Official Title: **A RANDOMIZED, DOUBLE-BLIND, DOUBLE-DUMMY, ACTIVE-COMPARATOR, MULTICENTER STUDY TO EVALUATE THE EFFICACY AND SAFETY OF RITUXIMAB VERSUS MMF IN PATIENTS WITH PEMPHIGUS VULGARIS**

NCT Number: NCT02383589

Document Date: SAP Version 4: 10-December-2018

SAP Version 1: 18-October-2019

STATISTICAL ANALYSIS PLAN

TITLE:	**A RANDOMIZED, DOUBLE-BLIND, DOUBLE DUMMY, ACTIVE-COMPARATOR MULTICENTER STUDY TO EVALUATE THE EFFICACY AND SAFETY OF RITUXIMAB VERSUS MYCOPHENOLATE MOFETIL IN PATIENTS WITH PEMPHIGUS VULGARIS**
PROTOCOL NUMBER:	WA29330
STUDY DRUG:	Rituximab
VERSION NUMBER:	4
IND NUMBER:	121595
EUDRACT NUMBER:	2014-000382-41
SPONSOR:	F. Hoffmann-La Roche Ltd.
PLAN PREPARED BY:	[redacted], Ph.D.
DATE(S) AMENDED:	Version 1: 4 November 2015 Version 2: 1 February 2018 Version 3: 12 July 2018
DATE FINAL:	See electronic date stamp below

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

Name	Reason for Signing	Date and Time (UTC)
[redacted]	Company Signatory	10-Dec-2018 20:35:42

CONFIDENTIAL

This is an F. Hoffmann-La Roche Ltd document that contains confidential information. Nothing herein is to be disclosed without written consent from F. Hoffmann-La Roche Ltd.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

1. BACKGROUND

Study WA29330 (PEMPHIX) aims to evaluate the efficacy and safety of rituximab (RTX) versus mycophenolate mofetil (MMF) in patients with Pemphigus vulgaris (PV). PV is a rare, serious autoimmune disease characterized by blisters and erosions of the skin and mucosal areas.

Study WA29330 is a multicentre, randomized, double-blind, double-dummy, active-comparator, two-arm, parallel-group study, which consists of three periods: a screening period of up to 28 days, a 52-week, double-blinded, treatment period followed by a 48-week safety follow up (SFU) period, that begins at the time of study treatment completion or discontinuation.

This preliminary Statistical Analysis Plan (SAP) describes the statistical methodology and data handling rules to be used for the Week 52 primary analysis. The primary reporting event will include all data up to Week 52 or early withdrawal. The SAP will be finalized prior to database lock and data unblinding for the primary analysis.

A final reporting event and study database lock will occur after all patients complete the 48-week SFU period or withdraw early from SFU, details of which will be documented in a separate SAP.

2. STUDY DESIGN

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is provided in Appendix 1. Additional details including the study flowchart, schedule of assessments, and corticosteroid treatment guideline can be found in the Appendices (Appendix 2, Appendix 3, and Appendix 4).

2.2 OUTCOME MEASURES

2.2.1 Primary Efficacy Outcome Measures

The proportion of patients (excluding telemedicine [TM] patients) who achieve a sustained complete remission without experiencing an event that constitutes treatment failure (as defined in Section 4.2) beforehand, during the 52-week treatment period.

Sustained complete remission is defined as achieving healing of lesions with no new active lesions (i.e., pemphigus disease area index [PDAI] activity score of 0) while on 0 mg/day prednisone or equivalent, and maintaining this response for a total of at least 16 consecutive weeks, during the 52-week treatment period. Patients with transient new lesions for 1 week or less that heal without additional systemic corticosteroid therapy will not be considered to have experienced treatment failure.

2.2.2 Secondary Efficacy Outcome Measures

The secondary efficacy outcome measures are as follows:

- Cumulative oral corticosteroid dose (prednisone or equivalent) over the treatment period
- Total number of disease flares during the treatment period
- Time to initial sustained complete remission
- Time to disease flare
  - Disease flare is defined as appearance of three or more new lesions a month that do not heal spontaneously within 1 week, or by the extension of established lesions, in a patient who has achieved disease control.
- Change in health-related quality of life (HRQoL), as measured by the dermatology life quality index (DLQI) score, from baseline to Week 52
- Duration of sustained complete remission
- Patients' impression of change in PV symptoms, as measured by the patient global impression of change (PGIC) score, during the treatment period
- Clinician impression of change in patients' PV symptoms, as measured by the clinician global impression of change (CGIC) score, during the treatment period.

The first five secondary efficacy outcome measures will be included in a hierarchical testing procedure to control for Type I error (See Section 4.7); the remaining secondary endpoints will be considered exploratory and will not be tested for Type 1 error.

2.2.3 Safety Outcome Measures

The safety outcome measures are as follows:

- Nature, frequency, and severity of adverse events (AEs), including serious adverse events and AEs leading to discontinuation
- Vital signs and clinical laboratory test results (including complete blood count and blood chemistry)
- Incidence of anti-drug antibodies(ADA)
- Circulating B cells, T cells, natural killer (NK) cells, plasma cells, and other leukocytes
- Plasma immunoglobulin (Ig) levels (total Ig, IgG, IgM, and IgA)
- Corticosteroid-related AEs in relation to corticosteroid exposure

2.2.4 Exploratory Outcome Measures

The exploratory outcome measures are as follows:

- Proportion of patients achieving a complete remission by Week 24 and by Week 52

Complete remission is defined as achieving wound healing with no new active lesions (i.e., PDAI activity score of 0) for at least 8 consecutive weeks during the 52-week treatment period.

- Proportion of patients achieving a partial remission by Week 24 and by Week 52

  Partial remission is defined as the presence of transient new lesions that heal within 1 week (while the patient is receiving minimal therapy, including topical corticosteroids). Minimal therapy is defined as $\leq$10 mg/day prednisone (or equivalent) for at least 8 consecutive weeks during the 24-week or 52-week treatment period.

- Pharmacokinetics/pharmacodynamics (PK/PD) of RTX

  PK/PD parameters include, but are not limited to, serum levels of RTX, peripheral CD19+B-cell counts, HACA, and autoantibody concentrations.

- To explore the PD of MMF in patients with PV
- Change in total PDAI activity score during the treatment period
- Change from baseline in anti-desmoglein-1 (Dsg1) and anti-desmoglein-3 (Dsg3) autoantibodies
- Change in health utilities as assessed by the European quality of life (EuroQol) 5-dimension questionnaire, 3-level version (EQ-5D-3L) score from baseline to Week 52
- Change in HRQoL, as measured by the Skindex-29 from baseline to Week 52
- Proportion of patients experiencing treatment failure from Week 12 to Week 52.

2.3 DETERMINATION OF SAMPLE SIZE

The primary efficacy endpoint of this study is the proportion of patients (excluding TM patients) achieving a sustained complete remission, as assessed at Week 52, that has been maintained for $\geq$ 16 consecutive weeks during the 52-week treatment period. With use of limited data available in the literature from randomized clinical trials of MMF and investigator-initiated trials of rituximab (Beissert et al. 2010), it is estimated that approximately 40% of patients with PV receiving MMF will achieve a sustained complete remission. It is estimated that patients receiving rituximab will induce a sustained complete remission rate of 65%. On the basis of these assumptions, a total of 122 patients randomized to either the rituximab arm or the MMF arm in a 1:1 ratio (61 patients in the rituximab arm and 61 patients in the MMF arm) will yield approximately 80% power in a two-sided test at the 5% significance level. To account for the 10 TM patients excluded from the primary efficacy analysis, approximately 132 patients will be randomized in total. The sample size calculation was performed using nQuery Advisor 7 using the two-group chi-squared test of equal proportions. The primary analysis method using a Cochran-Mantel-Haenszel test adjusting for stratification factors is expected to provide at least 80% power to detect the difference in proportions.

No adjustment will be made for dropouts. Patients withdrawing early from the study or meeting the criteria for treatment failure (including rescue therapy) will be assessed based on available data up until the point of withdrawal or treatment failure only for the primary endpoint. Patients with missing or incomplete data where a status cannot be determined after applying relevant imputation rules will be set to non-responders.

2.4 ANALYSIS TIMING

The primary analysis will be conducted when all patients have completed the Week 52 study visit or have withdrawn early from the study. The reporting event for the primary analysis will include data for all patients up to their Week 52 assessment visit within the treatment period, or data up to the point of withdrawal. Safety follow up (SFU) data will not be included within the primary analysis data-cut. All data up to the Week 52 study visit will form part of the Primary Week 52 Clinical Study Report (CSR).

A final reporting event and study database lock will occur after all patients complete the 48-week SFU period or withdraw early from SFU, details of which will be documented in a separate SAP and final CSR. The final CSR will include all available data up to study completion (last patient's last visit [LPLV] in SFU) or early withdrawal from SFU.

3. STUDY CONDUCT

3.1 RANDOMIZATION ISSUES

After completing the informed consent form, eligible patients will be enrolled into the study and randomized using a central Interactive Voice/Web Response System (IxRS). Patients will be randomly allocated to receive RTX plus MMF placebo or RTX placebo plus MMF in a 1:1 ratio. Randomization will be performed according to a stratified permuted block design using the two baseline stratification factors (duration of illness [diagnosed $\leq$ 1 year from screening versus. $>$ 1 year] and geographical region [North America vs Rest of World]).

A mock randomization schedule was generated by the IxRS vendor ([redacted]) and user acceptance testing was performed on the IxRS system for both voice and web based methods. Testing of the system by the Sponsor's study statistician confirmed the following: correct strata and block allocation, correct treatment assignment, and agreement between IxRS output and the IxRS datasets delivered to the Sponsor.

The independent Data Coordinating Center (iDCC) will perform one round of randomization checks early during the recruitment period. The checks will be similar to those performed during user acceptance testing and will be carried out no later than when 30% patients have been randomized. The iDCC will notify the Sponsor's study statistician of any potential issues. Further checks at additional time points may be performed if necessary.

3.2 DATA MONITORING

Safety data and data pertaining to trial conduct will be reviewed by an independent Data Monitoring Committee (iDMC). The iDMC will meet according to the schedule below, whichever occurs first:

- After the first 10 patients have completed at least 3 months of study treatment or have withdrawn early from the study
- After the first 5 TM patients have completed at least 3 months of study treatment or have withdrawn early

Thereafter, meetings will be held approximately every 4 or 6 months (frequency may be adjusted during the recruitment period), after 10 TM patients have completed at least 3 months of study treatment or withdrawn early from the study, or as requested by the iDMC.

Efficacy data will be available on request, to enable a benefit-risk assessment should there be a safety concern.

The Sponsor's study team will have no direct contact with the iDMC during the trial conduct except for the scheduled open sessions of the iDMC.

Further details can be found in the iDMC charter.

4. STATISTICAL METHODS

All analyses, summaries, and listings will be performed using SAS® software (Version 9.2 or later).

4.1 ANALYSIS POPULATIONS

All patient populations for the analysis of the 52-week, double-blind treatment period will be defined and documented prior to unblinding the data collected up to Week 52.

4.1.1 Randomized Population

The randomized population will include all patients randomized into the study.

4.1.2 Pharmacokinetic-Evaluable Population

The pharmacokinetic-evaluable population (PK population) will include all patients randomized to and received any dose of RTX given as study medication, have at least one post-dose PK sample that is evaluable, and have no major protocol deviations that would impact the PK results.

4.1.3 Safety Population

The safety population will include all patients who were randomized and received any part of an infusion of study drug or oral administration of study drug. Patients who received the incorrect treatment throughout the study from that assigned at

randomization will be reported under the treatment actually received. Patients who were not randomized but who received study drug will be included in the safety population and reported under the treatment actually received.

Patients receiving at least one dose of both active study treatments during the study (e.g., in error, or in the unlikely event of crossing over) will be reported under the original treatment received. This does not include MMF or RTX received as rescue therapy.

All safety analyses including AEs, in addition to analysis of previous and concomitant medications, previous and concomitant diseases, laboratory data, and compliance with study treatment, will be performed using the safety population.

4.1.4 Intent-to-Treat Population

The intent-to-treat (ITT) population will include all randomized patients who received any part of an infusion of study drug or oral administration of study drug.

Patients who prematurely withdraw from the study for any reason and for whom an assessment is not performed for whatever reason will still be included in the ITT analysis. Patients who received an incorrect therapy (i.e., different from that to which they were randomized) will be summarized in the treatment arm to which they were randomized.

Patients with no post-baseline data will still be included in the ITT population but will have their efficacy results imputed according to the imputation methods described in this SAP for all key efficacy assessments.

Patients who were not randomized but still received study treatment will be excluded from the ITT population and subsequent analyses.

Sensitivity analyses of the efficacy outcomes will be performed using the ITT population.

4.1.5 Modified Intent-to-Treat Population

The modified intent-to-treat population (mITT) population includes patients in the ITT population, excluding the 10 TM patients. This population will be used in the analyses of efficacy outcomes.

4.2 TREATMENT FAILURE

Treatment failure is the occurrence of any one of the following events:

- After disease control has been achieved, a two-step increase in prednisone (or equivalent) dose to ≥20 mg/day for more than 2 weeks to treat increased disease activity during the corticosteroid taper
- Failure to achieve disease control after receiving 240 mg/day oral prednisone (or equivalent) for 1 week

Disease control is defined as the time at which new lesions cease to form and established lesions begin to heal (beginning of the consolidation phase).

- Increase in MMF/matching placebo dosage above 2 g/day
- Initiation of any rescue therapy (addition of another immunosuppressive medication, additional RTX or MMF open-label doses, intravenous (IV) Ig, plasmapheresis, or another treatment or procedure) as determined by the investigator's best medical judgment

The clinical definitions of treatment failure above are aligned with "non-responder" classification for the primary analysis. Patients experiencing treatment failure prior to achieving sustained complete remission will be categorized as non-responders in the primary analysis, but will continue to be followed in the study per the schedule of assessments. Patients who achieve sustained complete remission without experiencing an event constituting treatment failure beforehand will not be categorized as non-responders, regardless of whether they subsequently meet the criteria during the remaining treatment period.

Patients who experience treatment failure will be analyzed using data up to the point of treatment failure for efficacy remission endpoints, including the primary endpoint for sustained complete remission.

From Week 12 through Week 52, patients who experience treatment failure are eligible to receive rescue therapy during the treatment period with another immunosuppressive medication, IV Ig, plasmapheresis, or another treatment or procedure as per the investigator's best medical judgment. An early withdrawal visit is not required for a patient to receive rescue therapy, and patients should continue to be followed in the study per the 52-week treatment period (see Appendix 1).

Patients who receive rescue therapy will be evaluated up to the point of rescue only for all main efficacy analyses. Methods for data handling and/or imputing post-rescue data are described throughout Section 4.7. Additional exploratory summaries on key efficacy assessments will be produced that will include rescue data. Safety analyses will include data up to the point of rescue only.

Listings will include all data, including data post-rescue treatment.

Depending on the proportion of patients receiving rescue, patients will be analyzed separately using descriptive statistics for key analyses on study conduct, efficacy, and safety. For relevant summaries, patients may be re-baselined at the time of rescue and data presented as pre- and post-rescue treatment. Rescue analyses will use data as observed. Subject to patient counts, a limited number of analyses may be summarized by type of rescue received (MMF given as rescue, RTX given as rescue, or other).

4.3 ANALYSIS OF STUDY CONDUCT

The following analyses of study conduct will be carried out on the safety population or mITT, as appropriate.

The patients excluded from each analysis population will be summarized, including the reason for exclusion.

A summary of enrollment by country and investigator site will be produced.

The number of patients who completed or discontinued from the study by Week 52, including the reason for discontinuation, will be summarized by treatment group.
A listing of early withdrawals will be produced.

A summary of the number of patients who entered the study and who failed to meet all inclusion/exclusion entry criteria will be produced. Major protocol deviations will be listed and summarized.

Patient duration in the 52-week treatment period will be summarized. Duration will be calculated as:

date of last visit − date of the first dose +1

Previous and concomitant medications will be summarised descriptively by mapped term and treatment group. A summary of rescue medications for PV will also be produced.

A summary of the number of patients who experience treatment failure along with the reason(s) for treatment failure will be produced. The number of patients receiving rescue will also be listed and summarised.

A summary of exposure to study treatment will be produced using the dose recorded on the electronic case report form (eCRF). The number of RTX/placebo RTX infusions and dose received at each infusion and overall will be summarised by treatment group. The mean daily dose, total cumulative dose, and duration of treatment of MMF/placebo MMF over the 52-week treatment period will be summarised.

4.4 ANALYSIS OF TREATMENT GROUP COMPARABILITY

Summary tables will be produced for the safety population for baseline demographic and disease characteristics. If it is found that there is a large difference in the number of patients in the safety population compared to the ITT population, additional summaries will be produced based on treatment as randomized.

Demographic and baseline disease characteristics will be summarized as described in the following sections.

4.4.1 Demographics

Demographics will be summarized, including:

- Gender (male/female)
- Age (years)
- Height (cm)
- Weight (kg)
- Body mass index (kg/m^2)
- Race
- Ethnicity
- Geographic region
- Smoking status
- Female reproductive status

4.4.2 Disease Characteristics

Baseline values will be summarized, including:

- Duration of PV (months)
- Prednisone dose (mg/day)
- Overall PDAI total activity score
- Site of disease (mucosal only, mucocutaneous, cutaneous only. Site of disease will be determined as follows:
  - Mucosal only: PDAI total mucosal activity score>0 and PDAI total skin activity score = 0 and PDAI total scalp activity score=0
  - Cutaneous only: PDAI total mucosal activity score=0 and at least one of PDAI total skin activity score or PDAI total scalp activity score is>0
  - Mucocutaneous: PDAI total mucosal activity score>0 and at least one of PDAI total skin activity score or PDAI total scalp activity score is>0
- Anti-Dsg-1
- Anti-Dsg-3
- CD19
- Prior MMF treatment
- Prior RTX treatment

In addition, any clinically significant ECG or chest X-ray abnormalities at baseline will be listed using the safety population.

Medical history data, including surgery and procedures, and baseline conditions will be summarized descriptively by treatment group using the safety population. A glossary

showing the mapping of investigator verbatim terms to diseases will be produced for the medical history data.

4.5 DATA-CUT FOR ANALYSES

For each patient, the day of first dose of study treatment will be designated study day 1. Each subsequent assessment point will be assigned a study day calculated as:

date of assessment − date of first dose+1.

Data beyond the Week 52 visit will not be included in the primary Week 52 CSR. For each patient, data will be cut-off and reported up to and including the earliest of the following:

- Week 52 visit (using upper visit window of Day 364+10 days)
- Date of early withdrawal

Patients who complete the 52-week treatment period but have a missing Week 52 visit, or where the Week 52 visit falls outside the reporting cut-off will be evaluated up to study Day 365 for all analyses involving a duration or exposure element. In such analyses, any references to Week 52 throughout this SAP will imply study day 365 for these patients.

