Protocol B7981037

A PHASE 2a, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY INVESTIGATING THE SAFETY OF RITLECITINIB (PF-06651600) IN ADULT PARTICIPANTS WITH ALOPECIA AREATA

Statistical Analysis Plan
(SAP)

Version: 5.0

Date: 04 April, 2023

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. VERSION HISTORY

Table 1. Summary of Major Changes in SAP Amendments

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
1.0 05 Aug 2020	Original 8 May 2020	Not Applicable	Not Applicable
2.0 28 Sep 2021	Amendment 1 20 Apr 2021	To match TLF	Add abbreviations to the analysis set
		For clarification	In section 5.2, elaborate the "Binary safety data"
		To assess the impact of COVID-19	In section 6.7, add COVID-19 related safety analysis
		For clarification	Add appendix 2, list of countries for regions
		For clarification	Add details to calculate cumulative dose/compliance by visit in Section 6.6.3 and appendix 3
		ECG analysis was missing in version 1	Add ECG analysis in Section 6.2.9
		For clarification	Added details to calculate cumulative dose and compliance in appendix 3.
		For clarification	In section 7, specified only data up to Month 9 will be included in the analysis for PCD
3.0 22 Nov 2021	Amendment 1 20 Apr 2021	For clarification	In section 7, update the scope of PCD CSR, and unblinding/blinding status of sponsors, investigators and participants after PCD
		For clarification	Insection 3.6.3, specified that only QTcF interval will be analyzed, instead of all ECG parameters
		Team decision	Delete the analysis of physical examinations, C-SSRS, and cumulative dose up to each visit
4.0 17 Aug 2022	Amendment 2 28 Apr 2022	Updated to incorporate changes in protocol amendment 2	In section 2, updated the study objects and endpoints, study design

Table 1. Summary of Major Changes in SAP Amendments

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
		Updated to incorporate changes in protocol amendment 2	Added exploratory efficacy endpoints in section 3.4
		Clarify according to PCD CSR	Add clarification in section 3.2 "For each participant, data were reported as the percentage of IENF with any number of swellings. The endpoint "axonal dystrophy" refers to the percentage of IENF with axonal swellings.", similar in section 6.2.2 "
		This analysis was added to the PCD CSR to assist with the interpretation of secondary endpoints	In section 6.2.3, the following analysis was added, "Descriptive Summary of absolute and change from baseline in IENFD normalized by age and gender."
		This analysis was was added to the PCD CSR to assist with the interpretation of the secondary endpoints	In section 6.2.4 "percent change from baseline in amplitude of wave V on BAEP at a stimulus intensity of 80dB at Months 6, 9, 9E, and 15E" was added
		Updated to incorporate changes in protocol amendment 2	Section 8 and Section 9 were added to specify the timing of each analysis and the endpoints included in each CSR.
		Updated to incorporate changes in protocol amendment 2	Section 11, the analysis visit window definition was updated to include TP3
5.0 04 April 2023	Amendment 2 28 Apr 2022	Update to include the definition of risk period for the purpose of summarizing	Section 3.6.1, add the definitions of risk periods for the summaries of AEs.

Table 1. Summary of Major Changes in SAP Amendments

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
		safety data. This update was made for consistency with the AA ritlecitinib program.	
		clarified that the Overall column in binary summary tables will also include the observation period	Section 8.2, Added table 4.
		Clarification of the statistical method used to analyze continuous endpoints	Section 5.2. General methods. Remove 'and subject as a random effect', and add instructions on how to run the model in SAS in case of convergence issues.
		Clarification that the linear mixed-effects model includes not only the primary timepoint, but also other timepoints	In Sections 6.1. and 6.2.1. added added 'which includes all post-baseline measurements throught …' to the statistical methods.
		Clarification of the statistical method used for the overall and AA SALT scores	Section 6.3.1. Removed the details describing the linear mixed-effects model,

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study B7981037. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

2.1. Study Objectives, Endpoints, and Estimands

The study objectives and corresponding endpoints are listed below in Table 2. Estimands are not applicable since there is no formal statistical hypothesis testing for this study.

Table 2. Study Objectives and Endpoints

Objectives	Endpoints
Primary – Safety	
• To assess I-V interwave latency on brainstem auditory evoked potentials (BAEPs) in adult participants with alopecia areata (AA) treated with PF-06651600.	• Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80 decibels (dB) at Month 9.
Secondary – Safety	
• To assess I-V interwave latency on BAEPs in adult participants with AA treated with PF-06651600.	• Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80dB at Months 6, 9E, and 15E.
• To assess axonal dystrophy in intraepidermal nerve endings over time in adult participants with AA treated with PF-06651600.	• Change from baseline in axonal dystrophy in skin punch biopsies at Month 9 and Month 15E.
• To assess intraepidermal nerve fiber density (IENFD) over time in adult participants with AA treated with PF06651600.	• Change from baseline in IENFD in skin punch biopsies at Month 9 and Month 15E.
• To assess wave V amplitude on BAEP over time in adult participants with AA treated with PF-06651600.	• Change from baseline in amplitude of wave V on BAEP at a stimulus intensity of 80dB at Months 6, 9, 9E, and 15E.
• To assess presence of wave V on BAEP over time in adult participants with AA treated with PF-06651600.	• Absence of wave V on BAEP at stimulus intensities ranging from 80dB to 40dB at Months 6, 9, 9E, and 15E.
• To evaluate the safety and tolerability of PF-06651600 in adult participants with AA.	• Incidence of treatment emergent adverse events (TEAEs), treatment-emergent serious adverse events (TESAEs), and adverse events (AEs) leading to discontinuation;[a] • Incidence of clinically significant abnormalities in vital signs and ECG; [a];

