Statistical Analysis Plan

Protocol Title:	A Phase 1b/2 Open-label Study Evaluating the Safety, Tolerability, Pharmacokinetics, Pharmacodynamics, and Efficacy of efavaleukin alfa in Adult Subjects With Steroid Refractory Chronic Graft Versus Host Disease
Short Protocol Title:	Safety and Efficacy of efavaleukin alfa in Subjects With Steroid Refractory Chronic Graft versus Host Disease
Protocol Number:	20160283
NCT Number:	NCT03422627
Authors:	[redacted]
Sponsor:	Amgen Inc. One Amgen Center Drive, Thousand Oaks, CA, 91320, USA
SAP Date:	**Document Version** — **Date** Original — 14 December 2017 Amendment 1 (v1.0) — 17 August 2020 Amendment 2 (v2.0) — 12 September 2022 **Amendment 3 (v3.0) — 31 January 2023**

Version Number	**Date**	**Summary of Changes, including rationale for changes**
Original	14 December 2017	NA
[Amendment 1 (v1.0)]	17 August 2020	Changes related to protocol amendment 2 (4 December 2018) and protocol amendment 3 (26 August 2019)
[Amendment 2 (v2.0)]	12 September 2022	The statistical analysis plan for this study was amended to reflect the changes in protocol amendment 4 (8 December 2020) and 5 (22 June 2021), and the changes due to the decision to terminate the study early without conducting phase 2. Editorial changes and minor clarifications were also implemented throughout the document. <u>Description of Global Changes</u> • Removed all the text and analysis related to Phase 2 part. • Removed interim analysis section for Phase 1b. • Added language on the additional extended treatment period beyond week 104 in Study Design section. • Updated the Study Design section and Sample Size section to include Cohort 5. • Updated Important Protocol Deviations section to include summaries for COVID-19 IPDs. • Events of interest and AEs identified by COVID-19 MedDRA queries are added to Adverse events and Disease-related Events section.
[Amendment 3 (v3.0)]	**31 January 2023**	**The statistical analysis plan was amended to reflect the following changes:** **<u>Description of Global Changes</u>**

		• **Added the decision made in Ad hoc DLRM** • **Added a few definitions like actual treatment period, baseline medication, planned treatment period, dose reduction period, anti-drug antibody analysis set** • **Added immunosuppressant medication in baseline characteristics** • **Added additional summaries for TEAE by antibody status, additional EOIs and systemic corticosteroid dose reduction**

Table of Contents

List of Table

List of Figures

List of Abbreviations and Definition of Terms

Abbreviation or Term	Definition/Explanation
AE	adverse event
ALT	alanine transaminase
ANC	absolute neutrophil count
ARDS	acute respiratory distress syndrome
AUC	area under the concentration-time curve
AUC_{last}	area under the concentration from time 0 to the time of the last quantifiable concentration
AUC_{tau}	area under the concentration-time curve over a dosing interval
BLRM	bayesian logistic regression model
BSA	body surface area
CD	cluster of differentiation
cGVHD	chronic graft versus host disease
C_{max}	maximum observed serum concentration
COVID-19	coronavirus disease
CPMS	clinical pharmacology, modeling and simulation
[redacted]	
CRF	case report form
CTCAE	common terminology criteria for adverse events
CV	co-efficient of variation
DILI	drug-induced liver injury
DLRM	dose level review meeting
DLT	dose limiting toxicity
DMP	data management plan
DOR	duration of response
[redacted]	
ECG	electrocardiogram
eCRF	electronic case report form
eGFR	estimated glomerular filtration rate
EOT	end of treatment
EOET	end of extended treatment
EOS	end of study
FEV1	forced expiratory volume in 1 second
[redacted]	

Abbreviation or Term	Definition/Explanation
Foxp3	forkhead box P3
GI	gastrointestinal
GSO-DM	global study operations-data management
GVHD	graft versus host disease
ICH	international council for harmonisation
IL-2	interleukin 2
Interactive Voice Response System (IVRS)	telecommunication technology that is linked to a central computer in real time as an interface to collect and process information
Interactive Web Response System (IWRS)	web based technology that is linked to a central computer in real time as an interface to collect and process information
IP	investigational product
IPD	important protocol deviations
KM	kaplan-meier
[redacted]	
MAD	multiple ascending dose
MedDRA	medical dictionary for regulatory activities
NI	no involvement
NIH	national Institutes of health
NK cells	natural killer cells
NKT cells	natural killer T cells
NRI	non-responder imputation
[redacted]	
NS	not significant
[redacted]	
OMRS	total score for all mucosal changes
PD	pharmacodynamics or progressive disease
PFT	pulmonary function test
PI	principal investigator
PK	pharmacokinetic
PKDM	pharmacokinetics and drug metabolism
[redacted]	
P-ROMQW	patient related outcome measures
QW	every week
Q2W	every 2 weeks

Abbreviation or Term	Definition/Explanation
RBC	red blood cell
[redacted]	
SAP	statistical analysis plan
SC	subcutaneous
[redacted]	
SFU	safety follow up
SOC	standard of care
Tcon	conventional T cells
T_{max}	time of maximum observed concentration
TPI	toxicity probability interval
Teff	T effector cells
Treg	regulatory T cells
ULN	upper limit of normal
USA	united states of america

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol for study 20160283, efavaleukin alfa dated 22 June 2021. The scope of this plan includes the analysis that is planned for phase 1b part and will be executed by the Global Biostatistical Sciences (GBS) department unless otherwise specified. Analyses described in the protocol for the phase 2 part will not be conducted as the study is being terminated early.

2. Objectives, Endpoints and Hypotheses

2.1 Objectives and Endpoints

Objectives	Endpoints
Primary	
• To evaluate the safety and tolerability of multiple ascending doses of efavaleukin alfa in subjects with steroid refractory chronic graft versus host disease (cGVHD)	• Incidence of dose limiting toxicities (DLTs) at first 4 weeks • Incidence of all treatment-related and treatment-emergent adverse events and serious adverse events
Secondary	
[redacted]	
Exploratory	
[redacted]	

Objectives	Endpoints
[redacted]	

2.2 Hypotheses and/or Estimations

Clinical hypothesis: efavaleukin alfa will be safe and well-tolerated in subjects with steroid refractory cGVHD.

There is no statistical hypothesis planned for the phase 1b part of the study.

3. Study Overview

3.1 Study Design

This is an open-label, multi-center phase 1b/2 study to evaluate the safety and efficacy of efavaleukin alfa in subjects with steroid refractory cGVHD.

The phase 1b part of the study will be conducted as a multiple ascending dose (MAD) study. Each dosing cohort will consist of between 3 and 6 DLT-evaluable subjects who will receive efavaleukin alfa SC either every week (QW) or every 2 weeks (Q2W) plus protocol permitted background therapy for 52 weeks. The DLT evaluation period is defined as 4 weeks after the first dose of study IP. At the discretion of the Sponsor, following discussion and agreement between the principal investigator and medical monitor, subjects responding to efavaleukin alfa (as assessed by the end of week 50), who wish to continue treatment, may continue to receive efavaleukin alfa treatment at their current dosing regimen for up to an additional 52 weeks. The remaining subjects will complete the week 52/EOT visit. All subjects who continue to receive efavaleukin alfa during the extended dosing period will be re-evaluated at week 76 for their response to treatment. Following discussion and agreement between the principal investigator and medical monitor, the Sponsor may decide to allow these subjects to continue treatment

through week 102 (Q2W dose) or week 103 (QW dose). Subjects participating in extended dosing will complete the week 104/end of extended treatment (EOET) visit. At the discretion of the Sponsor, following discussion and agreement between the principal investigator and medical monitor, subjects responding to efavaleukin alfa (as assessed by the end of week 104), who wish to continue treatment beyond week 104, may continue to receive efavaleukin alfa treatment at their current dosing regimen for up to an additional 156 weeks. Only subjects who are deemed eligible to continue extended dosing beyond week 102 or week 103 will complete dosing at week 104. The remaining subjects will complete the week 104/EOET visit. All subjects who continue to receive efavaleukin alfa during the extended dosing period will be re-evaluated every 6 months for their response to treatment. Following discussion and agreement between the principal investigator and medical monitor, the Sponsor may decide to allow these subjects to continue treatment through week 258 (Q2W dose) or week 259 (QW dose). Subjects participating in extended dosing through week 258 or 259 will complete the week 260/EOET visit. All subjects will complete the 6-week safety follow-up after the last dose of efavaleukin alfa.

