temperature will be summarized descriptively following the windowing convention for visits specified events. Subjects with markedly abnormal values for vital signs values will be tabulated (see Appendix F for details).

7.11.4 12-Lead ECGs

Not applicable

7.11.5 Other Observations Related to Safety

Not applicable

7.12 Interim Analysis

One interim analysis for futility and sample size re-estimation is planned for this study, which will occur after approximately 30% of the targeted primary endpoint events (ie, approximately 44 intestinal aGvHD/death events by Day +180 after allo-HSCT) have been accrued. The analysis will be performed by an independent statistical center (ie, an external vendor) in a manner that maintains the blinding to the sponsor study team, the investigators, and the subjects. The interim analysis results will be presented to the DMC but will not be shared with the sponsor. Based on the result, the DMC will make a recommendation to the Sponsor Liaison (who has been specified in DMC charter) and Sponsor Executive Committee (SEC) (who generally does not involve in daily trial management or attend the DMC meeting). The SEC and Sponsor Liaison may communicate information in the Open Report to the Sponsor's senior management and may inform them of the DMC-recommendation in which the SEC and Sponsor Liaison has reached a final decision agreeing with the recommendation. Specific criteria for DMC recommendations will be provided in the DMC charter. The futility stopping is considered nonbinding. Details on the DMC recommendation and communication between DMC and SEC will be documented in the DMC Charter.

If 148 primary endpoint events are accrued before all subjects have completed Day +180 or dropped out, the primary analysis for efficacy (ie, EOT analysis) may be performed at that time.

7.12.1 Analysis Endpoints

7.12.1.1 Primary Efficacy Endpoint

The efficacy endpoint of interest is intestinal aGvHD-free survival by Day +180 after allo-HSCT. Events are considered death due to any cause or Stage 1-4 intestinal involvement per Acute Graft versus-Host Disease Clinical Stage criteria. Subjects who have completed Day +180 or withdrawn from the study prior to Day +180 but without an observed primary efficacy event and subjects who are still ongoing and have not reached Day +180 by the time of IA will be censored at the date of last assessment/visit/contact, or Day +180, or interim data cut date, whichever occurs first.

The cumulative incidences of Intestinal aGvHD-free survival by Day +180 after allo-HSCT and the two-sided 95% confidence intervals will be provided by treatment group using the Kaplan-Meier method. Log-rank test statistic will be used to derive conditional power for interim decisions.

7.12.2 Assessment of Conditional Power

Conditional power for the primary efficacy endpoint will be assessed based on the log-rank test statistic using the observed intestinal aGvHD-free survival data at IA.

The calculation of the conditional power is based on B-value (Lan et. al. 1988). When data are monitored at t, B(t) is observed, and the conditional power assuming the trend indicated by the interim data is calculated as

$$CP_t = 1 - \Phi\left\{\frac{z_{(1-\frac{\alpha}{2})} - B(t)/t}{\sqrt{1-t}}\right\}$$

- t is information fraction at the time of IA relative to the originally planned sample size. $t = \frac{n1}{N}$, where $n1$ is the observed number of primary efficacy events at IA; and N is the originally planned number of primary efficacy events (ie, N=148).
- B(t) is B-value at IA. $B(t) = Z_t\sqrt{t}$, Z_t denotes a negative value of the observed log-rank test statistic LR_1 at the interim analysis, ie, $Z_t = -LR_1$. Note that the observed hazard ratio HR_1 at the interim analysis is approximately $HR_1 = \exp\left(\frac{LR_1}{\sqrt{\frac{n1}{4}}}\right)$.
- Φ is the standard normal cumulative distribution function.
- $\alpha = 0.05$.

7.12.3 Interim Decision Rules

The study design employs promising zone design, an adaptive sample size re-assessment approach (Mehta and Pocock, 2011) that strongly controls the overall type I error rate. The conditional power (see Section 7.12.2 for details) will be calculated based on primary efficacy endpoint, intestinal aGvHD-free survival. If the conditional power falls in the promising zone, the sample size will be increased according to a prespecified sample size adaptation rule; if the conditional power falls in the futility zone, the study may stop for futility; otherwise the study will continue with the sample size unadjusted. The DMC may also recommend that the study be stopped because of concerns about the safety for the study participants.