Table 2. Study Objectives and Endpoints

Objectives	Endpoints
	• Incidence of clinically significant abnormalities in clinical laboratory values. [a]
Secondary – Efficacy TP1-TP2	
• To evaluate response to PF-06651600 measured by the Severity of Alopecia Tool (SALT) in adult participants with AA.	• Change from baseline in overall and AA SALT score at Month 9 and other time points collected. [a]
• To evaluate the response to PF-06651600 measured by the Patient's Global Impression of Change (PGI-C) tool in adult participants with AA.	• PGI-C score response defined as greatly improved or moderately improved from baseline at Month 9 and other time points collected.
Tertiary/Exploratory-Efficacy TP3	
• To evaluate long-term response to PF 06651600 measured by the Severity of Alopecia Tool (SALT) in adult participants with AA.	• Change from baseline in overall and AA SALT[b] score at Month 51E and other time points collected.
• To evaluate the long-term response to PF-06651600 measured by the Patient's Global Impression of Change (PGI-C) tool in adult participants with AA.	• PGI-C score response defined as greatly improved or moderately improved at Month 51E and other time points collected.

a. Refer to Section 9 for specific safety summary included in each clinical study report
b. AA SALT is the amount of scalp hair loss due to AA.

2.2. Study Design

This is a Phase 2a, randomized, double-blind, placebo-controlled safety study designed to evaluate the safety and tolerability of PF-06651600, including the assessments of BAEP and IENF, in adults 18 to ≤50 years of age with ≥25% scalp hair loss due to AA.

Treatment Period 1 (TP1) and Treatment Period 2 (TP2):

The total duration of participation in the study for TP1 and TP2 will be approximately 26 months.

This includes up to a 5-week screening period, a 9-month placebo-controlled treatment phase, a 15-month active therapy extension phase, and a 4-week follow-up

period [for participants not continuing to TP3 or discontinuation of study intervention]. Screening will occur within 35 days prior to Day 1 of the study to confirm that selection criteria for the study are met. At Month 9, the placebo-controlled phase will end, and all participants may enter the active therapy extension phase of the study and receive PF- 06651600. For any participant discontinued for a neurological or auditory adverse event, the participant will be evaluated again at 6 months after discontinuation; the maximum duration of participation in the study for such participants could be approximately 32 months if the event occurred at the last visit.

Treatment Period 3 (TP3):

The total duration for TP3 will be of variable length for individual participants; assuming a participant does not require discontinuation per protocol. A participant may continue to receive PF-06651600 in TP3 for a maximum of an additional 36 months or until availability of commercial product in their country, or until the Sponsor terminates the study in that country, whichever occurs first.

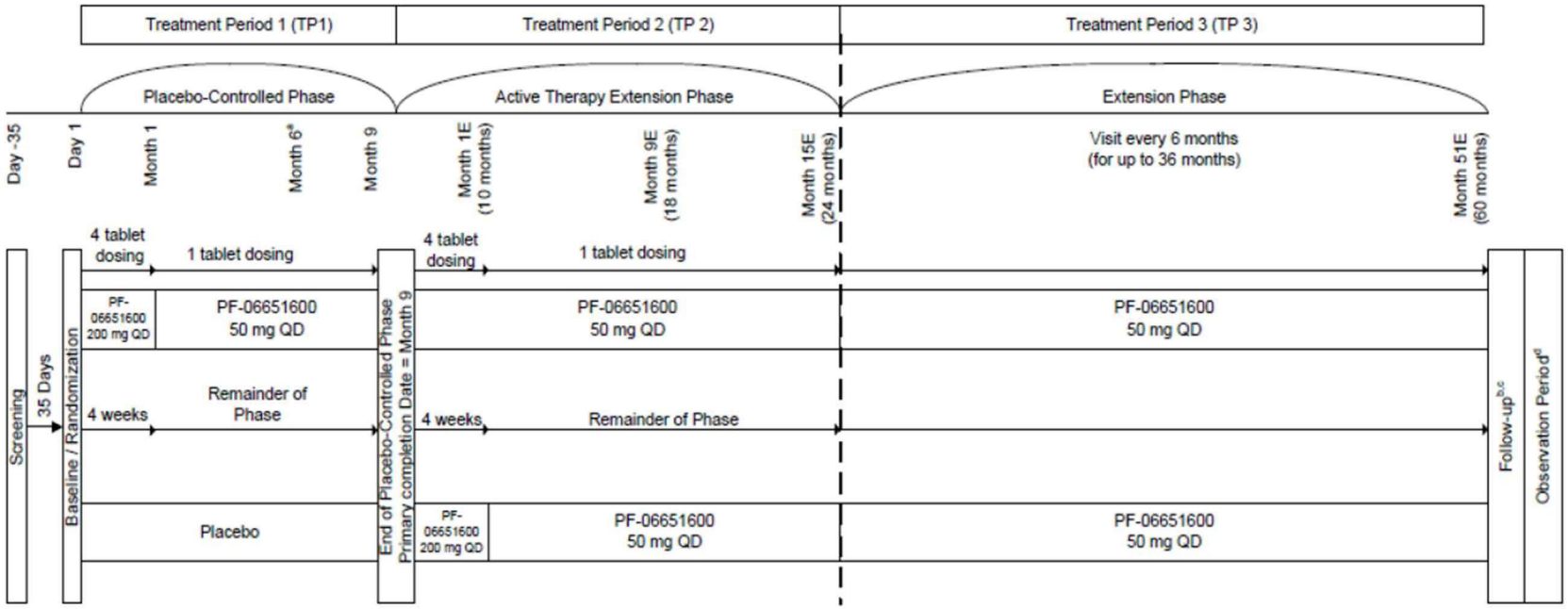
In TP3, participants will receive 50 mg QD PF-06651600. The total duration of participation in TP3 is approximately 37 months, including up to 36 months study intervention, and a follow up period of 4 weeks after completion or discontinuation of study intervention.

