Title: An Open-Label, Dose-Finding Study of Vedolizumab IV for Treatment of Steroid-Refractory Acute Intestinal Graft-Versus-Host Disease (GvHD) in Patients who Have Undergone Allogeneic Hematopoietic Stem Cell Transplantation (allo-HSCT)

NCT Number: NCT02993783

SAP Approve Date: September 13, 2018

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 information or company confidential information.

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: Vedo-2004

An Open-Label, Dose-Finding Study of Vedolizumab IV for Treatment of Steroid-Refractory Acute Intestinal Graft-Versus-Host Disease (GvHD) in Patients who Have Undergone Allogeneic Hematopoietic Stem Cell Transplantation (allo-HSCT)

PHASE 2a

Version**:** Final 1.0

Date**:** September 13, 2018

Prepared by:

Protected Personal Data

Based on:

Protocol Version: Amendment 2

Protocol Date: 27 November 2017

Approval Signatures

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

Study Title: An Open-Label, Dose-Finding Study of Vedolizumab IV for Treatment of Steroid-Refractory Acute Intestinal Graft-Versus-Host Disease (GvHD) in Patients who Have Undergone Allogeneic Hematopoietic Stem Cell Transplantation (allo-HSCT)

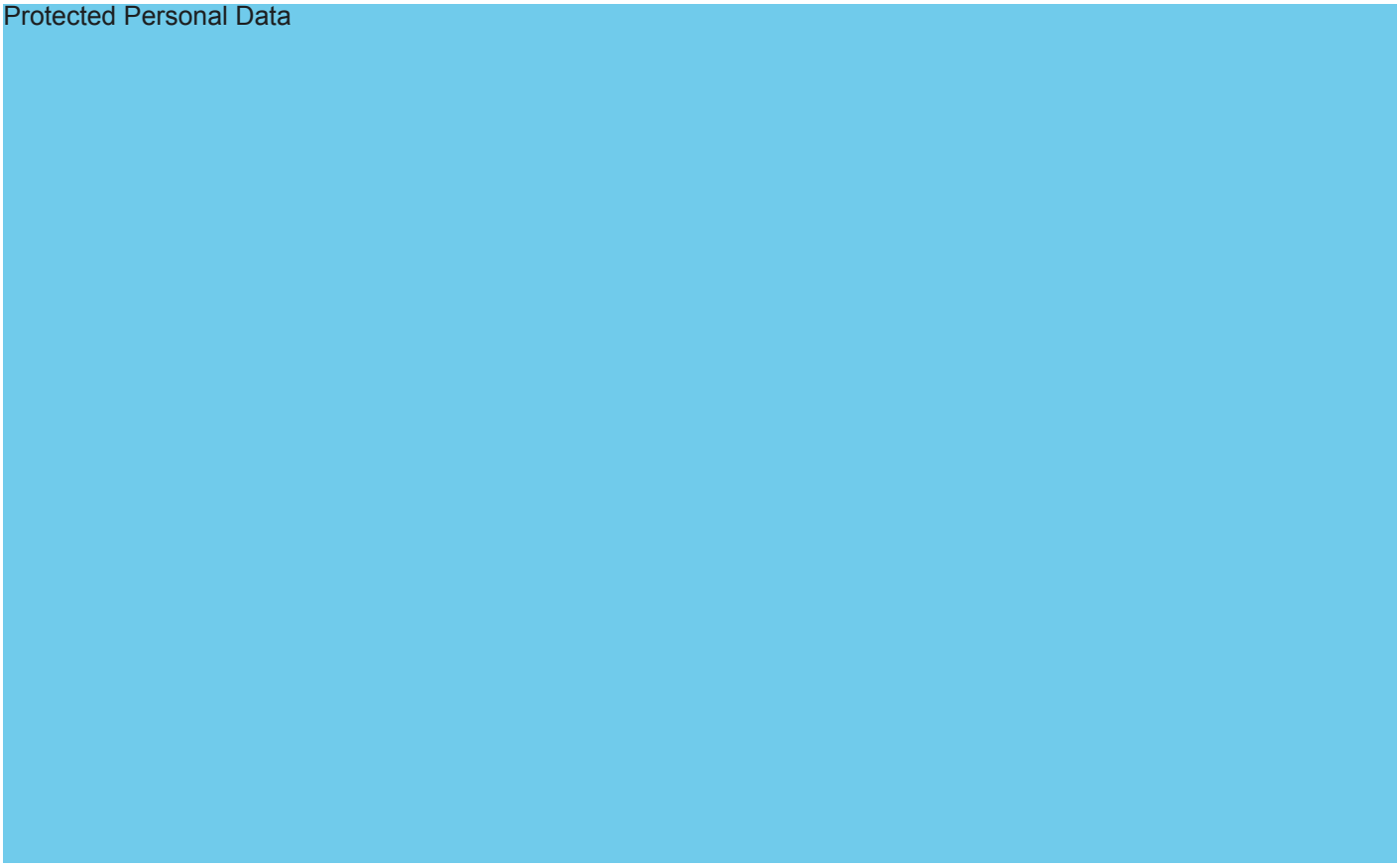


TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ANOVA	analysis of variance
AST	aspartate aminotransferase
BMI	body mass index
BUN	blood urea nitrogen
CPK	creatine phosphokinase
CRF	case report form
ECG	electrocardiogram
FAS	full analysis set
GGT	γ-glutamyl transferase
HIV	human immunodeficiency virus
ICH	International Conference on Harmonisation
IVRS	Interactive Voice Response System
LDH	lactate dehydrogenase
LLN	lower limit of normal
LOCF	last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
PD	pharmacodynamics
PK	pharmacokinetics
QOL	quality-of-life
PRO	patient-reported outcome
SAE	serious adverse event
SAP	statistical analysis plan
SDB	standard database
TLGs	tables, listings, and graphs
ULN	upper limit of normal
WHODrug	World Health Organization Drug Dictionary

1.0 OBJECTIVES

1.1 Primary Objectives

The primary objective is to describe the initial activity, tolerability, and safety and to identify a recommended dose and regimen of vedolizumab IV administered for treatment of steroid-refractory acute intestinal GvHD in patients who have undergone allo-HSCT.