The DMC will consider the following guidelines based on conditional power for making recommendations at the Interim Analysis. Of note the actual unblinded conditional power thresholds of promising zone design and sample size adaptation rule will be prepared by an independent design statistician in a separate document and will only be available to DMC and independent design statistician.

1. Futility Assessment: If the conditional power is below a prespecified cutoff, 10% (ie, CP <10%), the recommendation will be to stop the trial for futility. The sponsor may or may not follow the recommendation of futility stopping.
2. If the conditional power falls into an "unfavorable" zone, ie, above 10% but below a prespecified k2 (where k2 ≥50%) (eg, 10% ≤CP < 50%), the recommendation will be to continue the trial with no sample size modification.

3. If the conditional power falls into an "promising" zone, ie, above k2 but below a prespecified k3 (eg, 50% ≤CP < 90%), the recommendation will be to continue the trial with an increase in the sample size. The increase in the event size will be a step function up to maximum 35% increase from the original planned number of primary events (ie, approximately 52 additional events). The sponsor may or may not follow the recommendation of sample size increase.
4. If the conditional power falls into a "favorable" zone, ie, above k3 (eg, CP ≥90%), the recommendation will be to continue the trial with no sample size modification.

Note the actual k2 and k3 in the pre-specified sample size adaptation rule document prepared by independent design statistician may be different from the numeric examples above.

In addition to the above formal criteria, relevant safety data from the current study will be used to guide the potential modification of design of the study. The DMC may also recommend that the study be stopped because of concerns about the safety for the study participants. Potential modifications that the DMC can recommend include:

- Continue trial without modifications (favorable or unfavorable zone).
- Continue trial with sample size adjustment (promising zone).
- Terminate trial due to lack of efficacy (futility zone).
- Terminate trial due to safety (at the discretion of the DMC).

Any recommendation other than "continue trial without modifications" must be accompanied by justifications for the recommendations and other follow up requirements as deemed necessary.

7.12.4 Multiplicity Adjustment

The 50% conditional power (CP) principle will be implemented, which allows sample size increase only when the unblinded interim results are promising or the conditional power is greater than 50% (Chen, DeMets, and Lan, 2004; Mehta and Popock, 2011; Broberg, 2013). Thus, the conventional unweighted test statistics and critical values can be used without inflation of type I error rate. The interim analysis results will be presented to the DMC but will not be shared with the sponsor.

7.12.5 Data to be Presented

A list of TLFs to be generated for this planned IA with interim SSR are specified in Appendix B. Additional data-summaries or listings requested by the DMC may be generated.

7.13 Planned End of Treatment (EOT) Analysis and End of Study (EOS) Analysis

7.13.1 Planned EOT Analysis

The EOT analysis (ie, primary analysis) for efficacy and safety data was planned at the end of treatment (EOT), when all subjects have completed Day +180 or withdrawn from the study, or when the planned number of primary endpoint events as determined by the sample size adaptation rule in the interim analysis has been accrued per Protocol Amendment 7 (dated

18Sept2019). However, the sponsor has decided to terminate the enrollment of Vedolizumab-3035 due to the unexpected of COVID-19 pandemic and a shift in standard of care in transplant practice. Subsequently, the EOT analysis will not be performed as planned. The final analysis (ie, EOS analysis) will be performed after the final database lock at the end of study.

7.13.2 Planned EOS Analysis

The EOS analysis (ie, final analysis) will be performed after the final database lock at the end of the study. All available efficacy and safety data will be included in the EOS analysis.

7.14 Changes in the Statistical Analysis Plan

7.14.1 Revision History

Version	Date	Description of Revision
1.0	16Oct2019	N/A
2.0	27Aug2020	• Added COVID-19 related analyses in Section 7.3.2, 7.7, 7.8.4, 7.8.5 and 7.11.1 to assess the impact of COVID-19 • CCI [redacted] • Modified the classification of shift table in Section 7.11.2 to present low, normal and high values relative to normal range instead of NCI CTCAE grade based on the discussion within the study team • Modified prespecified cutoff value for futility assessment for Interim Analysis in Section 7.12.3 based on the discussion within the study team. • Updated the definition of aGvHD to include the criteria for adolescent subjects in Appendix G
3.0	22Oct2021	• Clarified the determination of Intestinal aGvHD event in Section 7.1.1; • Clarified the relationship between study day per protocol versus analysis day in ADaM in Section 7.1.2; • Removed PPS from Section 7.2; • Removed sensitivity analysis based on PPS from Section 7.8; • Removed subgroup analysis by age from Section 7.8.5; • Added references of EQ-5D and FACT-BMT scoring • Removed the EOT analysis from Section 7.13.