Observation Period:

If study intervention is permanently discontinued, the participant will be asked to remain in the study after the Follow-up visit without study intervention, and continue to comply with study visit schedules for approximately 2 years after the last dose of study intervention or until study end, whichever occurs first.

A study design schematic is presented in Figure 1.

Figure 1. Study Design Schematic



Abbreviations; E= active therapy extension phase; QD = once daily; TP = Treatment Period

a. Any participant with baseline overall SALT score ≤75 will be given the option to enter the active therapy extension phase at Month 6 if the overall SALT score at Month 6 has increased from baseline by ≥25 points.

b. After completion of TP2 (for participants not continuing to TP3) or discontinuation of study intervention, a Follow-up period of 4 weeks will occur. Participants in countries where PF-06651600 is not commercially available at the time their Month 24 visit will have the opportunity to enter TP3, of variable length for individual participants for a maximum of 36 months or until availability of commercial product in their country, or until the Sponsor terminates the study in that country, whichever occurs first

c. In TP3, after completion or discontinuation of study intervention, a Follow-up period of 4 weeks will occur.

d. If study intervention is permanently discontinued, the participant will be asked to remain in the study after the Follow-up visit for the Observation Period without study intervention and continue to comply with study visit schedules for approximately 2 years or until study end, whichever occurs first.

Note: If a participant discontinues for neurological or audiological events at any time during the study, a follow-up period of 6 months will occur.

Sample Size Determination

Approximately 60 adults 18 to 50 years of age will be randomized in a 1:1 allocation ratio to either the PF-06651600 200 mg QD/50 mg QD dose group or the placebo group.

For the primary safety endpoint, change from baseline in I-V interwave latency on BAEP at Month 9, the published standard deviations for I-V interwave latency on BAEP range from 0.1 to 0.3 ms.[1-3] Assuming a standard deviation of 0.2 ms, and assuming the standard deviation of change from baseline is similar to the standard deviation of actual score, the half-width of 95% confidence interval for the group will be 0.07 ms for 30 participants per group.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint – Safety in TP1

- Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80 decibels (dB) at Month 9.

3.2. Secondary Endpoints – Safety in TP1+TP2

- Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80dB at Months 6, 9E, and 15E;
- Change from baseline in axonal dystrophy in skin punch biopsies at Month 9 and Month 15E;
    - For each participant, data was reported as the percentage of IENF with any number of swellings. The endpoint “axonal dystrophy” refers to the percentage of IENF with axonal swellings.
- Change from baseline in IENFD in skin punch biopsies at Month 9 and Month 15E;
- Change from baseline in amplitude of wave V on BAEP at a stimulus intensity of 80dB at Months 6, 9, 9E, and 15E;
- Absence of wave V on BAEP at stimulus intensities ranging from 80dB to 40dB at Months 6, 9, 9E, and 15E;
- Incidence of treatmentemergent adverse events (TEAEs);
- Treatment-emergent serious adverse events (TESAEs);
- TEAEs leading to discontinuation;
- Incidence of clinically significant abnormalities in vital signs and ECG;
- Incidence of clinically significant abnormalities in clinical laboratory values.

3.3. Secondary Endpoints – Efficacy in TP1+TP2

- Change from baseline in overall and AA SALT score at Month 9 and other time points collected in TP1 and TP2;

- PGI-C score response defined as greatly improved or moderately improved from baseline at Month 9 and other time points collected in TP1 and TP2.

3.4. Exploratory Endpoints – Efficacy in TP3

- Change from baseline in overall and AA SALT score at Month 51E and other time points collected in TP3.

- PGI-C score response defined as greatly improved or moderately improved at Month 51E and other time points collected in TP3.

3.5. Baseline Variables

Baseline values will be defined as follows for both safety and efficacy endpoints:

- Baseline is defined as the last non-missing measurement obtained before the first dose in the placebo-controlled phase.

3.6. Safety Endpoints

The primary and secondary safety endpoints are defined above in Section 3.1 and Section 3.2.

3.6.1. Adverse Events

- The following risk period (RP) will be applied for TP1

  - For patients who entered TP1 and continued into TP2, RP is defined from date of 1st dose of TP1 to min (date of 1st dose of TP2 - 1, date of the last dose of TP1 + 35). This RP excluded AEs starting on the same day of 1st dose of TP2 or AEs starting on any day in TP2.

  - For patients who entered in TP1, permanently discontinued from study (and did not enter in TP2), RP is defined from date of 1st dose of TP1 to min (date of last contact, date of the last dose of TP1 + 35), where last contact date is min (date of discontinuation of study, date of data cut-off), if a subject dies, the death date will be the last contact date.

- The following risk period will be applied for TP2

  - For patients who entered TP2 and continued into TP3, RP is defined from date of 1st dose of TP2 to min (date of 1st dose of TP3 - 1, date of the last dose of TP2 + 35). This RP excludes AEs starting on the same day of 1st dose of TP3 or AEs starting on any day in TP3

  - For patients who entered in TP2 and, directly entered in the Observation Period or permanently discontinued from study (and did not enter in TP3), RP is defined from date of 1st dose of TP2 to min (date of last contact, date of the last dose of TP2 + 35), where last contact date is min (date of discontinuation of study, date of data cut-off), if a subject dies, the death date will be the last contact date. This RP extends 35 days into the observation period.