Five dose levels are planned: [redacted] μg every 2 weeks (Q2W) (cohorts 1a), [redacted] μg every week (QW) (cohorts 2a), [redacted] μg Q2W (cohort 3), [redacted] μg QW (cohort 4), and [redacted] μg Q2W (cohort 5). Dose levels 1a ([redacted] μg Q2W) and 2a ([redacted] μg QW) will be enrolled concurrently. Three subjects each will be assigned to cohorts 1a ([redacted] μg Q2W) and 2a ([redacted] μg QW) alternately (total of 6 subjects). After the last subject completes the DLT evaluation period, a dose level review meeting (DLRM) will occur and if deemed necessary, 3 additional subjects each will be assigned to cohorts 1a and 2a alternately (total of 6 subjects) to gain additional information. After the last subject completes the DLT evaluation period, a DLRM will occur and if dose escalation is deemed appropriate, subjects will be enrolled in cohorts 3 ([redacted] μg Q2W) 4 ([redacted] μg QW), and 5 ([redacted] μg Q2W) as follows: first, 3 subjects will be enrolled in cohorts 3. An internal safety review will be conducted by the Amgen Medical Monitor and Global Safety Officer after the first 3 subjects complete the DLT evaluation period. If concerning safety issues are identified, a DLRM will occur to assess if it is safe to proceed with additional enrollment in cohort 3 and concurrent enrollment in cohort 4. If no concerning safety issues are identified in these first 3 subjects, enrollment of the remaining 3 subjects in cohort 3 will proceed without a DLRM. Concurrently, enrollment of 6 subjects in cohort 4 will begin with alternate assignment of subjects between cohorts 3 and 4 until enrollment of cohort 3

has completed. Enrollment of 6 subjects in cohort 5 will begin after the DLRM for cohort 4 is complete. DLRMs for cohorts 3, 4 and 5 will occur after the last subject in that cohort completes the DLT evaluation period.

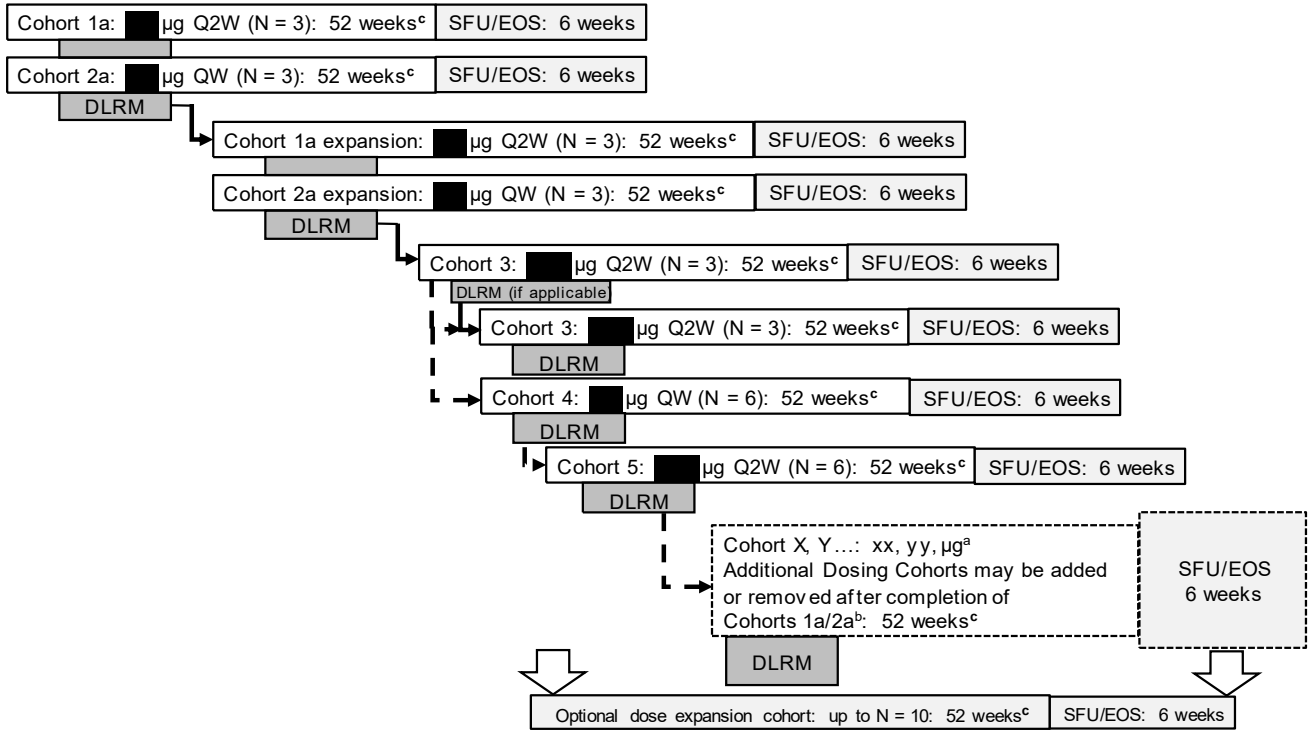
All planned dose levels for a given cohort may be adjusted at any time based on emerging data. Additional dosing cohorts may be added, removed or substituted at any time based on emerging data and the results of continuous modeling**. As a result of a recommendation from an ad hoc DLRM, Amgen discontinued enrollment to cohort 5 at the planned [redacted] mcg Q2W dose. Subjects who were on-going at the time of this decision had their dose reduced to [redacted] μg Q2W for their remaining visits.**

After there is at least 1 DLT observed at any dose level, a Bayesian logistic regression model (BLRM) (Bailey et al, 2009; Neuenschwander et al, 2008) will be used to inform dose escalation and the results of this analysis will be provided to the DLRM. Three toxicity probability intervals (TPI) of DLT will be defined: target TPI (20% to 30%), excessive TPI (30% to 60%) and unacceptable TPI (60% to 100%). Adverse events meeting DLT criteria after week 4 and PD information may also be included in the model, as available. The model will recommend an MTD as the dose with highest probability in the target TPI, but with less than 25% probability in excessive TPI and unacceptable TPI. Lower dose levels, intermediate dose levels or alternative dosing schedules may be considered based on all available information as long as they do not exceed the estimated MTD per the model. In addition, Amgen may add subjects to dose levels below the MTD in order to better characterize PK and PD. After DLT is observed, the BLRM model will be run after each dosing cohort and may be run at any time to continuously update the MTD information. For technical details of BLRM, please refer to Appendix C**.**

The study has been terminated early and therefore, the optional dose expansion cohort of phase 1b and the phase 2 part of the study is canceled.

The overall study design is described by a study schema in Figure 1. The endpoints are defined in Section 2.1.

Figure 1. Study Schema



DLRM = dose level review meeting; EOET = end of extended treatment; PD = pharmacodynamics; PK = pharmacokinetics; TPI = toxicity probability interval; SFU = safety follow-up; EOS = end of study; Q2W = every 2 weeks; QW = every week

[a] Doses for subsequent cohorts will be determined by results of Bayesian TPI and PK/PD modeling of previous cohorts.

[b] A DLRM is required for escalation to higher doses but not for allocation of additional subjects to doses equivalent to or lower than previously given that have been deemed tolerable.

[c] At the discretion of the Sponsor, following discussion and agreement between the principal investigator and medical monitor, subjects responding to efavaleukin alfa (as assessed by the end of week 50), who wish to continue treatment, may continue to receive efavaleukin alfa treatment at their current dosing regimen for up to an additional 52 weeks. The remaining subjects will complete the week 52/EOT visit. All subjects who continue to receive efavaleukin alfa during the extended dosing period will be reevaluated at week

76 for their response to treatment. Following discussion and agreement between the principal investigator and medical monitor, the Sponsor may decide to allow these subjects to continue treatment through week 102 (Q2W dose) or week 103 (QW dose). Subjects participating in extended dosing through week 104 will complete the week 104/EOET visit. At the discretion of the Sponsor, following discussion and agreement between the principal investigator and medical monitor, subjects responding to efavaleukin alfa (as assessed by the end of week 104), who wish to continue treatment beyond week 104, may continue to receive efavaleukin alfa treatment at their current dosing regimen for up to an additional 156 weeks. Only subjects who are deemed eligible to continue extended dosing beyond week 102 or week 103 will complete dosing at week 104. The remaining subjects will complete the week 104/EOET visit. All subjects who continue to receive efavaleukin alfa during the extended dosing period will be reevaluated every 6 months for their response to treatment. Following discussion and agreement between the principal investigator and medical monitor, the Sponsor may decide to allow these subjects to continue treatment through week 258 (Q2W dose) or week 259 (QW dose). Subjects participating in extended dosing through week 258 or 259 will complete the week 260/EOET visit. All subjects will complete the 6-week safety follow-up after the last dose of efavaleukin alfa.

