STATISTICAL ANALYSIS PLAN

Study Title:	Evaluation of the Efficacy and Safety of GS-5745 as Add-On Therapy to a Tumor Necrosis Factor Inhibitor and Methotrexate Regimen in Subjects with Moderate to Severe Rheumatoid Arthritis
Name of Test Drug:	Andecaliximab (GS-5745)
Study Number:	GS-US-373-1499
Protocol Version (Date):	Amendment 2 (23 September 2016)
Analysis Type:	Final Analysis
Analysis Plan Version:	1.0
Analysis Plan Date:	11 December 2017
Analysis Plan Author(s):	PPD

CONFIDENTIAL AND PROPRIETARY INFORMATION

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AE	adverse event
BMI	body mass index
CDAI	Clinical Disease Activity Index
CRF	case report form
CRP	C-reactive protein
CSR	clinical study report
CTCAE	common terminology criteria for adverse events
DAS28	disease activity score for 28 joint count
DB	double-blind
ECG	electrocardiogram
eCRF	electronic case report form
FAS	full analysis set
Gilead	Gilead Sciences, Inc.
HAQ-DI	health assessment questionnaire-disability index
HLGT	high-level group term
HLT	high-level term
ID	identification
LLT	lower-level term
LOQ	limit of quantitation
MCII	Minimal Clinically Important Improvement
MedDRA	medical dictionary for regulatory activities
MTX	methotrexate
OL	open-label
OLE	open-label extension
PhGA	physician's global assessment
PT	preferred term
PtGA	patient's global assessment
PK	pharmacokinetic
Q1, Q3	first quartile, third quartile
RA	rheumatoid arthritis
SAE	serious adverse event
SAP	statistical analysis plan
SC	subcutaneous
SD	standard deviation
SDAI	Simplified Disease Activity Index
SE	standard error
SOC	system organ class
TE	treatment-emergent

TEAE	treatment-emergent adverse event
TFLs	tables, figures, and listings
TNF	tumor necrosis factor
VAS	visual analog scale
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis methods and data presentations to be used in tables, figures, and listings (TFLs) in the clinical study report (CSR) for Study GS-US-373-1499. This SAP is based on the study protocol amendment 2 dated 23 September 2016 and the electronic case report form (eCRF). The SAP will be finalized before database finalization. Any changes made after the finalization of the SAP will be documented in the CSR.

On 5 June 2017, Study GS-US-373-1499 was terminated by the sponsor due to business consideration. Consequently, some of the statistical analyses and all the statistical testing planned in the study protocol will not be performed due to study early termination. This SAP describes the analyses to support a synoptic CSR.

1.1. Study Objectives

The primary objective of this study is as follows:

- To assess the efficacy of andecaliximab (GS-5745) versus placebo as an add-on therapy to a tumor necrosis factor (TNF) inhibitor and methotrexate in subjects with moderately to severely active rheumatoid arthritis (RA)

The secondary objectives of this study are as follows:

- To assess the safety and tolerability of andecaliximab versus placebo as an add-on therapy with a TNF inhibitor and methotrexate (MTX) in subjects with moderately to severely active RA

- To assess the pharmacokinetics (PK) of andecaliximab as an add-on therapy with a TNF inhibitor and MTX in subjects with moderately to severely active RA

The exploratory objectives of this study are as follows:

PPD

1.2. Study Design

This is a Phase 2, double-blind, placebo-controlled, randomized multi-center study evaluating the efficacy and safety of andecaliximab as add-on therapy in subjects with moderately to severely active RA who have had an inadequate response to anti-TNF biologic therapy.

Subjects will be stratified by Disease Activity Score for 28 joints using the C-reactive protein (CRP) (DAS28[CRP]), either DAS28(CRP) > 5.1 or DAS28(CRP) > 3.2 and ≤ 5.1. In addition, subjects will be stratified by the prior number of RA biologics used (< 3 or ≥ 3).

Subjects will receive subcutaneous (SC) injections of study drug (andecaliximab or placebo) once a week for 12 weeks in the blinded period. Eligible subjects who complete the blinded period of the study may participate in the optional open-label extension (OLE), where they will receive weekly SC injections of 300 mg of andecaliximab for 52 weeks, in addition to continuing their current TNF inhibitor and MTX.

A total of approximately 75 subjects will be randomized in this study. The study design is shown in Figure 1-1.

