Official Title: KPL-716-C201: A Phase 2a/b, Randomized, Double-Blind, Placebo-Controlled Study to Investigate the Efficacy, Safety, Tolerability and Pharmacokinetics of KPL-716 in Reducing Pruritus in Subjects with Prurigo Nodularis

NCT Number: NCT03816891

Document & Date: Statistical Analysis Plan for Phase 2b, version 1: 12-Aug-2022

Kiniksa Pharmaceuticals, Ltd.

STATISTICAL ANALYSIS PLAN (SAP)

for KPL-716-C201 Phase 2b

A Phase 2a/b, Randomized, Double-Blind, Placebo Controlled Study to Investigate the Efficacy, Safety, Tolerability and Pharmacokinetics of KPL-716 in Reducing Pruritus in Subjects with Prurigo Nodularis

Protocol Number: KPL-716-C201
IND Number: 132912

SAP Version: Version 1

Date: 12AUG2022

TABLE OF CONTENTS

Listing of Abbreviations

Abbreviation	Full description
ADA	Anti-drug antibodies (anti-KPL-716 antibodies)
AE	Adverse event
ANCOVA	Analysis of covariance
ATC	Anatomical Therapeutic Chemical
BMI	Body mass index
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CSR	Clinical study report
DB	Double Blinded
DBL	Database lock
ECG	Electrocardiogram
eCRF	Electronic case report form
EDC	Electronic data capture
EOT	End of treatment
EOS	End of study
ET	Early Termination
GAD-7	General Anxiety Disorder-7
HR	Heart rate
ICH	International Conference on Harmonization
IGA-CNPG-A	Investigator's Global Assessment for Prurigo Nodularis-Activity
IGA-CNPG-S	Investigator's Global Assessment for Prurigo Nodularis-Stage
LOCF	Last observation carried forward
LS	Least squares
MCMC	Markov Chain Monte Carlo Method
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
mITT	Modified intent-to-treat
MMRM	Mixed model for repeated measures
OLE	Open Label Extension
PD	Pharmacodynamic(s)

Abbreviation	Full description
PGIC	Patient Global Impression of Change
PHQ-9	Patient Health Questionnaire-9
PK	Pharmacokinetic(s)
PN	Prurigo Nodularis
PN-NAT	Prurigo Nodularis Nodule Assessment Tool
PN-IGA	Prurigo Nodularis Investigator Global Assessment
PP	Per protocol
PT	Preferred term
Q1	First quartile
Q3	Third quartile
Q4W	Every Four (4) Weeks
QoL	Quality of life
QT	Electrocardiographic interval from the beginning of the QRS complex to the end of the T wave
QTcB	QT interval corrected for heart rate using Bazett's method
QTcF	QT interval corrected for heart rate using Fridericia's method
SAE	Serious adverse experience/event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	Standard deviation
SE	Standard error
SOC	System organ class
TEAE	Treatment-emergent adverse event
TCS	Topical corticosteroid
VAS	Visual Analog Scale
WHO	World Health Organization
WHODRUG	World Health Organization Drug Dictionary
WI-NRS	Worst Itch Numeric Rating Scale
WOCF	Worst Observation Carried Forward

1. INTRODUCTION

This Statistical Analysis Plan (SAP) contains a detailed description of data presentations, statistical analysis methods, and data reporting specifications for the programming outputs and preparation of the Clinical Study Report (CSR) for the Phase 2b portion of study KPL-716-C201. The SAP for the phase 2a portion was already prepared separately.

The SAP is based on the protocol KPL-716-C201version 6 dated 19 Jul 2021, KPL-716-C201 protocol European version 1.0 dated 06AUG2021, and KPL-716-C201 protocol, UK-Specific version 2.0 dated 18AUG2021.

1.1. Study Objectives

1.1.1. Primary objective

- To evaluate the efficacy of subcutaneous (SC) KPL-716 vs. placebo in reducing pruritus in Prurigo Nodularis (PN) subjects experiencing severe pruritus

1.1.2. Secondary objectives

- To evaluate the effect of SC KPL-716 vs. placebo in improving sleep in PN subjects experiencing severe pruritus
- To evaluate the efficacy of SC KPL-716 vs. placebo in improving quality of life in PN subjects experiencing severe pruritus
- To evaluate the effect of SC KPL-716 vs. placebo in reducing disease severity in PN subjects experiencing severe pruritus
- To evaluate the safety and tolerability of SC KPL-716 vs. placebo in PN subjects experiencing severe pruritus
- To evaluate the pharmacokinetics (PK) of SC KPL-716 in PN subjects experiencing severe pruritus

1.1.3. Exploratory objectives

- To evaluate the immunogenicity of SC KPL-716 in PN subjects experiencing severe pruritus
- To evaluate the effect of KPL-716 vs. placebo on skin and blood biomarkers
- To evaluate the correlation between KPL-716 responsiveness and co-morbidities in PN

1.2. Study Design

This is a randomized, double-blind, placebo-controlled study to investigate the efficacy, safety, tolerability, PK, and immunogenicity of KPL-716 administered SC in subjects with moderate to severe PN experiencing severe pruritus.

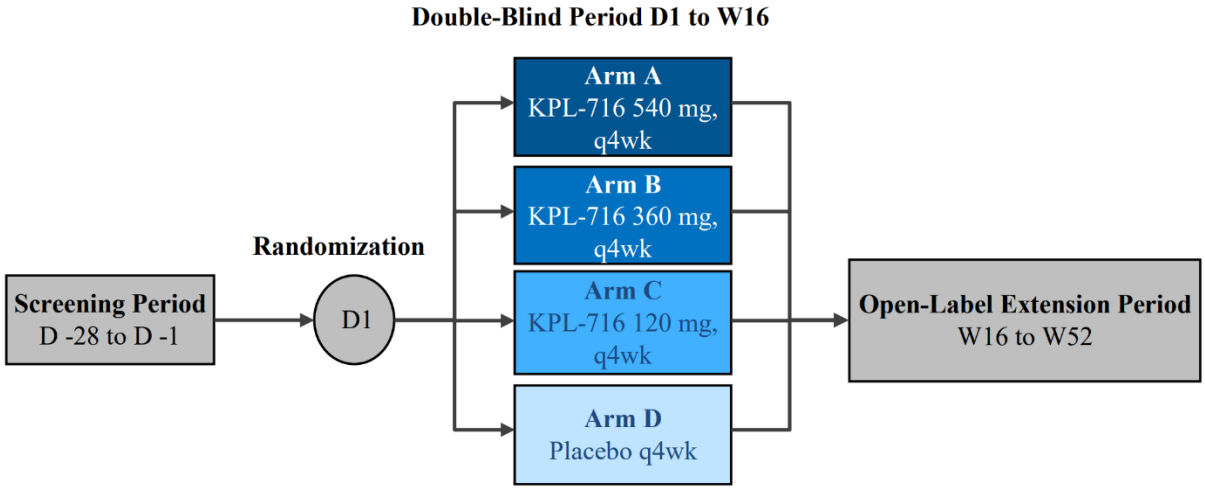
The Phase 2b study (Figure 1) consists of a 4-week screening period and a 16-week double-blind (DB) period, followed by a 36-week open-label-extension (OLE) period. Approximately 180

subjects with moderate to severe PN experiencing severe pruritus will be randomized (at 1:1:1:1 ratio) into one of the following 4 arms:

- Arm A: KPL-716, 540 mg, SC, every 4 weeks (Q4W)
- Arm B: KPL-716, 360 mg, SC, Q4W
- Arm C: KPL-716, 120 mg, SC, Q4W
- Arm D: Placebo, SC, Q4W

A total of 4 doses of study drug will be administered during the DB Period to measure the efficacy, safety, and PK of KPL-716.

Figure 1 Phase 2b Study Design Diagram



After the DB period, all subjects will have the option to receive KPL-716 360 mg SC, Q2W during the OLE period to evaluate the long-term safety and PK of KPL-716. For subjects who complete the DB period but choose not to participate in OLE, or who terminate earlier, an End of Study (EOS) visit will be required 4 weeks following the date of last study drug dose.

Subjects will follow the schedule below:

- Screening period (minimum 7 days and maximum 28 days)
- DB period (Day 1 [Baseline] to Week 16)
- OLE period (End of Week 16 to Week 52)

The allowable visit window from Week 2 to Week 16 is ± 2 days and thereafter to Week 52 is ± 3 days.

Randomization and Dosing:

Subjects will be randomized 1:1:1:1 into one of the following 4 arms: KPL-716 540 mg SC, Q4W; KPL-716 360 mg SC, Q4W; KPL-716 120 mg SC, Q4W; or placebo SC, Q4W. At the end of DB period, all subjects will have the option to enter the OLE period to receive KPL-716

360 mg SC Q2W. Stratification will be performed based on sex and years since first nodule observed (i.e., ≥10 years vs. <10 years).

Efficacy Assessments:

Efficacy in reduction of pruritus will be assessed via daily recording of the Worst Itch Numeric Rating Scale (WI-NRS). Efficacy in improvement of sleep will be assessed via Sleep Loss Visual Analog Scale (VAS). Impact on quality of life will be assessed via Itchy Quality of Life (ItchyQoL). Depression and anxiety will be assessed based on Patient Health Questionnaire-9 (PHQ-9) and General Anxiety Disorder-7 (GAD-7). Impact on disease severity will be followed by Prurigo Nodularis Investigator Global Assessment (PN-IGA), Prurigo Nodularis Nodule Assessment Tool (PN-NAT), Investigator's Global Assessment for Prurigo Nodularis-Stage (IGA-CNPG-S), Investigator's Global Assessment for Prurigo Nodularis-Activity (IGA-CNPG-A), Patient Global Impression of Prurigo Nodularis Severity (PGI-PN-S), and Patient Global Impression of Change (PGIC). Schedules of these efficacy assessments and related tools are included in the Protocol appendices.

- WI-NRS score: The daily diary contains one numerical rating scale to assess subjects' pruritus (WI-NRS) and one Yes/No question regarding the use of rescue medication for itch relief. Subjects will be asked to assign a numerical score to the intensity of their most severe (worst) pruritus using a scale from 0 to 10, with 0 indicating no pruritus and 10 indicating the worst imaginable pruritus. This pruritus NRS will be used to assess subjects' daily level of worst pruritus from the Screening Visit through the EOS Visit.
- PN-IGA: This is a novel exploratory tool for the overall assessment of PN disease severity based on the elevation of the nodules and number of elevated nodules. The PN-IGA will be performed by the Investigator. The PN-IGA utilizes a 5-point scale that ranges from 0 (clear) to 4 (severe). Subjects may present with smaller papular lesions, scarring, or post-inflammatory hyper- or hypopigmentation of the skin, which are not considered as part of this assessment. A PN-IGA score is assigned based on the appearance of the disease at the time of evaluation without referring to the baseline state. Every effort will be made for the same Investigator to assess PN-IGA over time.
- IGA-CNPG-S: This is a novel tool for the assessment of PN disease severity based on the number of palpable nodules. The IGA-CNPG-S will be performed by the Investigator. The IGA-CNPG-S utilizes a 5-point scale that ranges from 0 (clear) to 4 (severe). A score is assigned based on the appearance of the disease at the time of evaluation without referring to the baseline state. Every effort will be made for the same Investigator to assess IGA-CNPG-S over time.
- Sleep Loss VAS: Subjects will be asked to place a line perpendicular to the VAS line at the point that represents the intensity of their average sleeplessness experienced over the previous 3 nights using a scale from 0 to 10, with 0 indicating no sleeplessness and 10 indicating the worst imaginable sleeplessness. The Sleep Loss VAS is administered per the Schedule of Activities.
- ItchyQoL: This tool focuses on the impact of pruritus on daily activities and on the level of psychological stress. It contains 22 items. The frequency items are scored

using a 5-point Likert scale ranging from "never" to "all the time." The bother items are scored from 1 (not bothered) to 5 (severely bothered). The recall period in ItchyQoL is the past week. The ItchyQoL total score is defined as the sum of all 22 item scores. The ItchyQoL-symptom subscale score is calculated as the sum of items 1-6. The ItchyQoL-functional subscale score is calculated as the sum of items 7-13. The ItchyQoL-emotional subscale score is calculated as the sum of items 14-22. For each subscale, if more than 50% of items have missing values, the subscale will be set to missing and no total score will be calculated. If items with missing values are equal to or less than 50%, the subscale will be prorated based on items with non-missing values. Subscale can be rounded into integer after proration. Total scores can be classified as little (≤30), mild (31-50), moderate (51-80), or severe (81-110).