3.2 Sample Size

Approximately 40 subjects will be enrolled in phase 1b of the study, with about 30 subjects in the dose escalation cohorts and approximately 10 subjects in the optional dose expansion cohort. The total sample size may be higher than 40 subjects if, following a DLRM recommendation or Amgen decision to evaluate additional doses, dosing cohorts are added and/or existing cohorts are expanded; or subjects are replaced. Additional subjects may be enrolled in each cohort to enable all screened eligible subjects to participate in the study.

With 40 subjects receiving efavaleukin alfa, there is an 87% chance of detecting an adverse event with a true incidence rate of 5%. Without the optional dose expansion cohort (ie, with 30 subjects), the chance of detecting an adverse event with a true incidence rate of 5% reduces to 79%.

3.2.1 Replacement of Subjects

Subjects may be replaced if they receive fewer than 4 doses of efavaleukin alfa prior to discontinuing study participation.

3.2.2 Number of Sites

Approximately 14 sites in North America, Europe, and Asia will participate in the study. Sites that do not enroll subjects within approximately 3 months of site initiation may be closed.

3.3 Adaptive Design

The adaptive design elements implemented for this study is Bayesian logistic regression model to assist dose escalation.

4. Covariates and Subgroups

4.1 Planned Covariates

No covariates will be used for this study.

4.2 Subgroups

No subgroup analyses will be conducted for this study.

5. Definitions

5.1 Basic Definitions

Investigational Product (IP)

The term ‘investigational product’ is used in reference to efavaleukin alfa.

Actual Treatment Received

Actual treatment received is the IP the subject actually received regardless of the dose specified in the protocol for the cohort. In cases where a subject received multiple different efavaleukin alfa doses, the actual treatment received will be based on the highest efavaleukin alfa dose received.

Study Points of Reference

Baseline

For any variable, baseline is the last assessment taken prior to or on the first investigational product administration unless stated otherwise; for subjects who did not receive any investigational product, baseline is the last assessment on or before the enrollment date.

Study Day 1

Study day 1 is defined as the first day of administration of the investigational product or the day of enrollment if the subject does not receive any investigational product. The day prior to Study Day 1 is considered as Day – 1.

Study Day

Study day is defined as the number of days from Study Day 1.

Post study day 1: study day= (date of interest – date of Study Day 1) + 1

Pre study day 1: study day= (date of interest – date of Study Day 1)

5.2 Study Dates

Informed Consent Date

The date on which the subject signs the informed consent form.

Last Dose Date

The date on which a subject is administered the last dose of IP.

Subject Level End-of-Study (EOS) Date

End of study for each subject is defined as the date the subject last completed a protocol-specified procedure. The date will be recorded on the End of Study eCRF.

Progressive Disease Visit

Progressive disease visit should be performed any time during the study if the investigator suspects progressive disease and subject is experiencing worsening of cGVHD that requires the addition of new immunosuppressive medication(s) or an increase in systemic corticosteroid dose above baseline after week 4 of the study.

End of Treatment Visit

Subjects discontinuing treatment prior to week 52 for any reason (see Section 8.1 in protocol) will be asked to complete an End of Treatment visit within a window of ± 7 days of the next regularly scheduled visit after the last dose of efavaleukin alfa, consisting of all assessments included in the Week 52 visit. If the End of Treatment visit for a subject occurs < 7 days after a progressive disease visit, then the procedures completed at the End of Treatment visit should only be those NOT performed at the progressive disease visit.

Subjects participating in the extended dosing, subjects discontinuing treatment prior to week 102 (Q2W dosing) or week 103 (QW dosing) will be asked to complete an End of Extended Treatment visit within a window of ± 7 days of the next regularly scheduled visit after the last dose of efavaleukin alfa, consisting of all assessments included in the Week 104 visit.

Subjects participating in the extended dosing, subjects discontinuing treatment after week 104 but prior to week 258 (Q2W dosing) or week 259 (QW dosing) will be asked to complete an End of Extended Treatment visit within a window of ± 7 days of the next regularly scheduled visit after the last dose of efavaleukin alfa, consisting of all assessments included in the Week 260 visit.

Safety Follow-up/End of Study Visit

Upon completion of the treatment period or permanent discontinuation from the study treatment for any reason, a safety follow-up visit will be performed approximately 6 weeks (± 3 days) after the last dose of efavaleukin alfa. Women of childbearing potential will be required to complete a pregnancy test at this visit.

Study Visit

Since the actual visit for a subject may not exactly coincide with their scheduled visit date, the actual visit date is mapped to the analysis visit as per the analysis visit windows described in Appendix G.

5.3 Subject Disposition

Exposed to IP

Subjects are defined as exposed if they receive at least 1 dose of IP.

Opted for extended dosing period

Subjects are defined as opted for extended dosing period if subject has marked Yes for extended dosing on "**Extended Dosing Schedule**" eCRF.

Enrolled

A subject is considered enrolled when the investigator decides that the subject has met all eligibility criteria and has been enrolled via the Interactive Voice Response System/Interactive Web Response System (IVRS/IWRS). See protocol Section 6.4 for further details. The subject will be marked as enrolled on Enrollment eCRF and enrollment date will be included.

On-Study

Subjects are considered on-study when they have enrolled but have not ended the study.

Planned Treatment Period

The time period from study day 1 until end of study or, for cohort 5 subjects, until the day before the first dose of IP at the reduced dose ([redacted] mcg), whichever occurred first.

Dose Reduction Period

For cohort 5 subjects, the time period from the first dose of IP at the reduced dose ([redacted] mcg) until end of study, if applicable.

Completed IP

Subjects are defined as completing IP if the reason for ending IP on End of IP administration eCRF is marked as Completed. Based on whether subjects opted for extended dosing, these subjects will be stratified for summary purposes into those who 1) completed IP up to week 52 and did not opt for extended dosing, and 2) completed IP through extended dosing period.

Completed Study

Subjects are defined as completing study if the reason for ending the study on EOS eCRF is marked as Completed. Based on whether subjects opted for extended dosing, these subjects will be further stratified for summary purposes into those who 1) completed the study up to week 52 and did not opt for extended dosing, and 2) completed the study through extended dosing period.

5.4 Arithmetic Calculations

Change from baseline

Change from baseline is the arithmetic difference between post-baseline and baseline. i.e. post-baseline minus baseline.

Percent change from baseline

Percent change from baseline is the change from baseline divided by the baseline value times 100. If the change from baseline is not equal to 0 and the baseline value is 0 then percent change is not defined. If the change from baseline is equal to 0 and the baseline value is also 0 then percent change is 0.

Fold change from baseline

Fold change from Baseline equals the post-Baseline value divided by the Baseline value. If the change from baseline is not equal to 0 and the baseline value is 0 then fold change is not defined. If the change from baseline is equal to 0 and the baseline value is also 0 then fold change is 1.

Dose Administered

The dose administered (μg) will be calculated as follows using quantity administered (mL), amount of IP used in preparation (mL) and total volume of preparation (mL) as reported in the clinical database and packaged drug concentration (mg/mL) as reported in manufacturing lot file:

- If dilution required,

  Dose Administered = Quantity administered x packaged drug concentration x (amount of IP used / total volume of preparation) x 1000;

- If dilution not required,

  Dose Administered = Quantity administered x packaged drug concentration x 1000

If the dilution of IP is not required, the amount of IP used in the preparation of the dose and the total volume of preparation are not included in the calculation.

Number of days on IP

Number of days on IP is defined as last dose date – first dose date + 1

Estimated glomerular filtration rate (eGFR)

Estimated glomerular filtration rate at any visit will be calculated as follows using serum creatinine (umol/L) collected as part of chemistry labs at that visit and demographic information collected at screening:

For Japanese subjects:

eGFR (mL/min/1.73 m²) = 194 x [(Creatinine/88.4)$^{-1.094}$] * (Age$^{-0.287}$) * 0.739 (if female)

For non-Japanese subjects:

eGFR (mL/min/1.73 m²) = 175 x [(Creatinine/88.4)$^{-1.154}$] * (Age$^{-0.203}$) * 1.212 (if black) * 0.742 (if female)

[redacted]

5.5 Definitions

AUC_{last}

Area under the concentration-time curve from time 0 (time of investigational product administration) to the time of the last quantifiable concentration.

AUC_{tau}

Area under the concentration-time curve after the first and last doses.

C_{max}

Maximum observed serum concentration.

Dose Limiting Toxicity (DLT)

The DLT evaluation period for each subject is 4 weeks from the first dose of efavaleukin alfa. To be evaluable for a DLT, subjects must have received at least 2 doses of efavaleukin alfa, or have experienced a DLT within the DLT evaluation period. All subjects who experience a DLT will discontinue study therapy and complete safety follow up. Dose limiting toxicity events will be used to guide dose escalation decisions and to determine the maximum tolerated dose as described in Section 3.1.

These events will be marked as DLT on Events eCRF.

