Official title: A Long-Term extension study to demonstrate safety of Tildrakizumab in subjects with Psoriatic Arthritis who have previously completed study with Tildrakizumab.

Document: Statistical Analysis Plan

NCT number: NCT03552276

Document date: 11-Aug-2020

[redacted]

Statistical Analysis Plan (SAP)

Protocol Title:	A Multiple-Dose, Long-Term Extension Study to Demonstrate the Safety and Efficacy of Tildrakizumab in Subjects with Psoriatic Arthritis [redacted] who have previously completed studies with Tildrakizumab
Protocol Number:	CLR_18_07
Protocol Version, Date	V1.0 Final, 15 March 2018 Amendment 1, 22 July 2020
[redacted]	[redacted]
Document Version, Date:	Final v1.0, 11 Aug 2020

[redacted]

On behalf of:

Sun Pharma Global FZE

[redacted]

[redacted]

Statistical Analysis Plan (SAP)

[redacted]

[redacted]

SIGNATURE PAGE

[redacted]

REVISION HISTORY

Not Applicable.

[redacted]

TABLE OF CONTENTS

[redacted]

[redacted]

LIST OF ABBREVIATIONS

The following abbreviations will be used within this SAP.

Abbreviation or special term	Explanation
ACR	American College of Rheumatology
ADA	Anti-Drug Antibodies
AE	Adverse Event
AESI	Adverse Events of Special Interest
ASaT	All Subjects as Treated
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Curve
BMI	Body Mass Index
BSA	Body Surface Area
CASPAR	Classification of Psoriatic Arthritis
CI	Confidence Interval
cm	Centimetre
CRF	Case Report Form
CRP	C-reactive protein
CSR	Clinical Study Report
DAS	Disease Activity Score
DBL	Database Lock
DMC	Data Monitoring Committee
ECG	Electrocardiogram
ECI	Events of Clinical Interest
eCRF	Electronic Case Report Form
HAQ-DI	Health Assessment Questionnaire Disability Index
hsCRP	high sensitivity C-reactive protein
IA	Interim Analysis
IMP	Investigational Medicinal Product
IVRS	Interactive Voice Response System
kg	Kilogram
MACE	Major Adverse Cardiac Events

[redacted]

MedDRA	Medical Dictionary for Regulatory Activities
MDA	Minimal Disease Activity
n	Number of non-missing observations
PGA	Physician Global Assessment
PK	Pharmacokinetics
PsA	Psoriatic Arthritis
PT	Preferred Term
PtGA	Patient Global Assessment
Q1	Lower Quartile
Q3	Upper Quartile
q4wk	Every 4 weeks
q12wk	Every 12 weeks
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SD	Standard Deviation
SF-36	36-item Short Form Health Status Questionnaire
SI	Standard International
SOC	System Organ Class
SOP	Standard Operating Procedure
TEAE	Treatment Emergent Adverse Event
TLFs	Tables, Listings and Figures
ULN	Upper Limit of Normal
VAS	Visual Analog Scale

[redacted]

1 INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to provide a detailed description of the statistical methods, data derivations and data presentations to be employed for study protocol CLR_18_07 "A Multiple-Dose, Long-Term Extension Study to Demonstrate the Safety and Efficacy of Tildrakizumab in Subjects with Psoriatic Arthritis and [redacted] or [redacted] who have previously completed studies with Tildrakizumab" which was originally issued on 15 Mar 2018. The protocol was amended on 22 July 2020 with the title "A Long-Term Extension Study to Demonstrate Safety of Tildrakizumab in Subjects with Psoriatic Arthritis Who Have Previously Completed Study with Tildrakizumab".

Any deviations from this SAP will be described in the Clinical Study Report (CSR).

This SAP supersedes the statistical considerations identified in Protocols for CLR_18_07 and where considerations are substantially different, they will be identified as such in this document.

This SAP has been developed and approved prior to a data snapshot and unblinding of the clinical database for Protocol CLR_18_07.

[redacted]

2 STUDY OBJECTIVES

2.1 Primary objectives

Primary Safety Objectives

To assess the long-term safety of tildrakizumab when administered to PsA subjects by evaluation of:

- Incidence and intensity of all AEs,
- Changes in vital signs, laboratory assessments, electrocardiograms (ECGs), and Columbia-Suicide Severity Rating Scale (C-SSRS),
- Immunogenicity of multiple-dose administration of tildrakizumab in these subjects.

Primary Efficacy Objectives

To assess the long-term efficacy of multiple-dose administration of tildrakizumab in subjects with PsA by evaluation of:

- PsA subjects: The proportion of subjects achieving a 20% reduction from Baseline in American College of Rheumatology response criteria (ACR20) at measured time points.