- IGA-CNPG-A is a novel tool for the assessment of PN disease activity based on the proportion of nodules with excoriations and crusts. The IGA-CNPG-A will be performed by the Investigator. The IGA-CNPG-A utilizes a 5-point scale that ranges from 0 (clear) to 4 (severe). A score is assigned based on the appearance of the disease at the time of the evaluation without referring to the baseline state. Every effort will be made for the same Investigator to assess IGA-CNPG-A over time.
- PN-NAT: This is a novel exploratory tool for the evaluation of disease severity based on 5 components:
    1. estimate of the number of nodules over the whole body,
    2. estimate of hardness of nodules in the representative area,
    3. estimate of extent of excoriation over the whole body,
    4. distribution of nodules,
    5. exact nodules count in the representative area.

  Subjects may present with smaller papular lesions, scarring, or post-inflammatory hyper- or hypo-pigmentation of the skin, which are not considered as part of this assessment. A score is assigned to each component based on the appearance of the disease at the time of the evaluation without referring to the baseline state. Given its exploratory nature, subtotal score of the first three components will be calculated while distribution of nodules and exact nodule count in the representative area will be analyzed separately.
- The PHQ-9 is a subject-reported measure that assesses the subject's general mental health. The PHQ-9 contains 9 items scored on a Likert scale ranging from 0: "not at all" to 3: "nearly every day." The recall period is the last 2 weeks. The PHQ-9 will be administered at designated visits.
- The GAD-7 is a subject-reported survey that assesses the subject's level of anxiety. The GAD-7 contains 7 items scored on a Likert scale ranging from 0: "not at all" to 3: "nearly every day." The recall period is the last 2 weeks. The GAD-7 will be administered at designated visits.

- The PGI-PN-S is a subject-reported questionnaire that assesses the subject's impression of PN disease severity at the time of evaluation. It is scored from 0: "absent (no sign or symptom of prurigo nodularis)" to 4: "severe."
- The PGIC is a subject-reported survey that assesses the subject's impression of change in their PN disease since the beginning of the study. Scoring ranges from 1: "very much improved" to 7: "very much worse".

1.3. Study Endpoints

1.3.1. Efficacy Endpoints

1.3.1.1. Primary Endpoint

- Percent change from baseline in weekly average Worst Itch–Numeric Rating Scale (WI-NRS) at Week 16

1.3.1.2. Key Secondary Endpoints

- Proportion of subjects achieving at least a 4-point reduction from baseline in weekly average WI-NRS at Week 16
- Proportion of subjects achieving clear (0) or almost clear (1) in PN-IGA at Week 16
- Proportion of subjects achieving at least a 6-point reduction from baseline in weekly average WI-NRS at Week 16

1.3.1.3. Other Secondary Efficacy Endpoints

Related to pruritus:

- Change and percent change from baseline in weekly average of WI-NRS over time
- Proportion of subjects achieving at least a 6-point reduction from baseline in weekly average WI-NRS over time
- Proportion of subjects achieving at least a 4-point reduction from baseline in weekly average WI-NRS over time

Related to disease severity:

- Proportion of subjects achieving clear (0) or almost clear (1) in PN-IGA over time
- Proportion of subjects with at least 2-point improvement from baseline in PN-IGA over time
- Proportion of subjects achieving clear (0) or almost clear (1) in Investigator's Global Assessment for Prurigo Nodularis-Stage (IGA-CNPG-S) over time
- Proportion of subjects with at least 2-point improvement from baseline in IGA-CNPG-S over time

Related to sleep:

- Change and percent change from baseline in weekly average of Sleep Loss Visual Analog Scale (VAS) over time

Related to quality of life:

- Change and percent change from baseline in ItchyQoL over time

1.3.1.4. Exploratory Endpoints

- Proportion of subjects achieving clear (0) or almost clear (1) in IGA-CNPG-Activity (IGA-CNPG-A) over time
- Change from baseline in Prurigo Nodularis Nodule Assessment Tool (PN-NAT) over time
- Change from baseline in the total score of Patient Health Questionnaire (PHQ-9) over time
- Change from baseline in the total score of General Anxiety Disorder (GAD-7) over time
- Days and proportion of subjects using rescue medication over time
- Patient Global Impression of Prurigo Nodularis Severity (PGI-PN-S) over time
- Patient Global Impression of Change (PGIC) over time
- Change from baseline in skin and blood biomarkers over time

1.3.2. Safety Endpoints

- Incidence rate and severity of treatment-emergent adverse events (TEAEs)
- Incidence rate and severity of study drug-related TEAEs
- Vital signs, electrocardiogram (ECG), and clinical laboratory test results

1.3.3. Other Endpoints

- PK: Measurement of plasma concentrations of KPL-716
- Immunogenicity: Measurement of anti-drug (anti-KPL-716) antibodies (ADA)

1.4. Estimands

The primary estimands defined for main endpoints are summarized below in Table 1. More details are provided in Section 4.6.

Table 1: Summary of primary estimands for main endpoints

Endpoint Category	Estimands			
	Endpoint(s)[a]	**Analysis set**	**Intercurrent event(s) strategy and missing data handling**	**Analysis set-level summary**
Primary objective: To evaluate the efficacy of subcutaneous (SC) KPL-716 vs. placebo in reducing pruritus in prurigo nodularis (PN) subjects experiencing severe pruritus				
Primary endpoint	Percent change from baseline in weekly average Worst Itch–Numeric Rating Scale (WI-NRS) at Week 16	mITT	The intercurrent events will be handled as follows: • Discontinuation of study treatment before Week 16: Off-study treatment data up to Week 16 will be included in the analysis (treatment policy strategy). • Taking rescue/prohibited medications that impact efficacy analyses prior to Week 16: data will be set to missing values after the medication usage, and the participant's worst postbaseline observation before the time of the medication usage will be carried forward (WOCF) to impute missing endpoint value (hypothetical strategy) In addition, the missing data imputation rules are as follows: • Missing data at Week 16 will be imputed with Multiple Imputation (MI) method.	Analysis of Covariance (ANCOVA) model will be used with treatment and randomization stratification factors of sex and years since the first nodule observed (i.e., ≥10 vs. <10 years) as fixed effect factors, baseline as covariate.