The following DLT criteria apply to the study:

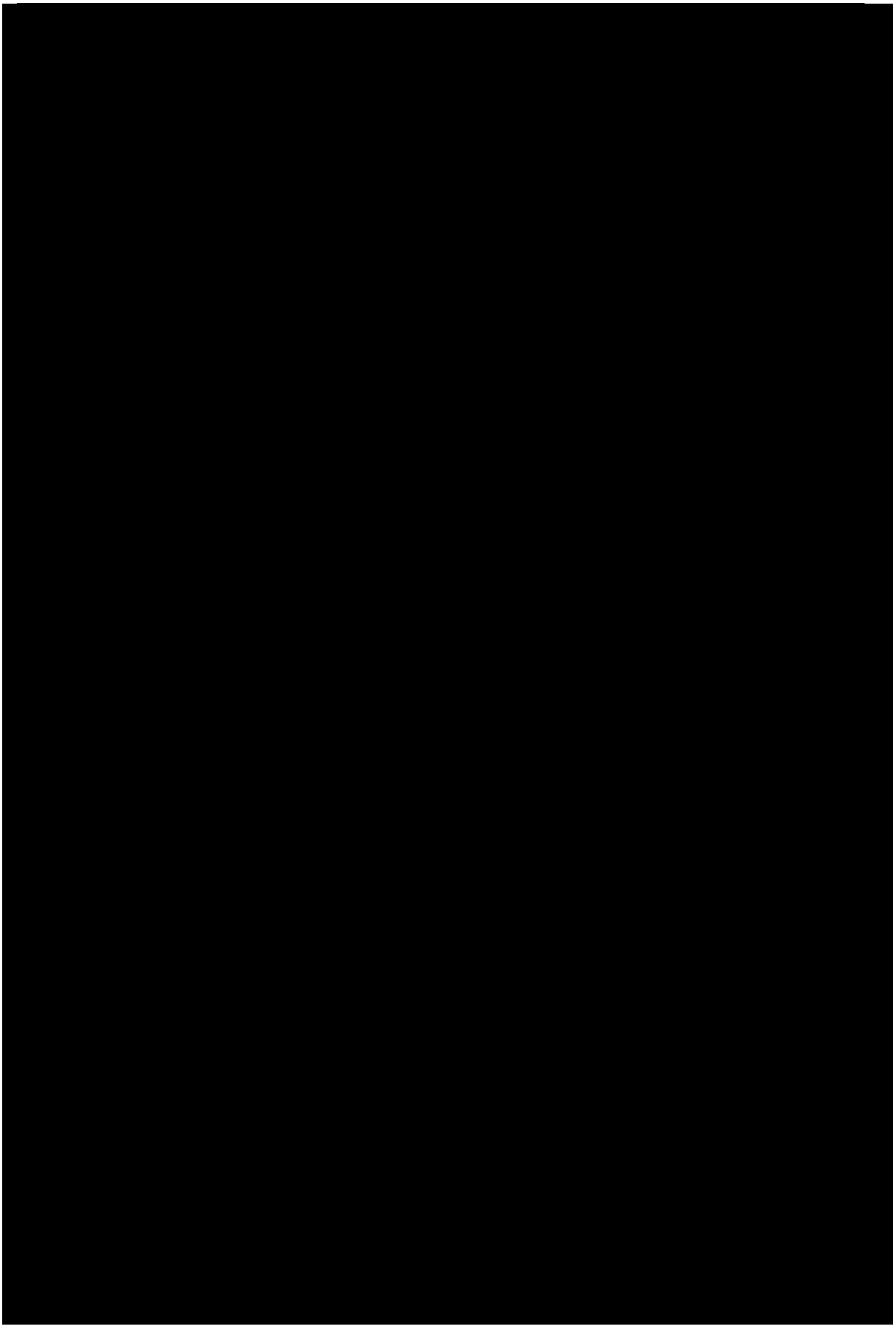
- Non-hematological toxicity ≥ grade 4 related to efavaleukin alfa. Non-hematological laboratory abnormalities without clinical significance will not be considered DLTs (based upon the investigator's discretion in conjunction with Amgen).
- Hematological toxicity ≥ grade 4 related to efavaleukin alfa defined as decrease in peripheral counts (ANC or platelets) persisting longer than 72 hours, as measured by 2 separate results, that are not related to malignant disease relapse, infection, or other etiologies.
- Constitutional events (i.e., fever, fatigue) ≥ grade 3 that are classified as serious adverse events by the investigator and related to efavaleukin alfa.
- Infection: Infection is considered an expected complication of cGVHD and its treatment. cGVHD itself increases the risk of infection, including life threatening infection. Therefore only grade 4 or 5 infections considered by the investigator to be related to efavaleukin alfa will be reviewed by the DLRM to determine whether the infection is considered a DLT.

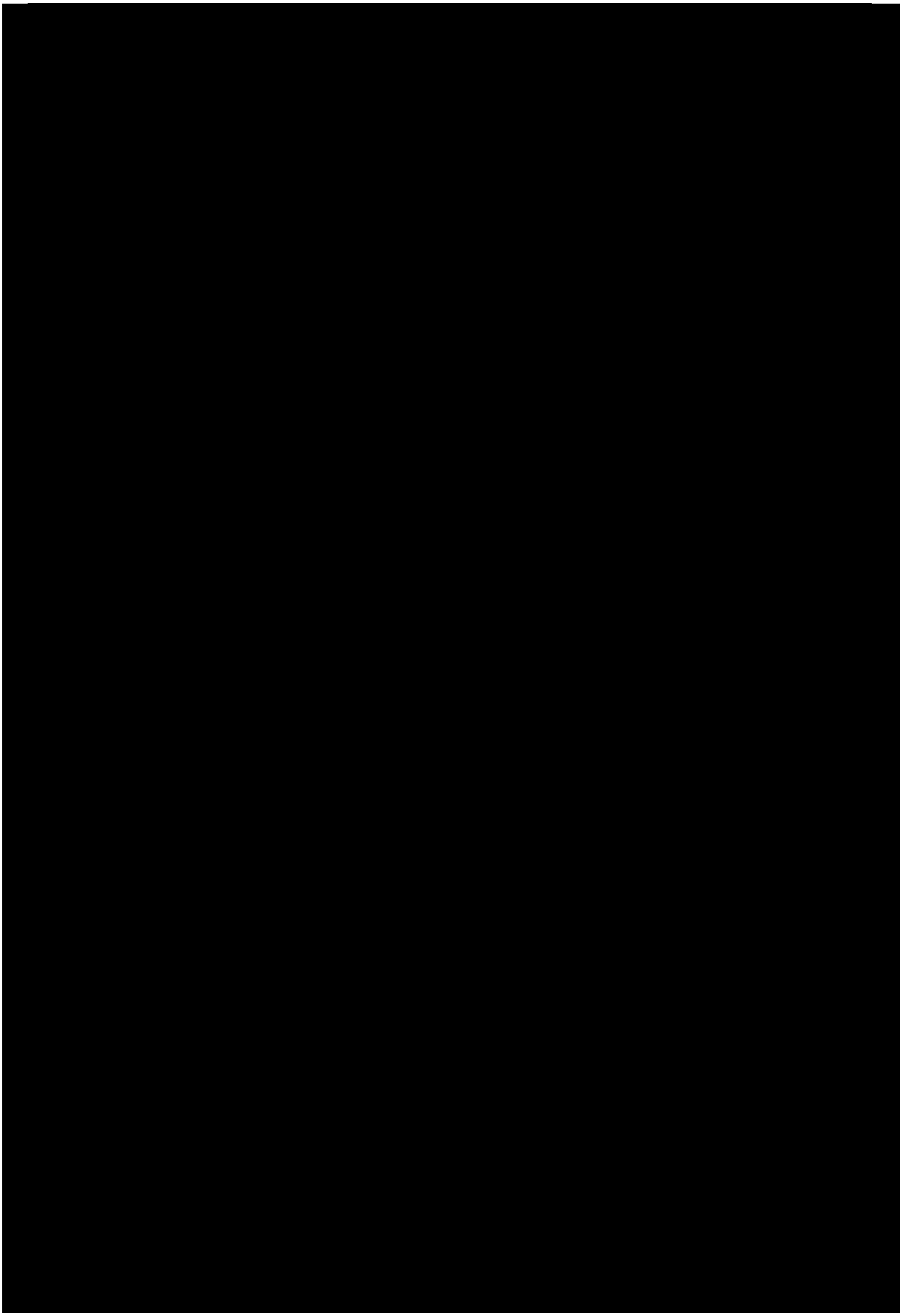
Subjects who develop an uncontrolled grade 4 life threatening infection (e.g., sepsis, ARDS), as assessed by the investigator, will discontinue efavaleukin alfa and remain in the study for safety follow-up (see Section 5.4) after completing the 4-week DLT evaluation period, subjects will be evaluated by standard safety reporting.