2.2 Secondary objectives

To evaluate long-term treatment outcomes of tildrakizumab in adults with PsA by evaluation of:

- ACR20, ACR50, ACR70, the components of ACR, and 36-item Short Form (SF-36) at other measured time points.

2.3 Exploratory objectives

- [redacted]

The above objectives were in the original protocol. In the amended protocol, efficacy objectives were removed as safety is the main focus after subjects switch to 100 mg tildrakizumab q12wk in an open-label fashion.

[redacted]

3 STUDY DESIGN

3.1 General study design

This is a long-term extension study of tildrakizumab in subjects with who have previously completed the treatment period of studies with tildrakizumab. Eligible PsA subjects who have previously completed the treatment period of Phase 2b study CLR_16_23 (hereafter referred to as the parent Phase 2 study) will be entered into this long-term extension study. Other subjects, including those completing the treatment period of Phase 3 PsA tildrakizumab studies, may also be eligible to enter this long-term extension study.

Up to 286 subjects with PsA could be enrolled from the parent Phase 2 study. Additional subjects may be enrolled from other parent studies. No randomization will occur for this long-term extension study. Treatment allocation will be based upon the dose regimen assigned at the end of the treatment period in the parent study, with all subjects receiving tildrakizumab in the long-term extension study. The eligible subjects will continue to be administered one of the following dose regimens – 200 mg Q4W, 200 mg Q12W or 100 mg Q12W to maintain the blind of the ongoing parent study to at least Week 52 of the long-term extension study. Thereafter, all subjects began migrating to receive 100 mg tildrakizumab SC injection Q12 weeks in an open-label fashion.

Note: eligibility for the long-term extension study will be based on response criteria at Week 52 [using Week 48 laboratory data from the parent study where relevant for calculations] regardless of timing of subject entry. Subject eligibility for continuation to the long-term extension study based on the clinical response criterion will be determined by the interactive voice response service (IVRS).

The parent Phase 2 study will retain a blind status at the time of the initiation of the extension study. One randomized dose group in the PsA study will be receiving tildrakizumab 200 mg q4 weeks at the end of the parent Phase 2 study. In order to allow continued q4 week dosing in the extension study, and to maintain the blinding status of the parent Phase 2 study (which will remain ongoing at initiation of first subject entered into the extension study), all subjects entered to this extension study will additionally attend interim visits q4 weeks for IMP administration (tildrakizumab 200 mg or dummy placebo). The interim dosing visits will also be used to record AEs and concomitant medications and any other evaluations as deemed necessary by the Investigator to ensure subject safety.

All scheduled visits will occur at the study site. However, should unusual circumstances such as public health emergency (i.e. pandemic with an infectious agent) or natural disasters arise, video conference or teleconference between the subject and site investigator will be permitted as deemed appropriate by the Sponsor provided no physical examination is needed for the visit. In this unusual circumstance, self-injection of study medication made by the subject at home is acceptable.

A decision to remove 1 or more treatment arms from the long-term extension study may occur as information from the parent study becomes available. Subjects already enrolled in the long-term extension study may be switched to selected dose regimens for continued evaluation at that time. This decision is anticipated to occur approximately 1 year after the start of the long-term study.

When all subjects have completed the PsA parent Phase 2 study, the blind will no longer need to be maintained for that cohort, and this long-term extension study will have the option to become open-label for those subjects. At that time, the dummy placebo administration will be removed. Note: the timing of switch to an open-label design may differ across the cohorts of subjects entered, dependent on the parent study. If other studies are also included in the long-term extension study, similar considerations for maintenance of blind will be necessary. Those subjects will be entered as separate cohorts and procedures such as continued double-dummy dosing regimens will be utilized, even after other parent Phase 2 cohorts have the option to switch to open-label.

In PsA Phase 2 study [CLR_16_23 PsA study], both tildrakizumab, 200 mg q12 weeks and tildrakizumab 100 mg, q12 weeks provided similar efficacies in the ACR20, ACR50, and ACR70 at Week 24. The lowest dose of 20 mg q12 weeks, did not demonstrate superiority over placebo in ACR70 and had numerically lowest ACR50 response rate among the active doses tested. The highest dose of tildrakizumab 200 mg q4 weeks, did not provide significant benefit over tildrakizumab 200 mg, q12 weeks. Therefore, the dose regimen of tildrakizumab 100 mg, q12 weeks is chosen for the long-term extension study going forward.

After all subjects switch to receive 100 mg tildrakizumab Q12W in an open-label fashion, there will be no efficacy assessments and the primary focus is the long-term safety of tildrakizumab.