8.0 REFERENCES

Chen, YHJ, DeMets, DL, Gordon Lan KK. Increasing the sample size when the unblinded interim result is promising. *Stat Med* 2004; 23(7):1023–1038.

Broberg, P. Sample size re-assessment leading to a raised sample size does not inflate type I error rate under mild conditions. *BMC Medical Research Methodology* 2013; 13:94.

Gernot, W, Werner, B. *Group Sequential and Confirmatory Adaptive Designs in Clinical Trials*. 2016. Switzerland: Springer International Publishing.

Mehta, CR, Pocock, SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. *Stat Med* 2011;30(28):3267-84.

Lan, KG, Wittes, J. The B-value: a tool for monitoring data. *Biometrics* 1988; 1:579-85.

Appendix A AEs of Special Interest

Based on the mechanism of action of vedolizumab, certain adverse events of special interest (AESIs) have been predefined. The categories of adverse events of special interest, and other planned analyses, are described below.

Events	MedDRA Terms or definitions
Infections	• SOC: Infections and Infestations expect PT of cytomegalovirus colitis. • Gastrointestinal infections HLGT in the Gastrointestinal SOC.
PML	• PT: Human polyomavirus infection in the Infections and Infestations SOC. • PT: JC virus infection in the Infections and Infestations SOC. • PT: JC virus test positive in the Infections and Infestations SOC. • PT: Leukoencephalopathy in the Infections and Infestations SOC. • PT: Polyomavirus test positive in the Infections and Infestations SOC. • PT: Progressive multifocal leukoencephalopathy in the Infections and Infestations SOC.
Malignancies including relapse of the primary disease	• SOC: Neoplasms benign, malignant and unspecified (incl cysts and polyps)
Liver injury	• Cholestasis and jaundice of hepatic origin SMQ (Broad) • Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions SMQ (Broad) • Hepatitis, non-infectious SMQ (Broad) • Liver related investigations, signs and symptoms SMQ (Narrow) • Liver infections SMQ (Broad)
Hypersensitivity Reactions	• Anaphylactic/anaphylactoid shock conditions SMQ (broad). • Angioedema SMQ (broad). • Hypersensitivity SMQ (broad).
Leukopenia or Lymphopenia	• PT: Lymphopenia • PT: Leukopenia neonatal • PT: Leukopenia • PT: Radiation leukopenia •
CMV colitis	• PT: Cytomegalovirus colitis in the Infections and Infestations SOC

Appendix B Table, Figures, and Listings for IA

A separate documented labeled as "DMC TLF shells" will summarize the IA report presentation details.

Number	Table Title	Blinded	Unblinded
1	Demographic and Other Baseline Characteristics	Y	Y
2	Summary of Primary Efficacy Endpoint and Conditional Power Assessment		Y

Appendix C Recommended Prophylaxis Treatment: A Combination of CNI (TAC or CYS) + MTX and ATG

The background aGvHD prophylaxis regimen described herein is recommended. Dose and regimen modification are allowed based on toxicity per investigator discretion, per institutional practice, and/or per local prescribing information.

CNI:

1. T***AC***

a) TAC treatment should start during conditioning, before infusion of the graft.

b) The goal TAC treatment should be to achieve a trough concentration of 5 to 20 ng/dL.

c) Recommended to keep at therapeutic levels through Day +100 or per institutional standard.

d) Taper off after Day +100 if no signs of GvHD are observed.

e) The goal should be to discontinue TAC treatment by Day +180 after allo-HSCT.

f) Doses may be modified or held based on toxicity and institutional practice.

2. CYS

a) CYS should start during conditioning, before infusion of the graft.

b) The goal CYS treatment should be to achieve 150 to 450 ng/L

c) Recommended to keep at therapeutic levels through Day +100 or per institutional standard.

d) Taper off after Day +100 if no signs of GvHD are observed

e) The goal should be to discontinue CYS treatment by Day +180 after allo-HSCT.

f) Doses may be modified or held based on toxicity and institutional practice.

