

PROTOCOL B3281006

**A PHASE 3, RANDOMIZED, DOUBLE-BLIND STUDY OF PF-05280586
VERSUS RITUXIMAB FOR THE FIRST-LINE TREATMENT OF PATIENTS
WITH CD20-POSITIVE, LOW TUMOR BURDEN, FOLLICULAR LYMPHOMA**

**STATISTICAL ANALYSIS PLAN
(SAP)**

Version: 3.0 (Amendment 2)

Author: PPD [REDACTED]

PPD [REDACTED]

Date: 13 November 2017

TABLE OF CONTENTS

LIST OF TABLES3

1. AMENDMENTS FROM PREVIOUS VERSION(S)4

2. INTRODUCTION5

 2.1. Study Design5

 2.2. Study Objectives5

3. INTERIM ANALYSES, FINAL ANALYSES AND UNBLINDING5

4. HYPOTHESES AND DECISION RULES6

 4.1. Statistical Hypotheses6

 4.2. Statistical Decision Rules6

 4.3. Sample Size7

 4.4. External Data Monitoring Committee7

5. ANALYSIS SETS8

 5.1. Intent-to-Treat Population8

 5.2. Modified Intent-to-Treat Population8

 5.3. Safety Population8

 5.4. ‘Per-Protocol’ Analysis Population8

 5.5. Response-Evaluable Population9

 5.6. Pharmacokinetic Analysis Set9

 5.7. Other Treatment Misallocations9

 5.8. Protocol Deviations9

6. ENDPOINTS AND COVARIATES10

 6.1. Endpoints10

 6.1.1. Primary Endpoint10

 6.1.2. Secondary Endpoints10

 6.2. Covariates11

7. HANDLING OF MISSING VALUES11

 7.1. Overall Response, Complete remission, and Partial Response Data11

 7.2. *Time to Event Data*11

 7.3. Pharmacokinetic Concentrations and Biomarker Data11

 7.4. Missing Dates12

8. STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES12

 8.1. Statistical Methods12

8.1.1. Analysis of Response Data	12
8.1.2. Analysis of Time to Event Data.....	12
8.2. Statistical Analyses	13
8.2.1. Standard Analyses	13
8.2.1.1. Patient Disposition	13
8.2.1.2. Patient Demographics	13
8.2.1.3. Medical History.....	13
8.2.1.4. Baseline Disease Characteristics.....	13
8.2.1.5. Duration of Follow-up.....	13
8.2.1.6. Drug exposure, Treatment Duration and Compliance	14
8.2.1.7. Prior and Concomitant Medications.....	14
8.2.2. Analyses of the Primary Efficacy Endpoint	14
8.2.3. Analyses of the Secondary Efficacy Endpoint	15
8.2.3.1. 1-Year Progression free survival.....	15
8.2.3.2. 1-Year Overall Survival	16
8.2.3.3. Time to Treatment Failure.....	16
8.2.3.4. Duration of Response.....	16
8.2.3.5. Complete Remission Rate at Week 26.....	16
8.2.4. Analyses of Pharmacokinetic Endpoints	17
8.2.5. Safety Analyses	17
CCI	17
8.2.5.2. Adverse Events.....	17
8.2.5.3. Laboratory Abnormalities	18
8.2.5.4. Vital Signs Abnormalities	18
8.2.6. Biomarker Analysis	18
8.2.7. Immunogenicity Assessment.....	19
8.2.8. Subgroup Analysis.....	19
9. APPENDICES	20
10. REFERENCES	40

LIST OF TABLES

Table 1 Summary of Major Changes in SAP Amendments	4
--	---

1. AMENDMENTS FROM PREVIOUS VERSION(S)

This version (3.0) is amendment 2 of the statistical analysis plan (SAP) for the study B3281006, based on Amendment 4 of the protocol B3281006 dated Apr. 19, 2016 [1]. This amendment is being made just prior to the database snapshot and unblinding associated with the primary completion date for the trial. The purpose is to detail analyses added since the last SAP version, and also address needed clarifications that have arisen during development of the analysis tables.

Table 1 Summary of Major Changes in SAP Amendments

SAP Version	Change and/or Rationale
1.0	Not Applicable
2.0	<ul style="list-style-type: none"> • Made editorial changes to align with Protocol Amendment 4. • Added equivalence criteria for the Japan regulatory authority. • Added additional detail on the unblinding plan for the study. • Added the mITT population for the analysis of biomarker data. • Defined Tier 1 adverse events. • Confirmed that the assessment of efficacy will be based on central review data. • Modified censoring rules. • Added analysis of anaphylactic reactions based on the criteria of Sampson. • Added subgroup analyses of efficacy. • Modified the time window for analysis.
3.0	<ul style="list-style-type: none"> • Clarified exclusion criteria for the Per Protocol Analysis Population. • Added potentially important protocol deviations. • Clarified the use of visit designations and dates in the imaging central reader dataset. • Added that the stratified Mantel-Haenszel method will be used to obtain the corresponding estimated treatment group difference for the analysis of response data (including the primary endpoint). • Updated the censoring rules for progression free survival (Table 2). • Added that serum drug concentrations will also be summarized by ADA status. • Added the summarization for Clinical Outcomes Associated with Immunogenicity (COAI). • Made clarifications to the section on immunogenicity assessment. • Added additional subgroup analyses for Ann Arbor Staging, and bone marrow involvement. Clarified that the efficacy analyses within the Japan subgroup of patients will not be stratified FLIPI2. • Updated the adverse event Tier-1 search terms, and cutoff for Tier-2. • Added the statistical specifications for programmatically determining which events are consistent with Sampson's Criteria.

