1.0 Title Page

Statistical Analysis Plan

Study M15-574

A Phase 4, Double-Blind, Randomised, Placebo-Controlled Multicenter Study to Assess the Safety and Efficacy of Adalimumab Used in Conjunction with Surgery in Subjects with Moderate to Severe Hidradenitis Suppurativa

Date: 12 Aug 2019

Version 3.0

2.0 Table of Contents

List of Figures

3.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analysis for adalimumab study Protocol Study M15-574. It provides high level summaries of the planned statistical analyses for key efficacy and safety endpoints, interim analysis and multiplicity control strategies. Further details and analysis conventions to guide the statistical programming work will be in a supplement document.

Analyses will be performed using SAS® version 9.4 (SAS. Institute, Inc., Cary, NC 27513) or higher using the UNIX operating system.

SAP will not be updated in case of future administrative or minor amendments to the protocol unless the changes have an impact on the analysis of the study data.

4.0 Study Background

4.1 Objective

The primary objective of this study is to assess the safety and efficacy of adalimumab prior to surgery in subjects with moderate to severe HS who are surgical candidates, using Hidradenitis Suppurativa Clinical Response (HiSCR, defined as at least a 50% reduction in the HS abscess plus inflammatory nodule [AN] count with no increase in abscess count and no increase in draining fistula count relative to Baseline) as the assessment measurement.

Secondary objectives of this study include: Assess the impact of adalimumab specifically on the planned HS surgical site before surgery, evaluate the safety and efficacy of adalimumab continued during the perioperative period and after surgery, and evaluate Patient Reported Outcomes (PRO) related to health status, HS-related symptoms (e.g., drainage, swollen skin), physical functioning, treatment satisfaction, and work/activity impairment. The pharmacokinetics (PK) and immunogenicity of adalimumab following subcutaneous (SC) injection in this HS surgical population will

also be assessed. The objective is to determine if the inflammatory status of the HS surgical subject alters the PK profile both prior to and after surgery.

4.2 Study Design

4.2.1 Study Design and Design Diagram

This interventional, randomised, double-blind, placebo-controlled study is designed to enroll 200 subjects to meet scientific and regulatory objectives without enrolling an undue number of subjects in alignment with ethical considerations. Therefore, if the target number of subjects has been enrolled, there is a possibility that additional subjects in screening will not be enrolled. Eligible subjects will have active HS disease with an axilla or unilateral inguinal region that requires excisional surgery plus two other anatomical regions with active HS lesions (at least one region at Hurley Stage II or III).

The study duration will include an up to 30-day Screening Period, an initial 12-week double-blind treatment pre-surgery period (Period A), a 2-week peri-surgery period with continuation of weekly double-blind study drug administration (Period B), and a subsequent 10-week double-blind treatment post-surgery period (Period C).

Screening Period: The duration of the screening period will be a minimum of 7 days and a maximum of 30 days during which time all of the inclusion and exclusion criteria will be evaluated, and at referring sites, an eligibility visit with the designated surgeon will take place. Subjects may enroll into the study after all screening procedures are complete and results are known and verified for eligibility at the Baseline visit (Day 1) to receive 160 mg of adalimumab or matching placebo.

Period A: A 12-week double-blind, placebo-controlled treatment period during which subjects are randomised in a 1:1 ratio to receive adalimumab or matching placebo, 160 mg initially at Day 1 (or given as two 40 mg injections per day for 2 consecutive days), followed by 80 mg 2 weeks later at Day 15 (given as two 40 mg injections on a single day). Two weeks later (Day 29) continue with a dose of 40 mg ew until Week 12. The randomisation will be stratified by baseline Hurley Stage (II versus III) and anatomical

location of the surgical site (i.e., axilla versus inguinal region). A subject's Hurley Stage is determined by the worst Hurley Stage (i.e., II versus III) across all affected anatomic regions. The projected size of the surgical excision established by the designated surgeon during the Screening Period (calculated from a tracing of the outer perimeter onto an acetate sheet or equivalent) will be recorded by the study physician in the CRF. On each clinical study visit day, all study procedures scheduled for that visit will be completed prior to study drug administration.