[redacted]

[redacted]

[redacted]

1 Subjects from the parent Phase 2 study who do not meet inclusion/exclusion criteria at Wk 52 of the parent study (using laboratory data from Wk 48 for efficacy response criteria), or who in the opinion of the Investigator would not benefit from continued treatment with tildrakizumab, will enter the 20-Wk wash-out phase in accordance with the parent protocol (Not shown). All eligible subjects will enter the long-term extension study, commencing at Wk 52 of the parent study (B/L) and receive tildrakizumab 100 mg Q12 wks regimen (Week 0). Subjects who completed the treatment period of their parent study and entered the wash-out phase prior to study site activation of the long-term extension protocol will also be eligible for inclusion in the long-term study (when available) provided they meet eligibility criteria.

2 B/L = Wk 0 of the long-term extension study, and will be the same as Wk 52 of the parent study.

3 The long-term extension study will remain double-blind for all subjects from the parent Phase 2 PsA study until such point that the PsA study is complete with database locked. At that time, the long-term study will have the option to become open-label for subjects entering from Phase 2 parent study. Note the timing of the switch to open-label may differ across the cohorts of subjects, dependent on the parent study. The IMP administration will continue Q12 weeks, while the safety assessments will be Q12 weeks in the first 2 years, followed by Q24 weeks in the last 2 years.

The study was originally planned to include subjects from both parent studies of Psoriatic Arthritis [redacted]

[redacted]

3.2 Randomization and blinding

3.2.1 Randomization

No randomization will occur for this long-term extension study. Treatment allocation will be based upon the dose regimen assigned at the end of the treatment period in the parent study, with all subjects receiving tildrakizumab in the long-term extension study.

3.2.2 Blinding

The parent Phase 2 studies will retain a blind status at the time of the initiation of the extension study. When all subjects have completed the PsA parent Phase 2 study, this long-term extension study will have the option to become open-label for those subjects. Open label dosing will proceed only after all subjects have completed at least 52 weeks of blinded dosing in the long-term extension study. If other studies are also included in the long-term extension study, similar considerations for maintenance of blind will be necessary. Those subjects will be entered as separate cohorts and procedures such as continued double-dummy dosing regimens will be utilized, even after other parent Phase 2 cohorts have the option to switch to open-label.

PK and anti-drug antibodies (ADA) data will be kept confidential before unblinding.

3.2.3 Unblinding

In an emergency, in which the Investigator must know a subject's treatment allocation to ensure the subject's safety, the Investigator will contact IVRS. When the Investigator contacts the system to break a treatment code for a subject, he/she must provide the requested subject identifying information and confirm the necessity to break the treatment code for the subject. The Investigator will then receive details of the IMP for the specified subject. The system will automatically inform the [redacted] Site Monitor, the medical monitor, and the [redacted] Project Manager that the code has been broken, but no treatment assignment will be communicated.

3.3 Study treatments and assessments

The maximum study duration from Baseline assessment to end of study is 208 weeks.

All subjects eligible for study participation will enter the extension study and continue to receive one of the 3 treatments as in the parent study. Study drug will be supplied as [redacted] and a subject at a visit is expected to receive two syringes packed in a kit. Below are the details:

- 200 mg q4 weeks arm: subjects will receive tildrakizumab 200 mg (two 1-mL injections of 100 mg/mL) at Baseline, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, and all subsequent 4-weekly time points up to Week 200

[redacted]

- 200 mg q12 weeks arm: subjects will receive tildrakizumab 200 mg (two 1-mL injections of 100 mg/mL) at Weeks 8, 20, 32, 44, 56, and all subsequent 12-weekly time points to Week 200, as well as placebo (two 1-mL injections) at interim 4-weekly time points commencing at Baseline of the Phase 2 parent study up to Week 196
- 100 mg q12 weeks arm: subjects will receive tildrakizumab 100 mg (one 1-mL injection of 100 mg/mL + 1 mL placebo) at Weeks 8, 20, 32, 44, 56, and all subsequent 12-weekly time points to Week 200, as well as placebo (two 1-mL injections) at interim 4-weekly time points commencing at Baseline of the Phase 2 parent study up to Week 196

If a subject misses a visit and/or a scheduled dose of IMP, the site must reschedule a visit to ensure the dose of IMP is taken as soon as possible within the visit window. If after 2 attempts to reschedule, the subject still is not able to take the dose, the Sponsor should be contacted to determine if the subject should be discontinued from the study.

Following the last subject's Week 52 visit or early termination prior to Week 52, all subjects will begin migrating to receive 100 mg tildrakizumab SC injection Q12 weeks in an open-label fashion.

A detailed description of procedures and assessments to be conducted during this study is summarized in the Schedule of Assessments in Table 1 below.