As a note, in this document any text taken directly from the protocol is *italicized*.

2. INTRODUCTION

2.1. Study Design

This is a double-blind, randomized, Phase 3 clinical trial evaluating the efficacy, safety, PK and immunogenicity of PF-05280586 (also described as rituximab-Pfizer in the protocol) versus rituximab-EU in patients with CD20-positive, low tumor burden, follicular lymphoma in the first-line treatment setting. Low tumor burden will be assessed according to the Groupe d'Etude des Lymphomes Folliculaires (GELF) criteria modified to allow slightly abnormal serum LDH and β 2-microglobulin levels. The hypothesis to be tested in this study is that the efficacy (ORR) of PF-05280586 is similar to that of rituximab-EU. Retrospective histological confirmation of CD20-positive FL will be obtained by a central pathology review. [Note that the central lab readings will be used in analysis although local lab readings may be used for patient enrollment.] Central imaging review will be performed for all disease assessments up through Week 52. The primary endpoint is Overall Response Rate (ORR) at Week 26 in accordance with the revised response criteria for malignant lymphoma.⁵ Secondary endpoints include safety, time to treatment failure (TTF), progression free survival (PFS), complete remission (CR), duration of response (DoR), overall survival (OS), selected peak and trough drug concentrations, CD19-positive B-cell depletion, and immunogenicity. Approximately 394 Patients will be randomized in a 1:1 ratio to receive PF-05280586 or rituximab-EU. Randomization will be stratified by low, medium, and high risk patients using the Follicular Lymphoma International Prognostic Index 2 (FLIPI2).⁷ During the study, patients will receive 4 weekly doses of PF-05280586 or rituximab-EU administered via intravenous infusion. The dose of PF-05280586 or rituximab-EU will be 375 mg/m² of body surface area. [Source: Protocol [1], Section 3]

2.2. Study Objectives

Primary Objectives

- *To compare the efficacy of PF-05280586 to rituximab-EU when administered as a first-line treatment to patients with CD20-positive, low tumor burden follicular lymphoma (LTB-FL).*

Secondary Objectives

- *To evaluate the safety of PF-05280586 and rituximab-EU;*
- *To evaluate the population pharmacokinetics of PF-05280586 and rituximab-EU;*
- *To evaluate the immunogenicity of PF-05280586 and rituximab-EU;*
- *To characterize CD19-positive B-cell depletion and recovery in patients receiving PF-05280586 and rituximab-EU.*

[Source: Protocol [1] Section 2]

3. INTERIM ANALYSES, FINAL ANALYSES AND UNBLINDING

There is no planned interim analysis in this study.

This double-blind study will be blinded to the patients and investigator/site staff, with the exception of the pharmacy staff preparing study treatment infusions. The study will be conducted in a double blinded fashion through Week 26 when the primary endpoint will be evaluated. Prior to Week 26 the patients, investigators and sponsor will be blinded to randomized study treatments. [Source: Protocol [1] Section 5.2]. The primary completion date (PCD) occurs when last patient completes their Week 26 visit (or withdraws from the study prior to Week 26). A data snapshot will be taken of the database by the cutoff date of the PCD and locked. After the PCD some members of the sponsor study team will be unblinded so that a clinical study report for the PCD snapshot data and submission documents can be generated. The patients, investigators and other blinded sponsor staff or their designees who interact with sites will continue to be blinded to individual study treatments until the end of the study. A detailed unblinding plan will be provided and maintained as a separate document.

The final analysis for all the data will be performed when the last patient last visit has occurred in the follow-up period (Week 52/early termination), and the database is then officially locked and released. The final clinical study report (CSR) will then be generated following the database release.

4. HYPOTHESES AND DECISION RULES

4.1. Statistical Hypotheses

The primary endpoint is the Overall Response Rate (ORR) at Week 26. The primary hypothesis for the primary endpoint is:

$$\text{TEST 1: } H_{0c}: \theta_1 - \theta_2 \geq D_{ub} \text{ vs. } H_{1c}: \theta_1 - \theta_2 < D_{ub}$$

$$\text{TEST 2: } H_{0d}: \theta_1 - \theta_2 \leq D_{lb} \text{ vs. } H_{1d}: \theta_1 - \theta_2 > D_{lb}$$

Where θ_1 is the ORR at Week 26 for patients randomized to PF-05280586, θ_2 is the ORR at Week 26 for patients randomized to rituximab-EU, D_{ub} is the largest acceptable difference for equivalence, and D_{lb} is the smallest acceptable difference for equivalence. In this study, $D_{ub}=16\%$ and $D_{lb}=-16\%$.