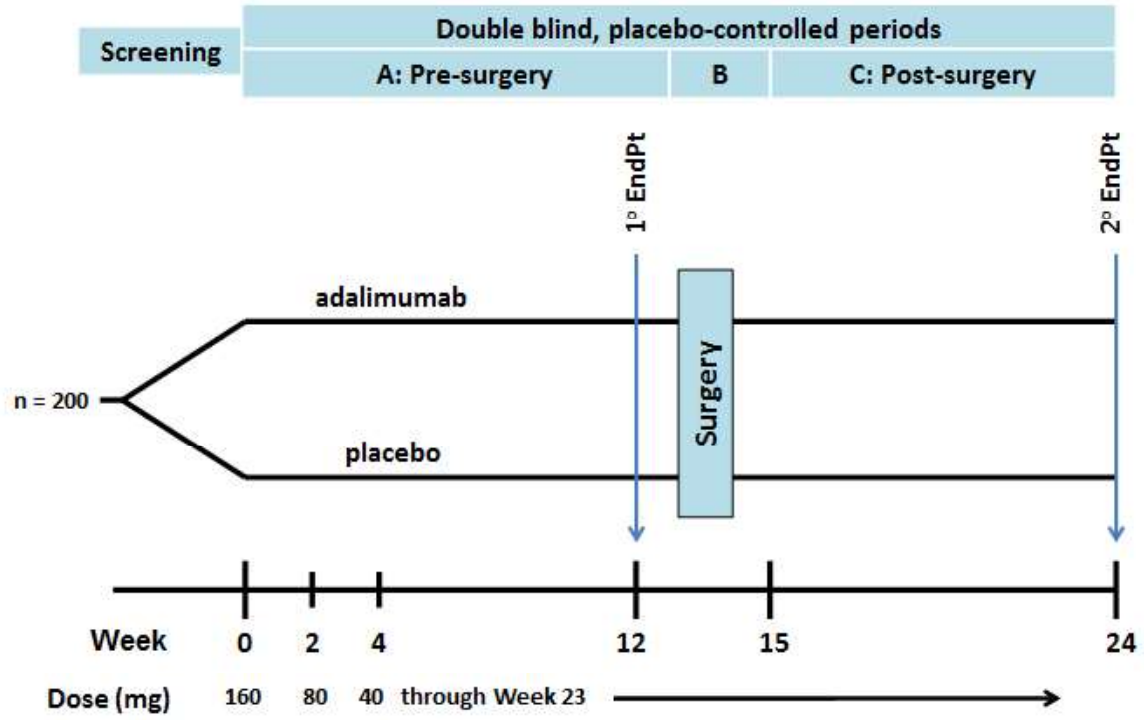
Period B: A 2-week double-blind, placebo-controlled treatment period consisting of Weeks 13 and 14 (the peri-operative period) in which subjects will continue the treatment on the first day of Weeks 13 and 14. The designated surgeon will measure the surface area of the actual surgery and record in source document. Surgery will occur during Week 13. The surgery and post-operative management (e.g., hospitalisation, surgical wound care) will be as per local practice.

Period C: A 10-week double-blind, placebo-controlled post-operative treatment period will occur from Week 15 through Week 24, during which the subject will continue the treatment assigned at Day 1 from Week 15 through Week 23. No study drug will be administered at Week 24, the final study visit. Subjects may begin commercial product (as prescribed by the subject's physician) after all Week 24 procedures have been completed.

Subjects will be contacted approximately 70 days following study drug discontinuation for an assessment of any new or ongoing AEs/SAEs, except for those subjects who continue on adalimumab not supplied in the context of the clinical trial after the end of study participation. These subjects are not required to complete the 70-day follow-up telephone call and any new AEs should be reported through the mechanism used for all post-marketing adverse experience.

The schematic of the study is shown in Figure 1. Further details regarding study procedures are located in study protocol Appendix C.

Figure 1. Study Schematic



See Section 5.2 of the protocol for information regarding eligibility criteria. Study sites and subjects will remain blinded for the duration of the study.

4.2.2 Variables used for Stratification at Randomisation

The randomisation will be stratified by baseline Hurley Stage (II versus III) and anatomical location of the surgical site (i.e., axilla versus inguinal region).

4.3 Endpoint

4.3.1 Primary Efficacy Endpoint

The primary efficacy variable is the proportion of subjects achieving HiSCR at Week 12. HiSCR is defined as at least a 50% reduction in the AN count with no increase in abscess count and no increase in draining fistula count relative to baseline.