Figure 1-1.   Study Schema
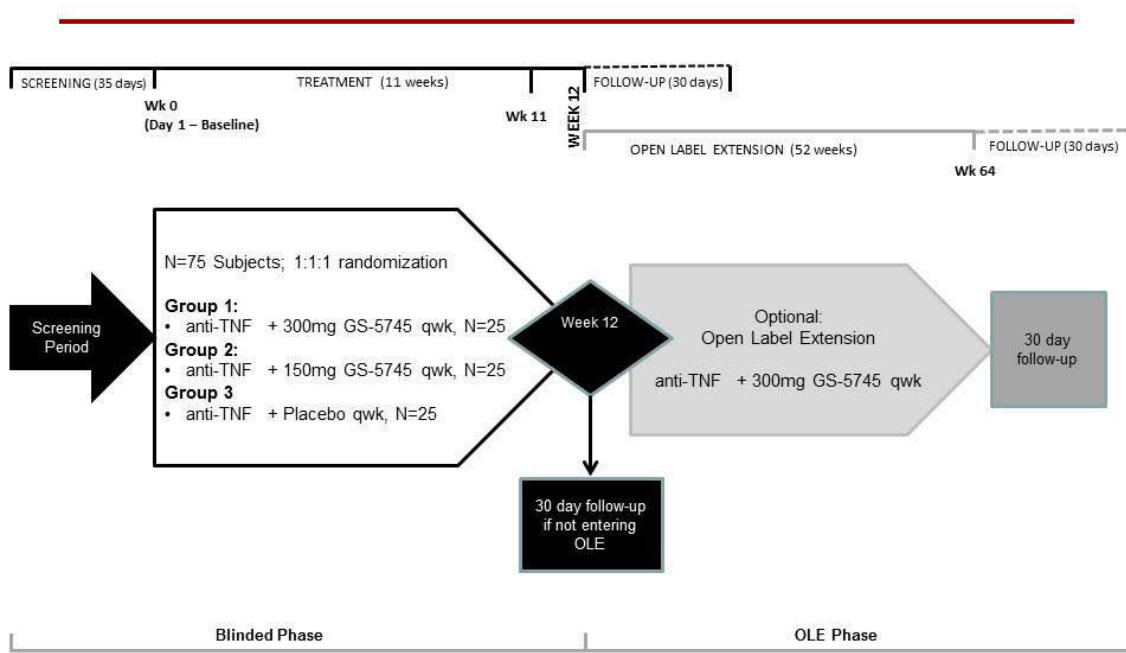
Following completion of screening assessments, eligible subjects will be randomized in a blinded fashion in a 1:1:1 ratio as follows:

- **Group 1:** andecaliximab 300 mg (2 x 1 mL andecaliximab [150 mg/ml] SC injections) administered weekly (N=25)

- **Group 2:** andecaliximab 150 mg (1 x 1 mL andecaliximab [150 mg/ml] + Placebo-to-match 1 x 1 mL SC injection administered weekly (N=25)

- **Group 3:** Placebo-to-match (2 x 1 ml SC injections) administered weekly (N=25)

At the Week 12 visit, eligible subjects who choose to participate in the OLE will start weekly, open-label (OL), SC injections of 300 mg of andecaliximab for 52 weeks, in addition to their current SC administration of a TNF inhibitor and continued MTX.

1.3.   Sample Size and Power

A total sample size of 75 subjects will be required for this study. Each of the 2 andecaliximab treated groups will be compared to the placebo group. A sample size of 25 per group will provide a power of 80% with a two-sided α level of 0.05 to detect a Minimal Clinically Important Improvement (MCII) in DAS28(CRP) change from Baseline of 1.2 at Week 12 between a andecaliximab treated group and the placebo group, assuming a common standard deviation of 1.35 and 15% early dropout rate.

The study was terminated after 15 subjects had been enrolled. Therefore, no inferential analyses will be performed due to the small sample size.

2.   TYPE OF ANALYSIS

2.1.   Interim Analyses

Administrative interim analyses of efficacy and safety data were performed by the Gilead internal unblinded team. The purpose of the analyses was to assess efficacy of andecaliximab for further planning and development as a single agent and in combination. This unblinded team was independent of the GS-US-373-1499 study team who were not directly involved in routine study conducts. People who were unblinded were documented in the unblinded form GF-27011A, and the justification was documented in GT-27011A per SOP-BM-27011. The Study Team has remained blinded to treatment assignments throughout the trial until the database has been locked and unblinded.

2.2.   Final Analysis

Due to early termination of the study by the sponsor, the final analysis of the data will be performed for all enrolled subjects as of 05 June 2017. Final analysis will occur after outstanding data queries have been resolved or adjudicated as unresolvable, and the data have been cleaned and finalized.

3.   GENERAL CONSIDERATIONS FOR DATA ANALYSES

Analysis results will be presented using descriptive statistics. For categorical variables, the number and percentage of subjects in each category will be presented; for continuous variables, the number of subjects (n), mean, standard deviation (SD) or standard error (SE), median, first quartile (Q1), third quartile (Q3), minimum, and maximum will be presented.

By-subject listings will be presented for all subjects in the All Randomized Analysis Set and sorted by treatment group, subject identification (ID) number, visit date, and time (if applicable). Data collected on log forms, such as AEs, will be presented in chronological order within the subject. The treatment group to which subjects were randomized will be used in the listings. Age, sex at birth, race, and ethnicity will be included in the listings.