1.2 Secondary Objectives

The secondary objectives are:

- To evaluate overall response to vedolizumab treatment at Day 28.
- To determine the nonrelapse mortality at 6 months after treatment with vedolizumab.
- To evaluate overall survival (OS).
- To determine the GvHD-free, relapse-free survival at 6 and 12 months after treatment with vedolizumab.
- To characterize the PK in patients treated with vedolizumab.

The additional/exploratory objectives are:

- Company Confidential Information

1.3 Additional Objectives

Not applicable.

1.4 Study Design

This is a phase 2a, open-label, dose-finding study designed to evaluate the safety, tolerability, and clinical activity of vedolizumab to treat patients who have developed acute intestinal GvHD that is refractory to primary steroid therapy. Clinical GvHD scoring will be used for assessment

of response to treatment. Patients with acute intestinal GvHD who have received no systemic therapy for the treatment of acute GvHD (prophylaxis acceptable) other than corticosteroids will be administered vedolizumab.

Patients will be randomized at a ratio of 1:1 to two treatment arms to receive either 300 mg or 600 mg of vedolizumab IV on Days 1, 15, 43, 71, and 99.

2.0 ANALYSIS ENDPOINTS

2.1 Efficacy Endpoints

2.1.1 Primary efficacy endpoint

The primary efficacy endpoint of the study is the proportion of subject with overall response (OR) at Day 28. A patient is considered to have an OR if there is either a complete response (CR), or a very good partial response (VGPR), or a partial response (PR), i.e. OR = PR + VGPR + CR. The definition of CR, VGPR, and PR are shown in the table below:

Response	Skin		Liver		Intestinal tract
CR	Skin involvement = No	and	Hepatic involvement = No	and	Intestinal tract involvement =No
VGPR	No rash, or residual erythematous rash involving <25% of the body surface without bullae.	and	Total serum bilirubin < 2 mg/dL or < 25% of baseline at enrollement (chemistry linked to eCRF)	and	Patient tolerates food or enternal feeding, and Predominantly formed stool (or fully formed), and no overt GI bleeding or abdominal cramping, and no more than occasional nausea/vomiting.
PR	Improvement of at least 1 GvHD stage since baseline	and	No progression (no worsening in stage)	and	No progression
	No progression	and	Improvement of at least 1 GvHD stage since baseline	and	No progression
	No progression	and	No progression	and	Improvement of at least 1 GvHD stage since baseline
	Improvement of at least 1 GvHD stage since baseline	and	Improvement of at least 1 GvHD stage since baseline	and	No progression
	Improvement of at least 1 GvHD stage since baseline	and	No progression	and	Improvement of at least 1 GvHD stage since baseline
	No progression	and	Improvement of at least 1 GvHD stage since baseline	and	Improvement of 1 at least GvHD stage since baseline
	Improvement of at least 1 GvHD stage since baseline	and	Improvement of at least 1 GvHD stage since baseline	and	Improvement of at least 1 GvHD stage since baseline

The stages for GvHD improvement are listed in Appendix B. Unless otherwise noted the Acute Graft-Versus-Host Disease Clinical Stage table is used (Przepiorka et al, 1995 [2]).

2.1.2 Secondary Efficacy Endpoints

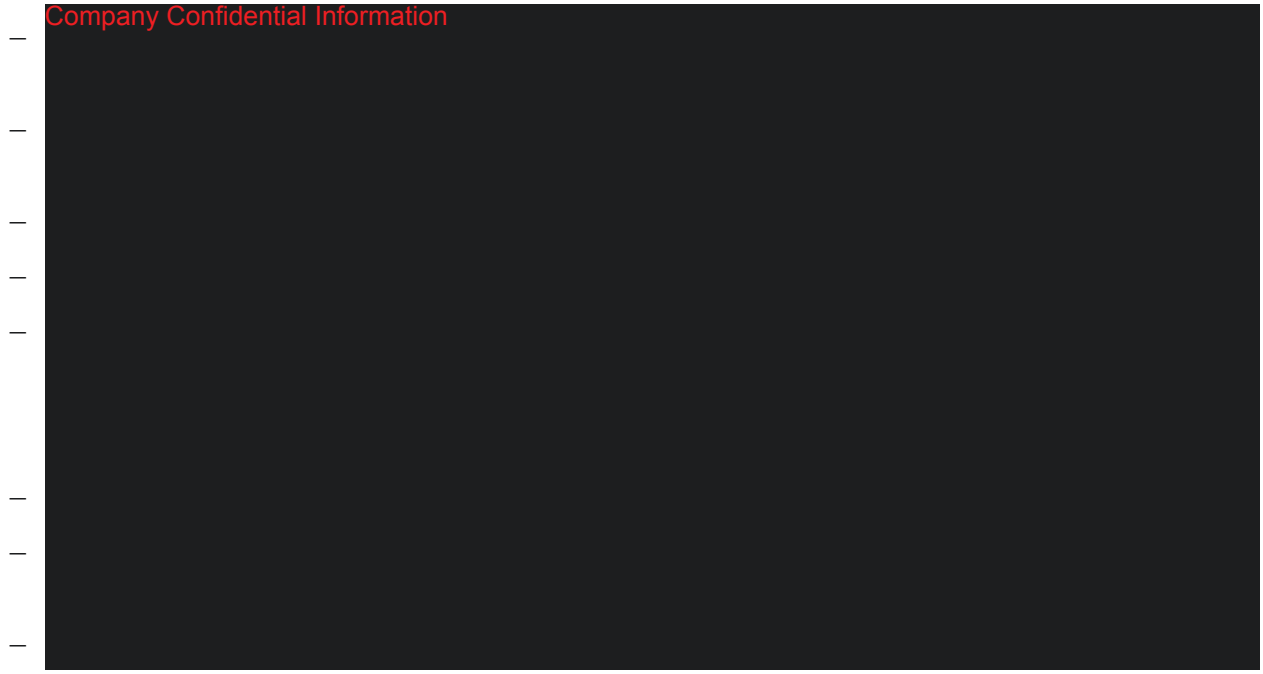
Secondary efficacy endpoints include:

- Proportion of subjects who have died in the absence of primary malignancy relapse after allo-HSCT at 6 months. The relapse of primary malignancy should be confirmed by the medical monitoring, and should be recorded in eCRF either as a reason to discontinue from study treatment, or as an adverse event.
- The proportion of subjects with acute GvHD CR at Day 28.
- The proportion of subjects with intestinal overall response at Day 28. Symptoms of acute intestinal GvHD will be measured using the BMT CTN-modified IBMTR index (See Appendix B). Intestinal Overall Response = CR, VGPR, or PR for intestine only, where
    - Intestinal CR = No intestinal tract involvement.
    - Intestinal VGPR = Patient tolerates food or enternal feeding, and Predominantly formed stool (or fully formed), and no overt GI bleeding or abdominal cramping, and no more than occasional nausea/vomiting.
    - Intestinal PR = Improvement of at least 1 GvHD stage since baseline
- Overall Survival (OS) at 6 and 12 months.
- The proportion of subjects alive without GvHD (acute or chronic) or relapse of primary malignancy at 6 and 12 months.
- The total dose of steroids administered (mg/kg/day of methylprednisolone or equivalent) from the start of the first IV infusion of vedolizumab through both 6 and 12 months.

2.1.3 Exploratory Efficacy Endpoint

In addition, the following exploratory endpoint will be assessed:

- Company Confidential Information
- [redacted]



2.2 Safety Endpoints

2.2.1 Primary Safety Endpoint

Primary safety endpoint is the number and percentage of patients who experience treatment emergent serious adverse events (TESAEs) from administration of the first dose of vedolizumab IV through Day 28.

2.2.2 Secondary Safety Endpoints

Secondary safety endpoints include the following:

- The number and percentage of patients who experience treatment-emergent adverse events (TEAEs) from administration of the first dose of vedolizumab IV through 18 weeks after administration of the last dose of vedolizumab IV.
- The number and percentage of patients who experience TESAEs from administration of the first dose of vedolizumab IV through 18 weeks after administration of the last dose of vedolizumab IV.
- Markedly abnormal values (MAV) of laboratory values and vital signs as required by Takeda standards (see Appendix D). The laboratory values and vital signs need to be confirmed by the medical monitor to determine whether there are significant changes from baseline.

2.2.3 Additional Safety Endpoints

Additional safety endpoints, adverse events of special interest (AESI) will be assessed, including hypersensitivity, infusion related reactions (occurred within a day), infection, PML, etc. See Section 4.11.1.5 and Appendix C for more details.

2.3 Pharmacokinetic Endpoints

Mean serum concentrations of vedolizumab before dosing (C_{trough}) on Day 99 will be assessed as a secondary endpoint.

2.4 Biomarker Endpoints

3.0 DETERMINATION OF SAMPLE SIZE

Approximately 38 evaluable patients were planned to be enrolled to identify an active and tolerable vedolizumab dose level with sufficient PK sampling to determine the PK parameters in this patient population. Sample size determination was mainly based on clinical considerations and on the primary objective of determining a recommended dose and regimen and to describe the initial activity, tolerability, safety, and activity of vedolizumab IV administered for the treatment of primary steroid-refractory acute intestinal GvHD.

4.0 METHODS OF ANALYSIS AND PRESENTATION

4.1 General Principles

All data will be provided in data listings sorted by treatment group and subject number. Tabular summaries will be presented by treatment group.

For any variable, unless otherwise noted, the baseline values are defined as the last non-missing measurement prior to or on the date of the first dose of study drug administration, which is defined as study Day 1.

A windowing convention will be used to determine the analysis value for a given study visit for the analyses of the data with scheduled visit days. Subjects disposition and the summary of tables for non-scheduled events (such as TEAE, concomitant medication use, etc.) will be based on the reported dates, not following the windowing convention.

Continuous data will be summarized using number of subjects with non-missing values, arithmetic mean, standard deviation, median, minimum, and maximum. Categorical data will be summarized using the number and percentage of subjects for each possible category where appropriate. The denominator for the proportion will be based on the number of subjects in each treatment group (column total) unless otherwise specified.

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.

All available safety, tolerability, efficacy, PK, and biomarker data will be included in data listings and tabulations. No imputation of values for missing data will be performed. The relevance of missing sample data will be assessed.

Data that are potentially spurious or erroneous will be examined according to standard data management operating procedures.

All statistical analyses will be conducted using SAS® version 9.2 or higher.

4.1.1 Study Definitions

The definitions of the study terms are provided in the table below:

Term	Definition
Overall Response (OR)	Defined for a patient with either a complete response (CR), or a very good partial response (VGPR), or a partial response (PR), i.e. OR = PR+VGPR+CR.
Complete Response (CR)	See Appendix A.
Very Good Partial Response (VGPR)	See Appendix A.
Partial Response (PR)	See Appendix A.
Overall Survival (OS)	Defined as the time from the date of enrollment to the date of death. Patients without documentation of death at the time of analysis will be

	censored at the date last known to be alive.
Acute GvHD	As reported in the eCRF.
Chronic GvHD	As reported in the eCRF.
Durable Overall Response	Defined as proportion of patients in each arm who achieve OR at Day 28 and maintain OR at Day 56.
Durable Intestinal Overall Response	Defined as proportion of patients in each arm who achieve intestinal OR at Day 28 and maintain OR at Day 56.
Overall Survival (OS)	Defined as the time from the date of enrollment to the date of death. Patients without documentation of death at the time of analysis will be censored at the date last known to be alive.

4.1.2 Definition of Study Days

Study Day 1 is defined as the designation date of the first intravenous [IV] infusion of vedolizumab. Other study days are defined relative to the Study Day 1 with Day 2 being the day after, and Day -1 being the day prior to Study Day 1.

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).

4.1.3 Definition of Study Visit Windows

For any variable, unless otherwise noted, the baseline values are defined as the last non-missing measurement prior to or on the date of the first dose of study drug administration, which is defined as study Day 1. All data will be categorized based on the scheduled and actual sampling time at which it was collected, allowing equally distanced windows (See Table 1) between the scheduled days when applicable. If a subject has more than one assessment visit within a window, the assessment closest to the target day will be used. In case of ties between the observations located on different sides of the target day, the later assessment will be used. In case when two or more assessments were done on the same day, the average value will be used.