Although SFU data will not be included in data displays for the Week 52 primary reporting event, any deaths at the time of the data-cut, occurring in SFU will be highlighted in the CSR.

4.6 VISIT WINDOWS

In general, data for assessments that are collected by scheduled visits will be mapped to visits that appear in the schedule of assessments per the protocol using the actual study day of assessment. Data mapped to scheduled visits will include all data collected up to the reporting cut-off date, and may include withdrawal visits and unscheduled visits. Data will never be mapped to visits for which the assessment was not scheduled in the protocol (i.e., results for an assessment conducted at an unscheduled visit will be mapped to the appropriate visit at which the assessment was scheduled). Visit windows will be continuous from the midpoint between two consecutive study visits, and will be dependent on the schedule of assessments for each variable independently. An example of the DLQI visit window is given in Table 1. Visit windows for each variable will be constructed in the same manner.

Table 1 Time Windows of Study Days to Assign Study Visit (for DLQI)

Study Visit	Scheduled Study Day (±3 days)	Time Window (Days)
Baseline	1	≤1
Week 12	84	>1 to ≤126
Week 24	168	>126 to ≤224
Week 40	280	>224 to ≤322
Week 52	364	>322 to ≤375

DLQI = dermatology life quality index.

Derived variables such as remission endpoints or assessments that are collected on an ongoing basis, such as flares and oral corticosteroid dose, will be presented using appropriate time intervals.

Data displays that are presented by visit will only include visits at which the assessment was scheduled to be collected. The methods for handling multiple values within a visit window are described in Section 4.10.3.

4.7 EFFICACY ANALYSIS

The primary and the first five ranked secondary endpoints (listed below) will be tested at an overall 5% significance level using two-sided hypothesis tests. The remaining secondary endpoints will be considered exploratory (Section 2.2.4) and not tested for significance.

To control for overall Type I error, the hypothesis testing on study endpoints will be conducted sequentially starting with the primary endpoint at 5% level of significance. If the null hypothesis relating to the primary endpoint is rejected, the first 5 secondary endpoints will then be tested in a fixed-sequence fashion in the following order:

1. Cumulative oral corticosteroid dose (prednisone or equivalent) over the treatment period
2. Total number of disease flares during the treatment period
3. Time to initial sustained complete remission
4. Time to disease flare
5. Change in health-related quality of life (HRQoL), as measured by the dermatology life quality index (DLQI) score, from baseline to Week 52

If at any step in the testing procedure, the p-value associated with a test is not <0.05, the remaining endpoints below the sequence will not be considered significant.

The mITT population will be the primary population used for the analysis of efficacy outcomes except for duration of sustained complete remission which will be on a subset

of mITT population (see Section 4.7.3.1). Additional analyses will be conducted for sensitivity or exploratory purposes, as described in the relevant sections below.

4.7.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of patients achieving sustained complete remission during the 52-week treatment period.

The following two-sided hypothesis will be tested at the 5% significance level:

$$H_0: p_{RTX} = p_{MMF}$$

$$H_1: p_{RTX} \neq p_{MMF}$$

where p denotes the proportion of patients achieving sustained complete remission status.

The primary method for comparing the proportion of patients achieving sustained complete remission will be the Cochran-Mantel-Haenszel test. The analysis will be stratified by the stratification factors applied at randomization. The proportion of patients achieving sustained complete remission in each treatment group will be presented along with the adjusted difference, 95% confidence interval for the difference, and p-value.

4.7.1.1 Sustained Complete Remission

A patient will be considered a responder for sustained complete remission if the following conditions are met for a minimum of 16 consecutive weeks during the 52-week treatment period:

- PDAI total activity score = 0
- Prednisone (or equivalent) dose = 0 mg/day

Patients who withdraw early from the study or meet the criteria for treatment failure will be evaluated using data up to the point of withdrawal or treatment failure only, whichever occurs first. Patients who are in sustained complete remission at the Week 52 visit will have their duration cut-off at the Week 52 visit for the purpose of calculating the minimum 16 week criteria for the primary analysis, however they will be censored in secondary analyses of duration.

Sustained complete remission is a composite endpoint derived using several sources of information. A description of each variable contributing to the composite endpoint is given below, including methods for data handling and dealing with missing data.

4.7.1.2 Pemphigus Disease Area Index

The PDAI is a scoring system ranging from 0 to 263 and assesses the number and sizes of lesions across cutaneous and mucosal areas of the body, where higher PDAI scores indicate greater disease activity (Appendix 5). The total PDAI score is made up of two components: total activity score (0 – 250) and total damage score (0–13). The total

damage score measures the extent of post inflammatory hyperpigmentation or erythema from resolving lesions. The trial objective seeks to achieve the absence of new or established lesions (i.e., total activity score=0), with lesser importance being placed on post-inflammatory hyperpigmentation left by lesions that have resolved. For this reason, only the PDAI total activity score will be used for the assessment of sustained complete remission, and will be denoted by PDAI^a throughout this SAP.

For calculating the duration, a patient with a PDAI$^a = 0$ at a visit who fails the criteria at the next visit will be assumed to fail at the date of the visit where PDAI$^a = 0$ was not met. That is, the patient is assumed to maintain the response during the whole interval between the two visits, as any clinically relevant increases in disease activity is expected to be captured by an increase in oral corticosteroids.

Missing Data

Missing PDAI^a scores at a visit will be imputed using the score from the previous or next visit, whichever is higher. If the PDAI^a score is missing at two or more consecutive visits then no imputation will be made, and it will be assumed that PDAI$^a > 0$ for those visits. This imputation will apply for missing overall total scores, and missing anatomical location scores.

4.7.1.3 Prednisone (or Equivalent) Dose

Patients entering the study should receive oral corticosteroids at a prednisone or equivalent dose of 60 or 80mg/day at Day 1. After Day 1, the dose will be tapered to a target dose of 0 mg/day as outlined in the protocol and will be documented in the eCRF. Dose adjustments are tolerated in this trial, however patients inadequately controlled after performing clinically relevant dose adjustments will be considered a treatment failure. Further details for treatment failure are provided in Section 4.2.

Oral corticosteroids will be monitored on an ongoing basis to ensure information is correctly recorded on the concomitant medication page of the eCRF. Any missing or ambiguous data will be queried and all efforts will be made to resolve data issues prior to database lock.

Full details of data handling methods is described in Section 4.7.2.1, this includes reporting conventions for prednisone equivalence and imputation methods for dealing with missing or partial data.

4.7.1.4 Treatment Failure

The primary method for dealing with treatment failures will be to exclude data from the first point a patient meets one of the criteria for treatment failure. The responder status (i.e., responder versus non-responder) for sustained complete remission will be determined using available data up to the point of treatment failure only.

4.7.1.5 Sensitivity Analyses

The primary analysis described in Section 4.7.1.1 will be repeated with the treatment failure criteria relaxed to include patients in whom rescue therapy is initiated. Patients who receive rescue therapy will continue to be analyzed using data up to the point of rescue only; however, patients who meet treatment failure criteria and do not receive rescue therapy will be analyzed using all available data during the 52-week treatment period.

A robustness analysis of patients who withdraw will be conducted using a tipping point analysis. The tipping point is defined as the difference in the number of missing events between the treatment groups that result in a change in the primary outcome conclusions (Yan et al. 2009). A two-dimensional plot will be produced to evaluate where the tipping point lies. The tipping point analysis will be used to assess what the outcome might have been had patients carried on in the trial. All other analysis methods will remain as described for the primary analysis, including the handling of patients who qualify for treatment failure. Patients who experience treatment failure and later withdraw will continue to be analyzed using data up to the point of failure, as this would have been their outcome had they continued in the study.

The primary analysis will also be repeated to include the patients who participated in the trial via TM (10 patients at one investigational site) in order to determine whether the trial conclusions are upheld and independent of the logistics of conducting study assessments, methods of interactions with patients, and method of data collection through TM.

4.7.1.6 Exploratory Analyses

Additional analyses may be carried out, for example:

- Logistic regression analysis to control for additional baseline factors (e.g., PDAI total activity score, prednisone (or equivalent) dose at baseline, location of PV lesions)
- Investigate any effects caused by changes to trial conduct made during the trial, such as changes to the treatment failure criteria between protocol versions 2 and 3
- Further analyses investigating rescue patients who receive an alternative treatment
- Descriptive statistics for the TM subset

4.7.2 Secondary Efficacy Endpoints

Hypotheses related to secondary endpoints will be tested in a fixed-sequence manner for the first five ranked secondary endpoints at the two-sided 5% significance level using the procedure discussed in Section 4.7.

4.7.2.1 Oral Corticosteroids

Oral corticosteroid exposure over the 52-week treatment period will be compared between RTX and MMF using a Wilcoxon Mann-Whitney test, assuming that the data will be non-normally distributed. If it is found that the data follows a normal distribution

using Shapiro-Wilk W Test, a t-test will be used to compare the difference in the two treatment groups.

Oral corticosteroid use over time will be further investigated; this includes data during the screening period. Summaries and plots will be generated, which will include:

- Mean and median daily dose per visit interval with error bars
- Total cumulative dose per visit interval
- Total number of days on oral corticosteroids over the 52-week period
- Categories of mean daily dose per visit interval (e.g., 0 mg/day, 0–<10 mg/day, ≥10 mg/day)

The mean daily dose for a given visit interval will be calculated as follows:

> sum of total daily dose across all days in the visit interval÷(end date of visit interval−start date of visit interval+1).

Data presented by specified visit intervals will use all available; if a patient withdraws during a time interval, no imputation will be made for the remaining days in the interval, and only available data up to the time of withdrawal will be used in the calculation. Rescue data will be included in descriptive summaries.

Additional summaries may be produced to investigate corticosteroid use for rescue patients.

A listing showing oral corticosteroid dosing for patients over the 52-week treatment period will be produced using all available data up to the exposure date, with rescue data being flagged.

Data Handling for Oral Corticosteroids

Oral corticosteroids will be recorded in the eCRF under the concomitant medications page. Dose modifications will be tracked by entering the start and stop dates for each dose received. Medications will be flagged as oral corticosteroids using the Sponsor's corticosteroid drug basket and the route administered information as recorded on the eCRF. The Sponsor will review the concomitant medications data on an ongoing basis to ensure medications are being correctly recorded, and are being flagged as appropriate.

For reporting purposes, all relevant oral corticosteroid entries will be standardized by converting to a prednisone equivalent dose, and will be reported under the term 'Prednisone'. A listing of all compounds reported as prednisone will be produced along with the conversion factor applied.

The study management team will monitor oral corticosteroids on an ongoing basis for data quality and to ensure all dose adjustments are being correctly captured. All efforts

will be made to resolve missing or ambiguous data prior to database lock. In the case that data queries cannot be resolved, data will be handled in a conservative manner as described in the sections below.

The intervals for each dose of oral corticosteroid captured during the taper are expected to be continuous and non-overlapping. Patients who are adequately controlled by their blinded treatment may discontinue oral corticosteroids altogether, but may re-start treatment as required to treat any increases in disease activity. With this in mind, the use of available data will be maximized and incorporated into the imputation of missing or partial dates as follows:

Missing Term	Imputation Method
Start Date	
Missing Day	Latest of the following: Previous End Date+1 First day of the month
Missing Month	Previous End Date+1
Missing Year	No imputation
End Date	
Missing Day	Earliest of the following: Next Start Date−1 Last day of the month
Missing Month	Next Start Date−1
Missing Year	No imputation

Within two consecutive dosing intervals, if the end date of the first dose interval is missing and the start date of the second dose interval is missing, a midpoint approach using the start date of the first interval and the end date of the second interval will be used to assign a cut-off end date, providing these dates are available.

A missing or 'ongoing' final end date will be imputed with the exposure date. A missing initial start date will be imputed with Day 1 providing the end date is after Day 1, otherwise no imputation will be made.

If the total daily dose is missing then the dose will be imputed using the dose from the previous dose or the next dose, whichever is higher.

In the event of overlapping dose intervals, the patient will be assumed to be taking the higher dose for the maximum amount of time possible and the lower dose for the minimum amount of time possible within the dates provided, such that the total daily dose will never be double counted.

4.7.2.2 Total Number of Flares

The total number of flares occurring during the treatment period will be evaluated by taking into account the number of flares experienced and the exposure period. Patient level flare and exposure data will then be aggregated, and the rate of flares compared between the RTX and MMF treatment groups. Rate of flare will be analyzed using a Poisson regression model with a log link and the log of each patient's duration in study as an offset variable. The model will include factors for treatment, region, duration of illness, baseline PDAI activity score, and baseline prednisone dose. The total number of flares, total exposure (weeks), and rate of flare will be presented by treatment group along with maximum likelihood estimates for the rate ratio, 95% confidence interval for the rate ratio, and p-value. A Pearson Chi-Square test will be used to assess over-dispersion of Poisson model. If an over-dispersion is detected based on this test, then a negative binomial model will be used. If the number of flares has an excess of zero counts according to the score test (van den Broek, 1995), a zero-inflated negative binomial model will be applied.

The exposure period for a patient will be the duration from Day 1 to the earliest of the following: early withdrawal, receipt of rescue, Week 52 visit.

Additional exploratory work may be undertaken to investigate the potential changes of flare rates over time, for example by assessing the number of flares over specified time intervals.

4.7.2.3 Time to Start of Sustained Complete Remission

Time to sustained complete remission is defined as:

date of first sustained complete remission response-date of first dose+1

The date of first sustained complete remission is the first date at which a patient meets both criteria below, providing it was maintained for a minimum of 16 consecutive weeks:

- PDAI total activity score=0
- Prednisone (or equivalent) dose=0 mg/day

Data will be censored as follows:

- Patients withdrawing early from the study without achieving sustained complete remission will be censored at the time of study withdrawal.
- Patients meeting the criteria for treatment failure who have not achieved sustained complete remission prior to treatment failure will be censored at the Week 52 visit.
- Patients completing the 52-week treatment period who have not achieved sustained complete remission will be censored at the Week 52 visit.

Time to first sustained complete remission will be summarized using Kaplan-Meier curves along with the median, 95% confidence interval for the median, 25th and 75th

percentiles for each treatment group if available. The log-log transformation will be used to compute the 95% confidence intervals.

A stratified log-rank test will be used to test the time to first sustained complete remission between the RTX and MMF treatment arms over the 52-week treatment period, adjusting for the stratification factors used at randomization. A sensitivity analysis may be carried out using the Wilcoxon test. The Wilcoxon test is expected to be more sensitive when the log-rank assumptions do not hold (e.g., convergence of the Kaplan-Meier curves).

The time to first sustained complete remission will be converted to weeks for the presentation of Kaplan-Meier curves and all associated statistics. The Proc Lifetest procedure in SAS will be used to generate the Kaplan-Meier curves and hypothesis testing.

4.7.2.4 Time to Disease Flare

Time to (first) flare is defined as:

date of first flare−date of first dose+1

For the purpose of the analysis, only flares occurring on or after Day 1 in patients who have previously achieved disease control will be analyzed. Flares occurring during the screening corticosteroid taper period will not be considered. Note that patients may not necessarily be in a state of disease control at Day 1; however, date of first dose (Day 1) will be used as the primary time origin as it is more reliable and less subjective than the date of disease control, which is determined by the investigator.

Time to disease flare during steroid taper and time to disease flare during CS withdrawal will be analyzed separately for the mITT population.

Data will be censored as follows:

- Patients withdrawing early from the study who do not experience a flare prior to withdrawal will be censored at the time of study withdrawal.
- Patients completing the 52-week treatment period who do not experience a flare will be censored at the date of study completion.

Time to first flare will be summarized using Kaplan-Meier curves along with the median, 95% confidence interval for the median, 25th and 75th percentiles for each treatment group if available. The log-log transformation will be used to compute the 95% confidence intervals.

A stratified log-rank test will be used to test the time to first flare between the RTX and MMF treatment arms over the 52-week treatment period, adjusting for the stratification factors used at randomization.

The time to first flare will be converted to weeks for the presentation of Kaplan-Meier curves and all associated statistics. The Proc Lifetest procedure in SAS will be used to generate the Kaplan-Meier curves and hypothesis testing.

Sensitivity Analyses

The analysis of time to disease flare will be repeated using disease control as the origin, defined as follows:

time to disease flare = date of first flare − date of disease control + 1

This sensitivity analysis aims to evaluate the onset of flares after the disease is adequately controlled. Patients withdrawing from the study who fail to achieve disease control prior to withdrawal will be excluded from the analysis.

All other methods including censoring methods will remain the same.

4.7.2.5 Dermatology Life Quality Index

The DLQI is a validated patient-reported outcome consisting of 10 questions assessing the impact of dermatologic diseases on patients' lives over the past seven days (Finlay and Khan, 1994). The DLQI questions are categorized into the following domains:

- Symptoms and feelings
- Daily activities
- Leisure
- Work and school
- Personal relationships
- Treatment

Responses to questions are ordinal in nature ranging from "Not at all" to "Very much". The DLQI will be completed and calculated as per the DLQI instructions for use and scoring manual (Appendix 6). The DLQI has a minimum score of 0 and a maximum score of 30, with higher scores indicating greater impairment on quality of life. A change in DLQI score of at least 4 points is considered clinically important (Basra et al, 2015).

Change from baseline to Week 52 in DLQI will be analyzed using a maximum likelihood-based repeated measures model. The analysis will adjust for treatment arm, region, duration of illness, baseline DLQI score, visit, and an interaction term for visit × baseline DLQI score. An appropriate variance structure will be used to model the within-patient errors. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom. Rescue data will be set to missing, otherwise data will be used in the analysis as observed. Missing DLQI scores will be handled within the framework of the repeated measures model, this includes data missing due to withdrawal or rescue.

The absolute value and change from baseline in DLQI at each visit will be summarized descriptively using observed data. Patients who withdraw or rescue will be summarized for visits up to withdrawal or rescue.

4.7.3 Exploratory Efficacy Endpoints

The following endpoints will provide additional information to assess treatment efficacy during the 52-week double blind treatment period. Exploratory endpoints will be summarized using frequency counts and percentages for categorical data and summary statistics for continuous data. Graphical displays will be used where relevant to show trends over time.

Treatment Failure:

The number and percentage of patients having treatment failure (defined in Section 4.2) from Week 12 to Week 52 will be summarized descriptively.

All data handling methods will follow those used for the derivation and classification of sustained complete remission.

4.7.3.1 Duration of Remission

Throughout this section, remission is referred to as a patient demonstrating the following criteria:

- PDAI[a] score = 0
- Prednisone (or equivalent) dose = 0 mg/day

Duration of remission will be calculated as the number of consecutive days a patient is in remission during the 52-week treatment period. Duration of remission will be analyzed using an analysis of covariance (ANCOVA) with logarithmic transformation (or any other suitable transformation depending on the distribution of the data) of duration of remission, adjusting for the stratification factors (duration of illness and geographical region) and presented as days. Patients who demonstrate fluctuating remission flare disease behaviour will be evaluated using their best response (i.e., longest duration) observed during the 52-week treatment period. For patients who withdraw or receive rescue therapy, the duration of remission will be calculated as the number of days in remission up until the point of withdrawal or receipt of rescue, respectively. Patients who never achieved remission will be assigned a duration of 0; this includes withdrawals.