4.3.2 Secondary Efficacy Endpoints

Ranked Secondary Efficacy Variables

1. Proportion of subjects achieving HiSCR-es (defined as the HiSCR excluding the HS surgical site) at Week 12
2. Proportion of subjects achieving HiSCR-es at Week 24
3. Percent change in surface area of the HS surgical site from Baseline to Week 12
4. Proportion of subjects at Week 12 that require a less extensive surgery than the surgical plan (determined at Baseline) or no surgery as determined by the designated surgeon

4.3.3 Other Efficacy Endpoints

Other efficacy variables to be analysed at each scheduled visit in Periods A and C, except for the ones included as the primary or ranked secondary variables that will be analysed for visits other than Weeks 12 and 24 include:

- Change from Baseline in CRP
- Proportion of subjects with DLQI = 0
- Proportion of subjects with DLQI = 0 or 1
- Change from Baseline in DLQI score
- Change from Baseline in EQ-5D index
- Change from Baseline in HS-PGA-SP

- Change from Baseline in HSIA
- Change from Baseline in HSSA-7d
- Change from Baseline in SF-12
- Change from Baseline in TSQM
- Change from Baseline in WPAI:SHP
- Proportion of subjects that experience a flare, defined as an at least 25% increase in AN count with a minimum increase of 2 relative to Baseline at each study visit, during Period A and during the study
- Proportion of subjects that experience at least 25% increase in each lesion type with a minimum increase of 2 relative to Baseline at each study visit, during Period A and during the study

For the proportion of subjects that experience a flare, and that experience at least 25% increase in each lesion type with a minimum increase of 2 relative to Baseline, analyses will be performed across all body regions at visits during Period A, and across body regions excluding the HS surgical site at all visits during the study.

Of note, HS PGIC is assessed for the validation of HS-PGA-SP for skin pain only and will not be analysed for efficacy.

4.3.4 Safety Endpoint

Adverse events, laboratory data, physical examinations, and vital signs will be collected, monitored, assessed, and recorded at the designated study visits listed in study protocol Appendix C and as described in study protocol Section 6.0 and Table 1.

4.3.5 Pharmacokinetic Endpoint

The pharmacokinetic endpoints will be analysed separately.

4.4 Sample Size Justification

The study is designed to enroll approximately 200 subjects. Assuming the true treatment difference in this study population is at least 20%, 100 subjects per arm will provide at

least 80% power to detect the treatment difference at the alpha level of 0.05. The study also has at least 80% power to demonstrate the point estimate of the treatment difference of at least 15%.

The response rates observed in the combined PIONEER I and II studies (Studies M11-313 and M11-810, respectively) for the HiSCR at Week 12 were 50.6% and 26.8% in the adalimumab ew group and placebo group, respectively. Due to the potential differences between the population of moderate-to-severe HS patients who are surgical candidates and the general moderate-to-severe HS patient population, a conservative estimate of 20% treatment difference is used for the power calculation.

4.5 Interim Analysis

No interim analysis is planned in this study.

4.6 Multiplicity Testing Procedures for Type-I Error Control

The statistical comparisons for the primary efficacy variable and the ranked secondary variables will be carried out in the hierarchical order under a two-sided significance level of 0.05.

4.7 Missing Data Imputation

Missing data will be imputed using the following methods for the efficacy analyses in the ITT Population defined in Section 5.0 below:

- Non-Responder Imputation (NRI): the NRI analysis will categorize any subject who has a missing value at a specific visit as non-responder for that visit. The only exception is when if the subject is a responder both before and after a specific visit window, then the subject will be categorized as a responder for the visit. The NRI will be the primary approach in the analyses of categorical variables. For analyses up to Week 12, NRI will be applied using evaluations within the period.
- Last Observation Carried Forward (LOCF): The LOCF analyses will use the completed evaluation from the previous visit within the particular period for

efficacy measures assessed to impute missing data at later visits in the same period. Baseline efficacy evaluations will not be carried forward. LOCF will be the primary approach in the analyses of continuous variables, and the secondary approach in the analyses of categorical variables.

- As-Observed: The As-observed analyses will not impute values for missing evaluations, and thus a subject who did not have an evaluation on a scheduled visit will be excluded from the As-observed analysis for that visit. As-observed analyses will be the secondary approach in the analyses of continuous variables.

Multiple Imputation (MI) for missing values will be performed for the primary efficacy variable. MI analysis will be carried out in three steps:

- Imputation of missing data. The imputation will be generated for each type of lesion count (abscess, inflammatory nodule, and draining fistula) as HiSCR components. The variables to be included in the imputation model are: Baseline lesion count for the corresponding type, smoking status, treatment group, stratification factors, and lesion count for the corresponding type at each visit from randomisation up to the end of the analysis period. For each component, 20 'complete' datasets will be generated from monotone regression model at each visit using SAS PROC MI. The seed for PROC MI will be set as the numeric form of the date when the first subject from this study is randomised. The imputed post-baseline lesion counts will be rounded to integers before the determination of HiSCR status. Subjects who discontinued due to lack of efficacy will be forced as non-responders.
- Analysis of imputed data sets. A Cochran-Mantel-Haenszel (CMH) test using Wilson-Hilferty transformation, stratified by stratification factors, will be used to analyse HiSCR response rates in each imputed dataset.
- Synthesis of imputation and analysis results. SAS PROC MIANALYZE will be used to generate the final inferences of the risk difference between treatment groups.