Table 1. Visit windows for scheduled visit related to efficacy and safety assessment

Scheduled Visit Day	Question-naires	Assessment of GvHD	Symptom-directed physical examination	Vital signs[b]	Hematology /chemistry	Blood samples for serum biomarkers	Weight
Screening[a]	≤ -1			≤ -1	≤ -1		≤ -1
+1		+1 – +4	+1 – +4	+1 – +4	+1 – +4	≤ +1	
+2						+2 – +4	
+7		+5 – +11	+5 – +11	+5 – +11	+5 – +11	+5 – +11	+1 – +11
+15		+12 – +18	+12 – +18	+12 – +18	+12 – +18	+12 – +16	+12 – +18
+18						+17 – +23	

+22		+19 – +25	+19 – +25	+19 – +25	+19 – +25		+19 – +25
+28	+1 – +35	+26 – +32	+26 – +32	+26 – +32	+26 – +35	+24 – +32	+26 – +32
+36		+33 – +39	+33 – +39	+33 – +39		+33 – +39	+33 – +39
+43	+36 – +57	+40 – +57	+40 – +57	+40 – +57	+36 – +57	+40 – +57	+40 – +57
+71	+58 – +85	+58 – +85	+58 – +85	+58 – +85	+58 – +85	+58 – +85	+58 – +85
+99	+86 – +109	+86 – +109	+86 – +109	+86 – +109	+86 – +109	+86 – +139	+86 – +109
4 months follow up	+110 – +135	+110 – +135		+110 – +135	+110 – +135		+110 – +135
5 months follow up	+136 – +165	+136 – +165		+136 – +165	+136 – +165		+136 – +165
6 months follow up	+166 – +225	+166 – +225		+166 – +225	+166 – +315	+140 - +315	+166 – +225
9 months follow up	+226 – +315	+226 – +315		+226 – +315			+226 – +315
12 months follow up/ EOS/ET	≥ +316	≥ +316	≥ +316	≥ +316	≥ +316	≥ 316	≥ +316

(a) Unless otherwise noted, the Screening visit must occur within 28 days before the day of the first dose of the study drug (Day 1). (b) Vital signs will be obtained before and within 60 minutes of completion of IV infusion of vedolizumab.

4.1.4 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 missing or partial:

- If the day is missing, the start day will be the first day of the month
- If the month is missing, the start month will be the month corresponding to 90 days prior to the first study medication date
- If the year is missing, the start year will be the year of the entry visit (or consent date, for those missing entry visit)
- If the entire date is missing, the start date will be the date of first study drug administration

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

- If the day is missing, the stop day will be the last day of the month reported
- If the month is missing, the stop month will be the month during which the last dose of induction treatment was administered

- If the year or the entire date is missing or if the medication is "continuing", the stop year will be the year in which the last dose of induction treatment was administered.

4.1.5 Conventions for Missing Scores in FACT-BMT

Company Confidential Information

[redacted]

4.2 Analysis Sets

The analysis sets used for analysis will include the following:

- **Safety population** is defined as the population of patients evaluable for vedolizumab safety is defined as all patients who receive any amount of study drug. Safety population will be used for the safety analysis.
- **Efficacy population** is defined as all patients from the safety population who have baseline efficacy assessment and at least one post-baseline efficacy assessment. Efficacy population will be used for efficacy analysis.
- **Per-protocol population** includes patients who satisfy all of the following:
    i. Met the inclusion criteria #3 and 5.

ii. Did not meet exclusion criteria #1, 3, 4, and 5.

iii. Received at least the first two does of study medication as assigned.

Per-protocol population will be used for sensitivity analysis.

- **PK population** is defined as patients from the safety set with at least 1 postdose PK sample collected. PK population will be used for the PK parameters analysis.
- **Biomarker population:** Defined as patients from the safety set with at least 1 biomarker sample collected. Biomarker population will be used for Biomarker analysis.

4.3 Disposition of Subjects

Summary of the subject disposition will be summarized by dose level including the following:

- Study Information, including date first subject signed ICF, date of last subject's last visit/contact, date of last subject's last procedure for collection of data for primary endpoint (when applicable).
- Summary of screen failures. If no subjects failed the screening, the number of screened and the number of eligible patients should be reported.
- Summary of eligibility for randomization into phase 2a.
- Summary of study subjects randomized by site, country, dose level and regimen.
- Summary of patients for each pre-defined population: efficacy population, safety population, PP population, PK population, and biomarker population.
- Summary of subjects completing 15 days.
- Summary of subjects completing 43 days.
- Summary of subjects completing 71 days.
- Summary of subjects completing 99 days.
- Summary of subjects completing 6 months.
- Disposition of subjects based on the reasons for discontinuation of treatment and for failing to complete the study.
- Significant protocol deviations, captured on the eCRF will be summarized descriptively.

4.4 Demographic and Other Baseline Characteristics

The following demographic and baseline characteristics will be summarized by dose level using Safety Population (See Table 2).

Table 2. Summary of the Demographic and Baseline Characteristics

Characteristic	*Summarized as*	*Categories and measurement units*
Age	Continuous	18—75 years
Age group	Categorical	<35 and >= 35 years <65, and ≥65 years
Gender	Categorical	Female/Male/Missing
Race	Categorical	White/ Black or African American/ Native Hawaiian or Other Pacific Islander/ Asian/ American Indian or Alaska Native/ Other/ Not reported
Ethnicity	Categorical	Hispanic or Latino Not Hispanic or Latino Not Reported
Weight	Continuous	kg
Height	Continuous	cm
Baseline ECOG Status	Ordinal Cetegorical	0, 1, 2, 3, 4, 5
Underlying disease diagnosis	Categorical	Myeloproliferative Neoplasm, Myelodysplastic or Myelofroliferative Neoplasm, Myelodysplastic syndrome, Acute myeloid leukaemia or related precursor neoplasm, Precursor lymphoid neoplasm, Mature B-cell neoplasm, Mature T-cell and NK-cell neoplasm
Underlying disease duration since diagnosis	Ordinal Categorical	< 1 years ≥1 and < 3 years ≥ 3 and, 7 years ≥ 7 years
Conditioning Regimen	Categorical	Myeloablative Transplant, Reduced-intensity Transplant
Donor-recipient Gender Match	Categorical	Female Subject-Female Donor Female Subject-Male Donor Male Subject-Female Donor Male Subject-Male Donor
Transplant characteristics:		
Donor Relationship to Subject	Categorical	Related/Not related
HLA Compatability	Categorical	Match/Mismatch
Source of stem cells	Categorical	Bone Marrow/Peripheral Blood Stem Cell