3.1.   Analysis Sets

Analysis sets define the subjects to be included in an analysis. Analysis sets and their definitions are provided in this section. The analysis set will be identified and included as a subtitle of each TFL.

For each analysis set, the number and percentage of subjects eligible for inclusion will be summarized by treatment group.

3.1.1.   All Randomized Analysis Set

All Randomized Analysis Set includes all subjects who are randomized in the study. This is the primary analysis set for by-subject listings.

3.1.2.   Full Analysis Set

The Full Analysis Set (FAS) includes all randomized subjects who received at least 1 dose of study drug. The study drug in this study is andecaliximab and placebo-to-match for andecaliximab. This is the primary analysis set for efficacy analyses.

3.1.3.   Safety Analysis Set

The Safety Analysis Set includes all subjects who received at least 1 dose of study drug. This is the primary analysis set for safety analyses.

3.2.   Subject Grouping

For analyses based on the All Randomized Analysis Set, FAS, subjects will be grouped according to the treatment to which they were randomized. For analyses based on the Safety Analysis Set, subjects will be grouped according to actual treatment received. The actual treatment received will differ from the randomized treatment only when their actual treatment differs from randomized treatment for the entire treatment duration.

3.3.   Strata and Covariates

No stratification factors or covariates will be used in the analyses supporting the synoptic CSR.

3.4.   Examination of Subject Subgroups

There will be no subgroup analyses for efficacy or safety.

3.5.   Multiple Comparisons

Adjustments for multiplicity will not be applied.

3.6.   Missing Data and Outliers

3.6.1.   Missing Data

In general, missing data will not be imputed unless methods for handling missing data are specified. Exceptions are presented in this document.

For missing last dosing date of study drug, imputation rules are described in Section 4.2.1. The handling of missing or incomplete dates for AE onset is described in Section 7.1.5.2.

3.6.2.   Outliers

Outliers will be identified during the data management and data analysis process, but no sensitivity analyses will be conducted. All data will be included in the data analyses.

3.7.   Data Handling Conventions and Transformations

In general, age (in years) on the date of the first dose of study drug will be used for analyses and presentation in listings. If an enrolled subject was not dosed with any study drug, the randomization date will be used instead of the first dosing date of study drug. For screen failures, the date the last informed consent was signed will be used for age calculation. If only birth year is collected on the case report form (CRF), "01 January" will be used for the unknown birth day and month for the purpose of age calculation. If only birth year and month are collected, "01" will be used for the unknown birth day.

Non-PK Data that are continuous in nature but are less than the lower limit of quantitation (LOQ) or above the upper LOQ will be imputed as follows:

- A value that is 1 unit less than the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "< x" (where x is considered the LOQ). For example, if the values are reported as < 50 and < 5.0, values of 49 and 4.9, respectively, will be used to calculate summary statistics. An exception to this rule is any value reported as < 1 or < 0.1, etc. For values reported as < 1 or < 0.1, a value of 0.9 or 0.09, respectively, will be used for calculate summary statistics.

- A value that is 1 unit above the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "> x" (where x is considered the LOQ). Values with decimal points will follow the same logic as above.

- The LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "≤ x" or "≥ x" (where x is considered the LOQ).

3.8.   Analysis Visit Windows

3.8.1.   Definition of Study Day

Study Day 1 is defined as the first dose date of study drug administration. The last dose date for the study drug will be the end date on study drug administration CRF for the record where the "study drug was permanently withdrawn" flag is "Yes".

Study Day will be calculated from the Study Day 1 and derived as follows:

- For postdose study days: Assessment Date – Study Day 1 + 1

- For days prior to the first dose: Assessment Date – Study Day 1

For the outputs based on the study period (double-blind and OLE), the following rule will be applied:

- If subject continues into the OLE then the double-blind (DB) period reports will include the data collected prior to the date of the first dose of study drug in the OLE. The OLE summaries will include the data from the date of the first dose of OL treatment, inclusive.

- If subject does not continue into the OLE then all data collected starting from Day 1 will be included in the DB period summaries.

3.8.2.   Analysis Visit Windows

Subject visits may not occur on protocol-specified days. Therefore, for the purpose of analysis, observations will be assigned to analysis windows.

In general, the baseline value will be the last nonmissing value on or prior to the first dose date of study drug.