According to a requirement from the regulatory authority in Japan, an additional analysis will be conducted to test equivalence using $D_{ub} = 14.9\%$ and $D_{lb} = -14.9\%$ if equivalence is established with the margins of $D_{ub}=16\%$ and $D_{lb}=-16\%$.

[Source: Protocol [1] Section 9.3.1]

4.2. Statistical Decision Rules

Equivalence will be considered established if the 95% confidence interval of the difference (PF-05280586 minus rituximab-EU) in ORR at Week 26 falls into the margins specified above.

4.3. Sample Size

The hypothesis to be tested in this study is that the difference between the ORR of PF-05280586 versus that of Rituximab-EU is within a pre-specified margin of -16% to 16% (the margin derivation is described below). A sample size of approximately 394 patients (~197 per treatment arm) provides approximately 93% power for achieving equivalence under the specified margin with 2.5% type I error rate assuming an ORR of 77% in both treatment arms.

Pfizer conducted an extensive, systematic literature search for rituximab and FL. The Ardeshta (2010)¹ study was the only randomized trial which compared the treatment of Rituximab alone with the treatment of “watchful waiting” (WW). In this study, at Month 7 the response rate to rituximab therapy (weekly for 4 weeks) was estimated to be 77% and the response rate in the WW arm was estimated to be 6%. The difference (rituximab - WW) was estimated to be 71% with the 95% confidence interval of (60%, 79%). Based on these results, the proposed margin of (-16%, 16%) will preserve at least 73% efficacy based on the lower bound of 95% CI in the ORR difference (rituximab-WW) as seen in the Ardeshta study.

[Source: Protocol [1] Section 9.1]

4.4. External Data Monitoring Committee

Safety monitoring will be conducted throughout the study by the Pfizer study team. In addition, this study will use an External Data Monitoring Committee (EDMC).

The EDMC will be a 3-member panel of external experts that will meet at approximate 3-month intervals throughout the course of the study unless safety concerns requiring their attention arise earlier. An EDMC liaison will be appointed; this is an individual who represents Pfizer to coordinate communications and facilitates access to Pfizer’s resources, but is not involved in the study design, study management, site management, data accrual, or study analysis. The SAP will outline plans for data review. An EDMC charter will outline the operating procedures of the committee, including a specific description of the scope of their responsibilities, and a communication plan. Records of EDMC meetings, interactions with Pfizer contacts, assessments and recommendations and materials reviewed will be maintained and kept proprietary and confidential by the EDMC.

The EDMC will be responsible for ongoing monitoring of the safety of patients in the study according to the Charter. The recommendations made by the EDMC to alter the conduct of the study will be forwarded to Pfizer for final decision. Pfizer will forward such decisions, which may include summaries of aggregate analyses of endpoint events and of safety data which are not endpoints, to regulatory authorities, as appropriate. In this instance, such disease-related efficacy endpoints are not reported individually as SAEs.

The EDMC will be responsible for the periodic review of accumulating safety data. The EDMC will advise the Sponsor regarding the safety of patients enrolled in the study.

Additionally, significant findings observed by the Study Team will be communicated to the EDMC for further review and advice.

[Source: Protocol [1] Section 9.8]

5. ANALYSIS SETS

5.1. Intent-to-Treat Population

The Intent-to-Treat (ITT) population is defined as all patients who are randomized. Patients will be analyzed according to the treatment they were randomized to receive, regardless of any errors of dosing. The ITT population will be used for the primary analysis for efficacy data.

5.2. Modified Intent-to-Treat Population

The modified Intent-to-Treat (mITT) population is defined as all patients who are randomized and receive at least 1 dose of any study drug. The mITT population will be used for the primary analysis for biomarker data analyses.

5.3. Safety Population

The safety population is defined as all patients who receive at least 1 dose of any study drug. Patients will be analyzed according to the treatment actually received. Safety population will be used for all safety related analyses such as AE, concomitant medication, laboratory tests, and vital signs.

[Source: Protocol [1] Section 9.2]

5.4. 'Per-Protocol' Analysis Population

The Per-Protocol (PP) Population is defined as all randomized patients who receive at least one dose of study treatment to which they are assigned, have adequate disease assessment at baseline as confirmed by central review, and have no important protocol deviations that would impact the efficacy assessments significantly, as determined by blinded medical review. The PP population will be used for sensitivity analyses of the efficacy. All decisions to exclude patients from the PP population will be made prior to PCD data snapshot and final database release.

[Source: Protocol [1] Section 9.2]

Specifically, a subject will be excluded from the PP Analysis Population if Reasons 1-2 apply.

- 1=Did not take the treatment to which the subject was randomized to;
- 2=No baseline or no adequate baseline disease assessment (based on central review), unless the subject died on/before Week 26.

A subject will be excluded from the PP Analysis Population if Reasons 3-5 apply, unless the subject either died on/before Week 26 or had an Overall Response of progressive disease (confirmed by central review) on/before the Week 26 time point.

- 3=For subjects that had an adequate baseline disease assessment (based on central review): No measurable disease at baseline as assessed by the central reader

selected after the adjudication step.