A separate analysis will evaluate the duration of sustained complete remission. This will be conducted on the subset of patients who met the primary endpoint of sustained complete remission, and will follow a similar method to the mITT analysis. The duration of sustained complete remission will begin on the first day that a patient entered remission, providing the response was maintained for at least 16 consecutive weeks. Thus, the minimum duration, in the duration of sustained complete remission analysis, will be 16 weeks.

4.7.3.2 Patient Global Impression of Change

The PGIC is a measure of perceived change in PV symptoms as assessed by the patient. The patient uses a 7-item categorical scale ranging from "very much improved" to "very much worse" to answer "How would you rate the change of your pemphigus vulgaris symptoms since you started taking the study drug?" (Appendix 7).

The PGIC categories will be summarized by visit up to Week 52. Rescue data will be set to missing from the point of rescue onwards; no imputation will be made for rescue data.

4.7.3.3 Clinician Global Impression of Change

The CGIC is a measure of perceived change in PV symptoms as assessed by the clinician. The clinician uses a 7-item categorical scale ranging from "very much improved" to "very much worse" to answer "How would you rate the change in the patient's pemphigus vulgaris symptoms since the patient started taking the study drug?" (Appendix 8).

Data for CGIC will be described in the same manner as PGIC, as described in Section 4.7.3.2.

4.7.3.4 Remission Endpoints

Several remission endpoints will be evaluated in this study to assess the degree of remission attained within each treatment arm. The following five remission endpoints will be evaluated:

- Sustained Complete Remission
- Complete Remission – off therapy
- Complete Remission – on minimal therapy
- Partial Remission – off therapy
- Partial Remission – on minimal therapy

Complete remission (CR) off-therapy may be regarded as a subset of sustained CR; a patient who achieves CR off-therapy may also achieve sustained complete remission. The four categories of CR on/off therapy and partial remission (PR) on/off therapy will be regarded as mutually exclusive events. A patient will fall into the category that is the highest degree of remission obtained.

All remission endpoints will be derived using PDAI[a], prednisone (or equivalent), treatment failure, or withdrawal date in the same manner as that used for sustained complete remission in Section 4.7.1.

4.7.3.5 Complete Remission

CR is defined as absence of new or established lesions on 0 mg/day prednisone or equivalent for at least 8 consecutive weeks. Two categories of complete remission will

be evaluated: off-therapy and on minimal therapy. A patient will be considered a complete responder if the following criteria are achieved and maintained for a minimum of 8 consecutive weeks:

- CR off-therapy: PDAI[a] = 0 and prednisone = 0 mg/day
- CR on minimal therapy: PDAI[a] = 0 and 0 < prednisone ≤ 10 mg/day

The number and percentage of patients achieving CR by Week 24 and Week 52 will be summarized descriptively. A patient is considered a responder at a visit if they have demonstrated and maintained a response by that visit.

All data handling methods will follow those used for the derivation and classification of sustained complete remission.

4.7.3.6 Partial Remission

PR is defined as having transient lesions that heal within a week (as determined by no increase in PDAI total activity score between visits) for a minimum of 8 consecutive weeks. Two categories of PR will be evaluated: off-therapy and on minimal therapy. A patient will be considered a partial responder if the following criteria are achieved and maintained for a minimum of 8 consecutive weeks:

- PR off-therapy: No increase in PDAI total activity score and prednisone = 0 mg/day
- PR on minimal therapy: No increase in PDAI total activity score and 0 < prednisone ≤ 10 mg/day

The number of patients achieving PR by Week 24 and Week 52 will be summarized descriptively.

All data handling methods will follow those used for the derivation and classification of sustained complete remission.

4.7.3.7 Pemphigus Disease Area Index

The absolute value and change from baseline in PDAI score will be summarized by visit for each treatment arm. Plots of mean total PDAI score over time will be produced.

The main method for dealing with missing PDAI data will be via imputation. Details of PDAI imputation methods are described in Section 4.7.1. Patients who receive rescue therapy will be analyzed up to the point of rescue; no imputation will be carried out for post-rescue data.

4.7.3.8 European Quality of Life 5-Dimension Questionnaire, 3-Level Version

The EuroQol 5-dimension questionnaire (EQ-5D) is a generic HRQoL measure used to generate a utility score for use in health economic modeling. This self-reported questionnaire consists of two pages comprising the EQ-5D descriptive system and the EuroQol visual analogue scale (EQ VAS). The EQ-5D-3L descriptive system includes

three response options for each domain ranging from “no problems” to “extreme problems.” The EQ-5D-3L comprises 5 dimensions of health on the following domains:

- Mobility
- Self-care
- Usual activities
- Pain/discomfort
- Anxiety/depression

The EQ VAS records the respondent’s self-rated health status on a vertical graduated (0 – 100) visual analogue scale. The EQ-5D-3L will be converted to a single index utility score as per the EQ-5D-3L user guide (Version 4.0, Appendix 9).

Absolute values and change from baseline up to Week 52 in EQ-5D-3L single index utility score and the EQ VAS will be analyzed using summary statistics. Plots may also be produced.

Additional post hoc analysis may be done, depending on requirements for economic modelling.

Missing data will be handled as per the EQ-5D-3L user guide. Rescue data will be set to missing from the point of rescue onwards, no imputation will be made for rescue data.

4.7.3.9 Skindex-29

The Skindex-29 is a HRQoL measure evaluating the impact of patients’ dermatologic condition over the previous 4 weeks. The Skindex-29 consists of 30 questions on a 5-point Likert-type scale ranging from “Never” to “All the time” and comprises the following four domains:

- Emotions
- Symptoms
- Functioning
- Treatment

The single item asking about treatment is not scored. All other responses will be transformed to a linear scale (0 – 100), with higher scores indicating higher impact of skin disease. Scores will be calculated per the Skindex Manual (Version 2.0 or later; Appendix 10).

The mean and change from baseline global score will be summarized by visit and treatment group. Plots may also be produced.

The overall (global) score will be set to missing if more than 25% of the responses are missing. Data for patients who rescue will be set to missing from the point of rescue onwards, no imputation will be made for rescue data.

4.7.3.10 Disease Control

The number and percentage of patients in a state of disease control will be summarized by visit up to Week 52. Patients who receive rescue therapy will be considered as not in control from the point of rescue onwards, regardless of the recorded value on the eCRF.

4.7.4 Subgroup Analyses

The proportion of patients achieving complete and partial remission by Week 24 and Week 52, and sustained complete remission by Week 52 will be summarized descriptively for the following subgroups (summaries will only be carried out for subgroups where there is a minimum of 20% of patients from the overall population):

- Gender (male, female)
- Age ($\leq$50 years, $>$50 years)
- Region (North America, Rest of World)
- Duration of illness (0 to $\leq$1 years, $>$1 to 2 years)
- Baseline prednisone dose (60 or 80 mg/day by Day 1)
- Site of disease (mucosal only, mucocutaneous, cutaneous only)
- Baseline PDAI[a] score (quartiles)

Additional descriptive analyses for subgroups may be carried out, as required.

4.8 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

Serum rituximab concentrations will be listed by patient and nominal time of collection. Summary statistics of serum rituximab concentrations will be presented by nominal time of collection including n, mean, SD, coefficient of variation, geometric mean, median, and range.

Individual and mean serum rituximab concentration versus time profiles will be plotted on both linear and logarithmic scales.

A population PK analysis will be conducted to estimate the typical PK parameter values and to characterize the between-patient variability of those parameters. The effect of several physiological covariates on between-patient variability will also be studied in this patient population.

Exploratory graphical and statistical techniques, including linear, nonlinear, and logistic regression, faceted group comparison plots, etc., will be used to explore potential relationships between treatment groups, and markers of rituximab exposure, safety,

and/or efficacy (e.g., peripheral CD19+ B-cell counts, ADA, and autoantibody concentrations).

Depending on the number of PK samples available from this study where the investigation of rituximab PK is an exploratory objective, the data may be pooled with data from previous studies. Once the data is available and reviewed, a separate PK analysis plan may be drafted, if appropriate.

Full methodology will be documented with the final results of the analyses in a separate report.

4.9 SAFETY ANALYSES

The safety analysis will be performed on the safety population as defined in Section 4.1.3. Analyses of safety data up to Week 52 will include rescue patients up to the point of rescue only. Post-rescue data for patients who receive rescue therapy will be summarized separately for selected key safety outputs.

4.9.1 Exposure to Study Medication

The number and percentage of patients who received RTX will be summarized by treatment course, along with summary statistics for the dose of RTX received at each infusion and overall.

Duration of rituximab exposure will be calculated as:

exposure date − date of first rituximab infusion + 1 where exposure date is the earliest of the following:
- Week 52 visit (using upper visit window of Day 364+10 days)
- Date of early withdrawal

RTX exposure data will be determined using the dosing information as recorded on the RTX administration log on the eCRF and the actual treatment received as per IxRS. Exposure to RTX will not include RTX given as rescue, which will be analyzed separately.

MMF exposure data will be determined using the dosing information as recorded on the MMF administration log on the eCRF and the actual treatment received as per IxRS.

Exposure to MMF will be calculated as:

date of the last MMF administration − date of the first MMF administration + 1

Exposure to MMF will not include MMF given as rescue, which will be analyzed separately.

The total cumulative dose will be calculated by summing the total daily dose over the treatment duration. Average daily dose will be calculated as:

total cumulative dose ÷ treatment duration

4.9.2 Adverse Events

Adverse events will be mapped and reported using the Medical Dictionary for Regulatory Activities (MedDRA) thesaurus using the most current version at the time of analysis (version 17.1 or later). Glossaries showing the mapping of investigator verbatim terms to coded events will be produced.

A listing of AEs occurring during screening (prior to Day 1) will be produced to capture AEs occurring during the oral corticosteroid taper.

Treatment-emergent AEs (TEAEs) will be summarized. All AEs referenced in the section below will be TEAEs. TEAEs are defined as those AEs with observed or imputed onset date on or after the start date of trial treatment. Only where the most extreme intensity is greater than the initial intensity (or if most extreme intensity is not missing and initial intensity is missing) will events with an onset date prior to the start of trial treatment be considered treatment-emergent. An AE with a completely missing start date will be assumed to be treatment-emergent unless the AE has a complete non-imputed end date that is prior to study Day 1.

Listings will show all AEs reported, including AEs occurring after rescue therapy, and any AEs that may have occurred during the oral corticosteroid dose adjustments in the screening period.

AEs will be summarized by system organ class (SOC) and preferred term (PT), and will be presented in order of descending frequency summed across the treatment arms within each SOC and PT.

The following will be summarized, and/or listings produced where required:

- AEs
- Serious Adverse Events (SAEs)
- AEs by relationship to study medication
- AEs leading to study drug dose modification or interruption
- AEs by common terminology criteria grade
- Infections
- Serious Infections
- Corticosteroid-related AEs
- Infusion related reactions (IRRs) (Roche Standard AE Grouped Term [AEGT] MedDRA Basket)
- Serious IRRs
- AEs leading to withdrawal

- AEs leading to death
- Opportunistic infections (OI) , including serious OI (Roche Standard AE Grouped Term [AEGT] MedDRA Basket for Opportunistic Infections)

If an imbalance in exposure is found between treatment groups, AE rates per 100 patient-years exposure may be calculated. AE rate per 100 patient-years is defined as:

(number of events÷total exposure)×100

The 95% CIs will be calculated based on exact X^2 distribution.

4.9.3 Deaths

Details of any deaths will be presented in the form of an individual patient listing.

4.9.4 Laboratory Data

4.9.4.1 Safety Laboratory Parameters

All laboratory data will be converted to SI units. The Safety Laboratory Data Standardization will be used to implement reference ranges and marked abnormalities for laboratory data where possible.

Summary statistics for the absolute and change from baseline values will be produced by visit for laboratory parameters.

Marked abnormalities will be summarized for laboratory parameters. Summaries by National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) grade will also be produced for lab assessments, including shift tables from baseline to highest grade post-baseline.

Shift tables from baseline to highest NCI-CTCAE grade post-baseline will be produced for liver function laboratory tests.

Positive pregnancy results will be listed, as identified using the serum test results, where the data will be taken from CRF “pregnancy report” page.

4.9.4.2 Immunoglobulins

The absolute values and change from baseline will be summarized by treatment group and visit for total Ig, IgA, IgG, and IgM.

The proportion of patients with a value of total Ig, IgA, IgG, and IgM which is less than the lower of limit of normal per visit will be produced.

A plot of the mean value over time will be produced for total Ig, IgA, IgG, and IgM. The plots will display error bars (mean ± standard deviation) around the mean.

4.9.4.3 Lymphocyte Populations

The following NK, T and B lymphocytes will be analyzed via flow cytometry:

- CD19
- CD3
- CD4
- CD8
- Naïve B cells (CD19+ CD27-)
- Transitional B cells ($CD24^{high}CD38^{high}$)
- Memory B cells (CD19 + CD27 +)
- NK cells (CD56+)

The absolute, change and percent change from baseline values will be summarized by visit and treatment group for the above lymphocyte populations.

The proportion of patients with a lymphocyte value less than the lower limit of normal by visit and treatment group will be produced for the following lymphocytes:

- CD3
- CD4
- CD8
- CD19

A plot of the mean and median values over the 52-week treatment period will be produced for the following:

- CD19
- CD3
- CD4
- CD8

Plots will display error bars (mean ± standard deviation) around the mean.

Patients who received rituximab rescue therapy will be analyzed separately.

4.9.4.4 Desmoglein-1 and Desmoglein-3 Autoantibody Titers

Plots of mean anti-Dsg1 and anti-Dsg3 titers over the 52-week treatment period will be produced by treatment group, along with summary statistics. Listings will include positive/negative status.

Patients who received rituximab rescue therapy will be analyzed separately.

4.9.4.5 Anti-Drug Antibodies

The proportion of patients that are ADA (anti-rituximab antibodies [referred to as HACA in the protocol]) positive/negative at each visit will be summarized by treatment group. A listing of ADA results will be produced.

Depending on patient counts, subgroup analyses on ADA positive patients may be carried out on key safety and efficacy variables. Such analyses may include AEs, SAEs, IRRs, CD19, and remission endpoints before/after ADA positivity.

4.9.5 Vital Signs

Vital signs are measured at baseline, Week 52, and at infusion visits immediately prior to an infusion and then every 30 minutes until the end of the infusion.

Summary tables for the absolute value and change from baseline in vital sign assessments will be presented by visit up to Week 52. Abnormal results will also be summarized. Vital sign assessments include respiratory rate, pulse rate, temperature, systolic and diastolic blood pressure. Body weight will be measured at baseline, Week 24, and Week 52. Body mass index will be calculated for these time points.

4.10 DATA HANDLING METHODS

4.10.1 Presentation of Data

Data displays include listings, tables, and plots. All data displays will be presented by treatment group.

Data presented in summary tables will be displayed using summary statistics. Continuous data will include number of patients, mean, standard deviation, median, minimum, and maximum. Categorical data will include the number and percentage of patients.

4.10.2 Definitions

Baseline

Baseline is defined as the last non-missing value prior to receipt of study treatment. Typically, baseline will be the pre-dose Day 1 assessment, but may include screening results if the pre-dose Day 1 result is not available or missing. If the Day 1 or screening result is missing then no imputation will be made and the baseline result will be set to missing.

Study Day

For each patient, the first day of study treatment will be designated study day 1. Each subsequent assessment will be assigned a study day calculated as:

date of assessment−date of first dose+1

4.10.3 Multiple Values

For summaries displayed by visit, multiple values for a given assessment within a time window will be handled as follows:

Safety and Efficacy

The nearest non-missing value to the nominal timepoint will be assigned to the visit. If two values are equidistant from the nominal timepoint then the latest value will be selected.

4.10.4 Missing Data

Methods for handling missing data for efficacy analyses are described in detail throughout Section 4.7. Missing or partial dates for AEs, concomitant medications (except oral corticosteroids), laboratory assessments, and medical history will be imputed as detailed in Appendix 11.

4.11 INTERIM ANALYSES

No efficacy interim analyses are planned.

5. REFERENCES

Basra MK, Salek MS, Camilleri L, Sturkey R, et al. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. Dermatology. 2015;230(1):27-33.

Beissert S, Mimouni D, Kanwar AJ, et al. Treating Pemphigus Vulgaris with Prednisone and Mycophenolate Mofetil: A Multicenter, Randomized, Placebo-Controlled Trial, Journal of Investigative Dermatology, 2010;130(8):2041-8.

Yan X, Lee S, Li N. Missing Data Handling Methods in Medical Device Clinical Trials, Journal of Biopharmaceutical Statistics, 2009;19(6):1085-98.

Finlay AY, Khan G. Dermatology Life Quality Index (DLQI)—a simple practical measure for routine clinical use. Clinical and experimental dermatology. 1994 May 1;19(3):210-6.

Van den Broek J. A score test for zero inflation in a Poisson distribution. Biometrics. 1995 Jun 1:738-43.

Appendix 1 Protocol Synopsis

PROTOCOL SYNOPSIS

TITLE:	**A RANDOMIZED, DOUBLE-BLIND, DOUBLE-DUMMY, ACTIVE-COMPARATOR, MULTICENTER STUDY TO EVALUATE THE EFFICACY AND SAFETY OF RITUXIMAB VERSUS MMF IN PATIENTS WITH PEMPHIGUS VULGARIS**
PROTOCOL NUMBER:	WA29330
VERSION NUMBER:	5
EUDRACT NUMBER:	2014-000382-41
IND NUMBER:	121595
TEST PRODUCT:	Rituximab (RO 0452294)
PHASE:	III
INDICATION:	Moderate-to-severely active pemphigus vulgaris
SPONSOR:	F. Hoffmann-La Roche Ltd

Objectives

Efficacy Objectives

The primary efficacy objective for this study is as follows:

- To evaluate the efficacy of rituximab compared with mycophenolate mofetil (MMF) in achieving sustained complete remission, evaluated by the Pemphigus Disease Area Index (PDAI; see Section 3.4.1.1), and assessed at Week 52 in patients with moderate-to-severely active pemphigus vulgaris (PV)

The secondary efficacy objectives for this study are as follows:

- To evaluate the efficacy of rituximab compared with MMF, as measured by the time to disease flare, the duration of sustained complete remission, the total number of disease flares during the treatment period, and the time to initial sustained complete remission
- To assess corticosteroid exposure over 52 weeks
- To assess the effect of rituximab compared with MMF on health-related quality of life (HRQoL), as measured by the Dermatology Life Quality Index (DLQI)
- To assess the effect of rituximab compared with MMF on patients' impression of PV symptoms, as measured by the Patient Global Impression of Change (PGIC) questionnaire
- To assess the effect of rituximab compared with MMF on clinician impression of patients' PV symptoms, as measured by the Clinician Global Impression of Change (CGIC) questionnaire

Safety Objectives

The safety objectives for this study are as follows:

- To evaluate the safety of rituximab compared with MMF, with a focus on AEs and safety laboratory values
- To evaluate corticosteroid-related AEs in relation to corticosteroid exposure

Appendix 1 Protocol Synopsis (cont.)