The analysis windows for joint count assessment, health assessment questionnaire - disability index (HAQ-DI), physician's global assessment of disease activity (PhGA), patient's global assessment of disease activity (PtGA), patient's pain assessment, on-treatment CRP, SF-36, weight and vital signs are provided in Table 3-1 The analysis windows for safety laboratory data is provided in Table 3-2

Table 3-1.   **Analysis Visit Windows for Joint Count Assessment, HAQ-DI, PhGA, PtGA, Pain, On-Treatment CRP, SF-36, Weight and Vital Signs for Double-Blind Treatment**

Nominal Visit	Analysis Visit	Nominal Study Day	Lower Limit	Upper Limit
Baseline	DB Baseline	1	(none)	1
Week 1	DB Week 1	8	2	18
Week 4	DB Week 4	29	19	43
Week 8	DB Week 8	57	44	71
Week 12	DB Week12	85	72	Prior to or on the first dose date of OL if applicable

Table 3-2.   **Analysis Visit Windows for Safety Laboratory Data for Double-Blind and Open-Label Treatment**

Nominal Visit	Analysis Visit	Nominal Study Day	Lower Limit	Upper Limit
Baseline	DB Baseline	1	(none)	1
Week 1	DB Week 1	8	2	18
Week 4	DB Week 4	29	19	57
Week 12	DB Week 12	85	58	Prior to or on the first dose date of OL if applicable
Week 13	OL Week 1	OL Day 8*	OL Day 2*	$\geq$ OL Day 8*

* OL Day (for post OL dose) is defined as: Assessment Date – First OL Dose Date + 1

Data collected on 30-day follow-up visit will be summarized as a separate visit, and labeled as "30-Day Follow-up Visit". Data obtained after last dose date plus 30 days will be excluded from the summaries, but will be included in the listings.

3.8.3.   Selection of Non-Efficacy Data in the Event of Multiple Records in an Analysis Visit Window

Depending on the statistical analysis method, single values may be required for each analysis window. For example, change from baseline by visit usually requires a single value, whereas a time-to-event analysis would not require 1 value per analysis window.

If multiple valid, nonmissing, continuous measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- In general, the baseline value will be the last nonmissing value on or prior to the first dose date of study drug, unless otherwise specified. If multiple measurements occur on the same day, the last nonmissing value prior to the time of first dose of study drug will be considered as the baseline value. If these multiple measurements occur at the same time or the time is

not available, the average of these measurements (for continuous data) will be considered the baseline value.

- For postbaseline visits:

  — The record closest to the nominal day for that visit will be selected.

  — If there are 2 records that are equidistant from the nominal day, the later record will be selected.

  — If there is more than 1 record on the selected day, the average will be taken, unless otherwise specified.

If multiple valid, nonmissing, categorical measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- For baseline, the last available record on or prior to the date of the first dose of study drug will be selected. If there are multiple records with the same time or no time recorded on the same day, the value with the lowest severity will be selected.

- For postbaseline visits:

  — The record closest to the nominal day for that visit will be selected.

  — If there are 2 records that are equidistant from the nominal day, the later record will be selected.

  — If there is more than 1 record on the selected day, the value with the worst severity will be taken, unless otherwise specified.

The rules for selecting efficacy data in the event of multiple records in an analysis visit window are specified in Section 6.1.

4.   SUBJECT DISPOSITION

4.1.   Subject Enrollment and Disposition

A summary of subject enrollment will be provided by treatment group for each country, investigator within a country, and overall. The summary will present the number and percentage of subjects enrolled. For each column, the denominator for the percentage calculation will be the total number of subjects analyzed for that column.

A summary of subject disposition will be provided by treatment group. This summary will present the number of subjects screened, the number of subjects randomized, and the number of subjects in each of the categories listed below:

- Safety Analysis Set

- Full Analysis Set

- Completed study drug

- Did not complete study drug with reasons for premature discontinuation of study drug

- Completed study

- Did not complete study with reasons for premature discontinuation from the study

For the status of study drug and study completion and reasons for premature discontinuation, the number and percentage of subjects in each category will be provided. The denominator for the percentage calculation will be the total number of subjects in the Safety Analysis Set corresponding to that column.

4.2.   Extent of Study Drug Exposure and Adherence

Extent of exposure to study drug will be examined by assessing the total duration of exposure to study drug.

4.2.1.   Duration of Exposure to Study Drug

Total duration of exposure to any study drug will be defined as (last dose date of any study drug - first dose date of any study drug + 1), regardless of any temporary interruptions in study drug administration and will be expressed in weeks using up to 1 decimal place (eg, 4.5 weeks).

If the last study drug dosing date is missing, the latest date among the study drug end date, clinical visit date, laboratory sample collection date, weight and vital signs assessment date that occurred during the on-treatment period will be used.

The total duration of exposure to any study drug will be summarized using descriptive statistics and using the number (ie, cumulative counts) and percentage of subjects exposed through the following time periods: Baseline (Day 1), Week 1 (Day 8), Week 4 (Day 29), Week 8 (Day 57), and Week 12 (Day 85).

Summaries will be provided by treatment group for the Safety Analysis Set. No formal statistical testing is planned.

4.2.2.   Adherence to Study Drug

A summary of adherence to study drug will not be provided.

A by-subject listing of study drug administration will be provided by subject ID number (in ascending order) and visit (in chronological order).

4.3.   Protocol Deviations

A by-subject listing will be provided for those subjects who did not meet at least 1 eligibility (inclusion or exclusion) criterion. The listing will present the eligibility criterion (or criteria if more than 1 deviation) that subjects did not meet and related comments, if collected.