4=No evaluable Week 26 assessment (based on central review, defined as the following: either missing, UE, or NA);

5=Any concomitant medication violation that could significantly impact the Week 26 Overall Response assessment, which occurred on/before the Week 26 time point.

5.5. Response-Evaluable Population

The response-evaluable population was defined as all patients in the ITT population who receive at least 1 dose of study drug, have adequate disease assessment at baseline, and at least 1 post baseline response assessment. The response-evaluable population will be used for the analysis of duration of response.

5.6. Pharmacokinetic Analysis Set

The pharmacokinetic analysis set (PKAS) will be the subset of patients from the safety analysis set who provide at least one post-dose pharmacokinetic concentration.

5.7. Other Treatment Misallocations

If a patient was:

- Randomized but not treated: the patient will be accounted for in the patient disposition table and listing. The patient will be reported under the randomized treatment group for efficacy analysis. The patient will not be included in safety analyses.
- Treated but not randomized: the patient will be reported under the treatment they actually received for the safety analyses. The patient will not be included in efficacy analyses.
- Randomized but received incorrect treatment: if patient received the incorrect treatment they will be reported under the treatment they actually received for safety analyses; and they will be reported under the randomized treatment group for efficacy analysis.

5.8. Protocol Deviations

Unexpected deviations that arise during study conduct and become known by the sponsor will be assessed by a blinded team on an ongoing basis. The determination of protocol deviations (PDs) and potentially important protocol deviations (IPDs) follows Pfizer standard operating procedures [3] [4]. A full list of PDs and IPDs will be determined by blinded data review prior to the PCD data snapshot and the database release, and will be included in the corresponding CSRs. As of this writing, the following protocol deviations (PDs) may be considered as IPDs that can be potentially excluded from the Per Protocol Analysis (see [Section 5.4](#) above):

- Patients who receive excluded concomitant medications or rescue medications during the treatment period;

- Patients who were randomized but took no treatment or the incorrect treatment;
- Patients who have no measurable disease at Screening based on central review.
- Patients who have a missing (or unevaluable) Screening imaging result or similarly at Week 26 (unless the patient withdrew from the study early or died) based on central review.

6. ENDPOINTS AND COVARIATES

6.1. Endpoints

6.1.1. Primary Endpoint

- *Overall Response Rate (ORR) at Week 26 of PF-05280586 and rituximab-EU based on central review in accordance with the revised response criteria for malignant lymphoma.* ORR is defined as the proportion of patients who achieved either complete remission (CR) or partial response (PR) in accordance with the revised response criteria for malignant lymphoma, with applicable clinical data incorporated into the radiographic response [5].

6.1.2. Secondary Endpoints

- *Safety characterized by type, incidence, severity, timing, seriousness, and relationship to study therapy of adverse events and laboratory abnormalities;*

CCI



- *Time to Treatment Failure (TTF);*
- *Progression-Free Survival (PFS);*
- *Complete Remission (CR) rate at Week 26;*
- *Duration of response;*
- *Overall survival;*
- *Peak and trough PF-05280586 and rituximab-EU concentrations;*
- *CD19-positive B-cell counts;*

- *Incidence of anti-drug antibodies (ADA), including neutralizing antibodies (NAb), and safety associated with immune response.*

[Source: Protocol [1] Section 2.2]

6.2. Covariates

The FLIPI2 categorization (low, medium and high) may be considered as stratification factor in efficacy analysis.

7. HANDLING OF MISSING VALUES

7.1. Overall Response, Complete remission, and Partial Response Data

Overall response (ORR), complete remission (CR) or partial response (PR) will be based on central review assessment in accordance with the revised response criteria for malignant lymphoma. For ORR, CR or PR, missing value is defined as no post-baseline response assessment either due to lost to follow-up or withdrawal by patient or other reasons. In the primary analysis, if a post-dose response assessment is missing, it will be counted as a non-responder instead of a missing value.

7.2. Time to Event Data

For the time-to-event endpoints (TTF, DOR, PFS, or OS), the missing data handling method will be censoring. Censoring mechanisms for these endpoints are described in [Section 8.2.3](#).

7.3. Pharmacokinetic Concentrations and Biomarker Data

Concentration below the limit of quantification

In all data presentations (except listings), concentrations below the limit of quantification (BLQ) may be set to zero. Other imputations (eg, $\frac{1}{2}$ LOQ) may also be considered in other analyses (eg, Pop-PK and PK/PD analyses), if deemed appropriate. In the listings BLQ values will be reported as “<LLQ”, where LLQ will be replaced with the value for the lower limit of quantification. The limits of quantification (LOQ) for various PK and biomarker concentrations will be noted in all tables and listings.

Deviations, missing concentrations and anomalous values

Patients with deviations from the protocol design that may affect their PK profile (eg, incomplete dosing due to injection reactions and wrong time, etc.) may be excluded from the PK analysis.

In drug concentration summary tables, statistics will be calculated having set concentrations to missing if one of the following cases is true:

1. A concentration has been collected as ND (ie, not done) or NS (ie, no sample);
2. A deviation in sampling time is of sufficient concern or a concentration has been flagged anomalous by the PK analyst.