The 3rd and 4th ranked secondary endpoints will only be analysed by OC approach. For subjects whose surgery is no longer needed due to HS improvement, the actual surgical

area would be treated as 0 and the type of surgery would be treated as a less-extensive surgery. Subjects who discontinued before surgery due to AE of worsening in HS, or exceeding protocol specified number of lesion interventions, or planned HS surgery performed prior to Week 13, will be forced as no change in the planned surgical area and no less-extensive surgery.

In all methods of handling missing data, efficacy assessments observed on or before the planned surgery cannot be used to impute missing values for post-surgery visits, except for the lesion count on body regions other than the surgical site.

Lesions that receive protocol-allowed intervention (e.g., contained within the excised surgical specimen, protocol-allowed minor interventions including incision and drainage) will be counted as permanently present from the date of the intervention.

5.0 Analysis Populations and Important Subgroups

5.1 Analysis Population

The Intent-to-Treat (ITT) Population is defined as all subjects who randomised at the Baseline visit. The ITT Population will be used for the efficacy analysis. Of note, the surgery-related endpoints during Periods B and C will be analysed among the subjects who undergo surgery. Subjects in the ITT population will be analysed by treatment group as randomised.

In order to evaluate the impact of major protocol deviations on the primary efficacy endpoints, additional sensitivity analyses may be performed on a Per-protocol Population, which excludes subjects with major protocol deviations that potentially affect the primary efficacy endpoints. If it is decided that this analysis should be performed, the criteria for exclusion of subjects from the Per-protocol Population will be fully defined in the classification plan and the exclusion of subjects from the Per-protocol Population will be finalized before blind break and before the database lock.

The Safety Population is defined as all subjects who are in the ITT Population and receive at least one dose of study drug. The Safety Population in each period will be used for safety analysis. Of note, the surgery-related safety endpoints during Periods B and C will be analysed among the subjects who undergo surgery. Subjects in the Safety Population will be analysed by treatment group as treated.

5.2 Subgroup

To evaluate the consistency of the efficacy over demographic and other baseline characteristics, summaries and analyses will be performed for selected subgroups for the primary endpoint and ranked secondary endpoints. The detailed subgroups will be outlined in Section 6.5 and Section 7.5.

6.0 Efficacy Analysis

6.1 General Considerations

The efficacy analysis will be conducted among the ITT Population. A sensitivity analysis for the primary efficacy endpoint may also be conducted among the Per-protocol Population, if defined.

Categorical variables will be analysed using Cochran-Mantel-Haenszel (CMH) test, stratified by baseline Hurley Stage and anatomical location of the surgical. Continuous variables will be analysed using ANCOVA method, with treatment, baseline Hurley Stage, anatomical location of the surgical, and Baseline values in the model.

6.2 Primary Efficacy Analysis

The primary endpoint to assess the efficacy of adalimumab prior to surgery in subjects with moderate to severe HS is:

- Proportion of subjects achieving HiSCR at Week 12

The primary null hypotheses are that there is no difference in proportion of subjects who achieve HiSCR at Week 12, between the adalimumab and placebo treatment groups. The null hypotheses will be tested under a two-sided significance level of 0.05.

6.3 Secondary Efficacy Analysis

The following null hypotheses for ranked secondary endpoints below will be tested in a hierarchical order using two-sided significance level of 0.05 only if the null hypothesis for the primary endpoint has been rejected:

1. Adalimumab is not different from placebo with respect to the proportion of subjects achieving HiSCR-es at Week 12
2. Adalimumab is not different from placebo with respect to the proportion of subjects achieving HiSCR-es at Week 24
3. Adalimumab is not different from placebo with respect to the percent change in surface area of the HS surgical site from Baseline to Week 12
4. Adalimumab is not different from placebo with respect to the proportion of subjects at Week 12 that require a less extensive surgery than the surgical plan (determined at Baseline) or no surgery as determined by the designated surgeon

6.4 Other Efficacy Analysis

Other efficacy endpoints will be compared between the adalimumab and placebo treatment groups among the ITT Population.