Endpoint Category	Estimands			
	Endpoint(s)[a]	**Analysis set**	**Intercurrent event(s) strategy and missing data handling**	**Analysis set-level summary**
Key secondary endpoint	Proportion of subjects achieving at least a 4-point reduction from baseline in weekly average WI-NRS at Week 16	mITT	The intercurrent events will be handled as follows: • Discontinuation of study treatment before Week 16: Off-study treatment data up to Week 16 will be included in the analysis (treatment policy strategy). • Taking rescue/prohibited medications that impact efficacy analyses prior to Week 16: Participants will be considered as non-responders (composite strategy). In addition, the missing data imputation rules are as follows: • Having missing data at Week 16: Participants will be defined as non-responders.	CMH test adjusted by randomization stratification variables: sex and years since the first nodule observed (i.e., ≥10 years vs. <10 years) will be used.

Endpoint Category	Estimands			
	Endpoint(s)[a]	**Analysis set**	**Intercurrent event(s) strategy and missing data handling**	**Analysis set-level summary**
Secondary objective: To evaluate the effect of SC KPL-716 vs. placebo in reducing disease severity in PN subjects experiencing severe pruritus				
Key secondary endpoint	Proportion of subjects achieving clear (0) or almost clear (1) from baseline in PN-IGA at Week 16	mITT	The intercurrent events will be handled as follows: • Discontinuation of study treatment before Week 16: Off-study treatment data up to Week 16 will be included in the analysis (treatment policy strategy). • Taking rescue/prohibited medications that impact efficacy analyses prior to Week 16 assessment: Participants will be considered as non-responders (composite strategy). In addition, the missing data imputation rules are as follows: • Having missing data at Week 16 assessment: Participants will be defined as non-responders.	CMH test adjusted by randomization stratification variables: sex and years since the first nodule observed (i.e., ≥10 years vs. <10 years) will be used.

[a] Additional secondary objectives/endpoints are not included in this table but will be handled with a similar strategy as the endpoint type (i.e., continuous, proportion) at other weeks, otherwise specified.

1.5. Determination of Sample Size

The Phase 2b sample size calculation is based on the primary efficacy endpoint (percentage change from baseline of WI-NRS at Week 16) using a two-sample t-test.

Approximately 180 subjects in total will be equally randomized to the 4 arms of Phase 2b. At Week 16, 53% and 30% reduction in weekly average of WI-NRS are assumed for KPL-716 and placebo arms, respectively. For each comparison, assuming treatment effect of 23% difference in weekly average WI-NRS reduction from baseline at Week 16 and standard deviation of 35%, and two-sided alpha of 0.05, 38 subjects in each arm will provide 80% power to detect the treatment effect. Approximately 45 subjects in each arm will be randomized assuming 15% drop out during the 16 weeks of treatment.

The sample size calculation was performed using nQuery version 8.5.

2. STUDY ANALYSIS SETS

The following analysis sets will be defined for Phase 2b.

2.1. Randomized Subjects

Randomized subjects will include all subjects with randomization number assigned.

2.2. Modified Intent-to-Treat (mITT) Analysis Set

All randomized subjects who receive at least one dose of KPL-716 or placebo and have at least 1 daily WI-NRS score in the double blind Treatment Period will be included in the modified intent-to-treat (mITT) analysis set. All mITT analyses will be based on each subject's randomized treatment assignment.

2.3. Safety Analysis Set

All randomized subjects who received at least one dose of KPL-716 or placebo will be included in the safety analysis set. Safety analyses will be based on the treatment that was administered to each subject.

2.4. Per Protocol (PP) Analysis Set

The PP analysis set includes all mITT subjects who have no important protocol deviations that may potentially bias efficacy analyses of the study. These deviations will be pre-specified prior to study unblinding/database lock.

2.5. OLE Analysis Set

OLE analysis set includes all randomized subjects who consented to OLE period and received at least one dose of KPL-716 for OLE period.

2.6. PK Analysis Set

The PK analysis set includes any subjects who received KPL-716 and who had at least one PK sample.

3. GENERAL CONSIDERATIONS

All analyses and summaries will be produced using SAS® version 9.4 (or higher).

For inferential statistical analyses, each arm of KPL-716 (Arm A KPL-716: 540 mg Q4W, Arm B KPL-716: 360 mg Q4W, Arm C KPL-716: 120 mg Q4W) will be compared to placebo. Unless otherwise specified, all tests will be at two-tailed using pre-specified level of significance 0.05.

Descriptive statistics (n, mean, standard deviation [SD], median, Q1, Q3, minimum, and maximum) will be presented for continuous variables. For continuous efficacy endpoints, least square (LS) mean, LS mean difference, standard error (SE), and 95% confidence interval (CI) will be calculated.

For categorical variables (including binary variables), counts and percentages will be presented. For binary efficacy endpoints, 95% CI for difference in proportion and p-values will be presented.

Subject listings in all randomized subjects, will be provided for all efficacy and safety data. In general, the subject listings will be sorted by treatment group, subject number and assessment date (and time, if applicable).

3.1. Baseline Value and Change from Baseline

The baseline value of efficacy parameters except is defined as the last non-missing value before randomization and prior to the first dose of study medication otherwise specified.

For daily collected efficacy endpoints (WI-NRS), baseline is calculated as weekly average of non-missing scores from 7 days prior to randomization (Day -7 to Day -1).

For safety endpoints, baseline is defined as the last non-missing value obtained prior to the first dose in DB Period otherwise specified.

3.2. Study Day and Treatment Day

For efficacy endpoints analyses, study day is defined as the number of days from randomization date to the event/visit date. It is calculated as follows:

- If the event date falls on or after randomization date,

  Study day = Event or visit date – randomization date + 1

- If the event date falls before randomization date,

  Study day = Event or visit date – randomization date

For safety endpoints analyses, treatment day is defined as the number of days from DB first dose date to the event/visit date. It is calculated as follows:

- If the event date falls on or after DB first dose date,

  Treatment day = Event or visit date –DB first dose date + 1

- If the event date falls before DB first dose date,

  Treatment day = Event or visit date –DB first dose date

3.3. Analysis Visits and Study Period

For efficacy analyses, DB period is defined from randomization up to the first OLE dose if subjects enter OLE treatment period, or end of study in subjects not enrolled in OLE period, unless specified. OLE period is defined from OLE first dose up to the end of study. If assessment time is available, assessments should be further differentiated by time compared to OLE first dose. If assessments are collected on the same day as OLE first dose date and no assessment times are available, those assessments should be counted into DB period.

For safety analyses, DB period is defined from first dose up to the first OLE dose if subjects enter OLE treatment period, or end of study if subjects not enrolled in OLE period.

Efficacy assessments

For the efficacy assessment, the reference date for the derivation of relative days of events or findings will be the randomization day. If a subject receives treatment dose prior to the randomization by mistake, the reference date of efficacy assessment will be the date of the first treatment dose administration for that subject.