Characteristic	*Summarized as*	*Categories and measurement units*
Cytomegalovirus IgG Antibody	Categorical	Positive/Negative
GvHD characteristics (Staging by organ involved)	Categorical	Organ: Liver, Intestinal, Skin Stage: 0, 1, 2, 3, 4
GvHD characteristics (Grade by organ involved)	Categorical	By IBMTR By modified Glucksberg

For the *categorical variables* the counts and proportions of each possible value will be tabulated by treatment group. No baseline laboratory values will be tabulated. Categories that contain zero observations will be reported as '0', not omitted. Missing and non-reported observation will be reported as such, not omitted. The denominator for proportion will be based on the number of subjects in each group (column total), unless otherwise specified.

Continuous variables will be tabulated by treatment group using the using the following:

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

No baseline comparison will be performed.

4.5 Medical History and Concurrent Medical Conditions

During the Screening period, a complete medical history will be compiled for each patient. The history will include the background of the patient's malignancy and prior therapies for it, and the background of the patient's transplant and post-transplant course.

Medical and Surgical History will be summarized by system organ class and preferred term.

4.6 Medication History and Concomitant Medications

In addition, concomitant medications used by the patient and therapeutic procedures completed by the patient will recorded in the eCRF from the first dose of study drug through 18 weeks after the last dose of study drug.

Concomitant medications will be coded using WHODrug. The number and percentage of patients taking concomitant medications will be tabulated by standardized medication name WHO drug generic term in the Safety Population, 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.

In addition, the number and percentage of subjects receiving blood transfusions will be tabulated in the Safety Population. GvHD prophylaxis at baseline may be summarized as well.

All concomitant medications administered and concomitant procedures will be provided in data listings.

4.7 Study Drug Exposure and Compliance

Vedolizumab will be administered or dispensed only to eligible patients under the supervision of the investigator or identified sub-investigator(s). The appropriate study personnel will maintain records of vedolizumab receipt and dispensing.

Exposure to vedolizumab will be summarized and reasons for discontinuation will be tabulated by dose level.

Compliance will be calculated as percentage of infusions (both completed and partial) out of the total number of infusions. "Partial infusions" is defined as an infusion with less than the scheduled amount.

Extent of Exposure for vedolizumab is calculated as the duration between the first and last dose of study drug +18 weeks, i.e.:

(Date of last dose of Vedolizumab) – (Date of first dose of Vedolizumab) + 1 + (126 days).

4.8 Efficacy Analysis

Statistical analyses will be primarily descriptive and graphical in nature. Efficacy signals will be evaluated using the initial cohorts and will be updated using the expansion cohort (if applicable) using Efficacy Population. In addition, the analysis for primary efficacy endpoint will be repeated using the PPS for sensitivity analysis.

Efficacy data will be summarized as observed. Missing data will not be imputed.

No formal statistical tests will be performed for this open-label study. The description of the statistical methods to analyze the endpoints is provided below. Additional exploratory analysis may be considered if confounding factors suspected. The proportion-based endpoints (percentage of subjects meeting endpoint criteria) will be summarized for each dosing cohort using descriptive statistics (count, percentage and associated 95% CI using normal approximation). If the number of responses is too small (i.e., ≤ 5), the exact method will be used instead (e.g., the Clopper-Pearson method).

Continuous endpoints will be descriptively summarized (number of subjects with non-missing values, arithmetic mean, median, SD, minimum, and maximum) by visit for each dosing cohort.

Time-to-event (survival) endpoints will be summarized for each dosing cohort. Subjects without documented events before reaching the end of study will be censored at the date of last assessment/visit/contact, whichever occurs last.

- 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.

4.9 Pharmacokinetic/Pharmacodynamic Analysis

The PK analysis set will be used for all PK analyses. Missing PK data will not be imputed. More details on the PK analysis will be described in a separate document, i.e., CPAP.

Measured serum concentrations of vedolizumab will be summarized using descriptive statistics by treatment group and visit/time separately. The mean serum concentration-time profile of vedolizumab will be plotted by treatment group and/or visit/time separately. Individual plot will be also presented.

The PK parameters (ex AUC, Cmax) will be summarized by treatment group and treatment period (Day -1, Day +13 and Day +42 after allo-HSCT) separately using descriptive statistics (non-missing values, mean, SD, SE, %CV, geometric mean, geometric mean %CV, median, minimum and maximum) as appropriate. Other PK parameters may be calculated if necessary. Additional details will be specified in the Clinical Pharmacology Analysis Plan (CPAP).