7.4. Missing Dates

In compliance with Pfizer standards, if the day of the month is missing for any date used in a calculation, the 1st of the month will be used to replace the missing date unless the calculation results in a negative time duration (eg, date of resolution cannot be prior to date of onset; if replacing resolution date with the 1st of the month results in a negative duration, the resolution date will be set to the onset date). Pfizer standards are similarly used if both month and day are missing (January 1 unless negative time duration). For overall time to event endpoints, if conventions result in a zero or negative duration, duration will be reset to 1 day.

If the start date is missing for an AE, the AE is considered to be treatment emergent unless the collection date is prior to the treatment start date.

8. STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

Whilst every effort has been made to pre-specify all analyses in this statistical analysis plan, if any additional exploratory analyses are found to be necessary, the analyses and the reasons for them will be detailed in the clinical study report (CSR).

Unless otherwise specified, the baseline value is defined as the value collected at the closest time prior to the start of the first study drug administration.

All response data (CR, PR, and ORR) will be categorized based on the scheduled visit at which it was collected. These visit designators are predefined values that appear in the central reader dataset based on predefined specification time windows used by the central reader. In addition, the dates corresponding to events of interest (CR, PR, PD, first response, etc.) are also taken directly from the central reader dataset. Analysis visit windows for laboratory, vital signs, immunogenicity, and biomarker data summary are located in [Appendix 1](#).

8.1. Statistical Methods

8.1.1. Analysis of Response Data

Descriptive statistics (frequency and percentage) for CR, PR, and ORR will be presented by treatment group. The stratified Miettinen and Nurminen method (1985) [2] will be used to obtain the 95% confidence interval for the difference (PF-05280586 minus rituximab-EU). The stratified Mantel-Haenszel method will be used to obtain the corresponding estimated treatment group difference (point estimate). The FLIPI2 categorization (low, medium and high) will be the stratification factor.

8.1.2. Analysis of Time to Event Data

Time to event (DOR, TTF, PFS and OS) curves between the two treatment groups will be compared with a log-rank test. Cox model will be used to estimate the hazard ratio and its 95% CIs for the treatment effect. The FLIPI2 categorization (low, medium and high) will be the stratification factor for both the log-rank test and Cox model. These endpoints will also be summarized using Kaplan-Meier survival curves. The Kaplan-Meier survival will also be employed to summarize the time to event data. For PFS and OS, 1-year rate will be estimated along with the corresponding 95% confidence interval based on the Brookmeyer and Crowley method [6].

8.2. Statistical Analyses

8.2.1. Standard Analyses

8.2.1.1. Patient Disposition

Patient disposition includes the number and percentage of patients for the following categories: patients in each of the study populations, patients discontinued from the treatment, primary reason to discontinue from the treatment, patients discontinued from the study, and primary reason to discontinue from the study.

A listing will present data concerning patient disposition.

8.2.1.2. Patient Demographics

Demographic characteristics will be summarized descriptively by treatment group using the ITT population.

8.2.1.3. Medical History

General medical history and prior medications will be listed for all patients.

Medical history will be summarized (frequency) for both treatment groups by the disease categories per MedDRA dictionary recorded in the database. A patient is counted only once within a category. Frequencies are based on the number of patients in the safety population per treatment group.

8.2.1.4. Baseline Disease Characteristics

The following baseline characteristics will be summarized. For continuous variables, summary statistics including N, mean, median, standard deviation and range will be provided; for categorical variables, number and percentage of patients in each category will be summarized.

- Duration since diagnosis, which is defined in years as $(\text{randomization date} - \text{date of diagnosis} + 1) / 365.25$.
- Ann Arbor Stage (I, II, III, or IV).
- Eastern Cooperative Oncology Group (ECOG) status of 0 to 1.
- FLIPI and FLIPI2 risk classification (low, medium, and high).
- Bone marrow aspirate (Number (percentage) of patients with bone marrow aspirate positive or negative for lymphoma cells).
- LDH and β 2-microglobulin levels (normal and abnormal).

8.2.1.5. Duration of Follow-up

The duration of follow-up is defined as time from the date of first dose of study treatment to the death or last known visit. If a patient dies before end of study, the duration equals the date of death minus study start date + 1. If a patient is alive at end of study, the duration equals the date when the patient last visit was minus study start data + 1.

8.2.1.6. Drug exposure, Treatment Duration and Compliance

Summary of drug exposure will be presented including number of doses, numbers and percentages of patients who had 1, 2, 3 or 4 doses for each treatment group in the safety population. Extent of treatment duration (days), which is calculated as (Last Dose Date of study drug – First Dose Date of study drug + 1), will also be presented.

For each patient, percent compliance will be calculated using the following formula:

Percent Compliance = # of doses actually administered / # of doses planned * 100%. The number of injections planned or actually administered is counted up to the conclusion date of the treatment period. Summary statistics will be provided to percent compliance by treatment group.

8.2.1.7. Prior and Concomitant Medications

Concomitant medications will be coded by preferred term using the World Health Organization (WHO) Drug Dictionary. The number and percentage of patients taking concomitant medications from the first dose through the end of the study will be tabulated by Anatomical Therapeutic Chemical (ATC) classification pharmacological subgroup and WHO drug generic term for each treatment group in the safety population. By-patient listing will also be presented for concomitant medications.