For the daily WI-NRS, weekly average will be aligned by study periods. For DB period, weekly average will be calculated based on every 7 days increment starting on Day 1. For week 16, only assessments before the first dose of OLE period will be included for weekly average calculation. Time windows are specified in Table 2 below.

Table 2: Time window of weekly average WI-NRS in DB period

Analysis week	Time windows referring to randomization date
Baseline	-7 to -1
Week 1	1 - 7
Week 2	8 - 14
Week 3	15 - 21
Week 4	22 - 28
Week 5	29 - 35
Week 6	36 - 42
Week 7	43 - 49
Week 8	50 - 56

Week 9	57 - 63
Week 10	64 - 70
Week 11	71 - 77
Week 12	78 - 84
Week 13	85 - 91
Week 14	92 - 98
Week 15	99 - 105
Week 16	106 - min(113, Day of OLE first dose)

The time window of weekly average NRS in OLE period will start from the day after the first dose of OLE period as specified in Table 3.

Table 3: Time window of weekly average WI-NRS in OLE period

Analysis week	Time windows referring to OLE first dose date
Week 17 - OLE	2 - 8
Week 18 - OLE	9 - 15
Week 19 - OLE	16 - 22
Week 20 - OLE	23 - 29
Week 21 - OLE	29 - 36
Week 22 - OLE	37 - 43
Week 23 - OLE	44 - 50
Week 24 - OLE	51 - 57
Week 25 - OLE	58 - 64
Week 26 - OLE	65 - 71
Week 27 - OLE	72 - 78
Week 28 - OLE	79 - 85
Week 29 - OLE	86 - 92
Week 30 - OLE	93 - 99
Week 31 - OLE	100 - 106
Week 32 - OLE	107 - 113
Week 33 - OLE	114 - 120
Week 34 - OLE	121 - 127
Week 35 - OLE	128 - 134
Week 36 - OLE	135 - 141

Week 37 - OLE	142 - 148
Week 38 - OLE	149 - 155
Week 39 - OLE	156 - 162
Week 40 - OLE	163 - 169
Week 41 - OLE	170 - 176
Week 42 - OLE	177 - 183
Week 43 - OLE	184 - 190
Week 44 - OLE	191 - 197
Week 45 - OLE	198 - 204
Week 46 - OLE	205 - 211
Week 47 - OLE	212 - 218
Week 48 - OLE	219 - 225
Week 49 - OLE	226 - 232
Week 50 - OLE	233 - 239
Week 51 - OLE	240 - 246
Week 52 - OLE	247 - EOS

For other efficacy variables, all available values of scheduled measurements will be assigned to the appropriate visit window according to Table 4. In the event of multiple measurements of the same test in the same window, the one closest to the targeted visit date will be used for the by-visit summary. If they are at the same number of days away from the target day, the latest one will be used.

Table 4: Time window for other efficacy variables in DB period

		Time windows			
Visit	Target day referring to randomization date	Sleep Loss VAS, GA-CNPG-S, IGA-CNPG-A, PN-NAT, PN-IGA, PGI-PN-S	ItchyQoL	PHQ-9, GAD-7	PGIC
Baseline*	1	≤1	≤1	≤1	
Week 2	15	2 - 22			
Week 4	29	23 - 43	2- 43		2- 43
Week 8	57	44 - 71	44 - 71	2 - 85	44 - 71
Week 12	85	72 - 99	72 - 99		72 - 99
Week 16	113	100 - OLE first dose	100 - OLE first dose	86 - OLE first dose	100 - OLE first dose

The visit mapping in OLE period will start from the day after the first dose of OLE period based on the time window specified Table 5.

Table 5: Time window for other efficacy variables in OLE period

		Time windows
Visit	Target day referring to OLE first dose date	Sleep Loss VAS, GA-CNPG-S, IGA-CNPG-A, PN-NAT, PN-IGA, PGI-PN-S, ItchyQoL, PHQ-9, GAD-7, PGIC
Week 18 - OLE	15	2 - 22
Week 20 - OLE	29	23 - 71
Week 32 - OLE	113	72 - 169
Week 48 - OLE	225	170 - 239
Week 52 - OLE	253	240 - EOS

Safety assessments

For the safety assessment, the reference date for the derivation of relative days of events or findings will be the DB first dose date. Selected safety variables will be summarized by the analysis window defined in Table 6 for the by visit descriptive analysis. In the event of multiple measurements of the same test in the same window, the one closest to the targeted visit date will be used for the by-visit summary. If they are at the same number of days away from the target day, the latest one will be used.

Table 6: Time window for safety endpoints in DB period

		Time windows		
Visit	Target day referring to DB first dose	Physical examination, Vital signs, Body weight, Clinical lab blood tests ADA	ECG Pregnancy test	Urinalysis
Baseline*	1	≤1	≤1	≤1
Week 2	15	2 - 22		
Week 4	29	23 - 43		

Week 8	57	44 - 71	2 - 85	
Week 12	85	72 - 99		
Week 16	113	100 - OLE first dose	86 - OLE first dose	2 - OLE first dose

* up to first dose date/time

The visit mapping in OLE period will start from the day after the first dose of OLE period using the time window specified in Table 7.

Table 7: Time window for safety variables in OLE period

		Time windows		
Visit	Target Day referring to OLE first dose date	Physical examination, Vital signs, Body weight, Clinical lab blood tests ADA	ECG	Urinalysis, Pregnancy test
Week 18 - OLE	15	2 - 22		
Week 20 - OLE	29	23 - 71		
Week 32 - OLE	113	72 - 169		2 - 169
Week 48 - OLE	225	170 - 239		170 - 239
Week 52 - OLE	253	240 - EOS	2 - EOS	240 - EOS

3.4. Missing Data Handling

No imputations will be performed for missing data for safety endpoints, unless otherwise specified. Missing data handling for efficacy endpoints are specified in Section 4.6.

3.5. Timing of Analyses

The primary efficacy analyses may be conducted after the last subject has completed the DB period. The final analyses including OLE data will be conducted after last subject has completed the last OLE visit and study.