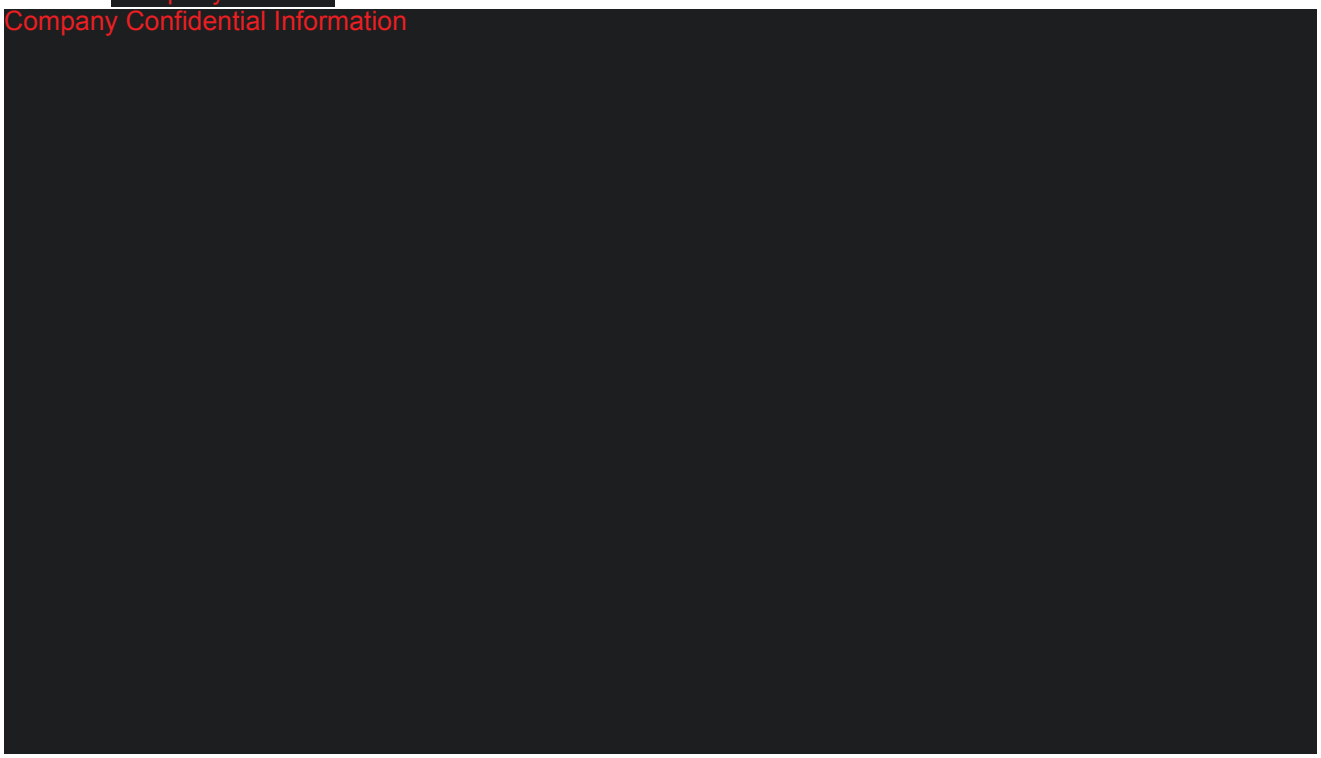
4.10 Other Outcomes

4.10.1 Biomarker Analysis

A multivariate analysis may be used as an exploratory analysis to generate hypotheses if necessary.

4.10.2 Company

Company Confidential Information



The effect AVA on P K, safety and efficacy may be evaluated as need.

4.11 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.

4.11.1 Adverse Events

Adverse events (AEs) will be coded by MedDRA (v20.0 or higher) and the type incidence, severity and relationship to study investigational product will be summarized. 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 (e.g., relatedness to study drug, action taken etc.) will be listed as appropriate.

A Treatment Emergent AE (TEAE) is defined as an AE that starts or worsens on or after Study Day 1 (defined as the day first dosed), and no more than 18 weeks/126 days after the last dose of study drug. The number of percentage of subjects with TEAEs will be summarized.

Specific adverse events will be counted once for each subject for calculating percentages. Key guidelines for determining the incidence of AEs are as follows:

- Where a patient has the same AE, based on preferred terminology, reported multiple times in a single analysis period, the patient will only be counted once at the preferred terminology level in AE tables.
- When a patient has the same AE, based on preferred terminology, reported multiple times in a single analysis phase, the following criteria, in order of precedence, will be used to select the event to be included in summary tables.
  - Relationship to study medication.
  - Intensity of event.
  - Onset date and time (where applicable).
- When reporting AEs by intensity, in addition to providing a summary table based on the event selection criteria detailed in Item 2 above, summary tables will also be provided based on the most intense event during the analysis phase – independent of relationship to study medication. For these tables, the following criteria, in order of precedence, will be used to select the event to be included in summary tables.
  - Intensity of event.
  - Onset date and time (where applicable).

4.11.1.1 Serious Adverse Events (SAE)

The number and percentage of subjects experiencing the SAEs will be summarized by system organ class (SOC), high level term (HLT), and preferred term (PT) and by treatment group. SAEs from administration of the first dose of vedolizumab IV through Day 28 and SAEs from administration of the first dose of vedolizumab IV through 18 weeks after administration of the last dose of vedolizumab IV will be summarized separately. All SAEs will be listed.

- SAEs by SOC, HLT and PT.
- SAEs by severity[*] and by SOC, HLT and PT.
- SAEs by relationship to the investigational product and by SOC, HLT, and PT.

4.11.1.2 TEAEs

The number and percentage of subjects experiencing the TEAEs will be summarized by SOC, HLT, and PT and by treatment groups.

- TEAEs by SOC, HLT and PT.
- TEAEs by severity[*] and by SOC, HLT and PT.

[*] Missing severity needs to be confirmed with the vendor. During the data cleaning proces the clinical cites need to be queried for missing severity.

- TEAEs by relationship to relationship to the investigational product and by SOC, HLT, and PT.
- The most commonly reported TEAEs (i.e., TEAEs occurring in ≥ 10% of all subjects) by PT in descending order of frequency.
- Grade 3 TEAEs.
- Grade 4 TEAEs.
- Grade 5 TEAEs.
- Grade 3 or higher drug-related TEAEs.
- All drug-related TEAEs.

4.11.1.3 Deaths

All on-study deaths recorded on the AE page, or death page (with a death date, cause of death, and outcome) of the CRF will be considered a death in the analysis.

All deaths will be listed, and summarized by PT.

4.11.1.4 TEAEs Resulting in Discontinuation of Study Drug

TEAEs resulting in permanent discontinuation of study drug will be listed and summarized by SOC, HLT and PT and by treatment group.

4.11.1.5 TEAEs of Special Interest

Based on the mechanism of action of Vedolizumab, certain adverse events of special interest have been predefined. These AEs of special interest will be summarized by SOC, HLT and PT and by treatment group.

The categories of AEs of special interest are as follows (see Appendix C for details):

- Gastrointestinal infections
- Malignancies
- Infections
- Infusion Related Reactions
- Hypersensitivity Reactions
- Suspected PML
- Liver Injury.

4.11.1.6 Other Observations Related to Safety

In addition, PML checklist data will be presented in data listings.

4.11.2 Clinical Laboratory Evaluations

Blood samples for analysis of the clinical chemistry and hematological parameters shown in Table 3 and urine samples for analysis of the parameters shown in Table 4 will be obtained as specified in the Schedule of Events as described in Study Protocol. Hepatitis and human immunodeficiency virus testing are to be performed only at the Screening visit.