Concomitant procedures (previous and concomitant nondrug treatments) will not be coded, but will be presented in a data listing in the safety population.

8.2.2. Analyses of the Primary Efficacy Endpoint

The primary efficacy endpoint is the Overall Response Rate (ORR) at Week 26 of PF-05280586 and rituximab-EU based on central review assessment. ORR is defined as the proportion of patients who achieved complete remission (CR) or partial response (PR) in accordance with the revised response criteria for malignant lymphoma, with applicable clinical data incorporated into the radiographic response [5].

Descriptive statistics (frequency and percentage) for CR, PR, and ORR will be presented by treatment group and visit. The 95% confidence interval of these response rates and the 95% confidence interval of the difference in the response rates between the two treatment groups will be constructed.

After all randomized patients have had the opportunity to complete their Week 26 visit and the assessment of response and once the PCD data snapshot has occurred, the primary efficacy analysis for equivalence will be performed. The point estimate for the difference in ORR between PF-05280586 and rituximab EU will be computed using the stratified Mantel-Haenszel method. The 95% confidence interval for the difference will be calculated using the asymptotic stratified method proposed by Miettinen and Nurminen (1985) [2]. The FLIPI2 categorization (low, medium, and high) will be considered as stratification factor.. The resulting treatment difference estimate and its associated 95% confidence interval will be used in the hypothesis test as described in [Section 4.1](#).

CR and PR at Week 26 will also be analyzed in the similar fashion as ORR.

These analyses will be primarily performed with the ITT population. The per-protocol population analyses may also be conducted as sensitivity analyses.

8.2.3. Analyses of the Secondary Efficacy Endpoint

Other secondary efficacy endpoints include:

- 1-Year progression free survival;
- 1-Year overall survival;
- Time to treatment failure (TTF);
- Duration of response;
- Complete remission (CR) rate at Week 26 based on central review assessment.

8.2.3.1. 1-Year Progression free survival

Progression-Free Survival (PFS) is defined as the time from date of randomization to first progression of disease (PD) or death due to any cause in the absence of documented PD. Censoring for the PFS endpoint is summarized in Table 2. Progression will be based on the central review assessments. The primary analysis for PFS will be based on the ITT population.

A log-rank test stratified by FLIPI2 risk will be used to compare the treatment groups with respect to PFS at a 2-sided alpha level of 0.05. PFS will also be summarized using the Kaplan-Meier method. The Kaplan-Meier estimates for the 1-year PFS rates, and the 2-sided 95% confidence interval of the rates using the Greenwood's formula will be reported. In addition, Cox model will be used to estimate the hazard ratio and its 95% CIs for the treatment effect.

Sensitivity analyses for PFS may be conducted based on 'Per-protocol' analysis set.

Table 2 Handling of Missing Assessments and Censoring Rules for PFS

Situation	Date of Progression or Censoring	Outcome
No baseline or no adequate baseline assessment, and no death	Date of Randomization	Censored
No post baseline or no adequate post baseline assessment, and no death	Date of Randomization	Censored
No death or disease progression	Date of last adequate assessment	Censored
Discontinued from study	Date of last adequate assessment	Censored
Disease progression or death	Date of death or first adequate assessment for progression, whichever is earlier	Progressed (event)

8.2.3.2. 1-Year Overall Survival

Time to death is defined as the time from date of randomization to death due to any cause. Patients will be censored for this endpoint on the date of the last recorded visit if they do not die. A log-rank test will be used to compare the treatment groups with respect to overall survival (OS) at a 2-sided alpha level of 0.05. The FLIPI2 categorization (low, medium, and high) will be considered as a stratification factor if appropriate. Overall survival will also be summarized using the Kaplan-Meier method. The Kaplan-Meier estimates for the 1-year OS rates, and the 2-sided 95% confidence interval of the rates using the Greenwood's formula will be reported. In addition, Cox model will be used to estimate the hazard ratio and its 95% CIs for the treatment effect. The primary analysis for overall survival (OS) will be based on the ITT population. The sensitivity analysis based on 'Per-protocol' population may also be performed.

8.2.3.3. Time to Treatment Failure

Time to Treatment Failure (TTF) is the time from date of randomization to first progression of disease based on central review, death due to any cause, or permanent discontinuation from treatment, or discontinuation from study for any reason, whichever comes first. The censoring mechanisms for TTF are similar to those described above for PFS with the exception that permanent discontinuation from treatment or discontinuation from study will be considered as treatment failure. A log-rank test stratified by FLIPI2 risk will be used to compare the treatment groups with respect to TTF at a 2-sided alpha level of 0.05. TTF will also be summarized using the Kaplan-Meier method. In addition, Cox model will be used to estimate the hazard ratio and its 95% CIs for the treatment effect. The primary analysis for TTF will be based on the ITT population. The sensitivity analysis based on 'Per-protocol' population may also be performed.