4. STATISTICAL METHODOLOGY AND ANALYSES

4.1. Subject Disposition

The subject disposition will be summarized by treatment and overall, and by study periods. All information will be presented in numbers and percentages in a summary table. The following information will be presented in the subject disposition table:

- Randomized subjects

- Safety analysis set
- mITT analysis set
- OLE analysis set
- Subjects who completed DB treatment
- Subjects discontinued DB treatment and the primary reasons for DB treatment discontinuation
- Subjects who completed DB period
- Subjects discontinued DB period and the primary reasons for DB period discontinuation
- Subjects who completed OLE treatment
- Subjects discontinued OLE treatment and the primary reasons for OLE treatment discontinuation
- Subjects who completed OLE
- Subjects discontinued OLE and the primary reasons for OLE discontinuation
- Duration of DB period in weeks: (DB period end date – randomization date +1)/7
- Duration of OLE period in weeks: (OLE period end date – DB period end date)/7
- Time on study in weeks: [max (DB period end date, OLE period end date) – randomization date+1]/7

4.2. Demographics and Baseline Characteristics

The demographics and baseline characteristics data will be summarized for the mITT and safety analysis sets. Descriptive statistics will be provided for continual variables. Numbers and percentages of subjects will be tabulated for categorical variables.

4.2.1. Demographics

The following demographic variables will be summarized:

Age (years), calculated as the number of years between the date of birth and the date of signing the Informed Consent form.

- Age (years)
- Sex (Male or Female), Childbearing potential status for female.
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Height (cm)
- Weight (kg)

- BMI (kg/m^2) = Weight(kg)/Height$(m)^2$

4.2.2. Baseline and Disease Characteristics

The following baseline characteristics will be summarized.

- Years since first PN nodules (calculated relative to first dose date)
- Years since diagnosis (calculated relative to first dose date)
- Elevated IgE: High (>200 IU/mL), normal (≤200 IU/mL)
- Co-morbidities includes, but not limited to, history of atopy, anxiety or depression, allergy, dermatologic conditions, lipid disorders, hypertension, type 2 diabetes, and chronic renal disease.
- Smoking status
- Treated with KPL-716 in previous KPL-716 studies (yes, no)

4.2.3. Medical History

The medical history is coded using Medical Dictionary for Regulatory Activities (MedDRA) of the version at database lock. The summary of medical history will be presented with number (%) by System Organ Class (SOC) and Preferred Term (PT). Each subject will be counted only once for each PT within a SOC per cohort. Similarly, for determination of MedDRA SOC incidences, subjects who experience multiple medical conditions under the same SOC will be counted only once for that SOC.

4.3. Prior and Concomitant Medications and procedures

4.3.1. Prior and Concomitant Medications

All medications will be coded using the World Health Organization (WHO) Drug Dictionary of the version at database lock.

Prior medications are defined as medications that started before the first dose of study drug. Concomitant medications are defined as medications that (1) started before the first dose of study drug and continued into the treatment period, or (2) started on or after the date of the first dose of study drug. The number (%) of subjects who took prior and concomitant medications will be summarized on the anatomical class (ATC level 3) and PT by treatment and overall subjects. For subjects who enter OLE, concomitant medications which start on or before OLE 1^{st} dose will be summarized into DB period, otherwise into OLE.

Prior Topical corticosteroid (TCS) regardless of indication, systemic corticosteroid, intralesional corticosteroid, and oral antihistamine will be summarized as well, of which identification will be based on medical review.

4.3.2. Prior and Concomitant Procedures

The procedures are coded using Medical Dictionary for Regulatory Activities (MedDRA) of the version at database lock. The prior and concomitant procedures are defined similarly as the prior and concomitant medications. The summary of prior and concomitant procedures will be

presented with number (%) by SOC and PT. Each subject will be counted only once for each preferred term within a SOC. Similarly, for determination of MedDRA SOC incidences, subjects who experience multiple procedures under the same SOC will be counted only once for that SOC.

4.4. Protocol Deviations

A protocol deviation is defined as a change, divergence, or departure from the approved study design or procedures defined in the protocol. Important protocol deviations are a subset of protocol deviations that may significantly impact the completeness, accuracy, and/or reliability of study data or that may affect the subject's rights, safety, or well-being.

Current ICH GCP guidelines request that the important protocol deviations must be listed in the clinical study report. These may include, but are not limited to:

- Subjects that are dosed on the study despite not satisfying the inclusion criteria.
- Subjects that develop withdrawal criteria whilst on the study but are not withdrawn.
- Subjects that receive the wrong treatment or an incorrect dose.

A summary table will be provided as the number (%) of subjects with at least one important protocol deviation and the number (%) of subjects in each category by study periods. Protocol deviations will be summarized in DB and OLE periods.

4.5. Treatment Exposure and Compliance

Descriptive statistics will be provided for the following.

- Duration of treatment in DB period in weeks:
  - For subjects who have NOT been enrolled in OLE period: [min (DB last dose date +27, DB period end date,) – DB first dose date+1]/7
  - For subjects who have been enrolled in OLE Period: [OLE first dosedate – DB first dose date]/7.
- Duration of treatment in OLE period in weeks:
  - [min (OLE last dose date +13, OLE period end date,) – OLE first dosedate + 1]/7.
- Duration of treatment over study in weeks:
  - For subjects who have NOT been enrolled in OLE period: [min (last DB last dose date +27, DB period end date) – DB first dose date +1]/7
  - For subjects who have been enrolled in OLE Period: [min (OLE last dose date +13, OLE period end date) – DB first dose date +1]/7.
- Number of doses received per subject will be summarized as 1 dose, 2 doses, 3 doses, 4 doses for DB period, and up to 17 doses for OLE period.
- Treatment compliance is defined as the number of administrations that the subject is compliant divided by the total number of administrations that the subject is planned to take from the first administration of study medications up to the actual last

administration of study medications. Note that a given administration will be considered noncompliant if the subject does not take the planned dose as required by the protocol. No imputation will be made for subjects with missing or incomplete data. Note that the treatment compliance may be summarized in DB and OLE periods separately.

All study drug exposure data will be listed including reasons for dose not administered.

4.6. Efficacy Analyses

All efficacy analyses will be performed in the mITT analysis set. Actual stratification factors based on EDC data will be used for efficacy analyses. The statistical analyses of the primary and key secondary endpoints will be repeated in PP analysis set. Analysis based on the mITT analysis set will be the primary analysis and the analyses based on other analysis sets will be considered as the supportive analyses.

Descriptive statistics will be provided for OLE data analyses in observed cases by all randomized subjects who enter OLE period as well as the DB treatment group, unless otherwise specified.

4.6.1. Multiplicity Adjustment

As this is a dose finding study, there is no multiplicity adjustment planned for the comparisons of multiple doses of KPL-716 versus placebo.