2014 cGVHD NIH response Consensus criteria algorithm (NIH Form A - 2014)

All subjects in the study will have their clinical response assessed using the NIH cGVHD Response assessment (NIH Form A) as per the 2014 cGVHD NIH Consensus Criteria. Subjects will undergo repeat detailed assessment of ocular, oral, cutaneous, musculoskeletal, gastrointestinal, hepatic, and pulmonary systems per [redacted]

[redacted]





[redacted]

Treatment Emergent AE (TEAE)

A treatment emergent adverse event is any adverse event categorized as adverse events starting on or after first dose of investigational product as determined by "Did event start before first dose of investigational product" equal to "No" or missing on the Events eCRF and up to and including 42 days after the last dose of investigational product excluding events reported after End of Study date on eCRF form, whichever is earlier.

Treatment-related Treatment Emergent Adverse Event (TRAE)

A treatment-related AE is any TEAE with the relationship flag on the Adverse Events eCRF indicating there is a reasonable possibility that the event may have been caused by investigational medicinal product. This assessment is performed by the study investigator. In the unlikely event that the relationship flag is missing and an event is TEAE, then it will be considered treatment-related and corresponding footnote will be included in the summary table.

Baseline medications

Baseline medication is defined as any medication with a start date on or before Study Day 1 and ongoing while on study.

Prior and Concomitant medications

Prior medication is defined as any medication with start date and end date prior to the first dose date of IP. Concomitant medication is defined as any medication with start date prior to the first dose date of IP but which continued to be taken after the first dose of IP or any medication with start date on or after the first dose date of IP and up to and including 42 days after the last dose date of IP.

Duration of Treatment-emergent Injection Site Reaction (ISR)

Duration = (min (ISR end date, EOS date) – ISR start date) + 1.

5.6 Demographic and Baseline Characteristics

Age

Age in years at the screening visit, which is collected in the eCRF.

Duration of cGVHD

Number of years from date of diagnosis of cGVHD (collected in chronic GVHD medical history eCRF) to the date of Study Day 1.

Duration of steroid refractory cGVHD

Number of years from date of diagnosis of steroid refractory cGVHD (collected in chronic GVHD medical history eCRF) to the date of Study Day 1.

6. Analysis Sets

6.1 Safety Analysis Set

The safety analysis set is defined as all subjects who have received at least 1 dose of efavaleukin alfa in the study. All analyses will be conducted on the safety analysis set unless noted otherwise.

6.2 DLT Analysis Set

The analysis of DLT will be restricted to DLT analysis set which consists of DLT-evaluable subjects in the study. The DLT evaluation period for each subject is 4 weeks from the first dose of efavaleukin alfa. To be evaluable for a DLT, subjects must have received at least 2 doses of efavaleukin alfa, or have experienced a DLT within the DLT evaluation period.

6.3 Pharmacokinetic/Pharmacodynamic Analyses Set(s)

6.3.1 Pharmacokinetic (PK) Concentration Analysis Set

The PK concentration analysis set consists of all subjects who have received at least one dose of efavaleukin alfa and have at least one quantifiable PK sample collected.

6.3.2 Pharmacodynamic (PD) Analysis Set

The PD analysis set consists of all subjects who have received at least one dose of investigational product and for whom at least one PD parameters have quantifiable baseline sample and at least one quantifiable post-baseline PD sample collected.

6.4 Anti-drug Antibody Analysis Set

The Anti-drug Antibody Analysis Set is defined as the subset of subjects in the Safety Analysis Set who had at least 1 evaluable anti-drug antibody test. Immunogenicity data will be analyzed according to the actual treatment received.

7. Planned Analyses

7.1 DLRMs

DLRMs: The study will have DLRMs for each cohort after the last subject in the cohort completes the DLT evaluation period. Section 3.1 details the timing and scope of each DLRM planned for the study. All available data for subjects up to and including the data snapshot date will be included in the analysis based on an “as-is” snapshot of the database without data locking. Data will be subject to ongoing checks for integrity,

completeness and accuracy in accordance with the Data Management Plan; however, there is no requirement to resolve outstanding data issues ahead of the DLRM snapshot. DLRM analysis will include subject-level listings and plots of available safety data (labs, and vitals) for each cohort and the BLRM analysis if required. The recommendations for dose escalation/dose modifications will be based on the DLRMs.

7.2 Final Analysis

A final analysis will be performed after the end of the study, i.e., after the last subject in phase 1b part of the study completes the study (including the treatment period and safety follow-up period), or early terminates. All analyses described in Section 9 will be performed at this time. Analyses will be based on all available data.

Data will be subject to ongoing checks for integrity, completeness and accuracy in accordance with the Data Management Plan with the expectation that all outstanding data issues are resolved ahead of the final lock.

8. Data Screening and Acceptance

8.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

8.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. The database will be subject to edit checks outlined in the Data Management Plan (DMP). See details of this section in the DMP.

8.3 Handling of Missing and Incomplete Data

Subjects may be missing specific data points for a variety of causes. In general, data may be missing due to a subject's early withdrawal from study, a missed visit, or inability to evaluate an endpoint at a particular point in time. The general procedures outlined below describe what will be done when a data point is missing.

8.3.1 Efficacy

No data imputation will be performed for missing efficacy endpoints.

8.3.2 Safety

The following imputation for missing or incomplete data will be performed if required: Incomplete or missing start/stop dates for adverse event and concomitant medication will be imputed as described in Appendix A.

Laboratory measurements that are below the lower quantification limits will be considered equal to the lower limit of quantification for all analyses unless explicitly noted otherwise. PD data that are below the lower limit of quantification will be considered equal to half of the lower limit of quantification for all analyses unless specified otherwise. PD data that are above the upper quantification limits will be considered equal to the upper limit of quantification for all analyses unless specified otherwise.

8.4 Detection of Bias

Lack of protocol compliance and the potential for biased statistical analyses will be examined by assessing the incidence of important protocol deviations in each cohort. The clinical study team will identify and document the criteria for important protocol deviations.

8.5 Outliers

Outlier data will not be excluded unless scientifically justified.

PK serum concentration data will be evaluated for outliers by visual inspection and decisions to re-assay individual samples will be made in accordance with standard PKDM practices.

8.6 Distributional Characteristics

Not Applicable

8.7 Validation of Statistical Analyses

Programs will be developed, maintained, and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software, for example the SAS System.

9. Statistical Methods of Analysis

9.1 General Considerations

Standard descriptive statistics will be provided**.** All binary endpoints will be summarized using the number and percentage of subjects. All continuous endpoints will be summarized using descriptive statistics including mean, standard deviation, minimum, maximum, and number of non-missing observations (n). Summaries will be provided at all scheduled time-points unless specified otherwise. Data collected at unscheduled visits or time-points will be included in the analysis unless stated otherwise. **All safety and efficacy analyses will be performed using the Safety Analysis Set based on subject's actual treatment received.**

Only critical subject-level data listings will be provided in the clinical study report.

9.2 Subject Accountability

The number and percent of subjects who are screened, enrolled, have received at least one dose of efavaleukin alfa, completed investigational product, discontinued from investigational product (including reasons for discontinuing), completed study, discontinued the study (including reasons for discontinuing) will be summarized overall and by treatment group. Separate subject disposition will be provided for the follow-up till week 52 and the extended follow-up.

Key study dates for the first subject enrolled, last subject enrolled, and last subject's end of study will be presented.

9.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's initial visit and updated during the IPD reviews throughout the study prior to database lock. If a snapshot is taken during the study rather than a database lock at the end of the study, categories should be updated during the IPD reviews throughout the study prior to database lock. These definitions of IPD categories, sub-category codes and descriptions will be used during the course of the study. The final IPD list is used to produce the summary of IPD table and the list of subjects with IPDs. The number of subjects with important protocol deviations and protocol deviations due to COVID-19 will be summarized in a table. A list of protocol deviations for subjects impacted by COVID 19 will also be provided.

Eligibility deviations are defined in the protocol. A list of deviations from eligibility criteria will also be generated.