- The following RP will be applied for Observation Period for patients who entered from TP2

  - For patients who entered the observation period from TP2, RP is defined from date of last dose + 1 to the last contact date, where the last contact date is min (date of discontinuation of study, date of data cut-off). If a subject dies, the death date will be the last contact date. This RP excludes AEs starting wile on-drug.

- The following RP will be applied for TP3

  - For patients who entered in TP3, RP is defined from date of 1st dose in TP3 to min (last contact date, last dose date + 35) where last contact date is max (date of discontinuation of study, date of subject's last visit). In the case of an interim data cut, the last contact date is min (date of discontinuation of study, date of data cut-off). If a subject dies, the death date will be the last contact date. This RP includes AEs that occure within 35 days after the last dose

- The following RP will be applied for Observation Period for patients who entered from TP3

  - For patients who entered the observation period from TP3, RP is defined from date of last dose + 1 to last contact date, where last contact date is max (date of discontinuation of study, date of subject's last visit). In the case of an interim data cut, the last contact date is min (date of discontinuation of study, date of data cut-off). If a subject dies, the death date will be the last contact date. This RP excludes AEs starting while on-drug.

- Supplemental CSR1 will include summaries of TEAEs as described in Table 4.

- Supplemental CSR2 will include summaries of TEAEs for the RP defined for TP3, and for the RP defined for the Observation period for patients who entered from TP3.

3.6.2. Laboratory Data

Below is a list of Laboratory Assessments to be performed at times specified in the schedule of activities in the protocol.

Hematology	Chemistry	Urinalysis	Other
Hemoglobin	BUN and creatinine	pH	At screening or Day 1 only:
Hematocrit	Glucose	Glucose (qual)	• FSH • Pregnancy test (β-hCG) • Hepatitis C antibody • HIV • HbA1C • HbsAg • HbcAb • HbsAb • HBVDNA • HCVAb • HCVRNA • QFT-G test or T-Spot test • Viral Screen • Fasting Lipid profile
RBC count	Calcium	Protein (qual)	
%Reticulocytes	Sodium	Blood (qual)	
MCV	Potassium	Ketones	
MCH	Chloride	Nitrites	
Platelet count	Total CO_2 (bicarbonate)	Leukocyte esterase	
WBC count	AST, ALT	Urobilinogen	
Neutrophils	Total bilirubin	Urine bilirubin	
Eosinophils	Direct and Indirect bilirubin	Microscopy	
Monocytes	Alkaline phosphatase	Urine culture	
Basophils	Uric acid		
Lymphocytes	Albumin		
	Total protein		
	Creatine Kinase		
	Follow up testing for potential DILI cases		

3.6.3. Electrocardiogram (ECG)

ECG measurements will be collected at screening, at the end of placebo-controlled phase and at the end of active therapy extension phase.

3.6.4. Vital Signs, including Height and Weight

Vital sign measurements are pulse rate, blood pressure, respiratory rate, and oral, tympanic or axillary temperature. They will be performed at times specified in the schedule of activities in the protocol.

Height will be collected at screening and weight will be collected at both screening and baseline. Weight will be collected at the end of the placebo-controlled phase and at the end of the active therapy extension phase.

3.6.5. Physical Examinations

Complete physical examinations consist of assessments of general appearance; skin; head, eyes, ears, nose, and throat (HEENT); mouth, heart; lungs; abdomen; extremities; neurologic and auditory function; back; and lymph nodes. In addition, dermatological full body exam

must be performed by the investigator, sub-investigator or a qualified health professional per local guidelines. Dermatological examinations should also include visual inspection of the breasts and external genitalia.

Targeted physical examinations should include skin, heart, lung, and abdomen, neurologic and auditory function, and examination of body systems where there are symptom complaints by the participant.

Physical examinations will be performed at times specified in the schedule of activities in the protocol.

In addition, at the times specified in the protocol, all patients will have an audiological evaluation including BAEP by an audiologist and a neurological evaluation by a specialist.

4. ANALYSIS SETS

For purposes of analysis, the following populations are defined:

Analysis Set	Description
Safety Analysis Set (SAS)	All participants taking at least 1 dose of study intervention. Participants will be analyzed according to the intervention they actually received.
Efficacy Analysis Set (EAS)	All randomized participants taking at least 1 dose of study intervention. Data will be summarized according to the intervention they were assigned.

5. GENERAL METHODOLOGY AND CONVENTIONS

The primary analysis for this study will be conducted following the primary completion date (PCD) at Month 9 and after the Month 9 data are cleaned and locked.

5.1. Hypotheses and Decision Rules

For the primary safety endpoint, there is no formal hypothesis testing. The analyses will focus on estimation. Summary statistics for each group with 95% confidence intervals will be provided.

No estimands will be defined since there is no formal statistical hypothesis testing for this study.

5.2. General Methods

In the placebo-controlled period, the following continuous endpoints: (change from baseline in I-V interwave latency on BAEP, change from baseline in amplitude of wave V, and change from baseline in overall and AA SALT score, will be analyzed using the linear mixed -effects model with baseline, treatment group, visit, and treatment group by visit interaction as fixed effects, with an unstructured covariance matrix assumption, if there are convergence issues, a compound symmetry matrix will be used. The Kenward-Roger degrees of freedom approximation will be used. For participants who switch to the active therapy extension phase at Month 6, only their data through Month 6 will be included in the analysis of the placebo-controlled period. Descriptive statistics based on observed data will also be provided. For TP2 and TP3, descriptive statistics will be presented.