Exploratory Objectives

The exploratory objectives for this study will include, but are not limited to, the following:

- To evaluate the efficacy of rituximab compared with MMF, as measured by the proportion of patients achieving complete or partial remission by 24 weeks and 52 weeks
- To explore the pharmacokinetics and pharmacodynamics of rituximab, and the pharmacodynamics of MMF, in patients with PV
- To evaluate the effect of rituximab as compared with MMF on the change in PDAI activity score
- To evaluate the effect of rituximab as compared with MMF on anti-desmoglein (anti-Dsg) autoantibody titers (anti-Dsg1 and anti-Dsg3) and other mechanistic studies of interest in patients with PV
- To assess the effect of rituximab compared with MMF on health utilities as measured by the European Quality of Life (EuroQol) 5-Dimension Questionnaire, 3-level version (EQ-5D-3L)
- To assess the effect of rituximab compared with MMF on HRQoL, as measured by the Skindex-29
- To evaluate the efficacy of rituximab compared with MMF, as measured by the proportion of patients experiencing treatment failure from Week 12 to Week 52

Study Design

Description of Study

This is a Phase III, randomized, double-blind, double-dummy, active-comparator, parallel-arm, multicenter study to evaluate the efficacy and safety of rituximab compared with MMF in patients with moderate-to-severely active PV requiring 60–120 mg/day oral (PO) prednisone or equivalent (1.0-1.5 mg/kg/day). Patients must have a confirmed diagnosis of PV within the previous 24 months (by skin or mucosal biopsy and immunohistochemistry) and evidence of moderate-to-severely active disease at screening (defined as at total PDAI activity score of $\geq$15).

This international study will be conducted at investigational sites throughout North America, Europe, the Middle East, and Latin America. It is expected that approximately 60 centers will participate. Approximately 132 patients will be randomized in a 1:1 ratio to receive either rituximab plus MMF placebo or rituximab placebo plus MMF. Randomization will be stratified by duration of illness (newly diagnosed [i.e., diagnosed within the 1 year prior to screening] vs. diagnosed greater than 1 year) and geographical region (North America [U.S./Canada] vs. rest of world).

The study will consist of three periods: a screening period of up to 28 days, a 52-week double-blind treatment period, and a 48-week safety follow-up (SFU) period that begins at the time of study treatment completion or discontinuation.

Rituximab (1000 mg or matching placebo) will be administered by intravenous (IV) infusion on Day 1 and Day 15, with repeat rituximab (or matching placebo) administration on Day 168 and Day 182 provided specific safety criteria have been met (see Section 4.6.4). MMF (500 mg or matching placebo) will be administered PO twice daily (every 12 hours [Q12H]), starting with a total dose of 1 g/day on Day 1. The MMF dose will be titrated to achieve a goal of 2 g/day in divided doses (1 g Q12H) by Week 2. Treatment with MMF (or matching placebo) will continue through Week 52.

All patients who withdraw from the treatment period or who complete the total 52-week treatment period must return for post Week 52 SFU assessments at 12, 24, 36, and 48 weeks after either the early withdrawal visit or study treatment completion, respectively. Thus, patients will be followed for approximately 1 year in the SFU period.

Appendix 1 Protocol Synopsis (cont.)

Telemedicine

For a small proportion of patients (approximately 10 patients) at one investigational site, the Sponsor is proposing the use of telemedicine (TM) consultation visits between the patient and the Principal Investigator to make the trial more accessible to this population of patients with a rare disease. The local dermatologist, a research nurse, and other medical personnel in the patient's immediate vicinity will participate throughout the trial as needed. Data collected via TM will be in accordance with the American Telemedicine Association 2012 Guidelines for Teledermatology (ATA 2012). The software platform to be used in the trial will allow efficient, high-quality data collection of all relevant clinical data (including the images of the skin and mucosa, prior medical records, notes, and laboratory tests) for easy and reliable review by the Principal Investigator. This type of assessment has become the standard for TM clinical care for the diagnosis and treatment of both mild and severe types of skin diseases. Images will be time-stamped and stored in a Health Insurance Portability and Accountability Act (HIPAA)-compliant advanced TM platform adapted for this specific trial use. Other safety and efficacy data sets will be transmitted through the TM platform to the Principal Investigator as well. The local dermatologist or other local primary care doctors may still provide standard-of-care services that do not require training on the protocol or investigational drug. All clinical trial-related medical decisions will be made by the Principal Investigator and or safety assessor.

Number of Patients

Approximately 132 patients with active moderate-to-severely active PV will be recruited into this study.

Target Population

Inclusion Criteria

Patients must meet the following criteria for study entry:

- Age 18–75 years
- Signed Informed Consent Form
- First confirmed diagnosis of PV within the previous 24 months, based on the presence of histological features of acantholysis via skin or mucosal biopsy and one of the following: tissue-bound immunoglobulin G (IgG) antibodies by direct immunofluorescence on the surface of affected epithelium or serological detection of serum Dsg3 autoantibodies against epithelial cell surface either by indirect immunofluorescence microscopy or by enzyme-linked immunosorbent assay
- Presence of moderate-to-severely active disease, defined as overall PDAI activity score of $\geq$15
- Receiving standard-of-care corticosteroids consisting of 60–120 mg/day PO prednisone or equivalent (1.0 – 1.5 mg/kg/day) and, in the judgment of the investigator, expected to benefit from the addition of immunosuppressive therapy
- For women who are not postmenopausal ($\geq$ 12 months of non–therapy-induced amenorrhea) or surgically sterile (absence of ovaries and/or uterus): agreement to remain abstinent or use two effective methods of contraception, including at least one method with a failure rate of $<$ 1% per year, during the treatment period and for at least 12 months after the last dose of study treatment

  Abstinence is acceptable only if it is in line with the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not acceptable methods of contraception.

  Barrier methods must always be supplemented with the use of a spermicide.

  Examples of contraceptive methods with a failure rate of $<$ 1% per year (highly effective contraceptive methods) include tubal ligation, male sterilization, hormonal implants, established, proper use of combined oral or injected hormonal contraceptives, and certain intrauterine devices.

- For men (including those who have undergone a vasectomy): agreement to remain abstinent or use a condom during the treatment period and for at least 12 months after the last dose of study treatment and agreement to refrain from donating sperm during this same period

  Abstinence is only acceptable if it is in line with the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not acceptable methods of contraception.

  In addition to male contraception, agreement to advise female partners of childbearing potential to use highly effective contraception during the study and for at least 12 months after the last dose of study treatment
- Agreement to avoid excessive exposure to sunlight during study participation
- Able to comply with the study protocol, in the investigator's judgment

Exclusion Criteria

Patients who meet any of the following criteria will be excluded from study entry:

- Diagnosis of pemphigus foliaceus or evidence of paraneoplastic pemphigus or other non-PV autoimmune blistering disease
- History of a severe allergic or anaphylactic reaction to humanized or murine monoclonal antibodies, or known hypersensitivity to any component of rituximab
- Known hypersensitivity or contraindication to MMF, mycophenolic acid, polysorbate, or oral corticosteroids
- Lack of peripheral venous access
- Pregnant or lactating, or intending to become pregnant during the study

  Women who are not postmenopausal (≥ 12 months of non–therapy-induced amenorrhea) or surgically sterile must have two negative results with a sensitivity of ≥25 mIU/mL: one from a serum pregnancy test at Day −8 to Day −10 of screening and another from a urine pregnancy test at Day 1 prior to randomization.
- Participated in another interventional clinical trial within 28 days prior to randomization
- Use of any investigational agent within 28 days or 5 elimination half-lives prior to randomization (whichever is the longer)
- Significant cardiovascular or pulmonary disease (including obstructive pulmonary disease)
- Evidence of any new or uncontrolled concomitant disease that, in the investigator's judgment, would preclude patient participation, including but not limited to nervous system, renal, hepatic, endocrine, malignant, or gastrointestinal disorders
- Any concomitant condition that required treatment with oral or systemic corticosteroids within 12 weeks prior to randomization
- Treatment with IV immunoglobulin (Ig), plasmapheresis, or other similar procedure within 8 weeks prior to randomization
- Treatment with immunosuppressive medications (e.g., azathioprine, MMF) within 1 week prior to randomization
- Treatment with cyclophosphamide within 12 weeks prior to randomization
- History of or currently active primary or secondary immunodeficiency, including known history of HIV infection and other severe Immunodeficiency blood disorders
- Known active infection of any kind (excluding fungal infections of nail beds) or any major episode of infection requiring hospitalization or treatment with IV anti-infectives within 4 weeks prior to screening, or completion of oral anti-infectives within 2 weeks prior to randomization

  Entry into this study may be reconsidered once the infection has fully resolved.

- History of or current cancer, including solid tumors, hematologic malignancies, and carcinoma in situ (except basal cell carcinoma and squamous cell carcinoma of the skin that have been excised and cured)
- Currently active alcohol or drug abuse, or history of alcohol or drug abuse within 24 weeks prior to screening
- Major surgery within 4 weeks prior to randomization, excluding diagnostic surgery
- Treatment with rituximab or a B cell–targeted therapy (e.g., anti-CD20, anti-CD22, or anti-BLyS) within 12 months prior to randomization
- Treatment with a live or attenuated vaccine within 28 days prior to randomization

  It is recommended that a patient's vaccination record and the need for immunization prior to study entry be carefully investigated.
- Aspartate aminotransferase (AST), alanine aminotransferase (ALT), or amylase $> 2.5 \times$ the upper limit of normal (ULN)
- Absolute neutrophil count (ANC) $< 1.5 \times 10^3/\mu L$
- Hemoglobin < 8.0 g/dL
- Positive test results for hepatitis B surface antigen (HBsAg), hepatitis B core antibody (HBcAb), or hepatitis C virus (HCV) serology at screening

Length of Study

Based on study enrollment projections, this study is estimated to take approximately 4 years to complete, from first patient in to last patient's last visit (LPLV), when the final patient completes the 48-week SFU period.

End of Study

The end of the study is defined as the date when the LPLV occurs. LPLV is expected to occur approximately 2 years after the last patient is enrolled, assuming a 52-week treatment period and a 48-week SFU period for the last patient.

Outcome Measures

Primary Efficacy Outcome Measure

The primary efficacy outcome measure for this study is as follows:

- Proportion of patients (excluding TM patients) who achieve a sustained complete remission without experiencing an event that constitutes treatment failure (as defined in Section 3.1.3), as measured at Week 52

  Sustained complete remission is defined as achieving healing of lesions with no new active lesions (i.e., PDAI activity score of 0) while on 0 mg/day prednisone or equivalent, and maintaining this response for a total of at least 16 consecutive weeks, during the 52-week treatment period.

  Patients with transient *new* lesions for 1 week or less that heal without additional systemic corticosteroid therapy will not be considered to have experienced treatment failure.

Secondary Efficacy Outcome Measures

The secondary efficacy outcome measures are as follows:

- Cumulative oral corticosteroid dose (prednisone or equivalent) over the treatment period
- Total number of disease flares during the treatment period
- Time to sustained complete remission
- Time to disease flare
  - Disease flare is defined as the appearance of three or more new lesions a month that do not heal spontaneously within 1 week or by the extension of established lesions in a patient who has achieved disease control.
- Change in HRQoL, as measured by the DLQI score from baseline to Week 52

- Duration of sustained complete remission
- Patients' impression of change in PV symptoms, as measured by the PGIC score during the treatment period
- Clinician impression of change in patients' PV symptoms, as measured by the CGIC score during the treatment period

Safety Outcome Measures

The safety outcome measures include, but are not limited to, the following:

- Nature, frequency, and severity of adverse events, including serious adverse events and adverse events leading to discontinuation
- Vital signs and clinical laboratory test results (including complete blood count and blood chemistry)
- Incidence of human anti−chimeric antibody (HACA)
- Circulating B cells, T cells, natural killer (NK) cells, plasma cells, and other leukocytes
- Plasma Ig levels (total Ig, IgG, IgM, and IgA)
- Corticosteroid-related adverse events in relation to corticosteroid exposure

Exploratory Outcome Measures

The exploratory outcome measures for this study will include, but are not limited to, the following:

- Proportion of patients achieving a complete remission by Week 24 and by Week 52

  Complete remission is defined as achieving wound healing with no new active lesions (i.e., PDAI activity score of 0) for at least 8 consecutive weeks during the 52-week treatment period.

- Proportion of patients achieving a partial remission by Week 24 and by Week 52

  Partial remission is defined as the presence of transient new lesions that heal within 1 week (while the patient is receiving minimal therapy, including topical corticosteroids). Minimal therapy is defined as ≤10 mg/day prednisone (or equivalent) for at least 8 consecutive weeks during the 52-week treatment period.

- Pharmacokinetics (PK)/pharmacodynamics (PD) of rituximab

  PK/PD parameters include, but are not limited to, serum levels of rituximab, peripheral CD19+ B-cell counts, HACA, and autoantibody concentrations.

- To explore the pharmacodynamics of MMF in patients with PV
- Change in total PDAI activity score during the treatment period
- Change from baseline in anti-Dsg1 and anti-Dsg3 autoantibodies
- Change in health utilities as assessed by the EQ-5D-3L score from baseline to Week 52
- Change in HRQoL as measured by the Skindex-29 from baseline to Week 52

Investigational Medicinal Products

Test Product-Rituximab

Patients randomized to the rituximab arm will receive treatment with 1000 mg IV rituximab (or matching placebo) on Day 1 and Day 15, with repeat rituximab (or matching placebo) administration on Day 168 and Day 182 provided specific safety criteria have been met (see Section 4.6.4).

Comparator-MMF

Patients randomized to the MMF arm will receive treatment with 500 mg MMF (or matching placebo) administered orally twice daily (Q12H) starting on Day 1. MMF dose will then be titrated to achieve a maximum dose goal of 2 g/day given as a divided oral dose (1 g Q12H) by Week 2. Slower titration will be allowed on the basis of tolerability. Treatment with MMF (or matching placebo) will continue through Week 52.

Rituximab—F. Hoffmann-La Roche Ltd
42/Statistical Analysis Plan WA29330

Appendix 1 Protocol Synopsis (cont.)

Non-Investigational Medicinal Products

In order to reduce the frequency and severity of infusion-related reactions, all patients will receive methylprednisolone 100 mg or saline solution prior to infusion of rituximab or rituximab placebo, respectively. This premedication will be administered by slow IV infusion, and administration should be completed at least 30 minutes prior to infusion of rituximab or rituximab placebo.

All patients should be premedicated with paracetamol/acetaminophen (1 g PO) and an antihistamine (diphenhydramine HCl 50 mg PO [or IV equivalent] or equivalent dose of a similar agent, or in accordance with local approved labeling) 30 to60 minutes prior to the start of the infusion of rituximab or rituximab placebo.

Statistical Methods

Primary Analysis

The proportion of patients (excluding TM patients) achieving a sustained complete remission (as defined in Section 3.4.1.1) in the rituximab and MMF arms will be compared using a stratified Cochran-Mantel-Haenszel test, adjusting for randomization stratification variables. Results will be summarized descriptively by treatment arm and expressed as proportions, corresponding adjusted 95% confidence intervals of the difference between response rates, and p-values.

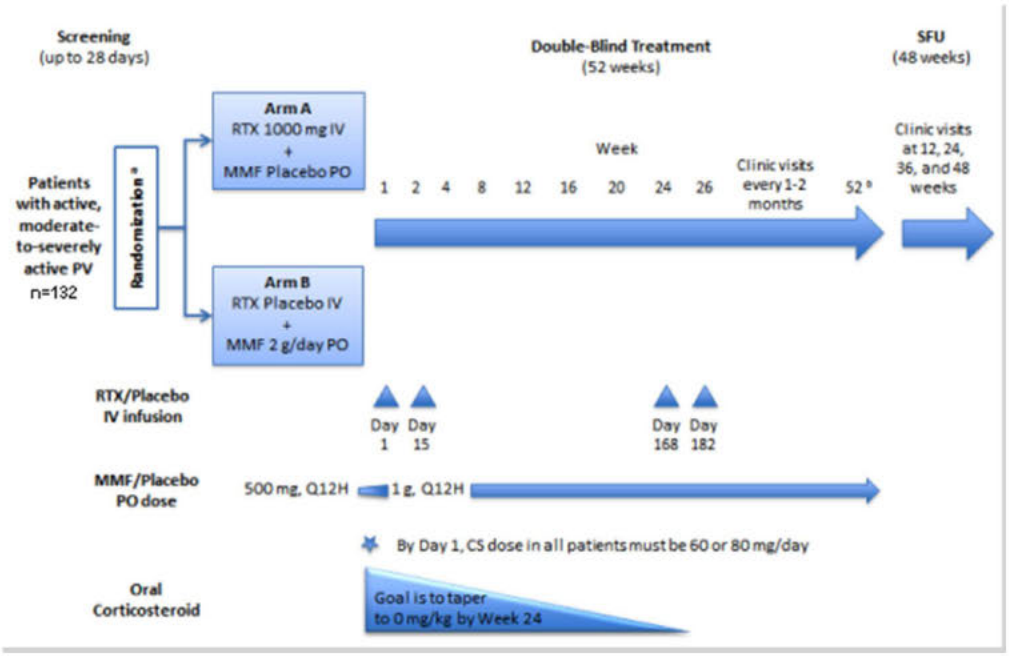
Patients who meet the pre-specified definitions of treatment failure before entering sustained complete remission (as defined in Section 3.1.3) will be deemed non-responders in the primary analysis.

Determination of Sample Size

The primary efficacy endpoint of this study is the proportion of patients (excluding TM patients) achieving a sustained complete remission, as assessed at Week 52, that has been maintained for $\geq$16 consecutive weeks during the 52-week treatment period. With use of limited data available in the literature from randomized clinical trials of MMF and investigator-initiated trials of rituximab, it is estimated that approximately 40% of patients with PV receiving MMF will achieve a sustained complete remission. It is estimated that patients receiving rituximab will induce a sustained complete remission rate of 65%. On the basis of these assumptions, a total of 122 patients randomized to the rituximab arm or the MMF arm in a 1:1 ratio (61 patients in the rituximab arm and 61 patients in the MMF arm) will yield approximately 80% power in a two-sided test at the 5% significance level. To account for the 10 TM patients excluded from the primary efficacy analysis, approximately 132 patients will be randomized in total.

No adjustment will be made to account for dropouts.

Appendix 2 Overview of Study Design



CS = corticosteroid (prednisone or equivalent); IV = intravenous; MMF = mycophenolate mofetil; PO = by mouth; PV = pemphigus vulgaris; Q12H = every 12 hours; RTX = rituximab; SFU = safety follow-up.