9.4 Demographic and Baseline Characteristics

Subject demographic and baseline disease characteristics will be summarized and include the following variables:

9.4.1 Demographic Characteristics

- Age (years) at screening (continuous summary statistics)
- Age categories (number and percent of subjects in 18 - 64, 65 - 74 years, ≥ 75 years)
- Sex (number and percentage of males and females)
- Ethnicity (number and percentage of Hispanic or Latino and Not Hispanic or Latino)
- Race (number and percentage of subjects in each race, or mixed race combination)

9.4.2 Baseline Characteristics

Physical Measurements

- Height and weight (continuous summary statistics**)**

Disease characteristics:

- Duration of cGVHD (continuous summary statistics)
- Duration of steroid refractory cGVHD (continuous summary statistics)
- Information collected on cGVHD medical history eCRF (e.g. underlying malignancy type for which subject received HSCT, Donor type, Stem cell source etc.) (number and percentage of subjects)

[redacted]

- NIH- Form A-2014 assessment of each organ system (number and percentage of subjects)
- Overall cGVHD severity as marked on Chronic Graft versus Host Disease Staging Evaluation eCRF (number and percentage of subjects)

Baseline immunosuppressant medication

- **Systemic Corticosteroids**
- **mTOR inhibitors**
- **Calcineurin inhibitors**

9.5 Efficacy Analyses

9.5.1 Analyses of Primary Efficacy Endpoint(s)

No primary endpoint is an efficacy endpoint.

9.5.2 Analyses of Secondary Efficacy Endpoint(s)

No secondary endpoint is an efficacy endpoint.

9.5.3 Analyses of Exploratory Efficacy Endpoint(s)

9.6 Safety Analyses

9.6.1 Analyses of Primary Safety Endpoint(s)

The primary endpoint is safety. The analysis will include the descriptive summary statistics for adverse events, disease-related events, laboratory test results, vital signs, and physical measurements. Details are described in respective section below.

9.6.2 Adverse Events and Disease-related Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 25.0 or later will be used to code all adverse events to a system organ class and a preferred term.

The subject incidence of adverse events will be summarized for all treatment-emergent, serious treatment emergent, treatment-related, those leading to withdrawal of investigational product, fatal and of special interest (if applicable). The identification of adverse events of special interest is a continuous process. Events may be identified and documented as the safety profile of the drug. The severity of each adverse event will be graded using CTCAE version 4.03 criteria.

Subject incidence of all treatment-emergent, serious treatment emergent, treatment-related, serious treatment-related, those leading to withdrawal of investigational product or other protocol-required therapies, and fatal adverse events will be tabulated by system organ class and preferred term in descending order of frequency. Treatment emergent, serious, and treatment-related events will also be presented by preferred term and worst severity grade using CTCAE version 4.03 criteria.

Overall summary of subject incidence of treatment-emergent adverse events and

by preferred term and worst severity grade will be provided by anti-AMG592 and [redacted] status. Subject incidence of events of interest (standardized MedDRA queries and/or Amgen Medical Queries) will also be summarized according to their categories, preferred term, and worst severity grade. Events of interest could include but are not limited to Hypersensitivity, Tachyarrhythmias, Tachypnea, Hematopoietic Cytopenia, **Eosinophilic Disorder**, Leucocyte Changes, Infection and Infestation, Cytokine Release Syndrome, **Drug Related Hepatic Disorder**, and Injection Site Reaction. Number of episodes and duration of treatment-emergent injection site reactions will be further summarized.

Subject incidence of treatment-emergent adverse events identified by COVID-19 standardized MedDRA queries and serious adverse events occurring on or after the COVID-19 infection will also be summarized.

Subject incidence of disease related events and fatal disease-related events will be tabulated by system organ class and preferred term. They are defined for this study in Appendix D.

9.6.2.1 Dose Limiting Toxicities

The analysis of the probability of DLT will be based on the Dose Escalation part of the study defined in Section 3.1. A listing and summary of the subject incidence of DLT will be provided should they occur. After there is at least 1 DLT observed at any dose level, a Bayesian logistic regression model (BLRM) (Bailey et al, 2009; Neuenschwander et al, 2008) will be used to inform dose escalation and the results of this analysis will be provided to the each DLRM. Three toxicity probability intervals (TPI) of DLT will be defined: target TPI (20% to 30%), excessive TPI (30% to 60%) and unacceptable TPI (60% to 100%). Adverse events meeting DLT criteria after week 4 and PD information may also be included in the model, as available. The model will recommend an MTD as the dose with highest probability in the target TPI, but with less than 25% probability in excessive TPI or unacceptable TPI. Lower dose levels, intermediate dose levels, or alternative dosing schedules may be considered based on all available information as long as they do not exceed the estimated MTD per the model.

Let $\boldsymbol{d}=(d_1, \ldots, d_k)$ denote the vector of all efavaleukin alfa doses to be studied in the study. The probability of a DLT at dose level $d_i, i=1, \ldots, k$ is assumed to follow a Bernoulli distribution with probability p_i where the logit of p_i increases linearly with the log of the standardized dose in the following 2-parameter logistic model:

$$\log\left[\frac{p_i}{1-p_i}\right] = logit(p_i) = \log[a] + \exp(\log[b])\log(\frac{d_i}{\max(\boldsymbol{d})})$$

where a and b are random variables and d_i is dose level represented by predicted [redacted] d_k is an additional dose assuming Treg concentration = [redacted] that's higher than the currently planned doses. Thus $\boldsymbol{d}$=[redacted]. If additional dose cohorts were added during study conduct (e.g., decision following a DLRM), the value of $\boldsymbol{d}$ will be updated.

The probability of each TPI and of a DLT will be summarized by dose along with the estimated dose-toxicity curve. A final estimate of the MTD will be estimated from the Bayesian model by utilizing all DLT-evaluable subjects. Please refer to Appendix C for more details on the BLRM and its operation characteristics. Before each DLRM, the above BLRM will be fitted with the prior distribution as specified in Appendix C. This model will only be used to aid DLRM dosing decision thus the recommendation of the model is non-binding in nature.

9.6.3 Laboratory Test Results

9.6.3.1 Chemistry and Hematology

The analyses of safety laboratory endpoints will include summary statistics of baseline, post-baseline lab values at all the time points, change from baseline to post-baseline, post-baseline maximum/minimum and change from baseline to post-baseline maximum/minimum by dose level.

Shifts tables indicating the change between the baseline and the worst post dose CTCAE grades based on CTCAE version 4.03 or later will be provided for selected laboratory parameters of interest. Subject incidence of worst post-baseline eosinophils counts by the laboratory normal range will be provided.

Chemistry and hematology laboratory tests performed by local laboratories will be included in the analyses where applicable in cases where central laboratory values were not obtainable (i.e. due to Covid restrictions).

A listing of subjects identified as potential Hy's law cases (ALT or AST of >3xULN associated with total bilirubin > 2xULN and ALP <2xULN at a visit) will be provided with relevant laboratory tests included in the listing.

9.6.4 Vital Signs

The analyses of vital signs will include summary statistics for baseline, post-baseline values at all the time points, change from baseline to post-baseline values at all the time points**,** post-baseline maximum/minimum and change from baseline to post-baseline maximum/minimum by dose level for systolic blood pressure, diastolic blood pressure, heart rate and temperature.

9.6.5 Physical Measurements

The analyses will include summary statistics of weight at scheduled time-points and by dose level.

9.6.6 Electrocardiogram

The ECG measurements will be performed as per standard of care for routine safety monitoring, rather than for purposes of assessment of potential QTc effect. Since these evaluations may not necessarily be performed under the rigorous conditions expected to lead to meaningful evaluation of QTc data; summaries and statistical analyses of ECG measurements are not planned, and these data would not be expected to be useful for meta-analysis with data from other trials.

9.6.7 Antibody Formation

If needed, exploratory analyses of the relationship between safety and immunogenicity may be performed**.**

9.6.8 Exposure to Investigational Product

The number of days on IP and the total dose of IP will be summarized using descriptive statistics. Number and percentage of subjects who missed IP doses due to COVID-19 measures will be summarized. The final IPD list will be used to identify such subjects.

A listing of the unique manufacturing lot numbers, and a listing of the subjects administered each manufacturing lot number will be provided

9.6.9 Exposure to Concomitant Medication

Number and proportion of subjects receiving medications/therapies of interest will be summarized by preferred term or category as coded by the World Health Organization Drug (WHODRUG) dictionary.

Number and proportion of subjects who experienced a systemic corticosteroid dose reduction from baseline in subjects who were on systemic corticosteroid at baseline will be summarized by treatment group through week 52.

9.7 Other Analyses

9.7.1 Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints

9.7.1.1 Analysis of Pharmacokinetic Endpoint

Pharmacokinetics analyses will be performed by Clinical Pharmacology, Modeling and Simulation (CPMS) group. The Pharmacokinetic Concentration Analysis set defined in Section 6.3.1 will be used in analyzing PK endpoints. Serum concentrations of efavaleukin alfa will be expressed in ng/mL. All concentrations below the limit of quantification or missing data will be labeled as such in the concentration data listings. Concentrations below the lower limit of quantification will be treated as zero in summary statistics.

Actual dosing and sampling times and nominal doses (unless subjects had dosing errors, dosing changes, or dosing interruptions) will be used for individual concentration-time plots and calculation of PK parameters for each subject. The reasons for excluding any sample from the analyses will be provided.