5.   BASELINE CHARACTERISTICS

5.1.   Demographics

Subject demographic variables will be summarized by treatment group and overall using descriptive statistics for continuous variables, and using number and percentage of subjects for categorical variables. The summary of demographic data will be provided for the Safety Analysis Set for the following:

- age (on the first dose date of any study drug) as a continuous variable

- sex at birth (male, female)

- race

- ethnicity (Hispanic or Latino, not Hispanic or Latino)

- country

A by-subject demographic listing, including the informed consent date, will be provided by treatment group and subject ID number in ascending order.

5.2.   Other Baseline Characteristics

Other baseline characteristics include:

- height (cm)

- weight (kg)

- body mass index (BMI; in kg/m^2)

- swollen joint count based on 66 joints (SJC66) and tender joint count based on 68 joints (TJC68)

- HAQ-DI

- patient's pain assessment

- DAS28(CRP)

- patient's global assessment of disease activity

- physician's global assessment of disease activity

- Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI)

- SF-36 (physical component summary score and mental component summary score)

- CRP (mg/L)

- stratification factor (DAS28[CRP] > 5.1 or DAS28[CRP] > 3.2 and ≤ 5.1; prior number of RA biologics used < 3 or ≥ 3)

- duration of RA (years)

  duration of RA (years) = ((first dose date) – (date of initial diagnosis) + 1 day) / 365.25. If the date of initial diagnosis is incomplete, then the following rules will be applied:

  — missing day: use the first of the month.

  — missing month: use January.

These baseline characteristics will be summarized by treatment group and overall using descriptive statistics for continuous variables and using number and percentage of subjects for categorical variables. The summary of baseline characteristics will be provided for the Safety Analysis Set.

A by-subject listing of other baseline characteristics will be provided by treatment group and subject ID number in ascending order.

5.3.   Medical History

The summaries of medical history will not be provided.

6.   EFFICACY ANALYSES

6.1.   General Considerations

The primary analysis set for efficacy analyses will be the FAS, defined in Section 3.1.2.

If multiple valid, nonmissing efficacy measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- The record closest to the nominal day for that visit will be selected

- If there are 2 records that are equidistant from the nominal day, or more than 1 record (with time known) on the selected day, the latest record will be taken

- If chronological order cannot be determined (eg, more than 1 record on the same day with time missing), the worst scenario will be selected

6.1.1.   Tender/Swollen Joint Counts (TJC/SJC)

TJC68 and SJC66 will be collected during the course of the study. The assessment for each joint will be from the following selections: Tender Only, Swollen Only, Tender and Swollen, Not Tender or Swollen, or Joint Non-Evaluable or Missing.

Individual joint with missing assessment will not be imputed. If at least half of the joints are assessed at a given visit, the *prorated* tender and swollen joint counts will be calculated using the following formula:

$$TJC68 = \frac{Total\ number\ of\ tender\ joints}{68 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 68\ joints)} \times 68$$

$$SJC66 = \frac{Total\ number\ of\ swollen\ joints}{66 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 66\ joints)} \times 66$$

If less than half of joints are assessed at a given visit, joint counts are treated as missing for that visit.

A more abbreviated assessment considering 28 joints as listed in Table 6-1 for both tenderness and swelling will also be conducted (as part of the TJC68 and SJC66 assessment), denoted as TJC28 and SJC28, respectively.

Table 6-1.   **Composition of the 28 Joints**

Joints	Number
Shoulder Joints (Left and Right)	2
Elbow Joints (Left and Right)	2
Wrist Joints (Left and Right)	2
Metacarpophalangeal Joints I-V (Left and Right)	10
Proximal Interphalangeal Joints I-V (Left and Right)	10
Knee Joints (Left and Right)	2

If there are non-evaluable or missing joints among the 28 joints, similar prorated tender and swollen joint counts will be calculated as follows:

$$TJC28 = \frac{Total\ number\ of\ tender\ joints}{28 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 28\ joints)} \times 28$$

$$SJC28 = \frac{Total\ number\ of\ swollen\ joints}{28 - (Number\ of\ nonevaluable\ or\ missing\ joints\ out\ of\ 28\ joints)} \times 28$$

If less than half of the 28 joints are assessed at a given visit, TJC28 and SJC28 are treated as missing for that visit.

6.1.2.   Global Assessment of Disease Activity

The Patient's Global Assessment of Disease Activity is a horizontal (0 – 100) visual analog scale (VAS) that ranges from "No Arthritis" to "Severe Arthritis."

The Physician's Global Assessment of Disease Activity is a horizontal (0 – 100) VAS that ranges from "No Disease Activity" to "Maximum Disease Activity."