For the binary endpoint PGI-C response, in the placebo-controlled phase, number and percentage of responders with 95% confidence interval (based on Clopper-Pearson method) by treatment group, and treatment difference with 95% confidence intervals (based on Chan and Zhang's exact method)[4] will be presented. Participants who discontinue from the study for any reason or who switch to active therapy extension phase before Month 9 will be considered as non-responders. Number and percentage based on observed data will also be provided. For TP2 and TP3, number and percentage of responders with 95% confidence interval (based on Clopper–Pearson method) will be presented based on observed data.

For all other data, descriptive summary statistics such as number and percentage will be presented for binary variables, and number, mean, standard deviation (or standard error of the mean), median, minimum, and maximum will be presented for continuous variables.

5.3. Methods to Manage Missing Data

- For the PCD CSR
  - For change from baseline in I-V interwave latency on BAEP, change from baseline in amplitude of wave V, and change from baseline in overall and AA SALT score during the placebo-controlled phase, the missing values will be handled in a linear mixed-effects model (Section 5.2).
  - For the PGIC response, participants who discontinue from the study for any reason or who switch to active therapy extension phase before Month 9 will be considered as non-responders at Month 9.
- For supplemental CSR 1 and CSR 2
  - All efficacy data will be summarized based on observed data. Missing data will not be imputed.

6. ANALYSES AND SUMMARIES

6.1. Primary Endpoint

- Primary endpoint: Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80 decibels (dB) at Month 9;
- Population: Safety;
- Statistical Method: a linear mixed-effects model in Section 5.2, which includes all the post-baseline measurements through Month 9, will be used. The estimated means, with standard errors and 95% confidence intervals for each treatment group based on the model will be presented. Descriptive statistics will also be provided based on the observed data. Data from right side and left side will be summarized separately.

6.2. Secondary Endpoints-Safety

6.2.1. I-V Interwave Latency on BAEP

- Endpoint: Change from baseline in I-V interwave latency on BAEP at a stimulus intensity of 80dB at Months 6, 9E, and 15E.

- Population: Safety.

- Statistical Method: Linear mixed-effects model in Section 5.2, which includes all the post-baseline measurements through Month 6, will be used. The estimated means, with standard errors and 95% confidence intervals for each treatment group based on the model will be presented. Descriptive statistics will also be provided based on the observed data. Descriptive statistics will be presented for data in the active therapy extension phase. Data from right side and left side will be summarized separately.

6.2.2. Percentage of IENF with Axonal Swellings in Skin Punch Biopsies

- Endpoint: Change from baseline in the percentage of IENF with axonal swellings in skin punch biopsies at Month 9/EOP (End of placebo-controlled phase/ beginning of active therapy extension phase) and Month 15E.

- Population: Safety.

- Statistical Method: Descriptive statistics for continuous variables.

6.2.3. IENFD in Skin Punch Biopsies

- Endpoint: Change from baseline in IENFD in skin punch biopsies at Month 9/EOP and Month 15E.

- Population: Safety.

- Statistical Method: Descriptive statistics for continuous variables. Descriptive Summary of absolute and change from baseline in IENFD normalized by age and gender will also be provided.

    - For each participant normalized IEFND data was obtained by dividing the IEFND by the value of the LLN of each age range and gender, and multiplying it by 100 to obtain percentages. LLN is given by the 5th percentile of the distribution of IENFD by gender and age: females in age ranges of 18-39, 40-59 and 60+ are 8.0, 5.4 and 3.5 respectively, and for males, 5.9, 4.6 and 3.3 respectively.

6.2.4. Amplitude of Wave V on BAEP

- Endpoint: Change and percent change from baseline in amplitude of wave V on BAEP at a stimulus intensity of 80dB at Months 6, 9, 9E, and 15E.

- Population: Safety.

- Statistical Method: For change from baseline, a linear mixed-effects model for Months 6 and 9 in Section 5.2 wil be used. The estimated means, with standard errors and 95% confidence intervals for each treatment group based on the model will be presented. Descriptive statistics based on observed data will also be provided. Descriptive statistics will be presented for data in the active therapy extension phase. Data from right side and left side will be summarized separately. For percent change from baseline, descriptive statistics will be presented.

6.2.5. Absence of Wave V on BAEP

- Summary: Proportion of subjects with an absence of wave V on BAEP at stimulus intensities ranging from 80dB to 40dB at Months 6, 9, 9E, and 15E.
- Population: Safety.
- Statistical Method: Descriptive statistics with number and percentage of subjects with the events only. Data from right side and left side will be summarized separately.

6.2.6. Incidence of Treatment-emergent Adverse Events (TEAEs)

- Endpoint: Incidence of subjects with treatment emergent adverse events.
- Population: Safety.
- Statistical Method: Descriptive statistics with number and percentage of subjects with the events only.

6.2.7. Incidence of Treatment-emergent Serious Adverse Events (TESAEs)

- Endpoint: Incidence of Treatment-emergent serious adverse events.
- Population: Safety.
- Statistical Method: Descriptive statistics with number and percentage of subjects with the events only.

6.2.8. Incidence of Adverse Events (AEs) Leading to Discontinuation

- Endpoint: Incidence of Adverse events (AEs) leading to discontinuation.
- Population: Safety.
- Statistical Method: Descriptive statistics with number and percentage of subjects with the events only.

6.2.9. Incidence of Clinically Significant Abnormalities in ECG (only QTcF) and Vital Signs

- Endpoint: Incidence of clinically significant abnormalities in ECG (only QTcF) and vital signs.