Individual concentration-time data will be tabulated and presented graphically. Summary of PK concentration over time and PK parameters will be provided. Mean concentration-time profiles for each dose will be provided. PK observations with missing concentration, missing dose, missing elapsed time will be excluded from PK statistical summary. PK parameters will be summarized for each dose using descriptive statistics, including subject numbers, mean, standard deviation, CV, median and range.

9.7.1.2 Analysis of Pharmacodynamic Endpoints

The Pharmacodynamic (PD) Analysis set defined in Section 6.3.3 will be used.

PD assessments include but are not limited to:

[redacted]

[redacted] Due to the skewed distribution of PD data, mean, standard deviation and standard error will be estimated based on log-transformed data, and the estimates will be back-transformed and presented, which are equivalent to geometric mean, geometric standard deviation and geometric standard error.

To maintain the integrity of PD results, data collected at an end of treatment or end of extended treatment visit which is completed outside the upper limit of protocol defined window (> 9 days after the last dose date for QW cohorts and > 16 days for Q2W cohorts) will not be included in the analysis. In addition, data collected at a visit after skipping one or more consecutive doses prior to that visit will be also excluded from analysis (the same window will be used as above i.e. (> 9 days after the last dose date for QW cohorts and > 16 days for Q2W cohorts).

9.7.1.3 PK/PD Analysis

Pharmacokinetic/pharmacodynamic modeling may be performed to explore relationship between efavaleukin alfa exposure and PD efficacy/safety endpoints by CPMS.

9.7.2 Analyses of Clinical Outcome Assessments

[redacted]

9.7.3 Analyses of Biomarker Endpoints

Not Applicable

10. Changes From Protocol-specified Analyses

The planned phase 1b interim analysis and phase 2 primary and final analyses are excluded from this analysis plan because the study is discontinued early. Instead, a final analysis will be conducted after the last subject in phase 1b part completes the study (including the treatment period and safety follow-up period).

11. Literature Citations / References

efavaleukin alfa Protocol 20160283. Thousand Oaks, CA: Amgen Inc

Bailey S, Neuenschwander B, Laird G, Branson M. (2009) A Bayesian case study in oncology phase I combination dose-finding using logistic regression with covariates. J Biopharm Stat. 19, 469-84.

Brookmeyer R and Crowley J. (1982) A Confidence Interval for the Median Survival Time. Biometrics, 38, 29-41.

Jagasia MH; Greinix HT; Arora M, et al. (2015) National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft versus Host Disease: I. The 2014 Diagnosis and Staging Working Group Report. Biol Blood Bone Marrow Transpl. 21(3), 389-401.

Kalbfleisch, J.D. and Prentrice, R.L. (1980) The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons.

Lee SJ; Wolff D; Kitko C; et al. (2014) Measuring Therapeutic Response in Chronic Graft-versus-host Disease. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host disease: IV. The 2014 Response.

Lee SJ; Cook EF; Soiffer R; Antin JH. (2002) Development and Validation of a Scale to Measure Symptoms of Chronic Graft-versus-Host Disease. Biol Blood Marrow Transplant. 8; 444-452.

Neuenschwander B, Branson M, Gsponer T. (2008) Critical aspects of the Bayesian approach to phase I cancer trials. Statistics in Medicine. 27, 2420-39

Ware JE. (2000) SF-36 health survey update. Spine. 25(24), 3130-3139.

12. Prioritization of Analyses

There is no prioritization of analyses.

13. Data Not Covered by This Plan

Exploratory data not included in this plan may be analyzed at a later date or may be analyzed by a different Amgen Department.

14. Appendices

Appendix A. Technical Detail and Supplemental Information Regarding Statistical Procedures and Programs

Imputation Rules for Partial or Missing Start Dates

The reference date for the following rules is the date of first dose of IP.

Start Date		Stop Date						
		Complete: *yyyymmdd*		**Partial: *yyyymm***		**Partial: *yyyy***		**Missing**
		< 1st dose	**≥ 1st dose**	**< 1st dose *yyyymm***	**≥ 1st dose *yyyymm***	**< 1st dose *yyyy***	**≥ 1st dose *yyyy***	
Partial: *yyyymm*	**= 1st dose *yyyymm***	2	1	n/a	1	n/a	1	1
	≠ 1st dose *yyyymm*		2	2	2	2	2	2
Partial: *yyyy*	**= 1st dose *yyyy***	3	1	3	1	n/a	1	1
	≠ 1st dose *yyyy*		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1=Impute the date of first dose;

2=Impute the first of the month;

3=Impute January 1 of the year;

4=Impute January 1 of the stop year;

Note: For subjects who were never treated (first dose date is missing), partial start dates will be set to the first day of the partial month or first day of year if month is also missing.

Imputation Rules for Partial or Missing Stop Dates

Initial imputation

- If the month and year are present, impute the last day of that month.
- If only the year is present, impute December 31 of that year.
- If the stop date is entirely missing, assume the event or medication is ongoing.

If the imputed stop date is before the start date, set stop date to missing.

If the imputed stop date is after the death date, impute as death date.

Imputation Rules for Partial or Missing Death Dates

If death year and month are available but day is missing:

- If yyyymm for the date last known to be alive equals yyyymm for death date, set death date to the day after the date last known to be alive.
- If yyyymm for the date last known to be alive is less than the yyyymm for death date, set death date to the first day of the death month.
- If yyyymm for the date last known to be alive is greater than yyyymm for death date, assume date last known to be alive is in error, set death date to the first day of the death month.

If month and day are missing and year of death is known:

- If yyyy for the date last known to be alive equals yyyy for death date, set death date to the day after last known to be alive date.
- If yyyy for the date last known to be alive is less than yyyy for death date, set death date to the first day of the death year.
- If yyyy for the date last known to be alive is greater than yyyy for death date, assume date last known to be alive is in error, set death date to the first day of the death year.

If a death date is totally missing:

- Set death date to the day after the date last known to be alive. If the date last known to be alive is a partial date, set it to the first day of the month last known to be alive or first day of the year last known to be alive if month is also missing.

Appendix B. Reference Values/Toxicity Grades

The Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 will be used and is available at the following location

Common Terminology Criteria for Adverse Events (CTCAE) | Protocol Development | CTEP (cancer.gov)

Appendix C. Bayesian Logistic Regression Model

The study will use BLRM to guide dose escalation. The MTD target toxicity probability interval for DLT is (0.20, 0.30), and TPIs of (0.30, 0.60) and (0.60, 1.00) are defined as excessive and unacceptable, respectively. The design seeks to identify a dose most likely to have a DLT rate in the target TPI, but with overdose control that limits the possibility the dose has an excessive or unacceptable DLT rate (Babb et al, 1998).

Let $\boldsymbol{d}=(d_1, \ldots, d_k)$ denote the vector of all efavaleukin alfa doses to be studied in the study. The probability of a DLT at dose level $d_i, i=1, \ldots, k$ is assumed to follow a Bernoulli distribution with probability p_i where the logit of p_i increases linearly with the log of the standardized dose in the following 2-parameter logistic model:

$$\log\left[\frac{p_i}{1-p_i}\right]=logit(p_i)=\log[a]+\exp(\log[b])\log(\frac{d_i}{\max(\boldsymbol{d})})$$

where a and b are random variables and d_i is dose level represented by predicted [redacted]). d_k is an additional dose assuming Treg concentration = [redacted] that's higher than the currently planned doses. Thus $\boldsymbol{d}$= [redacted]. If additional dose cohorts were added during study conduct (e.g., decision following a DLRM), the value of $\boldsymbol{d}$ will be updated.

[redacted]

A minimally informative prior distribution (Neuenschwander et al, 2008) was selected for $\theta=(\log[a], \log[b])$, where the probability that the true DLT rate is ≤ 0.20 at the lowest dose is 0.90 and the probability the true DLT rate is ≤ 0.50 at the maximum dose is 0.90. Median values for p_i were interpolated per the logistic model. For each d_i, 2 quantiles for $p_i$were selected from a minimum informative Beta distribution with the target median. This set of quantiles fully specified a target prior for theta. A bivariate normal distribution for θ was assumed where $(\log[a])$ has a normal distribution with mean ma and standard deviation sa, and $(\log[b])$ has a normal distribution with mean mb and standard deviation

sb, and r is the correlation between $\log[a]$ and $\log[b]$. Numerical integration with the R "integrate" function was used to calculate $\Pr[p_i \leq q(d_i)$ where $q(d_i)$ is a quantile for dose d_i. An optimal bivariate normal distribution was estimated that achieved the minimum sum of squared difference between achieved and specified quantiles across all doses using a non-linear optimization with the R "optim" function. The bivariate normal distribution prior solution has ma = -1.283, sa = 1.002, mb = 0.280, sb = 0.770, and r = -0.568.