MTX[a]

a) Administer at 10 to 15 mg/m^2 IV on Day +1, and 10 mg/m^2 IV on Days +3 and +6, and recommended fourth dose on Day +11 after allo-HSCT.

b) Doses may be modified or held based on toxicity and institutional practice.

c) Leucovorin rescue may be administered per institutional practice.

ATG

1. ATG-F

a) Administer up to 10 mg/kg/day (up to a cumulative dose of 30 mg/kg in divided doses) and per institutional practice.

b) NOT to be administered on the same day as study drug (eg, Day -1).

2. Thymoglobulin

a) Administer up to 2.5 mg/kg/day (up to a cumulative dose of 7.5 mg/kg in divided doses) and per institutional practice.

b) NOT to be administered on the same day as study drug (eg, Day -1).

[a] MMF may be substituted for MTX at investigator's discretion.

Appendix D Geographic Regions

Region		Countries	
North America	Canada	United States	
South America	Mexico	Brazil	Argentina
Western/ Northern Europe	Belgium	Sweden	Germany
	Italy	France	Netherlands
	Spain	Switzerland	United Kingdom
	Latvia	Portugal	Norway
Central Europe	Czech Republic	Hungary	Poland
	Austria	Serbia	Slovak Republic
	Greece		
Eastern Europe	Estonia	Israel	Russia
	Turkey	Romania	
Asia / Australia	Australia	Japan	Republic of Korea
	Hong Kong	Taiwan	Singapore

Appendix E Criteria for Identification of Markedly Abnormal Laboratory Values and Vital Sign Values.

Hematology—Criteria for Markedly Abnormal Values

Parameter	Low Abnormal	High Abnormal
Hemoglobin	<0.8 × LLN,	>1.2 × ULN
Hematocrit	<0.8 × LLN,	>1.2 × ULN
RBC count	<0.8 × LLN,	>1.2 × ULN
WBC count	$<2.0 \times 10^3/\mu L$	>1.5 × ULN
Platelet count	$<70 \times 10^3/\mu L$	$>600 \times 10^3/\mu L$

RBC=red blood cell, WBC=white blood cell. LLN=lower limit of normal, ULN=upper limit of normal.

Chemistry—Criteria for Markedly Abnormal Values

Parameter	Low Abnormal	High Abnormal
ALT	--	>3x ULN
AST	--	>3x ULN
GGT	--	>3x ULN
Alkaline phosphatase	--	>3x ULN
Total bilirubin	--	>2.0 mg/dL
Albumin	<2.5 g/dL	--
Total protein	<0.8x LLN	>1.2x ULN
Creatinine	--	>2.0 mg/dL
Sodium	<130 mEq/L	>150 mEq/L
Potassium	<3.0 mEq/L	>6.0 mEq/L
Bicarbonate	<8.0 mmol/L	--
Chloride	<75 mmol/L	>126 mmol/L
Calcium	<1.50 mmol/L	>3.25 mmol/L
Glucose	≤2.8 mmol/L	≥20 mmol/L
Phosphorous	<0.52 mmol/L	>2.10 mmol/L
CPK	--	>5x ULN

ALT=alanine aminotransferase, AST=aspartate aminotransferase, GGT=γ-glutamyl transferase, CPK=creatine phosphokinase, LLN=lower limit of normal, ULN=upper limit of normal.

Appendix F Criteria for Markedly Abnormal Values for Vital Signs

Parameter	Unit	Lower Criteria	Upper Criteria
Pulse	bpm	<50	>120
Systolic blood pressure	mm Hg	<85	>180
Diastolic blood pressure	mm Hg	<50	>110
Body temperature	^{o}C	<35.6	>37.7
	^{o}F	<96.1	>99.9

Appendix G Clinical Stages and Grades of GvHD

aGvHD Clinical Stage

Stage	Skin	Liver	Intestinal Tract[a]	
		Bilirubin: SI units (standard units)	Diarrhea/day	
			Subjects ≥18 yrs or <18 yrs and ≥50 kg	Subjects <18 yrs and <50 kg
1	Maculopapular rash <25% of body surface [b]	34-50 µmol/L (2-3 mg/dL)	>500 mL diarrhea/day	10-19.9 mL/kg/d
2	Maculopapular rash 25%-50% of body surface	51-102 µmol/L (3.1-6 mg/dL)	>1000 mL diarrhea/day	20-30 mL/kg/d
3	Rash >50% of body surface	103-255 µmol/L (6.1-15 mg/dL)	>1500 mL diarrhea/day	>30 mL/kg/d
4	Generalized erythroderma with bullous formation	>255 µmol/L (>15 mg/dL)	Severe abdominal pain, with or without ileus	Severe abdominal pain, with or without ileus

Source: Przepiorka et al., 1995 [45].