8.2.3.4. Duration of Response

Duration of Response (DOR) is defined as the time from date of the first documentation of overall response (CR or PR) to the first documentation of progressive disease (PD) or to death due to any cause in the absence of documented PD. The analysis for DOR will be based on central review assessment and the response-evaluable population.

The censoring mechanisms for DOR are similar with those described above for PFS with the exception that when a patient has missing response assessment(s) but remains as a CR or PR responder at the time of data analysis, the endpoint will be censored at the time of the last adequate assessment where CR or PR is declared. A log-rank test stratified by FLIPI2 risk will be used to compare the treatment groups with respect to DOR at a 2-sided alpha level of 0.05. In addition, Cox model will be used to estimate the hazard ratio and its 95% CIs for the treatment effect. The Kaplan-Meier survival estimates, together with the number of patients, percentage of patients to experience the event, and the number and percentage of patients censored will be summarized in a table by treatment group.

8.2.3.5. Complete Remission Rate at Week 26

Complete Remission (CR) will be summarized by treatment group and visit. CR will be analyzed in a similar fashion as for ORR as specified in [Section 8.2.2](#).

8.2.4. Analyses of Pharmacokinetic Endpoints

The drug concentration-time data will be summarized by descriptive statistics (mean, standard deviation, median, minimum, and maximum) according to treatment overall and by ADA status (positive and negative subgroups, see Section 8.2.7 for methods related to immunogenicity).

Population PK assessment will be conducted with the drug concentration-time data using the nonlinear mixed effect modeling approach in accordance with regulatory guidances. All patients from the PP population who are treated with PF-05280586 or rituximab-EU and provide at least one post-dose drug concentration measurement will be included in the population PK analysis. The population PK analysis will estimate typical value and variability for parameters including clearance (CL) and volume of distribution (Vd). Also, the influence of selected potential covariates on the PK parameters will be explored; the potential covariates to be explored will include drug product, selected demographics (eg, body weight, sex, age), and ADA status.

The detailed procedures for the population PK analysis, including the model implementation and evaluation, will be described in the Population Modeling Analysis Plan (PMAP). The results of the analysis will be summarized in a Population Modeling and Analysis Report (PMAR), separate from the clinical study report of this study.

[Source: Protocol [1] Section 9.5]

The PK data summary and analyses will be performed based on the pharmacokinetic analysis set (PKAS).

8.2.5. Safety Analyses

All patients treated with at least one dose of study treatment will be included in the safety analyses. All the tables, graphs and listings will follow Pfizer standard.

CCI

CCI

8.2.5.2. Adverse Events

Adverse events will be classified using the medical dictionary for regulatory activities (MedDRA) classification system. The severity of adverse events will be graded according to the NCI CTCAE version 4.03 whenever possible. Adverse events (treatment emergent adverse events; treatment-related adverse events; adverse events classified as NCI CTCAE Grade 3 or higher; and serious adverse events) will be summarized by body system and preferred term according to MedDRA terminology. A treatment emergent adverse event

(TEAE) is defined as any adverse event that occurs after the beginning of the study treatment or any pre-existing adverse event that worsens after the beginning of the study treatment.

Adverse events leading to death, adverse events leading to discontinuation of treatment and adverse events leading to discontinuation from study will be presented by treatment group.

Clinical Outcomes Associated with Immunogenicity (COAI) included IRRs, Sampson Criteria (Sampson et al., 2006 [7]), and potential hypersensitivity/anaphylaxis reactions to the drug. The Sampson criteria algorithm was developed by Pfizer and used to programmatically identify the cases retrospectively that fulfill the criteria (see [Appendix 3](#) for the Sampson algorithm details). The following endpoints were summarized for all patients in the safety Population, by visit and for the study overall, and for ADA positive and negative subgroups (see [Section 8.2.7](#) for methods related to immunogenicity):

- incidence of IRR AEs (as classified by the investigator);
- incidence of AEs that fulfill Sampson's Criteria;
- incidence of AEs belonging to the Standardized MedDRA Query (SMQ) groupings of Anaphylaxis and/or Hypersensitivity (broad + narrow PTs).

8.2.5.3. Laboratory Abnormalities

Hematology and chemistry laboratory data will be summarized by treatment and by visit. The laboratory results will be reported according to the NCI CTCAE severity grade. The frequencies of the worst severity grade observed will be displayed by study treatment. Shift tables from baseline values against each post-baseline visit may be provided to examine the distribution of changes in selected laboratory tests. For parameters for which an NCI CTCAE scale does not exist, the frequency of patients with values below, within, and above the normal ranges will be summarized by treatment and visit. Change from baseline will be additionally summarized by treatment group and visit.

8.2.5.4. Vital Signs Abnormalities

Vital sign data (heart rate, systolic and diastolic blood pressure, temperature, and respiratory rate) will be summarized by treatment and by visit. Change from baseline will be additionally summarized by treatment group and visit.

8.2.6. Biomarker Analysis

Summary statistics by treatment and visit will be provided for the biomarkers including CD19-positive B-cell counts, IgM and IgG, etc. Mean change (or percent change) from baseline will be also summarized by treatment and visit and presented in tabular form and/or graphically.

These analyses will be carried out with the mITT population using observed-case data.