4.6.2. Analyses of Primary Efficacy Endpoint

The primary analysis for the efficacy endpoints will be based on the mITT analysis set.

The following null hypothesis (H0) and alternative hypothesis (H1) will be tested for each KPL-716 dose against placebo:

- H0: No treatment difference between KPL-716 and placebo.
- H1: There is a treatment difference between KPL-716 and placebo.

The primary estimand for the primary endpoint is defined in Table 1, using. the treatment policy/hypothetical strategies. Off-study treatment data up to Week 16 will be included in the analysis. Data will be set to missing values after prohibited/rescue medication that impact efficacy analyses prior to Week 16, and the subject's worst post baseline observation before the time of the medication usage will be carried forward (WOCF) to impute missing endpoint value (for subjects whose postbaseline values are all missing, the baseline will be used to impute). Missing data at Week 16 will be imputed with Multiple Imputation (MI) method. This MI method will use subjects excluding those who used prohibited/rescue medications that impact efficacy analyses prior to Week 16 as input. See Appendix 7.1 for the sample SAS code for the imputation.

Each of the imputed complete data will be analyzed by fitting an ANCOVA model with treatment and randomization stratification factors of sex and years since the first nodule observed (i.e., ≥10 vs. <10 years) as fixed effect factors, baseline as covariate. The imputation number will be about 40. Statistical inference obtained from all imputed data will be combined using Rubin's rule. Descriptive statistics including number of subjects, mean, standard error, and least squares

(LS) mean percent changes (and standard error) will be provided. In addition, difference of each KPL 716 dose group against placebo in LS means and the corresponding 95% confidence intervals (CI) will be provided along with the p-values

Sensitivity analyses

The sensitivity analyses will be conducted for the same estimand for the primary endpoint. There are two sensitivity analyses planned for the primary efficacy endpoint to handle missing data at Week 16 after the WOCF method is applied for the subjects who use prohibited/rescue medications that impact efficacy.

- Sensitivity 1 (WOCF/LOCF): Missing data at Week 16 will be imputed with last observation carried forward (LOCF) method.
- Sensitivity 2 (WOCF/Observed): Missing data at Week 16 will not be imputed (observed) .

Same as the primary analysis, ANCOVA models will be used for the sensitivity analyses.

4.6.3. Analyses of Key Secondary Efficacy Endpoints

For the primary estimand for key secondary efficacy endpoints defined in Table 1, off-study treatment data up to Week 16 will be included in the analysis, and subjects will be considered as non-responders after usage of prohibited/rescue medications that impact efficacy analyses. Additionally subjects with missing data at Week 16 will be defined as non-responders. The analyses will use CMH test adjusted by randomization stratification variables: sex and years since the first nodule observed (i.e., ≥10 years vs. <10 years). Response rate difference adjusted to stratification factors between each KPL-716 dose group and placebo will be derived. In addition, odds ratio and the corresponding 95% confidence intervals (CI) will be provided along with the p-values.

4.6.4. Analysis of Other Efficacy Endpoints

The primary estimand for continuous endpoints of change and percent change from baseline specified as secondary and exploratory endpoints will be defined and analyzed as the primary estimand for primary efficacy endpoints. If some extreme small values are observed at baseline for a parameter, percent changes will not be summarized for that parameter due to inflated percent change. For example, preliminary data showed values of 0 for some subjects in sleep loss VAS.

The primary estimand for binary responder endpoints over time including week 16 will be defined and analyzed as key secondary efficacy endpoints above.

4.6.5. Subgroup Analysis

The primary and key secondary efficacy endpoints may be summarized by subgroups listed in Table 8 based on the mITT analysis set.

Table 8: Subgroups Defined at Baseline

Subgroup	Categories
Age group	<65 years, ≥65 years

Subgroup	Categories
Sex	Female, Male
Race	White, Black, Other
Ethnicity	Hispanic or Latino, Not Hispanic or Latino
Years since the first nodule observed	≥10 years, <10 years
Elevated IgE at Baseline	High (>200 IU/mL), normal (≤200 IU/mL)
Anxiety or depression at baseline	Yes, No

4.6.6. Rescue medications and prohibited medications

Rescue medications are allowed under specific guidance as specified in Pharmacy Manual Version 3.0, Dated 06 October 2021 (Table 9).

Table 9: Rescue medications listed in pharmacy manual

Tier 1	Antihistamines, oral: • Diphenhydramine • Cetirizine
Tier 2	• Hydrocortisone ointments, up to 2.5% • Desonide cream, lotion 0.05%
Tier 3	Non-Europe countries • Triamcinolone acetonide, 0.1% • Fluocinolone Acetonide, 0.025% Europe countries • Fluticasone propionate cream/ointment, up to 0.05%
Tier 4	Clobetasol propionate, 0.05%

For analysis purpose, TCS, oral/IV antihistamines, topical antihistamine, corticosteroid intralesional injection, and/or systemic corticosteroids, which were used to reduce pruritus or treat PN skin lesion on and after first dose date and have significant impact on efficacy interpretation, are defined as rescue medications. TCS which are deemed to have minimum impact on efficacy interpretation are not considered as rescue medications. For example, TCS used for injection site reactions which are not overlapping with the subject's PN lesions are not considered as rescue medication. Antihistamines targeting H2 receptor are not considered as rescue medications due to no impact on pruritis. Identification of rescue medications will be based on medical review in a blinded way before database lock.

Additionally, the protocol specified a list of medications which are prohibited from designated timepoints before Day 1 and throughout the study. Prohibited medication which impact on efficacy analyses will be identified based on blind medical review before database lock as well.

Number of subjects who used rescue/prohibited medications and days on rescue/prohibited medications will be summarized by study periods. Rescue/prohibited medications which overlap with both DB and OLE periods will be summarized based on proportion of days in each period. Use of rescue/prohibited medications over time will be summarized by weeks which will be defined consistently as weeks of WI-NRS. Time to 1st use of rescue/prohibited medications will be summarized based entire study period.

4.7. Safety Analyses

Safety analyses will consist of data summaries for clinical and laboratory parameters, vital sign, and for adverse events (AEs). Adverse events will be coded using MedDRA of the version at database lock. Laboratory parameters, vital signs, and ECG data will be summarized as descriptive statistics by treatment, visit, and study period. Safety summaries will be based on the safety analysis set.