Abbreviations: aGvHD, acute graft-versus-host disease; GvHD, graft-versus-host disease; SI, International System of Units.

[a] Staging of intestinal tract should be assessed based on subject's age and weight at the time of assessment.

[b] Use the "Rule of Nines" or burn chart to determine the extent of the rash.

Criteria for IBMTR Severity Index for aGvHD

	Skin		Liver		Intestinal Tract
Index	**Stage (max)**		**Stage (max)**		**Stage (max)**
A	1		0		0
B	2	*or*	1-2	*or*	1-2
C	3	*or*	3	*or*	3
D	4	*or*	4	*or*	4

Source: Adapted from Rowlings et al., 1997 [46].

Abbreviations: aGvHD, acute graft-versus-host disease; IBMTR, International Bone Marrow Transplant Registry Database.

aGvHD Grade (modified Glucksberg)

Grade	Skin	Liver	Intestinal tract
I	Stage 1-2	None	None
II	Stage 3 or →	Stage 1 or →	Stage 1
III	-	Stage 2-3 or →	Stage 2-4
IV	Stage 4 or →	Stage 4	-

Source: Przepiorka et al., 1995 [45].

Abbreviations: aGvHD, acute graft-versus-host disease; ECOG PS, Eastern Cooperative Oncology Group performance status.

Criteria for MAGIC Severity Index for aGvHD

Stage	Skin (active erythema only)	Liver (bilirubin)	Upper GI	Intestinal Tract[a] (stool output/day)	
				Subjects ≥18 yrs or <18 yrs and ≥50 kg	Subjects <18 yrs and <50 kg
0	No active (erythematous) GVHD rash	<2 mg/dL	No or intermittent nausea, vomiting or anorexia	<500 mL/day or <3 episodes/day	<10 mL/kg/d or <4 episodes/day
1	Maculopapular rash <25% of body surface (a)	2-3 mg/dL	Persistent nausea, vomiting or anorexia	500-999 mL/day or 3-4 episodes/day	10-19.9 mL/kg/d or 4-6 episodes/day
2	Maculopapular rash 25%-50% of body surface	3.1-6 mg/dL	-	1000-1500 mL/day or 5-7 episodes/day	20-30 mL/kg/d or 7-10 episodes/day
3	Maculopapular rash >50% of body surface	6.1-15 mg/dL	-	>1500 mL/day or >7 episodes/day	>30 mL/kg/d or >10 episodes/day
4	Generalized erythroderma (>50% BSA) plus bullous formation and desquamation >5% body surface	>15 mg/dL	-	Severe abdominal pain with or without ileus, or grossly bloody stool (regardless of stool volume)	Severe abdominal pain with or without ileus, or grossly bloody stool (regardless of stool volume)

Source: Harris et al, 2016 [47].

Abbreviations: aGvHD, acute graft-versus-host disease; BSA, body surface area; GI, gastrointestinal; GVHD, graft-versus-host disease; MAGIC, Mount Sinai Acute GVHD International Consortium.

Overall clinical grade (based upon most severe target organ involvement):

- Grade 0: No stage 1-4 of any organ.
- Grade I: Stage 1-2 skin without liver, upper GI or lower GI involvement.
- Grade II: Stage 3 rash and/or Stage 1 liver and/or Stage 1 upper GI and/or Stage 1 lower GI.
- Grade III: Stage 2-3 liver and/or Stage 2-3 lower GI, with Stage 0-3 skin and/or Stage 0-1 upper GI.
- Grade IV: Stage 4 skin, liver or lower GI involvement, with Stage 0-1 upper GI.

[a] Staging of intestinal tract should be assessed based on subject's age and weight at the time of assessment.

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'UTC')
PPD	Biostatistics Approval	01-Nov-2021 19:16 UTC