6.2.   Primary Efficacy Endpoints

6.2.1.   Definition of the Primary Efficacy Endpoints

The primary endpoint for the study is change from baseline in DAS28(CRP) at Week 12. The DAS28(CRP) score is calculated as follows:

$$DAS28(CRP) = 0.56\sqrt{TJC28} + 0.28\sqrt{SJC28} + 0.36\ln(CRP + 1) + 0.014 \times PtGA + 0.96,$$

where

> TJC28 = number of joints tender out of 28 joints

> SJC28 = number of joints swollen out of 28 joints

> CRP = CRP measurement in unit of mg/L

> PtGA = patient's global assessment of disease activity on a 0-100 VAS

Higher DAS28(CRP) value indicates more severe disease activity.

DAS28(CRP) will be calculated only for the time points with all measurements available, ie, missing DAS28(CRP) components will not be imputed.

6.2.2.   Statistical Hypotheses for the Primary Efficacy Endpoints

No statistical hypotheses will be tested due to early termination of the study.

6.2.3.   Analysis of the Primary Efficacy Endpoint

Due to early termination of the study by the sponsor, no formal statistical testing is planned for the final analysis. Primary efficacy endpoint data will be summarized using descriptive statistics by treatment group and visit.

6.3.   Secondary Efficacy Endpoints

The secondary efficacy endpoints are:

- Proportion of subjects that achieve DAS28(CRP) $\leq$ 3.2 at Week 12

- Proportion of subjects that achieve DAS28(CRP) < 2.6 at Week 12

- Plasma concentrations of andecaliximab

6.3.1.   Analysis Methods for Secondary Efficacy Endpoints

Due to the early termination of the study, no formal statistical testing is planned for the final analysis. Secondary efficacy endpoint data will be summarized using descriptive statistics by treatment group and analysis visit. The plasma concentrations of andecaliximab were not collected and will not be analyzed.

6.4.   Exploratory Efficacy Endpoints

PPD

6.4.1.   Definition of Exploratory Efficacy Endpoints

6.4.1.1.   SDAI and CDAI

PPD

6.4.1.2.   Health Assessment Questionnaire Disability Index (HAQ-DI)

The HAQ-DI score is defined as the average of the scores of eight functional categories (dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities), administered by the patient. Responses in each functional category are collected as: without any difficulty; with some difficulty; with much difficulty; unable to do a task in that area; and with or without aids or devices. The HAQ-DI score ranges from 0 (no disability) to 3 (completely disabled), when 6 or more categories are non-missing. Detailed algorithm for calculating HAQ-DI score is described in Appendix 1.

The HAQ-DI also includes a separate pain assessment and the subject will be requested to mark the severity of the pain in the past week on a 0-100 VAS, with 0 indicating "no pain" and 100 indicating "severe pain."

6.4.1.3.   36-Item Short-form Health Survey

The SF-36 (version 2) is a health related quality of life instrument consisting of 36 questions belonging to 8 domains in 2 components and covers a 4-week recall period:

- Physical component: 4 domains: physical functioning (10 items), role physical (4 items), bodily pain (2 items), and general health perceptions (5 items)

- Mental component: 4 domains: vitality (4 items), social functioning (2 items), role emotional (3 items), and mental health (5 items).

The physical component and mental component will be scored and presented.

6.4.2.   Analysis Methods for Exploratory Efficacy Endpoints

PPD

7.   SAFETY ANALYSES

7.1.   Adverse Events and Deaths

7.1.1.   Adverse Event Dictionary

Clinical and laboratory adverse events (AEs) will be coded using the current version of the Medical Dictionary for Regulatory Activities (MedDRA). System organ class (SOC), high-level group term (HLGT), high-level term (HLT), preferred term (PT), and lower-level term (LLT) will be provided in the AE dataset.

7.1.2.   Adverse Event Severity

Adverse events are graded by the investigator as Grade 1, 2, 3, 4, or 5 according to toxicity criteria specified in the protocol. The severity grade of events for which the investigator did not record severity will be categorized as "missing" for tabular summaries and data listings. The missing category will be listed last in summary presentation.

7.1.3.   Relationship of Adverse Events to Study Drug

Related AEs are those for which the investigator selected "Related" on the AE CRF to the question of "Related to Study Treatment." Relatedness will always default to the investigator's choice, not that of the medical monitor. Events for which the investigator did not record relationship to study drug will be considered related to study drug for summary purposes. However, by-subject data listings will show the relationship as missing.

7.1.4.   Serious Adverse Events

Serious adverse events (SAEs) will be identified and captured as SAEs if AEs met the definitions of SAE specified in the study protocol. SAEs captured and stored in the clinical database will be reconciled with the SAE database from the Gilead Drug Safety and Public Health (DSPH) Department before data finalization.