- Population: Safety.
- Statistical Method: Descriptive statistics with number and percentage of subjects with the events only.

6.2.10. Incidence of Clinically Significant Abnormalities in Clinical Laboratory Values

- Endpoint: Incidence of clinically significant abnormalities in clinical laboratory values.
- Population: Safety.
- Statistical Method: Descriptive statistics with number and percentage of subjects with the events only.

6.3. Secondary and exploratory Endpoints-Efficacy

6.3.1. Overall and AA SALT Scores

- Endpoint: Change from baseline in overall and AA SALT scores in all the timepoins collected.

The AA SALT score at each visit will be calculated using non-AA SALT scores as follows:

- the non-AA SALT score collected at Month 9 (or Month 6) or the Early Termination visit (as appropriate) will be used for visits in TP1.
- the non-AA SALT score collected at Month 15E or the Early Termination visit (as appropriate) will be used for visits in TP2, and
- the non-AA SALT score collected at Month 51E or the Early Termination visit (as appropriate) will be used for visits in TP3.
- The formula for deriving the AA SALT score is given by: AA SALT score = overall SALT score – non-AA SALT score.
- non-AA scores that were blank will be considered as zero in cases there was no medical history indicating AGA

- Population: Efficacy.
- Statistical Method: For placebo-controlled phase, a linear mixed-effects model in Section 5.2 will be used. The estimated mean, standard errors and 95% confidence intervals for each treatment group, and mean difference with 95% confidence interval will be presented at each time point. Descriptive statistics will be presented for data in TP2 and TP3.

6.3.2. PGI-C Response

- Summary: Proportion of subjects with PGI-C response defined as greatly improved or moderately improved at all the timepoints collected.
- Population: Efficacy.
- Statistical Method: For placebo-controlled phase, statistics including number and percentage with 95% confidence interval (based on Clopper-Pearson method) by treatment group, and treatment difference with 95% confidence interval (based on Chan and Zhang's exact method) will be presented. Missing data will be treated as non-responders. Summary based on observed data will also be provided. For TP2 and TP3, descriptive statistics including number and percentage with 95% confidence interval (based on Clopper–Pearson method) will be presented based on observed data.

6.4. Other Endpoints

Not applicable. No pharmacokinetic or pharmacodynamic parameters were evaluated in this study.

6.5. Subset Analyses

No subgroup analyses will be performed.

6.6. Baseline and Other Summaries and Analyses

6.6.1. Baseline Summaries

Demographics and medical history will be summarized for all participants. Baseline disease characteristics will also be summarized.

6.6.2. Study Conduct and Subject Disposition

Subjects evaluation, disposition, and discontinuation will be summarized for all participants.

6.6.3. Study Treatment Exposure

A summary of compliance and the number of doses received as well as the median total dose by visit and groups will be provided. Compliance (%) = (sum of actual number of doses / sum of planned number of doses) * 100 for each visit period. Where sum of actual number of doses = total duration of each visit period* actual dose (Tablet) and sum of planned number of doses = total duration of each visit period* planned dose (Tablet)

Duration of treatment in months will be summarized by the total number of days of dosing, calculated by (last dosing end date - first dosing start date + 1 - number of missed or zero doses days)/ 30.4375..

For a specific dosing record, if the dosing end date is missing and the participant is still in the study, then the CSR cut-off date will be applied.

6.6.4. Concomitant Medications and Non-Drug Treatments

Prior drug and non-drug treatments, concomitant drug and non-drug treatments will be summarized for all participants.

6.6.5. Alopecia areata history

Alopecia areata history will be summarized.

- Participants with a SALT score of <100% at Baseline (regardless of the category in the AA history CRF) will be categorized as non-AT/non-AU.
- Participants with a SALT score of 100% at Baseline (regardless of the category in the AA history CRF) will be categorized as AT/AU.
    - Within the AT/AU category, classification into the AT and AU subcategories will be based in the AA history CRF;
    - Participants with a SALT score of 100% at Baseline who were not classified as AT or AU in the AA history CRF are considered as "Not specified".

6.7. Safety Summaries and Analyses

Safety analysis will be based on the safety population.

Analyses for the primary and secondary safety endpoints are discussed in Section 6.1 and Section 6.2.

In general, safety data will be presented in tabular and/or graphical format and summarized descriptively, where appropriate. All safety endpoints will be listed and summarized in accordance with Pfizer Standards. Categorical outcomes (eg, AEs) will be summarized by subject counts and percentage. Continuous outcome (eg, blood pressure, pulse rate, etc.) will be summarized using N, mean, median, standard deviation, etc. Change from baseline in laboratory data, ECGs and vital signs will also be summarized. Subject listings will be produced for these safety endpoints accordingly. Clinically significant abnormalities in vital signs and clinical laboratory values are discussed in Sections 6.2.9 and 6.2.10.

In order to report the impact of COVID-19 on clinical trial popuations and study data, additional listings and summaries will be produced:

- Protocol deviations related to COVID-19
- Summary of discontinuations due to COVID-19 related reasons
- Discontinuation from study/study drug interruptions due to COVID-19 related adverse events
- COVID-19 related adverse events
- Number of days with missed doses due to COVID-19
- Missed visits summary for I-V interval latency on BAEP

6.7.1. Adverse Events

The safety data will be summarized in accordance with CDISC and Pfizer Standards (CaPS). All safety data will be summarized descriptively through appropriate data tabulations, descriptive statistics, categorical summaries, and graphical presentations.

Adjudication data will be reported separately.