[a] Administrations of the first dose of study treatment (Day 1) should occur within 24 hours following the baseline assessments. However, administration up to 72 hours will be allowed when necessary. The second infusion should occur on Day 15 ± 1 day.

[b] Primary efficacy endpoint evaluated by PDAI is to be measured at Week 52.

Appendix 3 Schedule of Assessments

Week	Screen (−28 to −1)	1	2	4	8	12	16	20	24	26	32	40	48	52	WD [a]	SFU [b]
Day		1 (BL and Rz)	15 ±1	28 ±3	56 ±3	84 ±3	112 ±3	140 ±3	168 ±3	182 ±1	224 ±3	280 ±3	336 ±3	364 ±3		
Informed consent	x															
Medical history, [c] previous medications, and baseline conditions	x															
Pregnancy test (serum) [d]	x															
Pregnancy test (urine) [d]		x		x	x	x	x	x	x	x	x	x	x	x		
Complete physical examination [e]	x															
Body weight	x								x					x	x	
Height	x															
Vital signs [f]	x	x [g]	x [g]						x [g]	x [g]				x	x	
ECG	x															
CXR [h]	x															
Efficacy and PRO/ClinRO																
Disease activity (PDAI)	x	x		x	x	x	x	x	x		x	x	x	x	x	x
Flare assessment	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Appendix 3 Schedule of Assessments (cont.)

Week	Screen	1	2	4	8	12	16	20	24	26	32	40	48	52		
Day	(−28 to −1)	1 (BL and Rz)	15 ±1	28 ±3	56 ±3	84 ±3	112 ±3	140 ±3	168 ±3	182 ±1	224 ±3	280 ±3	336 ±3	364 ±3	WD [a]	SFU [b]
DLQI [i]		x				x			x			x		x	x	x
PGIC [i]						x			x			x		x	x	x
CGIC [i]						x			x			x		x	x	x
EQ-5D-3L [i]		x				x			x			x		x	x	
Skindex-29 [i]		x				x			x			x		x	x	x
Prednisone (or equivalent) dose assessment/ adjustment [j]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Safety and Routine Laboratory Assessments																
Adverse events [k]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x [l]
Concomitant medications	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x [l]
Hematology [m]	x	x [n]	x [n]	x	x		x		x [n]	x [n]	x	x	x	x	x	
Blood chemistry [o]	x	x [n]	x [n]	x	x		x		x [n]	x [n]	x	x	x	x	x	
HBsAg, HBcAb, and HCV antibody	x															
Urinalysis [p]	x	x			x		x		x		x		x		x	

Appendix 3 Schedule of Assessments (cont.)

Week	Screen (−28 to −1)	1	2	4	8	12	16	20	24	26	32	40	48	52		
Day		1 (BL and Rz)	15 ±1	28 ±3	56 ±3	84 ±3	112 ±3	140 ±3	168 ±3	182 ±1	224 ±3	280 ±3	336 ±3	364 ±3	WD [a]	SFU [b]
Immunology and Other Labs																
Lymphocyte subtypes and FACS panel [q]		x [n]	x [n]	x	x		x		x [n]	x [n]		x		x	x	x
Quantitative Ig [r]		x					x		x			x		x		x
Autoantibodies [s]		x			x				x			x		x	x	x
HACA		x							x					x	x	x
PK samples		x [t]	x [t]	x	x		x		x [t]	x [t]	x	x	x	x	x	
Optional serum biomarker sample for RCR		x [u]		x	x		x		x [u]	x [u]				x	x	
Optional whole blood sample (RNA) for RCR		x [u]		x	x		x		x [u]	x [u]				x	x	
Optional whole blood sample (DNA) for RCR [v]		x														

Appendix 3 Schedule of Assessments (cont.)

Week	Screen	1	2	4	8	12	16	20	24	26	32	40	48	52		
Day	(−28 to −1)	1 (BL and Rz)	15 ±1	28 ±3	56 ±3	84 ±3	112 ±3	140 ±3	168 ±3	182 ±1	224 ±3	280 ±3	336 ±3	364 ±3	WD [a]	SFU [b]
Study Treatment Dispensed																
Methylprednisolone/ saline solution premedication [w]		x	x						x	x						
Rituximab/matching placebo [x]		x [y]	x						x	x						
MMF/matching placebo [z]		x	x	x	x	x	x	x	x		x	x	x			

ALT = alanine aminotransferase; AST = aspartate aminotransferase; BL = baseline; BUN = blood urea nitrogen; CGIC = Clinician Global Impression of Change; ClinRO = clinician-reported outcomes; CXR = chest X-ray; DLQI = Dermatology Life Quality Index; Dsg1 = desmoglein 1; Dsg3 = desmoglein 3; ECG = electrocardiogram; eCRF = electronic Case Report Form; EQ-5D-3L = EuroQol 5-Dimension Questionnaire, three-level version; FACS = fluorescence-activated cell sorting; HACA = human anti-chimeric antibodies; HBcAg = hepatitis B core antibody; HBsAg = hepatitis B surface antigen; HCV = hepatitis C virus; Ig = immunoglobulin; IV = intravenous; LDH = lactate dehydrogenase; MMF = mycophenolate mofetil; PDAI = Pemphigus Disease Area Index; PGIC = Patient Global Impression of Change; PK = pharmacokinetic; PRO = patient-reported outcome; Q12H = every 12 hours; RBC = red blood cell; RCR = Roche Clinical Repository; Rz = randomization; SFU = safety follow-up; WBC = white blood cell; WD = withdrawal visit.

Notes: All assessments and laboratory sampling should be performed within 3 days of the scheduled visit, unless otherwise specified. On treatment days, all assessments should be performed prior to dosing, unless otherwise specified.

a Patients who discontinue both study treatments early, and patients who experience treatment failure (as defined in Section 3.1.3) before Week 12, will be asked to return to the clinic for a treatment WD visit. If a patient withdraws at or during a scheduled or unscheduled visit, the visit should proceed as the withdrawal visit, that is, only the withdrawal visit assessments should be conducted. After the withdrawal visit, the patient will be asked to enter the SFU period of the study. An early withdrawal visit is not required for a patient to receive rescue therapy between Week 12 through Week 52, and patients should continue to be followed in the 52-week treatment period.

b Safety follow-up required at 12, 24, 36, and 48 weeks (± 1 week) after withdrawal from or completion of treatment (see Appendix 2- Schedule of Assessments for the Safety Follow-Up Period of protocol for Study WA29330). Required follow-up information will be collected every 3 months or until death, loss to follow-up, or study termination by the Sponsor.

Appendix 3 Schedule of Assessments (cont.)

c Medical history includes clinically significant diseases, surgeries, cancer history (including prior cancer therapies and procedures), reproductive status, smoking history, use of alcohol and drugs of abuse, and all previous medications (e.g., prescription drugs, over-the-counter drugs, herbal or homeopathic remedies, nutritional supplements) used by the patient within 6 months prior to the screening visit.

d All women who are not postmenopausal (≥ 12 months of non–therapy-induced amenorrhea) or surgically sterile will have a serum pregnancy test at screening. Urine pregnancy tests will be performed at specified subsequent visits. If a urine pregnancy test is positive, it must be confirmed by a serum pregnancy test. Prior to starting study drug, female patients of childbearing potential must have two negative pregnancy tests. The second test on Day 1 prior to randomization should be performed 8 to 10 days after the first test during the screening period.

e Includes evaluation of the head, eyes, ears, nose, and throat, and the cardiovascular, dermatological, musculoskeletal, respiratory, gastrointestinal, genitourinary, and neurologic systems. Record abnormalities observed at baseline on the eCRF. Record new or worsened clinically significant abnormalities on the Adverse Event eCRF.

f Includes respiratory rate, pulse rate, temperature, and systolic and diastolic blood pressure while the patient is in a seated position.

g Vital signs should be taken immediately prior to infusion and then every 30 minutes until completion of the rituximab or rituximab placebo infusion.

h If screening chest radiographs taken within the past 3 months show no clinically significant abnormality, further radiographic assessment is not required.

i The DLQI, PGIC, EQ-5D-3L, and Skindex-29 will be completed by the patient before any other non-PRO assessments and before the patient receives any disease-status information or study drug during that visit. The CGIC, where possible, will be completed prior to the completion of other non-PRO assessments.

j To be captured on the eCRF.

k After informed consent has been obtained but prior to initiation of study drug, only serious adverse events caused by a protocol-mandated intervention should be reported. In addition, any adverse events occurring during any corticosteroid dose adjustment during the screening period should be reported.

l See Appendix 2 (Schedule of Assessments for the Safety Follow-Up Period) of protocol for Study WA29330.

m Includes WBC count, RBC count, hemoglobin, hematocrit, platelet count, and differential (neutrophils, eosinophils, basophils, monocytes, lymphocytes, and other cells).

n Samples to be collected prior to administration of IV methylprednisolone or saline solution and 30 minutes following completion of rituximab or rituximab placebo infusion, respectively.

o Includes sodium, potassium, chloride, bicarbonate, glucose (except at 30 minutes following completion of study drug infusion at Weeks 1, 2, 24, and 26), BUN or urea, creatinine, total protein, albumin, phosphorus, calcium, total and direct bilirubin, alkaline phosphatase, ALT, AST, uric acid, and LDH.

p Dipstick for blood, protein, and glucose (microscopic examination at central laboratory if abnormal and applicable).

Appendix 3 Schedule of Assessments (cont.)

q See Section 4.5.6 for T- and B-cell assessments.

r Total Ig, IgG, IgA, and IgM.

s Anti-Dsg1 and anti-Dsg3 antibodies.

t On infusion days, PK samples will be collected prior to administration of IV methylprednisolone or saline solution and 30 minutes following completion of rituximab or rituximab placebo infusion, respectively.

u Samples to be collected prior to administration of IV methylprednisolone or saline solution for patients in the rituximab or rituximab placebo arms, respectively.

v The optional DNA sample can be collected prior to or after study drug administration.

w Methylprednisolone or saline solution should be administered by slow IV infusion to be completed at least 30 minutes prior to each infusion of rituximab or rituximab placebo, respectively.

x Rituximab at a dose of 1000 mg (or matching placebo) will be administered by IV infusion on Day 1 and Day 15, with repeat rituximab (or matching placebo) administration on Day 168 and Day 182 provided specific safety criteria have been met (see Section 4.6.4). See Appendix 3 (Procedures for the Intravenous Administration of Rituximab) of protocol for Study WA29330 for administration instructions.

y Administration of the first dose of study treatment (Day 1) should occur within 24 hours following the baseline assessments; however, administration up to 72 hours will be allowed when necessary. The second infusion should occur on Day 15 $\pm$ 1 day.

z MMF (500 mg or matching placebo) will be administered orally twice daily (Q12H) starting on Day 1. Patients will be initiated on MMF at 1 g/day or matching placebo in divided oral doses (Q12H). The MMF dose will be titrated to achieve a goal of 2 g/day in divided oral doses (1 g Q12H) by Week 2.

Appendix 4 Corticosteroid Treatment Guideline

Step	No. of Weeks	Daily Dose of Prednisone or Equivalent	Presence of Disease Activity [a]	
			No	Yes
1 (start) [b]	1 week	80 mg	Go to Step 2	Go to Step A
A	1 week	120 mg	Go to Step 2	Go to Step B
B	1 week	240 mg	Go to Step A	Discontinue taper [c]
2 (start) [b]	1 week	60 mg	Go 1 step forward	Go to Step 1
3	1 week	40 mg		Go 1 step back
4	2 weeks	30 mg		
5	2 weeks	20 mg		Go 2 steps back
6	2 weeks	15 mg		
7	2 weeks	12.5 mg		
8	2 weeks	10 mg		
9	2 weeks	7.5 mg		
10	2 weeks	5 mg		
11	2 weeks	2.5 mg	Stop corticosteroid	
12	During complete remission	0 mg		Restart at Step 5

Notes: On Day 1, start the oral corticosteroid regimen with an initial dosage of 60 or 80 mg/day prednisone (or equivalent), as outlined in the schedule. Increase the dose if disease activity is present after 1 week. If necessary, increase the dose again (to 240 mg/day) if disease activity is still present after the second week.

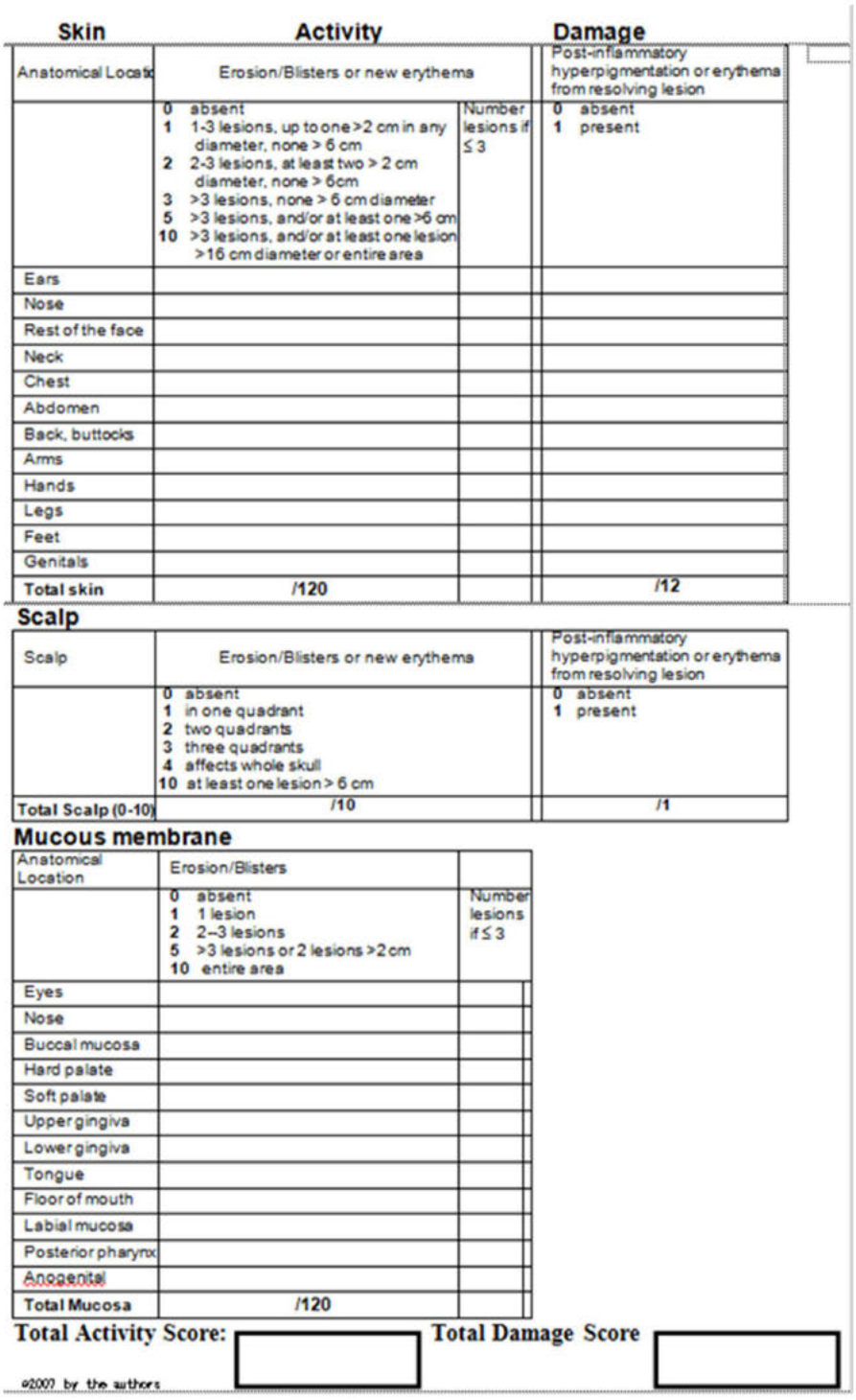
[a] Disease activity is defined as the appearance of new lesions or the extension of established lesions.

[b] Patients on 80 mg/d on Day 1 will start at Step 1. Patients on 60 mg/d on Day 1 will start at Step 2.

[c] Patients whose disease is not controlled after increasing the corticosteroid dose to 240 mg/day will be considered to have experienced treatment failure and may receive standard-of-care treatment per the investigator's best medical judgment.

Source: Modified from Mentink LF, Mackenzie MW, Tóth GG, et al. Randomized controlled trial of adjuvant oral dexamethasone pulse therapy in pemphigus vulgaris (PEMPULS Trial). Arch Dermatol 2006;142:570–6.