7.1.5.   Treatment-Emergent Adverse Events

7.1.5.1.   Definition of Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as one or both of the following:

- Any AEs with an onset date on or after the study drug start date and no later than 30 days after permanent discontinuation of study drug

- Any AEs leading to premature discontinuation of study drug

7.1.5.2.   Incomplete Dates

If the onset date of the AE is incomplete and the AE stop date is not prior to the first dosing date of study drug, then the month and year (or year alone if month is not recorded) of onset determine whether an AE is treatment emergent. The event is considered treatment emergent if both of the following 2 criteria are met:

- The AE onset is the same as or after the month and year (or year) of the first dosing date of study drug, and

- The AE onset date is the same as or before the month and year (or year) of the date corresponding to 30 days after the date of the last dose of study drug

An AE with completely missing onset and stop dates, or with the onset date missing and a stop date later than the first dosing date of study drug, will be considered to be treatment emergent. In addition, an AE with the onset date missing and incomplete stop date with the same or later month and year (or year alone if month is not recorded) as the first dosing date of study drug will be considered treatment emergent.

7.1.6.   Summaries of Adverse Events and Deaths

Treatment-emergent AEs will be summarized based on the Safety Analysis Set.

The AEs will be reported by the treatment period. If subject does not continue into the OLE then all TEAEs will be reported in the DB treatment period (Day 1 to Week 12) summaries. For subjects who continue into the OLE, TEAEs during the DB treatment period are the AEs that started prior to the first dose of OL treatment. The AEs with the onset date on or after the first dose of the OL treatment and no later than 30 days after permanent discontinuation of study drug will be reported in the OLE summaries.

7.1.6.1.   Summaries of AE incidence in Combined Severity Grade Subsets

For TEAEs described below, summaries will be provided by SOC, PT, treatment group, and study phase:

- All TE treatment-related AEs

- All TE SAEs

- All TEAEs leading to premature discontinuation of study drug

- TEAEs of interest:

    — injection site reactions (based on Gilead medical search terms provided by DSPH)

A brief, high-level summary of AEs described above will also be provided by treatment group and study phase.

In addition to the above summary tables, all TEAEs will be summarized by PT only, in descending order of total frequency. Summaries of all TEAEs by SOC, PT, and maximum severity will also be provided.

Multiple events will be counted only once per subject in each summary. Adverse events will be summarized and listed first in alphabetic order of SOC and PT in descending order of total frequency within each SOC. For summaries by severity grade, the most severe grade will be used for those AEs that occurred more than once in an individual subject during the study.

In addition, data listings will be provided for the following:

- All AEs, indicating whether the event is treatment-emergent

- SAEs

- Deaths

- AEs leading to premature discontinuation of study drug

7.2.   Laboratory Evaluations

A by-subject listing for laboratory test results will be provided by subject ID number and visit in chronological order for hematology and serum chemistry separately. Values falling outside of the relevant reference range and/or having a severity grade of 1 or higher on the common terminology criteria for adverse events (CTCAE) severity grade will be flagged in the data listings, as appropriate.

No formal statistical testing is planned.

7.2.1.   Summaries of Numeric Laboratory Results

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.8.3.

Descriptive statistics for laboratory tests will not be presented.

7.2.2.   Graded Laboratory Value

The CTCAE Version 4.03 will be used to assign toxicity grades (0 to 5) to laboratory results for analysis. Grade 0 includes all values that do not meet the criteria for an abnormality of at least Grade 1. For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (ie, increased, decreased) will be presented separately.

7.2.2.1.   Treatment-Emergent Laboratory Abnormalities

Treatment-emergent laboratory abnormalities are defined as values that increase at least
1 toxicity grade from baseline at any postbaseline time point, up to and including the date of last
dose of any study drug plus 30 days for subjects who permanently discontinued study drug.

If the relevant baseline laboratory value is missing, any abnormality of at least Grade 1 observed
within the time frame specified above will be considered treatment emergent.

7.2.2.2.   Summaries of Laboratory Abnormalities

The summaries (number and percentage of subjects) for treatment-emergent laboratory
abnormalities will be provided by lab test, treatment group and treatment period; subjects will be
categorized according to the most severe postbaseline abnormality grade for a given lab test. If
subject does not continue into the OLE then all treatment-emergent laboratory abnormalities will
be reported in the DB treatment period (Day 1 to Week 12) summary. For subjects who continue
into the OLE, treatment-emergent laboratory abnormalities during the DB treatment period are
the laboratory abnormalities that started prior to the first dose of OL treatment. The laboratory
abnormalities with the onset date on or after the first dose of the OL treatment and no later than
30 days after permanent discontinuation of study drug will be reported in the OLE summary.

A by-subject listing of treatment-emergent Grade 3 or higher laboratory abnormalities will be
provided by subject ID number and visit in chronological order. This listing will include all test
results that were collected throughout the study for the lab test of interest, with all applicable
severity grades displayed.

7.2.3.   Liver-related Laboratory Evaluations

Liver-related laboratory evaluations will be summarized as part of chemistry results.

7.3.   Body Weight and Vital Signs

Body weight and vital signs will not be summarized or listed.

7.4.   Prior and Concomitant Medications

Prior and concomitant medications will not be summarized or listed.

7.5.   Electrocardiogram Results

Electrocardiogram (ECG) data will not be summarized or listed.