[Source: Protocol [1] Section 9.5]

8.2.7. Immunogenicity Assessment

For the immunogenicity data, the percentage of patients with blood samples positive for ADA and NAb will be summarized for each treatment group. For patients who are positive for ADA, the magnitude (titer), isotype, time of onset, and duration of ADA response will also be described, if data permit. In addition, efforts will be made, as appropriate, to examine possible effect of the ADA on clinical data such as PK and safety.

Because the observed incidence of ADA is highly dependent on multiple factors including the assays used for ADA detection, timing of sample collection and immune status of the patients, the incidence of ADA observed in the planned study may differ from the incidence reported in historical clinical trials.

[Source: Protocol [1] Section 9.5]

This analysis will be carried out with the safety population.

8.2.8. Subgroup Analysis

Subgroup analysis on the primary endpoint, the ORR at Week 26 (ITT), may be conducted by age, gender, race and region, as well as by baseline FLIPI2 categorization, Ann Arbor Staging, and bone marrow involvement, if deemed necessary and appropriate.

To support the PMDA submission in Japan, selected analyses will be repeated for the subgroup of Japanese patients enrolled in this study. Given the fewer number of patients within the Japanese subgroup relative to the study overall, efficacy analyses will not be stratified by FLIPI2.

9. APPENDICES

Appendix 1. Definition and Use of Visit Windows in Reporting Lab, Vital Signs, Immunogenicity, and Biomarker

Note that Day 1 in the table below is taken as the first day of dosing with study drug. It may not be the same as the first study date which is the randomization date. Also note that Day 0 does not exist, so Day -1 is the day before Day 1. Also the relative days (rel_day) from Day 1 are defined as the visit date minus first dosing date plus one. Unless otherwise specified in specific analyses, the analysis visit windows will follow the rules in the following table.

Analysis Visit No.	Analysis Visit Label	Target Day	Visit Window
1	Screening	N/A	$-28 \leq \text{rel_day} \leq -1$
2	Baseline	1	$\text{Rel_day} = 1^*$
3	Week 2	8	$1 < \text{rel_day} \leq 11$
4	Week 3	15	$12 \leq \text{rel_day} \leq 18$
5	Week 4	22	$19 \leq \text{rel_day} \leq 25$
6	Week 5	29	$26 \leq \text{rel_day} \leq 57$
7	Week 13	85	$58 \leq \text{rel_day} \leq 130$
8	Week 26	176	$131 \leq \text{rel_day} \leq 221$
9	Week 39	267	$222 \leq \text{rel_day} \leq 312$
10	Week 52	358	$\text{rel_day} \geq 313$

* Baseline analysis visit window may be considered as $\text{Rel_day} \leq 1$ in some analyses (eg, those involving change from baseline). That is, in case that Day 1 observation is missing, the last non-missing observation before or on the first dosing date may be considered as the baseline.

Appendix 2. Tier-1 Search Terms

MedDRA SMQ PTs	Tumor lysis syndrome (broad + narrow terms)
-------------------	---

PTs with the HLGT	Encephalopathies
PTs with the HLGT	Central nervous system infections and inflammations
PTs with the HLT	Polyomavirus infections
PT	Encephalitis
PT	Encephalitis Enteroviral
PT	Herpes simplex encephalitis
PT	Encephalitis Viral
PT	Encephalopathy
PT	Progressive Multifocal Leukoencephalopathy
PT	JC Virus infection

PTs with the SOC	Infections and infestations
---------------------	-----------------------------

PT	Infusion related reaction
PT	where AECLAS=23 on the eCRF drop down box was selected by the investigator to indicate an infusion related reaction
MedDRA SMQ PTs	Anaphylactic reaction (broad + narrow terms)
MedDRA SMQ PTs	Hypersensitivity (broad + narrow terms)

PT	Neutropenia
PT	Neutrophil count decreased

PT	Febrile neutropenia
PT	Neutropenic sepsis
PT	Neutropenic infection

CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



CCI



10. REFERENCES

1. Pfizer Clinical Protocol B3281006 Amendment: A Phase 3, Randomized, Double-blind Study of PF- 05280586 Versus Rituximab for the First-line Treatment of Patients with CD20- Positive, Low Tumor Burden, Follicular Lymphoma, 19 Apr. 2016.
2. Miettinen O. and Numinen M., Comparative Analysis of Two Rates, *Statistics in Medicine*, 1985, 4(2), pp. 213-226.
3. Pfizer Standard Operating Procedure CT40-GSOP Version 1.0: Management of protocol deviations.
4. Pfizer Clinical and Medical Controlled Document Global Work Instructions CT40-WI- GL01 Version 1.0: Reporting of protocol deviations in the clinical study report.
5. Cheson BD et al: (2007) Revised response criteria for malignant lymphoma. *J Clin Oncol* 25:579-86.
6. Brookmeyer R, Crowley JJ. , A confidence interval for the median survival time. *Biometrics* 1982; 38:29-41
7. Sampson HA et. al., Second symposium on the definition and management of anaphylaxis: summary report-second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium. *Ann Emerg Med*. 2006 Apr;47(4):373-80.