6.7.2. Laboratory Data

Laboratory data will be listed and summarized in accordance with the Pfizer reporting standards. Summaries of subjects meeting pre-specified monitoring and discontinuation criteria will be created using methods for categorical data.

6.7.3. Vital Signs

Vital signs will be summarized at baseline and all-available post-baseline visits.

Height will be collected pre-dose and weight will be collected at pre and post dose.

6.7.4. Electrocardiogram

ECG parameters QTcF and change from baseline will be summarized at baseline and all available post-baseline visits. Categorization of QTcF will also be provided.

7. INTERIM ANALYSES

There is no formal interim analysis planned in this study. However, if required, interim analyses may be performed for study monitoring for internal decision making, due to regulatory requests, or to support regulatory submissions. As no statistical hypotheses will be tested in this study, there are no issues of protecting the Type I error rate.

This study uses an External Data Monitoring Committee (E-DMC). The E-DMC will be responsible for ongoing monitoring of safety of subjects in the study according to the DMC charter. The recommendations made by DMC to alter the conduct of the study will be forwarded to the appropriate Pfizer personnel for final decision. Pfizer will forward such decisions, which may include summaries of aggregate analyses of safety data, to regulatory authorities, as appropriate.

Independent safety adjudication committees will be used for opportunistic infections, cardiovascular events, neurological, including audiological events, and malignancy events.

8. TIMING AND TREATENT GROUPS DISPLAYED FOR EACH ANALYSIS

There will be 3 Clinical Study Reports (CSRs) for this study: one PCD-CSR, and two supplemental CSRs.

8.1. PCD-CSR

The PCD-CSR will include the primary analysis for this study, which will be conducted following the PCD at Month 9 (TP1), and after the Month 9 data are cleaned and locked for all participants. The following data will be included:

- Efficacy data through Month 9 (TP1)
- All safety data collected through the PCD data cut-off date

For duration of treatment, concomitant medication, lab abnormality, lab values meeting discontinuation criteria, vital signs/QTcF values meeting pre-specified criteria and AE tables, the following groups will be summarized:

Table 3. Groups

Period	Groups
Placebo-controlled phase (TP1)	• Ritlecitinib 200/50 mg QD • Placebo
Active therapy extension phase (TP2)	• Ext Ritlecitinib 50 mg QD • Ext Ritlecitinib 200/50mg QD
Overall (TP1+TP2)	• Ritlecitinib 200/50/50 mg QD • Any Ritlecitinib • Any Ritlecitinib + Placebo (Total)

For tables summarizing data by visit (such as BAEP, skin punch biopsy endpoints and efficacy endpoints), which only include data from TP1 in the PCD-CSR, the following groups will be summarized:

- Ritlecitinib 200/50 mg QD
- Placebo

For other tables such as dispositions, demographic and baseline characteristics, medical history, primary diagnosis, prior medications and compliance the groups to be summarized will be:

- Ritlecitinib 200/50 mg QD
- Placebo → Ritlecitinib 200/50 mg QD.

8.2. Supplemental CSR 1

The supplemental CSR 1 will be prepared after the last participant completes the 4-week follow-up visit, or the Month 15E visit (TP2) if the last participant is enrolled in TP3, whichever is later. This CSR will include:

- Efficacy through Month 15E (TP1+TP2). Efficacy data from participants while they are on the Observation Period will not be included in the summaries of efficacy. Instead, listings and plots may be provided.
- All safety data collected through the cut-off date for this CSR (including the Follow-up Period and the Observation Period).

For summaries in this CSR, the groups will be the following:

- For duration of treatment, concomitant medication, lab abnormality, lab values meeting discontinuation criteria, vital signs/QTcF values meeting pre-specified criteria and AE tables.

Table 4. Groups

Period	Groups
Placebo-controlled phase (TP1)	• Ritlecitinib 200/50 mg QD • Placebo
Active therapy extension phase (TP2)	• Ext Ritlecitinib 50 mg QD • Ext Ritlecitinib 200/50mg QD
Overall (TP1+TP2+TP3+Observation period)	• Ritlecitinib 200/50/50 mg QD • Any Ritlecitinib • Any Ritlecitinib + Placebo (Total)

- For all the other tables, groups will be:
  - Ritlecitinib 200/50/50 mg QD
  - Placebo → Ritlecitinib 200/50 mg QD.

8.3. Supplemental CSR 2

The supplemental CSR 2 will be focused on safety and efficacy data from TP3 and will be prepared after the last participant completes the last visit or discontinues from the study.

- Disposition date of TP2 (or of 4 weeks follow-up period if participants didn't enter TP3) will be used to identify records in TP2 and TP3.
- All safety data in TP3, regardless of whether participants are in the Observation Period or not, will be included in summary tables.
- Only one group "All participants" including all participants will be used in all summaries.
- Efficacy data of participants while in the Observation Period will not be included in summary tables that present TP3 data by visit. Instead, listings and plots may be provided.

9. ENDPOINTS INCLUDED IN EACH CSR

Efficacy analysis

PCD CSR (Month 9; TP1)	
Endpoint	**Endpoints**
Primary	• There are no primary efficacy endpoints.
Secondary	• PGI-C response • Overall and AA SALT.
Exploratory	• There are no exploratory efficacy endpoints included in this CSR
Supplementary CSR 1 (Month 15E; TP2)	
Endpoint	**Statistical Analysis Methods**
Primary	• There are no primary efficacy endpoints.
Secondary	• PGI-C response • Overall and AA SALT
Exploratory	• There are no exploratory efficacy endpoints included in this CSR.