Appendix 5 Pemphigus Disease Area Index



Skin — Activity

Anatomical Locati	Erosion/Blisters or new erythema		Post-inflammatory hyperpigmentation or erythema from resolving lesion
	0 absent 1 1-3 lesions, up to one >2 cm in any diameter, none > 6 cm 2 2-3 lesions, at least two > 2 cm diameter, none > 6cm 3 >3 lesions, none > 6 cm diameter 5 >3 lesions, and/or at least one >6 cm 10 >3 lesions, and/or at least one lesion >16 cm diameter or entire area	Number lesions if ≤ 3	0 absent 1 present
Ears			
Nose			
Rest of the face			
Neck			
Chest			
Abdomen			
Back, buttocks			
Arms			
Hands			
Legs			
Feet			
Genitals			
Total skin	/120		/12

Scalp

Scalp	Erosion/Blisters or new erythema	Post-inflammatory hyperpigmentation or erythema from resolving lesion
	0 absent 1 in one quadrant 2 two quadrants 3 three quadrants 4 affects whole skull 10 at least one lesion > 6 cm	0 absent 1 present
Total Scalp (0-10)	/10	/1

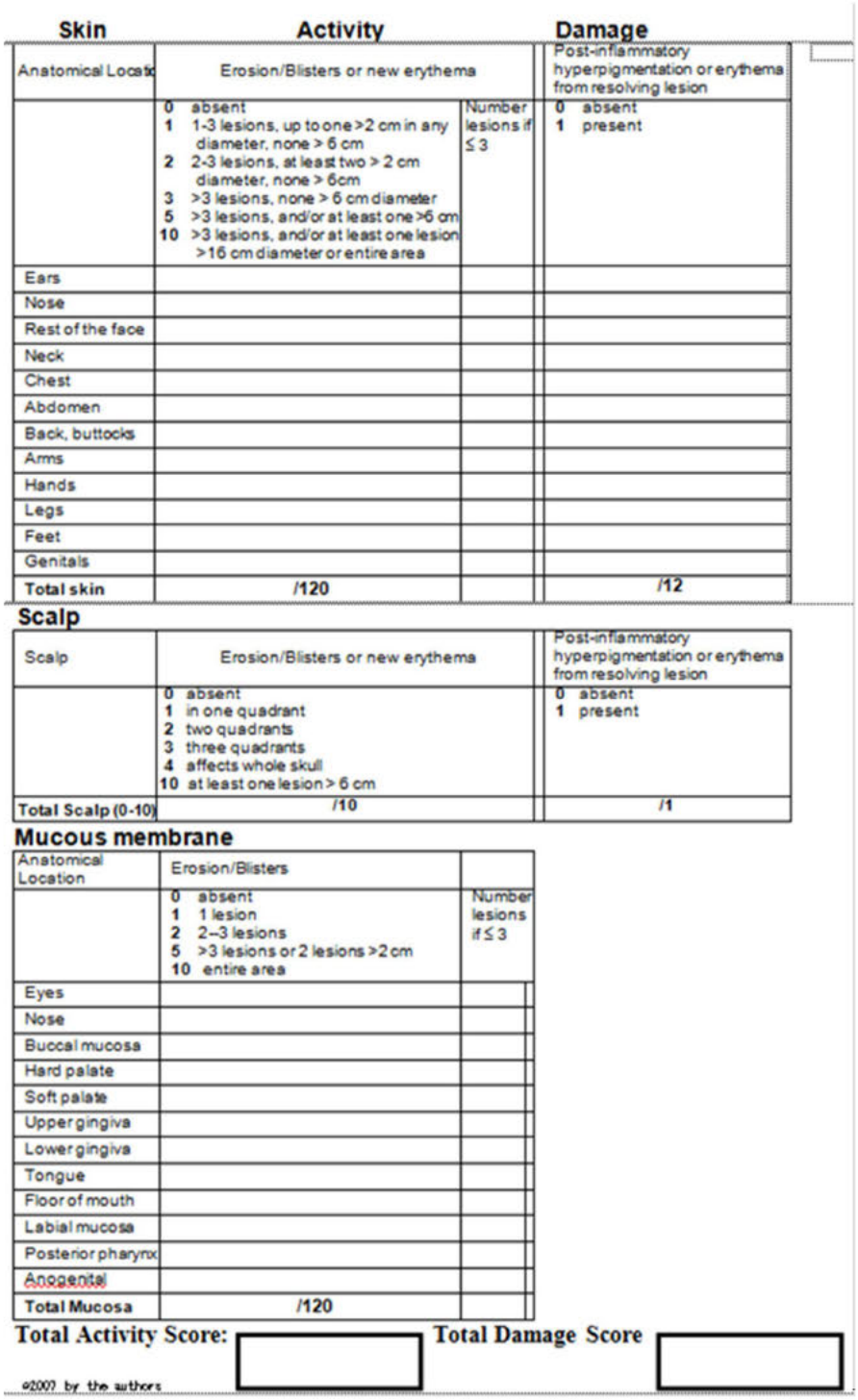
Mucous membrane

Anatomical Location	Erosion/Blisters	
	0 absent 1 1 lesion 2 2--3 lesions 5 >3 lesions or 2 lesions >2 cm 10 entire area	Number lesions if ≤ 3
Eyes		
Nose		
Buccal mucosa		
Hard palate		
Soft palate		
Upper gingiva		
Lower gingiva		
Tongue		
Floor of mouth		
Labial mucosa		
Posterior pharynx		
Anogenital		
Total Mucosa	/120	

Total Activity Score: **Total Damage Score**

©2007 by the authors

Appendix 5 Pemphigus Disease Area Index (cont.)



Skin — Activity — Damage

Anatomical Locatio	Erosion/Blisters or new erythema		Post-inflammatory hyperpigmentation or erythema from resolving lesion
	0 absent 1 1-3 lesions, up to one >2 cm in any diameter, none > 6 cm 2 2-3 lesions, at least two > 2 cm diameter, none > 6cm 3 >3 lesions, none > 6 cm diameter 5 >3 lesions, and/or at least one >6 cm 10 >3 lesions, and/or at least one lesion >16 cm diameter or entire area	Number lesions if ≤ 3	0 absent 1 present
Ears			
Nose			
Rest of the face			
Neck			
Chest			
Abdomen			
Back, buttocks			
Arms			
Hands			
Legs			
Feet			
Genitals			
Total skin	/120		/12

Scalp

Scalp	Erosion/Blisters or new erythema	Post-inflammatory hyperpigmentation or erythema from resolving lesion
	0 absent 1 in one quadrant 2 two quadrants 3 three quadrants 4 affects whole skull 10 at least one lesion > 6 cm	0 absent 1 present
Total Scalp (0-10)	/10	/1

Mucous membrane

Anatomical Location	Erosion/Blisters	
	0 absent 1 1 lesion 2 2--3 lesions 5 >3 lesions or 2 lesions >2 cm 10 entire area	Number lesions if ≤ 3
Eyes		
Nose		
Buccal mucosa		
Hard palate		
Soft palate		
Upper gingiva		
Lower gingiva		
Tongue		
Floor of mouth		
Labial mucosa		
Posterior pharynx		
Anogenital		
Total Mucosa	/120	

Total Activity Score: [ ] **Total Damage Score** [ ]

©2007 by the authors

Appendix 5 Pemphigus Disease Area Index (cont.)

The Pemphigus Disease Area Index (PDAI) lesion counts (number of lesions ≤ 3) for the Skin and Mucous membrane are assigned a score used in calculating the Total Skin and Total Mucosa Activity scores, respectively, when the anatomical location activity score is 1 or 2.

Skin

When an anatomical location is scored a 1 ($1-3$ lesions, up to one >2 cm in any diameter, none >6 cm), record the number of lesions at that site, ranging from $1-3$ lesions. Incorporate the lesion count into the scoring by giving each location a score of 1 if 1 lesion is present, a score of 1.3 if 2 lesions are present, and a score of 1.6 if 3 lesions are present.

When an anatomical location is scored a 2 ($2-3$ lesions, at least two >2 cm diameter, none >6 cm), record the number of lesions at that site, ranging from $2-3$ lesions. Incorporate the lesion count into the scoring by giving each location a score of 2 if 2 lesions are present and a score of 2.3 if 3 lesions are present.

Skin anatomical location activity score	**Number of lesions if ≤ 3**	**Final score**
1 ($1-3$ lesions, up to one >2 cm in any diameter, none >6 cm)	1	1
	2	1.3
	3	1.6
2 ($2-3$ lesions, at least two >2 cm diameter, none >6 cm)	2	2
	3	2.3

Mucous Membrane

When an anatomical location is scored a 1 (1 lesion), record the number one for number of lesions at that site. Incorporate the lesion count into the scoring by giving each location a score of 1 when 1 lesion is present.

When an anatomical location is scored a 2 ($2-3$ lesions), record the number of lesions at that site, ranging from $2-3$ lesions. Incorporate the lesion count into the scoring by giving each location a score of 2 if 2 lesions are present and a score of 2.3 if 3 lesions are present.

Appendix 5 Pemphigus Disease Area Index (cont.)

Mucous membrane anatomical location activity score	Number of lesions if ≤3	Final score
1 (1 lesion)	1	1
2 (2 – 3 lesions)	2	2
	3	2.3

Source: Rosenbach M, Murrell DF, Bystryn JC, et al. Reliability and convergent validity of two outcome instruments for pemphigus. J Invest Dermatol 2009;129:2404– 10; Victoria Werth, personal communication.

Appendix 6 Dermatology Quality of Life Index (DLQI) Questionnaire

DERMATOLOGY LIFE QUALITY INDEX

Hospital No: Date: **DLQI**
Name: Score:
Address: Diagnosis:

The aim of this questionnaire is to measure how much your skin problem has affected your life OVER THE LAST WEEK. Please tick one box for each question.

1.	Over the last week, how **itchy**, **sore**, **painful** or **stinging** has your skin been?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	
2.	Over the last week, how **embarrassed** or **self conscious** have you been because of your skin?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	
3.	Over the last week, how much has your skin interfered with you going **shopping** or looking after your **home** or **garden**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
4.	Over the last week, how much has your skin influenced the **clothes** you wear?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
5.	Over the last week, how much has your skin affected any **social** or **leisure** activities?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
6.	Over the last week, how much has your skin made it difficult for you to do any **sport**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
7.	Over the last week, has your skin prevented you from **working** or **studying**?	Yes No	❒ ❒	Not relevant ❒
	If "No", over the last week how much has your skin been a problem at **work** or **studying**?	A lot A little Not at all	❒ ❒ ❒	
8.	Over the last week, how much has your skin created problems with your **partner** or any of your **close friends** or **relatives**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
9.	Over the last week, how much has your skin caused any **sexual difficulties**?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒
10.	Over the last week, how much of a problem has the **treatment** for your skin been, for example by making your home messy, or by taking up time?	Very much A lot A little Not at all	❒ ❒ ❒ ❒	Not relevant ❒

Please check you have answered EVERY question. Thank you.

©AY Finlay, GK Khan, April 1992 www.dermatology.org.uk, this must not be copied without the permission of the authors.

Appendix 7 Patient Global Impression of Change

How would you rate the change of your pemphigus vulgaris symptoms since you started taking the study drug?

☐1 Very much improved
☐2 Much improved
☐3 Minimally improved
☐4 No change
☐5 Minimally worse
☐6 Much worse
☐7 Very much worse

Appendix 8 Clinician Global Impression of Change

How would you rate the change in the patient's pemphigus vulgaris symptoms since the patient started taking the study drug?

☐1 Very much improved
☐2 Much improved
☐3 Minimally improved
☐4 No change
☐5 Minimally worse
☐6 Much worse
☐7 Very much worse

Appendix 9 European Quality of Life 5-Dimension Questionnaire, 3-Level Version

Health Questionnaire

English version for the US

US (English) © 1998 EuroQol Group. EQ-5D™ is a trade mark of the EuroQol Group

Appendix 9 European Quality of Life 5-Dimension Questionnaire, 3-Level Version (cont.)

By placing a checkmark in one box in each group below, please indicate which statements best describe your own health state today.

Mobility

I have no problems in walking about	❑
I have some problems in walking about	❑
I am confined to bed	❑

Self-Care

I have no problems with self-care	❑
I have some problems washing or dressing myself	❑
I am unable to wash or dress myself	❑

Usual Activities (*e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities	❑
I have some problems with performing my usual activities	❑
I am unable to perform my usual activities	❑

Pain/Discomfort

I have no pain or discomfort	❑
I have moderate pain or discomfort	❑
I have extreme pain or discomfort	❑

Anxiety/Depression

I am not anxious or depressed	❑
I am moderately anxious or depressed	❑
I am extremely anxious or depressed	❑

2

US (English) © 1998 EuroQol Group. EQ-5D™ is a trade mark of the EuroQol Group

Appendix 9 European Quality of Life 5-Dimension Questionnaire, 3-Level Version (cont.)

By placing a checkmark in one box in each group below, please indicate which statements best describe your own health state today.

Mobility

I have no problems in walking about	❑
I have some problems in walking about	❑
I am confined to bed	❑

Self-Care

I have no problems with self-care	❑
I have some problems washing or dressing myself	❑
I am unable to wash or dress myself	❑

Usual Activities (*e.g. work, study, housework, family or leisure activities)*

I have no problems with performing my usual activities	❑
I have some problems with performing my usual activities	❑
I am unable to perform my usual activities	❑

Pain/Discomfort

I have no pain or discomfort	❑
I have moderate pain or discomfort	❑
I have extreme pain or discomfort	❑

Anxiety/Depression

I am not anxious or depressed	❑
I am moderately anxious or depressed	❑
I am extremely anxious or depressed	❑

2

US (English) © 1998 EuroQol Group. EQ-5D™ is a trade mark of the EuroQol Group

Appendix 9 European Quality of Life 5-Dimension Questionnaire, 3-Level Version (cont.)

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

Your own health state today

3

US (English) © 1998 EuroQol Group. EQ-5D™ is a trade mark of the EuroQol Group

Skindex

Skindex-29

Skindex-16

Scaling and Scoring

Version 2.0: February 2012

Written by:
MAPI Research Trust
27 rue de la villette
69003 Lyon
France
Phone: +33 (0) 4 72 13 65 75
Fax: +33 (0) 4 72 13 66 82
E-mail: contact@mapi-trust.org

Author's address:
Mary-Margaret Chren, MD
Professor
University of California at San Francisco
2340 Sutter Street, N412
Mail Code: Box 0808
San Francisco, CA 94143-0808
USA
Fax: +1 (415) 476-6751
Email: chrenm@derm.ucsf.edu

Copyright © 2012 MAPI Research Trust-All rights reserved
Not to be reproduced in whole or in part without written permission of MAPI Research Trust

Appendix 10 Skindex-29 Manual (cont.)

MAPI RESEARCH TRUST

The Skindex-29 and Skindex-16 are composed of respectively 29 and 16 items investigating three domains.

Domains and Clusters

Skindex-29:

Domains	Number of Items	Cluster of Items	Item reversion	Direction of Domains
Emotions	10	3; 6; 9; 12; 13; 15; 21; 23; 26; 28	No	Higher score = higher impact of skin disease
Symptoms	7	1; 7; 10; 16; 19; 24; 27		
Functioning	12	2; 4; 5; 8; 11; 14; 17; 20; 22; 25; 29; 30		

Skindex-16:

Domains	Number of Items	Cluster of Items	Item reversion	Direction of Domains
Emotions	7	5-11	No	Higher score = higher impact of skin disease
Symptoms	4	1-4		
Functioning	5	12-16		

Scoring of Domains

Skindex-29:

Item scaling	5-point Likert-type scale
Weighting of items	No
Range of scores	From 0 (no effect) to 100 (effect experienced all the time)
Scoring Procedure	All responses are transformed to a linear scale of 100. Transformed item scores: never = 0 rarely = 25 sometimes = 50 often = 75 all the time = 100 Scale scores: A scale score is the mean of a patient's responses to the items in a given scale. Note: Item 18 is a single item, not included in scoring.

Copyright © 2012 Mapi Research Trust-All rights reserved

Appendix 10 Skindex-29 Manual (cont.)

Interpretation and Analysis of missing data	• If responses to more than 25% of items are missing overall, the questionnaire is eliminated. • If any scale has more than 25% of the responses missing, the scale is missing. • Scale scores are the average of non-missing items in a given scale.
Interpretation of multiple answers for one item	An item with multiple answers is considered missing.

Skindex-16:

Item scaling	7-point Likert-type scale
Weighting of items	No
Range of scores	From 0 (no effect) to 100 (effect experienced all the time)
Scoring Procedure	All responses are transformed to a linear scale of 100. Item scores: If the raw score is 1-7, change to 0-6 scale by subtracting 1 from the raw Skindex score. Then multiply each raw score by 16.6667. Scale scores: A scale score is the mean of a patient's responses to the items in a given scale.
Interpretation and Analysis of missing data	• If any scale has more than 25% of the responses missing, the scale is missing. • Scale scores are the average of non-missing items in a given scale.
Interpretation of multiple answers for one item	An item with multiple answers is considered missing.

Frequently Asked Questions

Will the questionnaire be excluded if responses to items are the same for an entire page (like the Skindex-61)?

No, for Skindex-29 and Skindex-16, a subject could legitimately respond the same to all items on a page.

We couldn't find specific information on the sum score?

The psychometric validity of averaging or summing all the items or the scales ("composite score") has not been established. We have used it largely to simplify the presentation of results, and we generally try to report in text that consistent results were seen in analyses with subscale scores. When we have used a composite or sum score, we have calculated it as the average of the three scale scores.

Copyright © 2012 Mapi Research Trust-All rights reserved

Appendix 10 Skindex-29 Manual (cont.)

For Skindex-29, is item #18 used for scoring?

No, this item is not scored as part of the instrument. We included it because we were also interested in side-effects of treatment, which is a different topic from quality-of-life effects of the skin condition.

REFERENCE(S):

Chren MM, Lasek RJ, Flocke SA, Zyzanski SJ. Improved discriminative and evaluative capability of a refined version of Skindex, a quality-of-life instrument for patients with skin diseases. Arch Dermatol. 1997 Nov;133(11):1433-40

Chren MM, Lasek RJ, Quinn LM, Covinsky KE. Convergent and discriminant validity of a generic and a disease-specific instrument to measure quality of life in patients with skin disease. J Invest Dermatol. 1997 Jan;108(1):103-7

Chren MM, Lasek RJ, Sahay AP, Sands LP. Measurement properties of Skindex-16: a brief quality-of-life measure for patients with skin diseases. J Cutan Med Surg. 2001 Mar-Apr;5(2):105-10. Epub 2001 Mar 21

Copyright © 2012 Mapi Research Trust-All rights reserved

Skindex29
©MMChren,1996

DERMATOLOGY SURVEY

Review Copy
Do not use without permission

This survey concerns the skin condition which has bothered you the most during the past four weeks.

All answers are confidential. Your name will NOT appear on your survey, be put in your chart or be seen by your doctor.

This research is part of a study being conducted by the *San Francisco Veterans Affairs Medical Center*. Thank you for your participation.

Appendix 10 Skindex-29 Manual (cont.)

Skindex29
©MMChren,1996

These questions concern your feelings over the past 4 weeks about the skin condition that has bothered you the most. Check the answer that comes closest to the way you have been feeling.

HOW OFTEN DURING THE PAST FOUR WEEKS DO THESE STATEMENTS DESCRIBE YOU?	NEVER	RARELY	SOMETIMES	OFTEN	ALL THE TIME
1. My skin hurts	☐1	☐2	☐3	☐4	☐5
2. My skin condition affects how well I sleep	☐1	☐2	☐3	☐4	☐5
3. I worry that my skin condition may be serious	☐1	☐2	☐3	☐4	☐5
4. My skin condition makes it hard to work or do hobbies	☐1	☐2	☐3	☐4	☐5
5. My skin condition affects my social life	☐1	☐2	☐3	☐4	☐5
6. My skin condition makes me feel depressed	☐1	☐2	☐3	☐4	☐5
7. My skin condition burns or stings	☐1	☐2	☐3	☐4	☐5
8. I tend to stay at home because of my skin condition	☐1	☐2	☐3	☐4	☐5
9. I worry about getting scars from my skin condition	☐1	☐2	☐3	☐4	☐5
10. My skin itches	☐1	☐2	☐3	☐4	☐5
11. My skin condition affects how close I can be with those I love	☐1	☐2	☐3	☐4	☐5
12. I am ashamed of my skin condition	☐1	☐2	☐3	☐4	☐5
13. I worry that my skin condition may get worse	☐1	☐2	☐3	☐4	☐5
14. I tend to do things by myself because of my skin condition	☐1	☐2	☐3	☐4	☐5
15. I am angry about my skin condition	☐1	☐2	☐3	☐4	☐5
16. Water bothers my skin condition (bathing, washing hands)	☐1	☐2	☐3	☐4	☐5
17. My skin condition makes showing affection difficult	☐1	☐2	☐3	☐4	☐5
18. I worry about side-effects from skin medications / treatments	☐1	☐2	☐3	☐4	☐5
19. My skin is irritated	☐1	☐2	☐3	☐4	☐5
20. My skin condition affects my interactions with others	☐1	☐2	☐3	☐4	☐5

Please turn to next page

These questions concern your feelings over the past 4 week about the skin condition that has bothered you the most. Check the answer that comes closest to the way you have been feeling.