Table 3. Clinical Chemistry and Hematology Tests

Hematology	Serum Chemistry	
Hematocrit	Albumin	Chloride
Hemoglobin	Alkaline phosphatase (ALP)	Gamma glutamyl transferase (GGT)
Leukocytes with differential	Alanine aminotransferase (ALT)	Glucose
Neutrophils (absolute neutrophil count [ANC])	Aspartate aminotransferase (AST)	Lactate dehydrogenase (LDH)
Platelet (count)	Bilirubin (total)	Magnesium
	Blood urea nitrogen (BUN) (a)	Phosphate
	Calcium	Potassium
	Carbon dioxide (CO_2)	Sodium
	Creatinine	Urate

(a) Blood urea (in mmol/L)=BUN(in mg/dL)×0.357.

Table 4. Clinical Urinalysis Tests

Urinalysis	
Bilirubin	pH
Glucose	Protein
Ketones	Specific gravity
Leukocytes	Turbidity and color
Nitrite	Urobilinogen
Occult blood	

The analysis of the laboratory data will include parameters tabulation for each assessment visit by dose level, 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 (hematology, chemistry, and LFTs). Shift tables for laboratory parameters will be generated based on changes in NCI CTCAE grade from baseline to the worst postbaseline value.

Elevated hepatic parameters, and subject mappings for the number and percentage of subjects with elevated hepatic parameters summarized.

4.11.3 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.

4.12 Interim Analysis

The protocol prespecified one interim analysis when all patients in the dose finding phase have reached Day 28. Takeda has decided to terminate the Vedolizumab-2004 study early. Therefore, the performance of the Bayesian interim analysis is not applicable. There will not be separate analyses for interim or final database lock.

4.13 Changes in the Statistical Analysis Plan

1. Additional efficacy endpoints that were not defined in the protocol were added in Section 2.1.3, including durable overall response rate, durable intestinal overall response rate, time to overall response, and time to intestinal overall response.
2. Additional safety endpoints AESIs were added in Section 2.2.3 and Appendix C.
3. The protocol defined Efficacy population includes all patients from the safety population who have completed their Day 28 assessment. The Efficacy population was revised to include all patients from the safety population who have baseline efficacy assessment and at least one post-baseline efficacy assessment. Refer to Section 4.2 for details.
4. Sponsor has decided to terminate the study early. There will not be separate analyses for interim or final database lock.

5.0 REFERENCES

[1] Martin PJ, Bachier CR, Klingemann HG, McCarthy PL, Szabolcs P, Uberti JP, et al. Endpoints for clinical trials testing treatment of acute graft-versus-host disease: a joint statement. Biol Blood Marrow Transplant 2009;15(7):777-84.

[2] Przepiorka D, Weisdorf D, Martin P, Klingemann HG, Beatty P, Hows J, et al. 1994 Consensus Conference on Acute GVHD Grading. Bone Marrow Transplant 1995;15(6):825-8.

[3] Rowlings PA, Przepiorka D, Klein JP, Gale RP, Passweg JR, Henslee-Downey PJ, et al. IBMTR Severity Index for grading acute graft-versus-host disease: retrospective comparison with Glucksberg grade. Br J Haematol 1997;97(4):855-64.

6.0 APPENDICES

Appendix A. Definitions for Assessing Response in Treatment of Acute GvHD

Terminology	Definition
CR	Resolution of all signs and symptoms of acute GvHD.
VGPR	Skin: No rash, or residual erythematous rash involving <25% of the body surface, without bullae (excluding residual faint erythema and hyperpigmentation). Liver: Total serum bilirubin concentration <2 mg/dL or <25% of baseline at enrollment. Gut: • Patient tolerates food or enteral feeding. • Predominantly formed stools. • No overt gastrointestinal bleeding or abdominal cramping. • No more than occasional nausea or vomiting.
PR	Improvement of 1 GvHD stage in 1 or more organs without progression in any organ.

Source: Martin PJ, et al., 2009 [1]

Appendix B. Clinical Stages of Graft-Versus-Host Disease

Acute Graft-Versus-Host Disease Clinical Stage

Stage	Skin	Liver Bilirubin: SI Units (Standard Units)	Intestinal Tract Diarrhea/Day
1	Maculopapular rash <25% of body surface (a)	34-50 µmol/L (2-3 mg/dL)	>500 mL diarrhea/day
2	Maculopapular rash 25%-50% of body surface	51-102 µmol/L (3.1-6 mg/dL)	>1000 mL diarrhea/day
3	Rash >50% of body surface	103-225 µmol/L (6.1-15 mg/dL)	>1500 mL diarrhea/day
4	Generalized erythroderma with bullous formation	>255 µmol/L (>15 mg/dL)	Severe abdominal pain, with or without ileus

Source: Przepiorka et al, 1995 [2]
SI=International System of Units (Système Internationale d'Unités).
(a) Use the "Rule of Nines" or burn chart to determine the extent of the rash.

Criteria for IBMTR Severity Index for Acute Graft-Versus-Host Disease

	Skin			Liver			Intestinal Tract	
Index	Stage (max)	Extent of Rash		Stage (max)	Total Bilirubin (µmol/L)		Stage (max)	Volume of Diarrhea (mL/day)
A	1	<25%		0	<34		0	<500
B	2	25-50%	*or*	1-2	34-102	*or*	1-2	550-1500
C	3	>50%	*or*	3	103-255	*or*	3	>1500
D	4	Bullae	*or*	4	>255	*or*	4	Severe pain and ileus

From Rowlings et al, 1997 [3].
IBMTR=International Bone Marrow Transplant Registry Database.