Supplementary CSR 2 (Month 51E; TP3)	
Endpoint	**Statistical Analysis Methods**
Primary	• There are no primary efficacy endpoints.
Secondary	• There are no secondary efficacy endpoints included in this CSR.
Exploratory	• PGI-C response • Overall and AA SALT score

Safety

PCD CSR (Month 9; TP1)	
Endpoint	**Statistical Analysis Methods**
Primary	• Change from baseline in I-V interwave latency on BAEP at Month 9
Secondary	• Change from baseline in I-V interwave latency on BAEP at Month 6 • Change from baseline in IENFD at Month 9 • Change from baseline in percentage of IENF with axonal swellings in skin punch biopsies at Month 9 • Change and percent change from baseline in amplitude of wave V at Months 6 and 9 • Absence of wave V on BAEP at Months 6 and 9 • The incidence rates of safety events, significant clinical abnormalities, vital signs and ECG
Supplementary CSR 1 (Month 15E; TP2)	
Endpoint	**Statistical Analysis Methods**
Primary	• There are no primary safety endpoints included in the supplemental CSR 1.
Secondary	• Change from baseline in I-V interwave latency on BAEP • Change from baseline in IENFD • Change from baseline in percentage of IENF with axonal swellings from skin punch biopsies • Change and percent change from baseline in amplitude of wave V • Absence of wave V on BAEP • The incidence rates of safety events, significant clinical abnormalities, vital signs and ECG

Supplementary CSR 2 (Month 51E; TP3)	
Endpoint	**Statistical Analysis Methods**
Primary	• There are no primary safety endpoints included in the supplemental CSR 2.
Secondary	• Incidence rates of safety events, significant clinical abnormalities and vital signs

10. REFERENCES

1. Waliszewska-Prosól, M, et al. Visual and brainstem auditory evoked potentials in HCV-infected patients before and after interferon-free therapy — A pilot study. International Journal of Infectious Diseases 80 (2019) 122–128.

2. Mohamed, E, et al. Auditory system dysfunction in patients with vitiligo: is it a part of a systemic autoimmune process? The Egyptian Journal of Otolaryngology. 2017, 33:594–602.

3. Baweja, P, et al. Changes in brainstem auditory evoked potentials among North Indian females with Type 2 diabetes mellitus. Indian Journal of Endocrinology and Metabolism / Nov-Dec 2013 / Vol 17 | Issue 6.

4. Chan, IS, Zhang, Z. Test-based exact confidence intervals for the difference of two binomial proportions. Biometrics. 1999;55(4):1202-9.

11. APPENDICES

11.1. Appendix 1 Definition and Use of Visit Windows in Reporting

Visit windows will be used for efficacy variables and for any safety data that are displayed or summarized by study visit.

Visit Label	Target Day	Analysis Visit Window
Baseline	1	Up to Day 1
Month 1	31	Days 2 to 60
Month 3	91	Days 61 to 136
Month 6	181	Days 137 to 226 but <=Day 1E
Month 9	271	Days 227 to 316 but <=Day 1E
Month 3E	91E	Days 2E to 136E
Month 6E	181E	Days 137E to 226E
Month 9E	271E	Days 227E to 316E
Month 12E	361E	Days 317E to 406E
Month 15E	451E	Day 407E to 541E
Month 21E	631E	Day 542E to 721E
Month 27E	811E	Day 722E to 901E
Month 33E	991E	Day 902E to 1081E
Month 39E	1171E	Day 1082E to 1261E
Month 45E	1351E	Day 1262E to 1441E
Month 51E	1531E	Day 1442 to last dose + 14 days
28-Day Follow-up	28 days after last dose	Based on nominal visit
6-Month Follow-up	182 days after discontinuation	Based on nominal visit

E=relative day to the first dose at active therapy extension phase.

If more than one observation falls within a visit window, the observation that is the closest to the target date for that visit will be used for the summary tables. If two visits are equally distant from the target day, the data from the later visit will be used.

Data collected in TP2, from the first dose in TP2 till the end of follow-up or the first dose in TP3, will be used to derive analysis visits in TP2. Data collected in TP3, from the first dose in TP3 or the Observation Period will be used to derive analysis visits in TP3.

Some endpoints (eg, BAEP) are not scheduled to be assessed at each visit and some measurement may fall outside the scheduled visit windows. In this case, for variables where mixed model methodology is used, the data from the unscheduled 'windowed' visits will not be included in mixed model analysis. However, these visits will be summarized descriptively in the observed cases analysis, using the visit timepoint corresponding to their actual visit window (for example a participant with a visit on Day 59 will be summarized at Month 1 in the observed case summaries).

11.2. Appendix 2 List of Countries by Region

Region	Countries
North America	Canada, United States
Europe	Poland
Rest of World	Australia, Argentina, Chile, Colombia, Mexico

11.3. Appendix 3 Study Days Used to Define Each Visit Period in the Calculation of By Visit Table for Compliance.

Month 1: Day 1 – Day 31

Month 3: Day 32 to Day 91

Month 6: Day 92 to Day 181

Month 9: Day 182 to Day 1E -1

Month 3E: Day 1E to Day 91E

Month 6E: Day 92E to Day 181E

Month 9E: Day 182E to Day 271E

Month 12E: Day 272E to Day 361E

Month 15E: Day 362E to Day 451E

Month 21E: Day 452E to Day 631E

Month 27E: Day 632E to Day 811E

Month 33E: Day 812E to Day 991E

Month 39E: Day 992E to Day 1171E

Month 45E: Day 1172E to Day 1351E

Month 51E: Day 1352E to Day 1531E