4.7.1. Adverse Events

Adverse events will be mapped to PT and SOC using the most up to date MedDRA of the version at database lock. A treatment-emergent AE (TEAE) is defined as any event not present before exposure to study drug or any event already present that worsens in either intensity or frequency after exposure to study drug. Furthermore, if an AE cannot be determined as treatment-emergent due to incomplete/missing date, conservatively, it will be considered as treatment emergent and included in the summary tables. For subjects who enter OLE, TEAEs which occur (based on AE start dates) before OLE first dose date will be summarized into DB period, otherwise into OLE period. If AE occurring time is available, time should be compared to OLE first dose date and time for AE period derivation.

The following TEAE summaries will be summarized as an overview.

- Any TEAEs
- Drug-related TEAEs
- TEAE by maximum severity (mild, moderate, severe)
- Serious TEAEs (SAEs)
- Drug related serious TEAEs
- TEAEs leading to dose interruption
- TEAEs leading to treatment discontinuation
- TEAEs leading to study discontinuation
- Death
- TEAE of special interest

4.7.2. Clinical Laboratory Tests

Blood and urine samples will be collected for clinical laboratory evaluations (including clinical chemistry, hematology, urinalysis, and serology) at the time points specified in Protocol

Appendix 1. The data collected in different units were converted to SI units (the International System of Units) for summary.

The following summaries by study periods will be provided by laboratory category of chemistry, hematology, and urinalysis.

- Numeric laboratory values will be summarized as descriptive statistics for both actual value and change from baseline (post baseline minus baseline) by week. If lab values are recorded as <xx, the limits of xx will be used for summary.
- Shift table summarizing subject incidence of laboratory normal range (Low, Normal, High) at baseline contrasted with lowest or highest post baseline. Parameters with bidirectional abnormality (low and high) will be presented by post baseline lowest and highest.

All laboratory data will be listed.

4.7.3. Vital Signs

All vital signs including weight, pulse rate, body temperature, respiration rate, and systolic and diastolic blood pressure will be summarized. Descriptive statistics will be presented for the observed value and the change from baseline by treatment and visit if applicable. For repeated assessments at a visit, the average of these values will be taken for the summary at the visit.

4.7.4. Electrocardiograms

Numeric 12-lead ECG parameters will be summarized. Descriptive statistics will be presented for the observed value and the change from baseline at each scheduled visit. Incidence of a normal to abnormal shift for overall ECG interpretation from baseline to worst post baseline will also be summarized by periods. Findings of other categorical parameters will be listed. For repeated assessment, if interpretations are the same, the last value will be taken for numerical assessment for summaries. If interpretations are different, the last value with the abnormal interpretation will be used for summary.

QT interval corrected for heart rate using Fridericia's method - Below incidence will be summarized:

- QTcF interval >500 msec
- QTcF change from the baseline (pre-dose) is >60 msec

4.7.5. Physical Examination

Clinically significant new or worsened physical examination abnormalities following treatment will be recorded as AEs and will be reflected in the summary of AEs.

4.7.6. COVID-19 infection

If there are new COVID-19 infections during the trial, their incidence and severity will be analyzed per treatment group to explore any potential dose-relationship of COVID-19 susceptibility, and the impact of COVID-19 infection on efficacy and safety as well. Gender difference on COVID-19 infection may be explored if a significant number of events are observed.

If a COVID-19 outbreak occurs at a particular investigational site, a sensitivity analysis may be conducted to evaluate the impact on trial results.

4.8. Immunogenicity Analysis

Pre-dose anti-drug antibody (ADA) blood samples will be collected according to schedule in protocol appendix 1.

ADA status will be categorized as follows.

1. Treatment-emergent ADA, including treatment induced or boosted [increased by one quartile (with all titers collected) from baseline]
2. Prior existing ADA (without treatment boosted)
3. ADA negative at every assessment.

For subjects with treatment-emergent ADA, the maximum post-baseline titer level (low, medium, high) will be summarized. The 3 levels (low, medium, high) will be divided based on 25^{th} and 75^{th} percentiles from all ADA titers pooled.

Subjects with Neutralizing antibody will be summarized for subjects with treatment-emergent ADA and non-treatment-emergent prior existing ADA.

A summary table will be provided for number and percent of subjects with positive or negative anti-KPL-716 antibody by visit. An individual's ADA status and titer level (if positive) at each visit will be provided. The number and percent of subjects with anti-KPL-716-antibody positive at any visit vs. those with negative at every visit will also be provided.

Details are in a separate analysis plan and analysis results are also reported in a separate report.

4.9. Pharmacokinetic (PK) Analysis

The PK analysis plan and report will be prepared separately.

4.10. Pharmacodynamic Biomarker Analysis

The pharmacodynamic analysis plan and report will be done separately.

5. INTERIM ANALYSES AND DMC MEETING

There is no interim analysis planned and no DMC meeting scheduled. The primary analyses may be performed when all randomized subjects finish DB period.

6. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

There are no changes from the statistical analyses in protocol.

7. APPENDIX

7.1. Multiple Imputation Sample SAS Code

1. Data after use of rescue/prohibited medications will be set to missing and imputed with worst post baseline scores.
2. Impute intermediate missing data. 40 datasets with a monotone missing pattern will be obtained, induced by Markov Chain Monte Carlo (MCMC) method on subjects who have not taken the prohibited medications and/or rescue medications prior to Week 16. Variables in MCMC modeling includes actual stratification factors: sex, years since first PN diagnosis, treatment, baseline score, and percent change scores of weeks 1-16.

```
proc mi data=dataset seed=201 nimpute=40 out=dat_mc;
   mcmc impute=monotone;
var sex years since_first_diagnosis baseline treatment_group
percent_change_weeks 1-16;
run;
```

3. For each the imputed dataset with monotone missing pattern in Step 2, the remaining missing data will be imputed using the regression method for the monotone pattern with adjustment for covariates including sex, years since first PN diagnosis, treatment, baseline score, and scores of weeks 1-16 as covariates..

```
proc mi data=dat_mc nimpute=1 seed=201 out=dat_mi;
by _imputation_;
class sex years since_first_diagnosis treatment_group;
monotone method=reg;
var sex years since_first_diagnosis baseline treatment_group
percent_change_week 1-16;
run;
```

8. REFERENCES

1. Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. J Biopharm Stat. 2013;23(6):1352-1371.

Signature Page for [redacted]

Approval Task	[redacted] Clinical 12-Aug-2022 13:40:25 GMT+0000

Approval Task	[redacted] Clinical 13-Aug-2022 03:47:53 GMT+0000

Signature Page for [redacted]