Appendix C

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
Gastrointestinal Infections	The infection reports retrieved from the clinical database were evaluated for GI infections during the reporting period using the MedDRA 20.0 search criteria of: - Abdominal and gastrointestinal infections HLT in the Infections and Infestations SOC. - Gastrointestinal infections HLT of the Gastrointestinal SOC.
Malignancies	SOC: NEOPLASMS BENIGN, MALIGNANT AND UNSPECIFIED (INCL CYSTS AND POLYPS)
Infections	The infection reports retrieved from the clinical database, not classified as GI infections, or suspected PML will be searched using the MedDRA 20.0 search criteria of: Infections and Infestations SOC The following reports will then be excluded: Abdominal and gastrointestinal infections HLT. URTI HLT. Bronchitis PT. Influenza PT. Human polyomavirus infection PT. JC virus infection PT. JC virus test positive PT. Leukoencephalopathy PT. Polyomavirus test positive PT. Progressive multifocal leukoencephalopathy PT.
Infusion Related Reactions	Analysis for these AEs will occur on two levels: • Investigator defined Infusion Related Reactions (as indicated on the AE CRF). • All AEs that occur on or one calendar day after the infusion date.
Hypersensitivity Reactions	Anaphylactic/anaphylactoid shock conditions SMQ (broad). Angioedema SMQ (broad). Hypersensitivity SMQ (broad).
Suspected PML	The protocol incorporates an active screening program in order to identify and manage any case of PML. This program is known as the Risk Assessment and Minimization for PML (RAMP). The clinical database will be searched for suspected PML reports received within the Infection and Infestation SOC using the MedDRA 20.0 search criteria of: • Human polyomavirus infection PT.

Events	MedDRA Terms or definitions
	• JC virus infection PT. • JC virus test positive PT. • Leukoencephalopathy PT. • Polyomavirus test positive PT. • Progressive multifocal leukoencephalopathy PT.
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)

Appendix D
Criteria for Identification of Markedly Abnormal Laboratory Values
Hematology—Criteria for Markedly Abnormal Values

Parameter	Low Abnormal	High Abnormal
Red blood cells	<0.8 x LLN	>1.2 x ULN
White blood cells	<0.5 x LLN	>1.5 x ULN
Neutrophils	<0.5 x LLN	>1.5 x ULN
Eosinophils	--	>2 x ULN
Basophils	--	>3 x ULN
Lymphocytes	<0.5 x LLN	>1.5 x ULN
Monocytes	--	>2 x ULN
Hemoglobin	<0.8 x LLN	>1.2 x ULN
Hematocrit	<0.8 x LLN	>1.2 x ULN
Platelets (conventional)	$<75 \times 10^3/\mu L$	$>600 \times 10^3/\mu L$
Platelets (SI)	$<75 \times 10^9/L$	$>600 \times 10^9/L$
HbA1c (conventional)	--	>7%
HbA1c (SI)	--	>0.07

Criteria for Abnormal Changes from Baseline of Vital Signs

Vital Sign	Criterion Value	Change Relative to Baseline
Systolic arterial blood pressure	>180 mm Hg	Increase of ≥20 mm Hg
	<85 mm Hg	Decrease of ≥20 mm Hg
Diastolic arterial blood pressure	>110 mm Hg	Increase of ≥15 mm Hg
	<50 mm Hg	Decrease of ≥15 mm Hg
Pulse	>120 bpm	
	<50 bpm	
Body Temperature	>37.7 degrees Celsius	
	<35.6 degrees Celsius	
Weight		Change of greater than 7% body weight

Both the criterion value and the change from Baseline must be met.

Urinalysis – Criteria for Markedly Abnormal Values

Parameter	Low Abnormal	High Abnormal
pH	N/A	N/A
Specific Gravity	N/A	N/A
Protein	N/A	N/A
Glucose	N/A	N/A
Blood	N/A	N/A
Nitrite	N/A	N/A

Chemistry—Criteria for Markedly Abnormal Values

Parameter	Low Abnormal	High Abnormal
ALT	--	>3 x ULN
Alkaline phosphatase	--	>3 x ULN
AST	--	>3 x ULN
Total bilirubin (conventional)	--	>2.0 mg/dL
Total bilirubin (SI)	--	>34.2 μmol/L
Direct bilirubin	--	>2 x ULN
GGT	--	>3 x ULN
Total protein	<0.8 x LLN	>1.2 x ULN
Albumin (conventional)	<2.5 g/dL	--
Albumin (SI)	<25 g/L	--
Creatinine (conventional)	--	>2 mg/dL
Creatinine (SI)	--	>177 μmol/L
Blood urea nitrogen (conventional)	--	>30 mg/dL
Blood urea nitrogen (SI)	--	>10.7 mmol/L
Potassium (conventional)	<3.0 mEq/L	>6.0 mEq/L
Potassium (SI)	<3.0 mmol/L	>6.0 mmol/L
Sodium (conventional)	<130 mEq/L	>150 mEq/L
Sodium (SI)	<130 mmol/L	>150 mmol/L
Glucose (conventional)	<50 mg/dL	>350 mg/dL
Glucose (SI)	<2.8 mmol/L	>19.4 mmol/L
Calcium (conventional)	<7.0 mg/dL	>11.5 mg/dL
Calcium (SI)	<1.75 mmol/L	>2.88 mmol/L
Thyroid Stimulating Hormone (TSH)	<0.8 x LLN	>2.0 x ULN
Vitamin B12 (conventional)	<125 pg/mL	--
Vitamin B12 (SI)	<92 pmol/L	--
Folate (conventional)	<2.2 pg/dL	>17.5 pg/dL
Folate (SI)	<5.0 nmol/L	>39.7 nmol/L
Parathyroid hormone (PTH)	<0.8 x LLN	>2.0 x ULN
Free thyroxine (Free T4)	<0.8 x LLN	>2.0 x ULN
Rapid plasma reagin (RPR)	positive	positive
Bicarbonate	<8.0 mmol/L	--
Chloride	<75 mmol/L	>126 mmol/L
Phosphorous	<0.52 mmol/L	>2.10 mmol/L
CPK	--	>5x ULN

ALT=alanine aminotransferase, AST=aspartate aminotransferase, GGT=γ-glutamyl transferase, LLN=lower limit of normal, N/A=not applicable, ULN=upper limit of normal

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'UTC')
Protected Personal Data	Biostatistics Approval	17-Sep-2018 15:54 UTC
	Statistical Approval	17-Sep-2018 16:02 UTC
	Clinical Science Approval	17-Sep-2018 16:07 UTC
	Clinical Pharmacology Approval	17-Sep-2018 17:55 UTC
	Pharmacovigilance Approval	17-Sep-2018 17:59 UTC