The operating characteristics of the BLRM model were evaluated via simulation with cohort size of 3. All simulated studies started with an initial dose strength of [redacted] and subsequent doses were selected based on the following rules:

- After each cohort, the next dose is the one with the highest probability of the target TPI, but with a less than 0.25 probability of an excessive or unacceptable TPI.
- If ≥ 1 subject has a DLT at a dose, then no higher dose can be evaluated unless there are at least 6 evaluable subjects at this lower dose.

No skipping of planned dose is allowed.

The dose escalation will be stopped given one or more of the following conditions:

- The model recommends the same dose > 3 times (not necessarily sequentially)
- A maximum of 24 subjects is evaluated.

The design was evaluated for three possible dose-response scenarios consistent with the design's model: "Low", "Mid", and "High" MTD. The possible dosing strength in the simulation is [redacted], where [redacted] and [redacted] was added to simulate the "High MTD" scenario where the true MTD exceeds the currently planned dose. See Table 2 for the probability profiles of the three scenarios (acceptable MTDs (between 0.2 to 0.3) are marked bold)

Table 2. Probability Profile of Low, Mild and High Scenario

Dosing Strength	Probability of DLT		
	Low MTD	Mid MTD	High MTD
[redacted]	0.2	0.1	0.01
	0.3	0.2	0.05
	0.3	0.2	0.05
	0.4	0.3	0.1
	0.5	0.4	0.2
	0.6	0.5	0.3

For each dosing profile, [redacted] simulated trial data was generated and the BLRM model was used to guide the dose escalation. The simulation results are summarized in Table 3

Table 3. Summarized Simulation Results for BLRM

	Low MTD	**Mid MTD**	**High MTD**
N	14.8 (6, 21)*	19.7 (18, 24)	23.0 (24, 24)
MTD dose	129.1 ([redacted])	200.4 ([redacted])	416.8 ([redacted])
% select acceptable MTD	50.6%	62.2%	86.8%
% all dose too toxic	39.2%	12.0%	0%
Study # of DLTs	4.1 (3, 5)	4.2 (4, 5)	2.8 (2, 4)
Study prob. of DLT	33.1%	23.3%	11.9%

* all continuous data in this table is reported in mean (Q1, Q3)

Appendix D. Disease Related Events Table

Organ System	Disease Related Event
Skin	Scleroderma (superficial or fasciitis), lichen planus, vitiligo, scarring alopecia, hyperkeratosis pilaris, contractures from skin immobility, nail bed dysplasia
Mucous membranes	Lichen planus, non-infectious ulcers, corneal erosions/non-infectious conjunctivitis, xerostomia, keroconjunctivitis sicca
GI tract	Esophageal strictures, steatorrhea, anorexia, malabsorption, weight loss, abdominal pain
Liver	None
GU tract	Vaginal stricture, lichen planus
Musculoskeletal /Serosa	Non-septic arthritis, myositis, myasthenia, polyserositis, contractures from joint immobilization
Hematologic	None
Lung	Bronchiolitis obliterans

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Appendix G. Analytical Windows

The last visit with non-missing assessment prior to first dose will be defined as a baseline visit (unless specified otherwise) and the analysis visit name will be “Baseline”. For any visit up to Day 1 pre-dose which is not a baseline visit, the analysis visit will be ‘Pre-Analysis’. For Day 1 post-dose visits, and safety follow-up visit, the analysis visit will be same as the scheduled visit. For any unscheduled visit post safety follow-up visit, the analysis visit will be considered as safety follow-up visit. For post-dose assessments on Day 29 (the day corresponding to Week 4 dosing) and Day 113 (the day corresponding to week 16 dosing), the analysis visit will be same as the scheduled visit. Remaining visits (scheduled visit, unscheduled visit, progressive disease visit, or end of treatment visit) after Day 1 for any assessment will be mapped to the analysis visit based on visit windows defined in the tables below for that assessment. Only visits applicable for that assessment and phase of the study as per the Schedule of Activities tables in the study protocol will be considered for analysis.

Data collected through central lab and local lab will both be mapped into analysis visit windows. The local lab data can only be used in analysis when there is no central lab data available in the analysis visit window.

If more than one visit with non-missing data falls within an analysis visit window, the visit closest to the target day will be considered for analysis. If two visits are equidistant from the target day, the latest visit (or time if on the same day) will be considered. If more than one evaluation has same date and time (for chemistry, hematology or pharmacodynamics results), the value with the smallest accession number will be considered.

Analysis visit windows for selected assessments are included in tables below. For remaining assessments, no visit window will be applied. Day 29 and Day 113 used in Study day window calculation in the tables below is the respective study day when Week 4 dose and Week 16 dose was received. If that dose is missing, then 29 and 113 is used for respective calculation.

Vitals

Analysis Visit	Target Day	Study Day Window

Week 1, Day 2	2	2
Week 1, Day 3	3	3
Week 1, Day 4	4	4-6
Week 1, Day 8	8	7-9
Week 1, Day 11	11	10-12
Week n, Day 7*n+1 (n = 2, 3, 4, 5, 6, 7,8, 9, 10, 11, 12, 13, 14)	7n+1	Target day -2, Target day +4
Week 15, Day 106	106	104, Day 113-3
Week 16, Day 113 Pre-dose	113	Day 113-2, Day 113
Week 16, Day 114	114	Day 113 + 1
Week 16, Day 115	115	Day 113 + 2
Week 16, Day 116	116	Day 113 + 3
Week 17, Day 120	120	Day 113 + 4, 124
Week 18, Day 127	127	125-131
Week 20, Day 141	141	132-148
Week n, Day 7*n+1 (n = 22, 24, 26,.. to 258 in increments of 2)	7n+1	Target day -6, Target day +7
Week 260, Day 1821	1821	>1814

Chemistry and Hematology, Weight

Week 1, Day 2	2	2-3
Week 1, Day 8	8	4-12
Week n, Day 7*n+1 (n = 2, 3, 4, 5, 6, 7)	7*n+1	Target day -2, Target day +4
Week 8, Day 57	57	55-71
Week n, Day 7*n+1 (n=12, 16, 20, …, to 104 in increments of 4)	7*n+1	Target day -13 days, Target day + 14 days
Week 108, Day 7**57**	757	744, 764
Week n, Day 7*n+1 (n=110, …, to 258 in increments of 2)	7*n+1	Target day -6 days, Target day + 7 days
Week 260, Day 1821	1821	1815, SFU -1 if SFU* is available, otherwise >1814

*study day corresponding to safety follow-up visit assessment

Pharmacodynamic assessments

Week 1, Day 2	2	2-3
Week 1, Day 4	4	4-6
Week 1, Day 8	8	7-9
Week 1, Day 11	11	10-12
Week 2, Day 15	15	13-19

Week 3, Day 22	22	20, Day 29 - 3
Week 4, Day 29	29	Day 29 -2, Day 29
Week 4, Day 30	30	Day 29 +1
Week 4, Day 32	32	Day 29 + 2, Day29 +4
Week 5, Day 36	36	Day 29 + 5, 40
Week n, Day 7*n+1 (n = 6,7,8,9,10,11)	7*n+1	Target day -2, Target day +4
Week 12, Day 85	85	83, 92
Week 14, Day 99	99	93, Day 113 -7
Week 16, Day 113	113	Day 113- 6, Day 113
Week 16, Day 114	114	Day 113 +1
Week 16, Day 116	116	Day 113+ 2, Day 113 +4
Week 17, Day 120	120	Day 113 +5, 124
Week 18, Day 127	127	125- 134
Week n, Day 7*n+ 1 (n= 20, 22, ..to 50 in increments of 2)	7*n+1	Target day -6 days, Target day + 7 days
Week 52, Day 365	365	359-379
Week n, Day 7*n+ 1 (n= 56, 60, 64, …, to 104 in increments of 4)	7*n+1	Target day - 13 days, Target day + 14 days
Week 108, Day 7**57**	757	744, 764

Week n, Day 7*n+1 (n=110, …, to 258 in increments of 2)	7*n+1	Target day -6 days, Target day + 7 days
Week 260, Day 1821	1821	>1814

Antibody assessments

Week 2, Day 15	15	2- 22
Week 4, Day 29	29	23- 43
Week n, Day 7*n+ 1 (n= 8, 12, 16,…, to 104 in increments of 4)	7*n+1	Target day -13 days, Target day + 14 days
Week 108, Day 7**57**	757	744, 764
Week n, Day 7*n+1 (n=110, …, to 258 in increments of 2)	7*n+1	Target day -6 days, Target day + 7 days
Week 260, Day 1821	1821	>1814