Review Copy
Do not use without permission

Appendix 10 Skindex-29 Manual (cont.)

Skindex29
©MMChren,1996

HOW OFTEN DURING THE PAST 4 WEEK DO THESE STATEMENTS DESCRIBE YOU?	NEVER	RARELY	SOMETIMES	OFTEN	ALL THE TIME
21. I am embarrassed by my skin condition	☐1	☐2	☐3	☐4	☐5
22. My skin condition is a problem for the people I love	☐1	☐2	☐3	☐4	☐5
23. I am frustrated by my skin condition	☐1	☐2	☐3	☐4	☐5
24. My skin is sensitive	☐1	☐2	☐3	☐4	☐5
25. My skin condition affects my desire to be with people	☐1	☐2	☐3	☐4	☐5
26. I am humiliated by my skin condition	☐1	☐2	☐3	☐4	☐5
27. My skin condition bleeds	☐1	☐2	☐3	☐4	☐5
28. I am annoyed by my skin condition	☐1	☐2	☐3	☐4	☐5
29. My skin condition interferes with my sex life	☐1	☐2	☐3	☐4	☐5
30. My skin condition makes me tired	☐1	☐2	☐3	☐4	☐5

Review Copy
Do not use without permission

Appendix 10 Skindex-29 Manual (cont.)

Skindex29
©MMChren,1996

SKINDEX 29 SCORING (note that item #18 is not used in scoring)

SCALE	ITEMS
Emotion	3-6-9-12-13-15-21-23-26-28
Symptoms	1-7-10-16-19-24-27
Functioning	2-4-5-8-11-14-17-20-22-25-29-30

Review Copy
Do not use without permission

DERMATOLOGY SURVEY

This survey concerns the skin condition which has bothered you the most during the past week.

REVIEW COPY
DO NOT USE WITHOUT PERMISSION

All answers are confidential. Your name will NOT appear on your survey, be put in your chart or be seen by your doctor.

This research is part of a study being conducted by *University of California at San Francisco Department of Dermatology*. Thank you for your participation.

Skindex-16_English/USA_Original version
Skindex16_AU1.1_eng-USori

Skindex16
©MMChren,1997

THESE QUESTIONS CONCERN THE SKIN CONDITION WHICH HAS BOTHERED YOU THE MOST DURING THE PAST WEEK

During the past week, how often have you been bothered by:	Never Bothered ↓			•			Always Bothered ↓
1. Your skin condition itching	☐0	☐1	☐2	☐3	☐4	☐5	☐6
2. Your skin condition burning or stinging	☐0	☐1	☐2	☐3	☐4	☐5	☐6
3. Your skin condition hurting	☐0	☐1	☐2	☐3	☐4	☐5	☐6
4. Your skin condition being irritated	☐0	☐1	☐2	☐3	☐4	☐5	☐6
5. The persistence / reoccurrence of your skin condition	☐0	☐1	☐2	☐3	☐4	☐5	☐6
6. Worry about your skin condition (For example: that it will spread, get worse, scar, be unpredictable, etc)	☐0	☐1	☐2	☐3	☐4	☐5	☐6
7. The appearance of your skin condition	☐0	☐1	☐2	☐3	☐4	☐5	☐6
8. Frustration about your skin condition	☐0	☐1	☐2	☐3	☐4	☐5	☐6
9. Embarrassment about your skin condition	☐0	☐1	☐2	☐3	☐4	☐5	☐6
10. Being annoyed about your skin condition	☐0	☐1	☐2	☐3	☐4	☐5	☐6
11. Feeling depressed about your skin condition	☐0	☐1	☐2	☐3	☐4	☐5	☐6
12. The effects of your skin condition on your interactions with others (For example: interactions with family, friends, close relationships, etc)	☐0	☐1	☐2	☐3	☐4	☐5	☐6
13. The effects of your skin condition on your desire to be with people	☐0	☐1	☐2	☐3	☐4	☐5	☐6
14. Your skin condition making it hard to show affection	☐0	☐1	☐2	☐3	☐4	☐5	☐6
15. The effects of your skin condition on your daily activities.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
16. Your skin condition making it hard to work or do what you enjoy	☐0	☐1	☐2	☐3	☐4	☐5	☐6

Have you answered every item? Yes ☐ No ☐

Skindex-16_English/USA_Original version
Skindex16_AU1.1_eng-USorl

REVIEW COPY
DO NOT USE WITHOUT PERMISSION

Skindex16
©MMChren,1997

SKINDEX 16 SCORING

SCALE	ITEMS
Symptoms	1-2-3-4
Emotion	5-6-7-8-9-10-11
Functioning	12-13-14-15-16

Item scores transformed to 0 – 100 scale.
Scale Score: Average of items in given scale

REVIEW COPY
DO NOT USE WITHOUT PERMISSION

Skindex-16_English/USA_Original version
Skindex16_AUT.1_orig-USen

Appendix 11 Imputation Rules for Partial or Missing Dates

	Start Date/Time	End Date/Time
Concomitant Medications	Partially Missing **Day**: First day of the month (i.e., 01) **Month**: January **Year**: No imputation **Seconds**: 00 **Minutes**: 00 **Hours**: 00 Completely Missing **Date**: No imputation **Time**: 00:00:00 **Note**: if end time is available and start/end dates are equal and the imputed start time is after end time then impute start time with end time	Partially Missing **Day**: Last day of the month **Month**: December **Year**: No imputation **Seconds**: 59 **Minutes**: 59 **Hours**: 23 Completely Missing **Date**: No imputation **Time**: 23:59:59
Medical History	Partially Missing **Day**: First day of the month (i.e., 01) **Month**: January **Year**: No imputation **Seconds**: 00 **Minutes**: 00 **Hours**: 00 Completely Missing **Date**: No imputation **Time**: 00:00:00 Note: if end ime is available and start/end dates are equal and the imputed start time is after end time then impute start time with end time	Partially Missing **Day**: Last day of the month **Month**: December **Year**: No imputation **Seconds**: 00 **Minutes**: 00 **Hours**: 00 Completely Missing **Date**: No imputation **Time**: 00:00:00

Appendix 11 Imputation Rules for Partial or Missing Dates (cont.)

	Start Date/Time	End Date/Time
Laboratory	Partially Missing **Day**: No imputation **Month**: No imputation **Year**: No imputation **Seconds**: 23 **Minutes**: 59 **Hours**: 59 Completely Missing **Date**: No imputation **Time**: 23:59:59	Partially Missing **Day**: No imputation **Month**: No imputation **Year**: No imputation **Seconds**: 23 **Minutes**: 59 **Hours**: 59 Completely Missing **Date**: No imputation **Time**: 23:59:59
Adverse Events	Partially Missing **Day**: If the MM-YYYY of start date is not equal to the MM-YYYY of first treatment date then set to 01-MM-YYYY. If the MM-YYYY of start date is equal to the MM-YYYY of first treatment date then if the end date is ≥ date of first treatment set to max(01-MM-YYYY, date of first treatment) , otherwise if end date < date of first treatment set to 01-MM-YYYY. **Month**: Set to max(01-01-YYYY, date of first treatment). **Year**: Set to start date of first treatment. **Seconds**: 59 **Minutes**: 59 **Hours**: 23 Completely Missing Set to start date of first treatment. Note: For all imputations above, if the resulting imputed start date is after the end date then set start date to end date.	Partially Missing **Day**: Set to last day of the month, unless the outcome is death in which case impute with min(last day of month, date of death) **Month**: December, unless he outcome is death in which case impute with min(DD-12-YYYY, date of death) **Year**: If the outcome is death then impute with the date of death, otherwise no imputation. **Seconds**: 59 **Minutes**: 59 **Hours**: 23 Completely Missing If the outcome is death then impute with the date of death, otherwise no imputation.

STATISTICAL ANALYSIS PLAN
48-WEEK SAFETY FOLLOW-UP

TITLE:	**A RANDOMIZED, DOUBLE-BLIND, DOUBLE DUMMY, ACTIVE-COMPARATOR MULTICENTER STUDY TO EVALUATE THE EFFICACY AND SAFETY OF RITUXIMAB VERSUS MYCOPHENOLATE MOFETIL IN PATIENTS WITH PEMPHIGUS VULGARIS**
PROTOCOL NUMBER:	WA29330
STUDY DRUG:	Rituximab
VERSION NUMBER:	1
IND NUMBER:	121595
EUDRACT NUMBER:	2014-000382-41
SPONSOR:	F. Hoffmann-La Roche Ltd.
PLAN PREPARED BY:	[redacted], Ph.D.
DATE FINAL:	See electronic date stamp below.

STATISTICAL ANALYSIS PLAN APPROVAL

Date and Time(UTC)	Reason for Signing	Name
18-Oct-2019 13:14:46	Company Signatory	[redacted]

CONFIDENTIAL

This is an F. Hoffmann-La Roche Ltd document that contains confidential information. Nothing herein is to be disclosed without written consent from F. Hoffmann-La Roche Ltd.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

1. BACKGROUND

Study WA29330 (PEMPHIX) aims to evaluate the efficacy and safety of rituximab (RTX) versus mycophenolate mofetil (MMF) in patients with pemphigus vulgaris (PV). PV is a rare, serious autoimmune disease characterized by blisters and erosions of the skin and mucosal areas.

Study WA29330 is an international, multicentre, randomized, double-blind, double-dummy, active-comparator, two-arm, parallel-group study, which consists of three periods: a screening period of up to 28 days; a 52-week, double-blinded, treatment period; and a 48-week safety follow up (SFU) period that begins at the time of study treatment completion or discontinuation. All patients who withdraw early from the 52-week treatment period or who complete the 52-week treatment period, must return for post-study treatment SFU assessments at 12, 24, 36, and 48 weeks after either the early withdrawal visit or study treatment completion, respectively.

A separate primary Statistical Analysis Plan (SAP) which describes the statistical methodology and data handling rules used for the 52-week treatment period was finalized before primary database freeze, the analyses completed, and Clinical Study Report (CSR) finalized. This SAP describes the timing and data handling rules to be used for the 48-week SFU period prior to final database lock.

2. ANALYSIS TIMING

The SFU analysis will be conducted when all patients have completed the 48-week SFU period or withdraw early from SFU (Appendix 1). The final CSR will include all available data up to study completion (last patient's last visit [LPLV] in SFU) or early withdrawal from SFU.

3. STATISTICAL METHODS

All analyses, summaries, and listings will be performed using SAS® software (Version 9.2 or later).

3.1 ANALYSIS POPULATIONS

The primary analysis populations are defined in the primary SAP, Section 4.1.

The rescue population (RP) includes patients who received at least one rescue therapy during the study (i.e., between Week 12 to Week 52 of the study treatment period).

The analysis populations in SFU will be the subset of patients in the primary analysis populations who entered the SFU period.

3.2 DATA-CUT FOR ANALYSES

SFU period begins at the time of study treatment completion (for patients who complete the 52-week treatment period) or discontinuation.

For each patient, start of SFU period will be earliest of the following:

- Cut-off for Week 52 + 1 day (i.e., Day 375 + 1 day)
- Date of early treatment discontinuation + 1 day

Each subsequent assessment point will be assigned a study day calculated as:

date of assessment − date of first dose + 1.

For patients who had their Week 52 visit after the Week 52 cut-off date (Day 365 + 10 days), the Week 52 data will be included in the analysis of the SFU period.

3.2.1 Rescue Patients

Rescue patients are patients who received rescue therapy during the study (i.e., Week 12 to Week 52). There is no rescue therapy in the SFU.

The analysis of safety will be performed separately for rescue patients after the point of rescue therapy initiation.

The analysis of exploratory efficacy will be performed on the patients who haven't received rescue therapy during the 52-week treatment period, and the data for rescue patients will be excluded.

3.3 VISIT WINDOWS

In general, data for assessments that are collected by scheduled visits will be mapped to visits that appear in the schedule of assessments per the protocol using the actual study day of assessment. Data will not be mapped to visits for which the assessment was not scheduled in the protocol (i.e., results for an assessment conducted at an unscheduled visit will be mapped to the appropriate visit at which the assessment was scheduled). Visit windows will be continuous from the midpoint between two consecutive study visits, and will be dependent on the Schedule of Assessments for each variable independently. Since patients start the SFU period at different times, the start day of SFU will be considered in defining the SFU visit windows (See Table 1).

Table 1 Time Windows in Study Days Assigned to the SFU Study Visit

Study Visit	Time Window (Days)
Start of SFU (BL_{SFU})	BL_{SFU}
Week 12	$\geq BL_{SFU}$ to $\leq BL_{SFU}+126$
Week 24	$>BL_{SFU}+126$ to $\leq BL_{SFU}+210$
Week 36	$>BL_{SFU}+210$ to $\leq BL_{SFU}+294$
Week 48	$>BL_{SFU}+294$ to $\leq$ LPLV

LPLV = last patient last visit; SFU = safety follow-up.

Data displays that are presented by visit will be visits at which the assessment was scheduled to be collected. The methods for handling multiple values within a visit window are described in the Primary SAP.

3.4 ANALYSIS OF STUDY CONDUCT

The following analyses of study conduct will be carried out on the safety, intent-to-treat, modified intent-to-treat population (mITT) and rescue populations, as appropriate (See primary SAP, Section 4.1).

The number of patients who completed or discontinued from the study during the SFU period, including the reason for discontinuation, will be summarized by treatment group. A listing of early withdrawals will be produced.

Concomitant medications given for PV will be summarised descriptively by mapped term and treatment group.

Demographic and baseline disease characteristics will be summarized for patients who entered SFU, as described in the primary SAP, Section 4.1.

Summary statistics for duration of SFU period per treatment group will be produced. The duration of SFU period for a patient will be the duration from BL_{SFU} to the last available SFU assessment for that patient.

3.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

Serum rituximab concentrations will be listed by patient and nominal time of collection. Summary statistics of serum rituximab concentrations will be presented by nominal time of collection including n, mean, SD, coefficient of variation, geometric mean, median, and range.

3.6 SAFETY ANALYSES

The safety analysis will be performed on the safety population as defined in the primary SAP, Section 4.1.3, for patients who entered SFU.

Rescue patients entered SFU will be analyzed separately.

Serious adverse events (SAEs) and all infectious adverse events, irrespective of causality, will be analyzed in the SFU period. Pregnancy information will be reported for the duration of the SFU period.

3.6.1 Adverse Events

Adverse events (AEs) will be mapped and reported using the MedDRA thesaurus using the most current version at the time of analysis (version 17.1 or later). Glossaries showing the mapping of investigator verbatim terms to coded events will be produced.

AEs with observed or imputed onset date, on or after the start date of the SFU period, will be summarized and listed.

AEs will be summarized by system organ class (SOC) and preferred term (PT), and will be presented in order of descending frequency summed across the treatment arms within each SOC and PT.

Listings will show AEs reported, and highlight the telemedicine patients, the patients who received rescue therapy in the 52-week treatment period and the patients who received rituximab in the SFU period.

The following will be summarized, and/or listings produced where required:

- SAEs
- Infections
- Serious Infections
- Symptoms of Serious Infusion Related Reactions (IRRs)
- SAEs leading to withdrawal

3.6.2 Deaths

Details of any deaths will be presented in an individual patient listing.

3.6.3 Laboratory Data

3.6.3.1 Immunoglobulins

The absolute values will be summarized by treatment group and visit for total immunoglobulin (Ig), IgA, IgG, and IgM.

The proportion of patients with a value of IgA, IgG, and IgM which is less than the lower limit of normal at each visit will be produced. Plot of mean and median IgA, IgG and IgM will be produced.

3.6.3.2 Lymphocyte Populations

The following natural killer (NK) cells, T, and B lymphocytes will be analyzed via flow cytometry:

- CD19

- CD3
- CD4
- CD8
- Naïve B cells ($CD19+\ CD27-$)
- Transitional B cells ($CD24_{high}CD38_{high}$)
- Memory B cells ($CD19+CD27+$)
- NK cells ($CD56+$)

The absolute values will be summarized by visit and treatment group for the above lymphocyte populations.

The proportion of patients with a lymphocyte value less than the lower limit of normal by visit and treatment group will be produced for the following lymphocytes:

- CD3
- CD4
- CD8
- CD19

3.6.3.3 Desmoglein-1 and Desmoglein-3 Autoantibody Titers

Summary statistics for anti-desmoglein−1 (Dsg-1) and anti-Dsg-3 titers during the SFU period will be produced by treatment group. A listing of anti-Dsg-1 and anti-Dsg-3 titers will be produced.

3.6.3.4 Anti-Drug Antibodies

The proportion of patients that are ADA (anti-drug antibodies [referred to as HACA in the protocol]) positive/negative at each visit will be summarized by treatment group. A listing of ADA results will be produced.

3.7 EXPLORATORY ANALYSES

The exploratory efficacy analysis will be performed on the mITT population as defined in primary SAP, Section 4.1.5.

3.7.1.1 Oral Corticosteroids

Oral corticosteroid exposure over the SFU period will be summarized by treatment group.

Oral corticosteroid use over time will be further investigated. Summaries will be generated which will include:

- Total cumulative dose per visit interval
- Total number of days on oral corticosteroids over the SFU period

The mean daily dose for a given visit interval will be calculated as follows:

sum of total daily dose across all days in the visit interval ÷ (end date of visit interval − start date of visit interval + 1).

All available data presented by specified visit intervals will be used; if a patient withdraws during a time interval, no imputation will be made for the remaining days in the interval, and available data up to the time of withdrawal will be used in the calculation.

A listing showing oral corticosteroid dosing for patients over the SFU period will be produced.

See primary SAP, Section 4.7.2.1 for Data Handling for Oral Corticosteroids

3.7.1.2 Pemphigus Disease Area Index

The absolute value of Pemphigus Disease Area Index (PDAI) activity score will be summarized by visit for each treatment arm.

The main method for handling missing PDAI data will be via imputation. Details of PDAI imputation methods are described in the primary SAP, Section 4.7.1.2.

3.7.1.3 Total Number of Flares

The total number of flares occurring during the SFU period will be summarized and analyzed as described in the primary SAP, Section 4.7.2.2.

3.7.1.4 Dermatology Life Quality Index

The Dermatology Life Quality Index (DLQI) is a validated patient-reported outcome. (See primary SAP, Section 4.7.2.5 for more details)

The absolute value (DLQI score) at each visit will be summarized descriptively using observed data.

3.7.1.5 Skindex-29

The Skindex-29 is a health-related quality of life (HRQoL) measure evaluating the impact of patient's dermatologic condition over the previous 4 weeks (See primary SAP, Section 4.7.3.9 for more details).

The Skindex-29 score will be summarized by visit and treatment group.

Each scale score is the mean of the items in that scale (domain). The overall score is the mean for the 29 items, i.e., excluding item 18.

The overall (global) score will be set to missing if more than 25% of the responses are missing.

3.7.1.6 Patient Global Impression of Change

The Patient Global Impression of Change (PGIC) is a measure of perceived change in PV symptoms as assessed by the patient (See primary SAP, Section 4.7.3.2 for more details).