7.6.   Other Safety Measures

A data listing will be provided for subjects who become pregnant during the study.

7.7.   Changes From Protocol-Specified Safety Analyses

Only limited safety summaries are provided due to low number of subjects enrolled in the study
prior to the study termination.

8.   PHARMACOKINETIC ANALYSES

No PK analysis will be conducted due to the study early termination.

9.   BIOMARKER ANALYSIS

No biomarker analysis will be conducted due to the study early termination.

10. REFERENCES

Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clinical and experimental rheumatology 2005;23 (5 Suppl 39):S100-8.

11.   SAP REVISION

Revision Date (DD MMM YYYY)	Section	Summary of Revision	Reason for Revision

12.   APPENDIX

Appendix 1.   Health Assessment Questionnaire Disability Index (HAQ-DI)

The HAQ-DI score is defined as the average of the scores of eight functional categories (dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities), usually administered by the patient. Responses in each functional category are collected as 0 (without any difficulty) to 3 (unable to do a task in that area), with or without aids or devices.

The highest score for questions in each category (range 0 to 3) determines the score for the category, unless aids or devices are required. Dependence on equipment or physical assistance increases a lower score (i.e., scores of 0 or 1) to the level of 2 to more accurately represent underlying disability. The eight category scores are averaged into an overall HAQ-DI score on a scale from 0 (no disability) to 3 (completely disabled) when 6 or more categories are non-missing. If more than 2 categories are missing, the HAQ-DI score is set to missing. The HAQ-DI can be treated as a continuous measure.

The HAQ-DI score using aids (and/or devices) is computed by taking the maximum score of the questions in each category (range: [0, 3]) and whether or not aids/devices are used (0 or 1):

A =   max(dressing & grooming category questions, 2*aids indicator) +
      max(rising category questions, 2*aids indicator) +
      max(eating category questions, 2*aids indicator) +
      max(walking category questions, 2*aids indicator) +
      max(hygiene category questions, 2*aids indicator) +
      max(reach category questions, 2*aids indicator) +
      max(grip category questions, 2*aids indicator) +
      max(usual activities category questions, 2*aids indicator)

HAQ-DI = A/(total number of categories with at least 6 non-missing)

The following table shows the contribution of the 43 questions used to calculate the HAQ-DI:

| HAQ-DI Category: | Category questions: At least 6 categories must have scores to compute the HAQ-DI. | | HAQ-DI Category Score with Aids/Devices Calculation: |
	Category Questions	Aids/Devices Indicators	
Dressing / Grooming	HAQ0101, HAQ0102 (DRESS, HAIR)	HAQ0114, HAQ0119 (DRSG, GROOM)	Using each question with a scale of 0-3, calculate the category score as the maximum of the category questions. If the Aids/Devices indicator is "No", no need to adjust the category score. If the Aids/Devices indicator is "Yes" and the category score is <2, then the category score with the Aids/Devices is set to 2. If the Aids/Devices indicator is "Yes" and the category score is ≥2, then the category score with Aids/Devices is the calculated category score without adjustment. For example: The Dressing/Grooming category score is 2 if subject answered 1 for both questions 1 and 2 and "Yes" for both question 14 and 18. In the HAQ-DI score calculation, questions on other device/aids will not be used.
Arising	HAQ0103, HAQ0104 (STAND, BED)	HAQ0116, HAQ0120 (CHAIR, ARISING)	
Eating	HAQ0105,HAQ0106, HAQ0107 (MEAT, LIFT,MILK)	HAQ0115, HAQ0121 (UTENSIL, EAT)	
Walking	HAQ0108, HAQ0109 (WALK, STEPS)	HAQ0110, HAQ0111, HAQ0112, HAQ0113, HAQ0122 (CANE, WALKER, CRUTCH, WHEEL, WALKING)	
Hygiene	HAQ0123, HAQ0124, HAQ0125 (WASH, BATH, TOILET)	HAQ0134, HAQ0135, HAQ0137, HAQ0139, HAQ0142 (RAISEAT, BATHBAR, BATHSEAT, LONGBATH, HYGIENE)	
Reach	HAQ0126, HAQ0127 (REACH, BEND)	HAQ0138, HAQ0143 (LONGRCH, REACH)	
Grip	HAQ0128, HAQ0129, HAQ0130 (OPENCAR, JAR, FAUCET)	HAQ0136, HAQ0144 (JAROPEN, GRIP)	
Activity	HAQ0131, HAQ0132, HAQ0133 (SHOP, INCAR, CHORES)	HAQ0145 (ERRAND)	

Handling Missing Data: If no more than 2 categories have missing category scores, then the HAQ-DI is the mean of the non-missing category scores. Otherwise, the HAQ-DI score is set to missing.

If any of the category questions are missing, but the aids/device indicator is non-missing, the category score can still be computed. However, if all category questions and its aids/device indicators are missing, then the category score is considered missing.