6.5 Efficacy Subgroup Analysis

Primary and ranked secondary efficacy variables will also be analysed with respect to demographics and Baseline characteristics in addition to the stratification factors. The subgroups are defined as follows:

- Age group (< 40, 40 – 64, ≥ 65, if less than 10% of subjects in the ≥ 65 group, that group will be combined with 40 – 64 group.)
- Sex (male, female)
- Race (White, non-White)
- Duration of HS (by median)
- Weight (by median);
- BMI category: normal (< 25), overweight (25 – < 30), obese (30 – < 40), Morbid obesity (≥ 40)
- Current smoking status at Baseline (yes, no)
- Baseline CRP level (by median)
- Baseline AN count (≤ 5, 6 – 10, 11+)
- Baseline AN count (< median, ≥ median)
- Prior HS surgery history (yes, no)
- Time from prior HS surgery to the first dose of study drug (< median, ≥ median)

7.0 Safety Analysis

7.1 General Consideration

Safety analyses will include adverse events (AEs), laboratory, and vital sign measurements. Safety summaries will be provided using the safety populations.

Missing safety data will not be imputed.

7.2 Analysis of Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as any event with an onset that is after the first dose of study drug and with an onset date within 70 days after the last dose of study drug in the analysis period, or prior to the first dose in the subsequent period for subjects who entered in to the subsequent period.

Pre-treatment AEs will be summarized separately.

All TEAEs, serious adverse events (SAEs), AEs leading to discontinuation and AEs of Safety Interest will be summarized. The number and percentages of subjects experiencing TEAE will be tabulated using the Medical Dictionary for Drug Regulatory Activities (MedDRA®) system organ class and preferred term.

Summaries (including percentages and event per 100 patient-year) of SAEs, deaths, AEs leading to discontinuation, and AEs of Special Interest will be provided as well. No statistical tests will be performed.

7.3 Analysis of Laboratory Data

Mean change from Baseline in laboratory variables will be summarized. Shift tables for changes from Baseline according to the normal range will be provided. Frequencies and percentages of subjects with post baseline values meeting Criteria for Potentially Clinically Significant Laboratory values (i.e., the Common Toxicity Criteria (CTC) of Grade 3 or higher, as well as being a higher grade than the baseline CTC grade) will be summarized. For the assessments of laboratory data, values observed more than 70 days after the last dose of study drug will be excluded.

7.4 Analysis for Vital Signs

Changes from Baseline to post-baseline visits will be summarized. The number and percentage of subjects meeting the criteria for potentially clinically significant vital sign values will be summarized.

7.5 Safety Subgroup Analysis

The AE overview and AE by SOC and PT in each period will also be analysed with respect to the following Baseline characteristics. The subgroups are:

- Current smoking status at Baseline (yes, no)
- HbA1c ($< 7\%$, $\geq 7\%$)
- BMI (< 40, ≥ 40)
- Prior HS surgery history (yes, no)

8.0 Summary of Changes

The following changes are made in SAP version 2:

- Clarified multiple imputation details in Section 4.7.

The following changes are made in SAP version 3, before the study database lock and before the study unblinding:

- Clarified that HS PGIC is assessed to validate skin pain only and will not be analysed in Section 4.3.3.
- Clarified that the flare endpoint, and the endpoint about at least 25% increase in each lesion type with a minimum increase of 2 relative to Baseline, will be analysed across all body regions and across body regions excluding the surgical site in Section 4.3.3.
- Clarified details about missing data imputation in Section 4.7.
- Emphasized that lesions removed by protocol allowed interventions will be treated as permanently existing in Section 4.7.
- Clarified that efficacy subgroup analysis will be performed on selected subgroups in addition to stratification factors in Section 6.5.
- Removed 'in each period' in Section 7.5 for clarification purposes, since there is only one study period for safety analysis.

Document Approval

Study M15574 - Statistical Analysis Plan Version 3 - 12Aug2019 (E3 16.1.9)

Version: 1.0 **Date:** 02-Oct-2019 05:10:04 PM **Company ID:** 10022019-00F9F684301991-00001-en

Signed by:	Date:	Meaning Of Signature:
[redacted]	12-Aug-2019 06:47:22 PM	Author
	12-Aug-2019 07:41:58 PM	Approver
	13-Aug-2019 06:59:23 PM	Approver
	01-Oct-2019 07:14:49 PM	Approver
	02-Oct-2019 05:10:03 PM	Verifier and Approver