The PGIC categories will be summarized by visit.

3.7.1.7 Clinician Global Impression of Change

The CGIC is a measure of perceived change in PV symptoms as assessed by the clinician (See primary SAP, Section 4.7.2.6 for more details).

The CGIC categories will be summarized by visit.

3.8 DATA HANDLING METHODS

Data handling methods are similar to the primary analysis (See primary SAP, Section 4.10).

Appendix 1 Schedule of Assessments for the Safety Follow-Up Period

All patients who withdraw from both study treatments or complete the study will enter the safety follow-up period. Please note that for patients whose peripheral CD19+ B cells remain depleted (below the lower limit of normal or pre-rituximab baseline level, whichever is lower), safety follow-up will not be extended beyond the standard protocol-defined 48 weeks after treatment withdrawal or completion. CD19+ B-cell counts and other blinded laboratory values will be provided to the investigator at the end of the safety follow-up period.

Assessment	Week			
	12 (±1 week)	24 (±1 week)	36 (±1 week)	48 (±1 week)
Safety Assessments to Be Performed for All Patients				
Adverse events [a]	x	x	x	x
Concomitant medications [b]	x	x	x	x
HACA				x
Autoantibody testing [c]	x			x
Immunology and Other Labs				
Lymphocyte subtypes and FACS panel [d]	x	x	x	x
Quantitative Ig [e]	x	x	x	x
PK sample [f]	x	x	x	x
Disease Status and Activity Limitation Questions				
Disease activity (PDAI) and flare assessment	x	x	x	x
DLQI [g]	x			x
PGIC [g]	x			x
CGIC [g]	x			x
Skindex-29 [g]	x			x
Prednisone (or equivalent) dose assessment/adjustment [h]	x	x	x	x

CGIC = Clinician Global Impression of Change; DLQI = Dermatology Life Quality Index; Dsg1 = desmoglein 1; Dsg3 = desmoglein 3; eCRF = electronic Case Report Form; FACS = fluorescence-activated cell sorting; HACA = human anti–chimeric antibodies; Ig = immunoglobulin; PDAI = Pemphigus Disease Area Index; PGIC = Patient Global Impression of Change; PK = pharmacokinetic; PV = pemphigus vulgaris; SFU = safety follow-up.

[a] Only serious adverse events and all infectious adverse events, irrespective of causality, will be reported in the SFU period. For all serious infectious adverse events reported, quantitative Ig, and $CD19^+$ cells counts should be determined within 1 week of the adverse event becoming serious. Pregnancy information will be reported for the duration of the SFU period.

[b] During the SFU period, reporting of concomitant medications should be limited to subsequent PV treatments, use of biologics, immunosuppressive medications, and cell-depleting agents.

[c] Anti-Dsg1 and anti-Dsg3 antibodies.

[d] See protocol Section 4.5.6 for T- and B-cell assessments.

[e] Total Ig, IgG, IgA, and IgM.

[f] PK samples will be collected only in patients who withdraw early from study treatment.

[g] The DLQI, PGIC, and Skindex-29 should be completed before all other assessments during the study visit and prior to providing patient with any disease status information. The CGIC, where possible, should be completed prior to the completion of other non-PRO study assessments.

[h] To be captured on the eCRF.

Handling and Reporting of Laboratory Safety Data

This document describes methods used to standardize laboratory safety data for analysis and reporting purposes and presents for laboratory parameters that are frequently assessed in Roche/Genentech studies the range of values and changes from baseline for each laboratory test beyond which a value might be clinically relevant.

Standardization Methods

In clinical trials, samples for a particular laboratory parameter may be analyzed by several laboratories often using different methods. This can result in values being reported in different units with different reference ranges. Thus, to make the values comparable for analysis and reporting purposes, the following standardization process is applied to the data.

Conversion of values to Système International (SI) units

Where applicable, laboratory test values were converted from the investigator reported units to Système International (SI) units.

Transformation of data

For those tests that use common methodologies, a 'Roche Standard Reference Range' (adopted from the New England Journal of Medicine [Kratz and Lewandrowski 1998, Kratz et al. 2004]) was applied since their normal ranges show only slight variation between different laboratories. For other tests where different methodologies are used between laboratories, and thus the reference ranges differ widely, values were linearly transformed to a 'Roche Standard Reference Range' before analysis and reporting using the following formula (Smith and Givens 1993):

(1)

$$R_T = S_L + [(R_U - I_L)/(I_H - I_L)]) * (S_H - S_L)$$

R_T	transformed value
R_U	untransformed value
I_L, I_H	investigator limits, low and high
S_L, S_H	reference limits, low and high

Note: Values for females (where there are sex–dependent reference ranges) were transformed to the male reference range by replacing I_L, I_H by the reference limits for females, and S_L, S_H by the reference limits for males.

This linear transformation did not affect the relative position of a value to its reference range. For example, if an untransformed value was above the investigator's range, the transformed value was also above the Roche Standard Reference Range.

In cases where the transformed values could be negative, as could happen for low values where the lower limit of the reference range is zero or nearly zero, but the lower limit of the investigator range was markedly greater than zero, the investigator lower limit was replaced by zero for the transformation. This was considered to be clinically justifiable for tests where no abnormal pathology is evident at low values.

NOTE: For simplicity in technical implementation, standard reference ranges were expressed in the format “lower limit–upper limit” regardless of whether the actual limits were included in the published range (in some cases the reference ranges were one-sided).

Examples:

- Creatinine – Published reference range was <133 (µmol/L); implemented reference range was 0–133. A value equal to 133 was not flagged as high (outside the reference range).
- HDL Cholesterol -- Published reference range was ≥1.03 and <1.55 (mmol/L); implemented reference range was 1.03–1.55. A value equal to 1.55 was not flagged as high.

Marked Laboratory Test Value Abnormalities

For each laboratory parameter, a Marked Reference Range which was wider than the Roche Standard Reference Range was defined, above and/or below which a value was considered to be potentially clinically relevant. In addition, for each laboratory test, a percentage change, as an increase or decrease from the baseline value, was defined which represented a clinically relevant change from baseline. The ranges used to define a Marked Abnormality and the required change from baseline are shown in Table 1.

A Marked Abnormality was therefore defined as a test result:

- which is outside of the Marked Reference Range,

  AND

- which also represents a clinically relevant change from baseline of at least the designated amount during, or within a specified window after the end of, trial treatment (e.g., 30 days)

If the baseline value was zero and the percentage increase from baseline could not be calculated, the second condition was ignored (conservative approach).

Baseline was defined as the last valid value within a window specified by the project team (e.g., 30 days) before starting trial treatment. If no baseline was available, then the midpoint of the Standard Reference Range was used.

For crossover studies or parallel studies where treatment interruption longer than a specified period is of clinical significance: In general, if the wash-out interval was sufficient between 2 consecutive treatment periods, the user could derive a new baseline for the later of the 2 consecutive periods, as agreed within the project team.

NOTE: For calculating means and mean changes over time, missing values were not included. If no baseline was available, no change from baseline was calculated.

Table 1 Overview of Laboratory Parameters for Safety Reporting

Laboratory Test Class Laboratory Test	Specimen	SI Unit	Significant Decimals	Standard Reference Range [f, g, h]	Marked Reference Range	Direction of Change	Clinically Relevant Change from Baseline
HEMATOLOGY							
Hematocrit **(PARAMCD = HCRIT)**	Blood	fraction of 1	2	M 0.41–0.53 F 0.36–0.46 [a]	0.31–0.56	Increase Decrease	≥15% ≥15%
Hemoglobin **(PARAMCD = HGB)**	Blood	g/L	0	M 135–175 F 120–160 [a]	110–200	Increase Decrease	≥15% ≥15%
Leukocytes (WBC) **(PARAMCD = WBC)**	Blood	10^9/L	1	4.5–11.0	3.0–18.0	Increase Decrease	≥30% ≥30%
Platelets **(PARAMCD = PLATE)**	Blood	10^9/L	0	150–350	100–550	Increase Decrease	≥50% ≥30%
Mean corpuscular hemoglobin (MCH) (PARAMCD = MCH)	Blood	pgl	1	26.0–34.0	21.2–40.2	Increase Decrease	≥15% ≥15%
Mean corpuscular hemoglobin concentration (MCHC) (PARAMCD = MCHC)	Blood	g/L	0	310–370	260–430	Increase Decrease	≥15% ≥15%
Mean corpuscular volume (MCV) (PARAMCD = MCV)	Blood	fL	0	80–100	66–115	Increase Decrease	≥15% ≥15%
Erythrocytes (RBC) (PARAMCD = RBC)	Blood	10^12/L	2	M 4.50–5.90 F 4.00–5.20 [a]	3.80–6.10	Increase Decrease	≥15% ≥15%

Laboratory Test Class Laboratory Test	Specimen	SI Unit	Significant Decimals	Standard Reference Range [f, g, h]	Marked Reference Range	Direction of Change	Clinically Relevant Change from Baseline
DIFFERENTIALS [i]							
Bands (PARAMCD = BANDS)	Blood	10^9/L	2	0–0.70	0–1.40	Increase	≥30%
Bands (PARAMCD = BANDSF)	Blood	fraction of 1	2	0–0.10	0–0.15	Increase	≥30%
Basophils (PARAMCD = BASOS)	Blood	10^9/L	2	0–0.20	0–0.40	Increase	≥100%
Basophils (PARAMCD = BASOSF)	Blood	fraction of 1	2	0–0.03	0–0.06	Increase	≥100%
Lymphocytes (PARAMCD = LYMPH)	Blood	10^9/L	2	1.00–4.80	0.70–7.60	Increase Decrease	≥30% ≥30%
Lymphocytes (PARAMCD = LYMPHF)	Blood	fraction of 1	2	0.22–0.44	0.10–0.72	Increase Decrease	≥30% ≥30%
Monocytes (PARAMCD = MONOS)	Blood	10^9/L	2	0–0.80	0–1.70	Increase	≥100%
Monocytes (PARAMCD = MONOSF)	Blood	fraction of 1	2	0.04–0.11	0–0.22	Increase	≥100%
Neutrophils (PARAMCD = NEUTR)	Blood	10^9/L	2	1.80–7.70	1.50–9.25	Increase Decrease	≥20% ≥20%
Neutrophils PARAMCD = NEUTRF)	Blood	fraction of 1	2	0.40–0.70	0.37–0.90	Increase Decrease	≥20% ≥20%
Eosinophils (PARAMCD = EOSIN)	Blood	10^9/L	2	0–0.45	0–0.90	Increase	≥100%
Eosinophils (PARAMCD = EOSINF)	Blood	fraction of 1	2	0–0.08	0–0.16	Increase	≥100%

Laboratory Test Class Laboratory Test	Specimen	SI Unit	Significant Decimals	Standard Reference Range [f, g, h]	Marked Reference Range	Direction of Change	Clinically Relevant Change from Baseline
COAGULATION							
Prothrombin Time, normalized ratio (PTINR) (PARAMCD = PTINR)	Plasma	fraction of 1	2	0.70–1.30	≤2.00	Increase	≥30%
Partial Thromboplastin Time, activated (PTT) (PARAMCD = PTT)	Plasma	sec	1	22.1–34.1	0–45	Increase	≥40%
Fibrinogen (PARAMCD = FIB)	Plasma	g/L	2	1.50–4.00 [b]	≥1.30	Decrease	≥30%
HEART FUNCTION							
Lactate Dehydrogenase (PARAMCD = LDH)	Blood	U/L	0	0–190 [c]	0–420	Increase	≥50%
Creatine Kinase (CPK) total (PARAMCD = CPK)	Blood	U/L	0	M 60–400 [b] F 40–150 [a]	≤800	Increase	≥50%
CPK (MB Fraction) (PARAMCD = CPKMB)	Blood	µg/L	0	0–7 [c]	0–7	Increase	>0
LIVER FUNCTION							
AST (SGOT) (PARAMCD = AST)	Blood	U/L	0	0–35 [c]	0–80	Increase	≥50%
ALT (SGPT) (PARAMCD = ALT)	Blood	U/L	0	0–35 [c]	0–110	Increase	≥50%
Alkaline Phosphatase (PARAMCD = ALKPH)	Blood	U/L	0	0–120 [c]	0–220	Increase	≥50%
Total Bilirubin (PARAMCD = TBILI)	Blood	µmol/L	1	0–17.0 [c]	0–34.0	Increase	≥75%
Direct bilirubin (PARAMCD = DBILI)	Blood	µmol/L	1	0–5.1 [c]	0–14.0	Increase	≥75%

Laboratory Test Class Laboratory Test	Specimen	SI Unit	Significant Decimals	Standard Reference Range [f, g, h]	Marked Reference Range	Direction of Change	Clinically Relevant Change from Baseline
Gamma Glutamyl Transferase (PARAMCD = GGT)	Blood	U/L	0	0–94[c]	0–190	Increase	≥50%
RENAL FUNCTION							
Blood Urea Nitrogen (BUN) (PARAMCD = BUN)	Blood	mmol/L	1	3.6–7.1	0–14.3	Increase	≥50%
Creatinine (PARAMCD = CREATN)	Blood	µmol/L	0	0–133	0–154	Increase	≥50%
THYROID FUNCTION							
Triiodothyronine (T3) (PARAMCD = T3)	Blood	nmol/L	2	0.92–2.78	0.74–3.30	Increase Decrease	≥20% ≥20%
Thyroxine (T4) (PARAMCD = T4)	Blood	nmol/L	0	58–140	30–164	Increase Decrease	≥20% ≥20%
Free T4 (PARAMCD = T4FREE)	Blood	pmol/L	0	10–36	5–40	Increase Decrease	≥20% ≥20%
Thyroid-Stimulating Hormone (TSH) (PARAMCD = TSH)	Blood	mU/L	1	0–5.0 [c]	0–10.0	Increase	≥30%
PROTEIN							
Albumin (PARAMCD = ALBUM)	Blood	g/L	1	35.0–55.0	≥30	Decrease	≥20%
Total Protein (PARAMCD = TPROT)	Blood	g/L	0	55–80	50–87	Increase Decrease	≥20% ≥20%
LIPID CHEMISTRY							
Triglycerides (PARAMCD = TRIG)	Blood	mmol/L	2	0–1.80	0–2.83	Increase	≥100%
Cholesterol	Blood	mmol/L	2	0–6.18	0–8.30	Increase	≥30%

Laboratory Test Class Laboratory Test	Specimen	SI Unit	Significant Decimals	Standard Reference Range [f, g, h]	Marked Reference Range	Direction of Change	Clinically Relevant Change from Baseline
(PARAMCD = CHOLES)							
LDL-Cholesterol (PARAMCD = LDL)	Blood	mmol/L	2	0–4.13	0–5.4	Increase	≥30%
HDL-Cholesterol (PARAMCD = HDL)	Blood	mmol/L	2	1.03–1.55	≥0.65	Decrease	≥30%
ELECTROLYTES							
Chloride (PARAMCD = CHLOR)	Blood	mmol/L	0	98–106	95–115	Increase Decrease	≥7% ≥7%
Potassium (PARAMCD = POTAS)	Blood	mmol/L	1	3.5–5.0	2.9–5.8	Increase Decrease	≥20% ≥20%
Sodium (PARAMCD =SODIUM)	Blood	mmol/L	0	136–145	130–150	Increase Decrease	≥7% ≥7%
Bicarbonate (PARAMCD = BICARB)	Blood	mmol/L	0	21–28	17–30	Increase Decrease	≥20% ≥20%
MISCELLANEOUS							
Calcium (PARAMCD = CALCUM)	Blood	mmol/L	2	2.20–2.60	2.00–2.90	Increase Decrease	≥10% ≥10%
Phosphorus inorganic (PARAMCD = PHOSAT)	Blood	mmol/L	2	1.00–1.40	0.75–1.60	Increase Decrease	≥30% ≥30%
Blood Glucose (fasting) (PARAMCD = FASTGL)	Blood	mmol/L	2	4.20–6.40	2.80–11.10	Increase Decrease	≥75% ≥75%
Uric Acid (PARAMCD = URACID)	Blood	µmol/L	0	M 150–480 F 90–360 [a]	0–600	Increase	≥50%

Laboratory Test Class Laboratory Test	Specimen	SI Unit	Significant Decimals	Standard Reference Range [f, g, h]	Marked Reference Range	Direction of Change	Clinically Relevant Change from Baseline
URINALYSIS MICROSCOPIC [d]							
Urinary spec. gravity [e] (PARAMCD =USG)	Urine	No unit	3	1.001–1.035	not defined	not defined	not defined

NOTE: Laboratory tests in **bold** are proposed as the minimal standard. All other tests are optional.

[a] For reporting and analysing, the data for females are linearly transformed to the Reference Range for males according to formula (1).

[b] For reporting and analyzing, the data are linearly transformed to the Reference Range according to formula (1).

[c] For reporting and analysing, the data are linearly transformed to the Reference Range according to formula (1). To prevent transformed values becoming negative, the investigator lower limit is replaced by zero.

[d] To be done if dipstick is positive and microscopic examination needed for specific quantitation/confirmation.

[e] Specific gravity is a measurement of solute load in the urine. It is increased markedly by both proteinuria and glycosuria. A value >1.030 in the absence of proteinuria or glycosuria is most commonly caused by radiographic contrast. In volume depletion specific gravity is usually >1.020. A fixed value of 1.010 (isosthenuria) is characteristic of chronic renal impairment. A fixed value of 1.000–1.005 occurs in diabetes insipidus. Specific gravity has an important influence on urine microscopy because cells lyse more rapidly in dilute urine, which may lead to considerable inaccuracies in quantitation of cells. Cells swell and rupture in dilute urine and shrink in concentrate urine. Casts are rarely seen at low urine osmolality (Johnson and Feehally 2000). Conclusion: increase or decrease in urine specific gravity clinically relevant in context with other parameters from urinalysis.

[f] Kratz and Lewandrowski 1998.

[g] New England Journal of Medicine 1992 SI conversion guide.

[h] Kratz et al. 2004.

[i] Harrison's Textbook of Internal Medicine.

References

Harrison's Textbook of Internal Medicine. Absolute Values of Differentials

Johnson RJ, J. Feehally J., Comprehensive Clinical Nephrology, 2000, Mosby

Kratz A, Ferraro M, Sluss PM, Lewandrowski KB. Case records of the Massachusetts General Hospital. Weekly clinicopathological exercises. Laboratory reference values. N Engl J Med. 2004;351:1548-63. Erratum in: N Engl J Med. 2004;351:2461.

Kratz A, Lewandrowski KB. Case records of the Massachusetts General Hospital. Weekly clinicopathological exercises. Normal reference laboratory values. N Engl J Med. 1998;339:1063-72. Erratum in: N Engl J Med 1999;340:1300.

New England Journal of Medicine. 1992 SI conversion guide

Smith A and Givens SV. Dealing with and Defining Abnormalities in Laboratory Data, Drug Info J 1993;27:771−8