[redacted]

Table 1: Schedule of Assessments for PsA Subjects

In the original protocol:

[redacted]

Visit[a]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	0Y BL					1Y 4W				2Y		~2Y 6M		~3Y		~3Y 6M		~4Y	EoT/ FU
Week (± 7D)	0[b]	8	20	32	44	56	68	80	92	104	116	128	140	152	164	176	188	200	208[c]
Written informed consent	X																		
Inclusion/Exclusion criteria	X																		
Demographic information[d]	X																		
Physical examination	X					X				X				X				X	X
Vital signs	X	X	X	X	X	X		X		X		X		X		X		X	X
AEs/Concomitant medications[e]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Medical/Medication history[e]	X																		
ECG	X					X				X				X				X	X
IMP administration[f]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Hematology/chemistry/urinalysis[g]	X					X				X				X				X	X
Lipids[h]	X					X				X				X				X	X
Urine pregnancy test[i]	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
hsCRP and ESR[i]	X	X		X		X				X				X				X	X
Anti-drug antibodies	X					X				X				X				X	X
PK sample[h]	X					X				X				X				X	X

[redacted]

Visit[a]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	0Y BL					1Y 4W				2Y		~2Y 6M		~3Y		~3Y 6M		~4Y	EoT/ FU
Week (± 7D)	0[b]	8	20	32	44	56	68	80	92	104	116	128	140	152	164	176	188	200	208[c]
Tender and Swollen Joint counts	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
PGA of disease activity VAS	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
PsA pain VAS	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
PtGA of disease activity VAS	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
HAQ-DI	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
SF-36	X					X				X				X				X	X
C-SSRS[j]	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X

Abbreviations: AE = adverse event; BL = Baseline; C-SSRS = Columbia-Suicide Severity Rating Scale; D = days; ECG = electrocardiogram; ESR = erythrocyte sedimentation rate; EoT = End of Treatment; FU = Follow-up; HAQ-DI = Health Assessment Questionnaire Disability Index; hsCRP = high sensitivity C-reactive protein; IMP = investigational medicinal product; M = months; PGA = Physician Global Assessment; PK = pharmacokinetic; PsA = psoriatic arthritis; PtGA = Patient Global Assessment; SF-36 = 36-item Short Form; VAS = Visual Analog Scale; W = weeks; Y = year(s).

a Interim visits will occur at 4-week intervals (± 7 days) between the scheduled visits for 12-weekly IMP administration. Collection of AEs and concomitant medications will also be performed, and any other unscheduled safety assessment deemed appropriate in the opinion of the Investigator to ensure subject safety. A decision to remove 1 or more treatment arms from the long-term extension study may occur as information from parent studies becomes available. Subjects already enrolled in the long-term extension study may be switched to selected dose regimens for continued evaluation at that time. This decision is anticipated to occur approximately 1 year after the start of the long-term study. In addition, following completion and database lock of the

[redacted]

parent PsA Phase 2 study, the long-term extension study will have the option to become open-label for these subjects, and interim visits for IMP administration will continue only for subjects receiving a 4--weekly dosing regimen.

b The BL visit of the study Visit 1 (Week 0, Day 1) will be the same as Visit 16, Week 52 of parent PsA study (or later for subjects who entered the wash-out phase from the parent study due to the timing of study site activation of the long-term extension study).

c Subjects who withdraw from IMP at any time will complete the EoT/FU assessments approximately 4 weeks after administration of the last dose of IMP.

d To include body weight only. Other demographic data including date of birth, sex, ethnicity, and subject height will be obtained from the parent study.

e Report all AEs, SAEs, and concomitant medications that occur after signing of the Informed Consent Form for the long-term study. Any occurring prior to signing the Informed Consent Form for the long-term study should be recorded under Medical/Medication history.

f IMP administration will occur on Day 1 of Week 0.

g Blood samples are to be collected (pre-dose where applicable), after ECG and vital sign measurements.

h At visits where lipids are assessed, the blood samples are to be collected (pre-dose where applicable) after 8 hours fasting, following ECG and vital sign measurements.

i Urine pregnancy tests and ESR will be performed at the study site using materials supplied by the central laboratory.

j Evaluations will continue from the parent study using the version with questions referring to ‘since the last visit’. No lifetime version will be repeated at BL

[redacted]

[redacted]

Visit[a]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	1Y BL					2Y				3Y					4Y				EoT/ FU
Week (± 7D)	0[b]	8	20	32	44	56	68	80	92	104	116	128	140	152	164	176	188	200	208[c]
Written informed consent	X																		
Inclusion/Exclusion criteria	X																		
Demographic information[d]	X																		
Physical examination	X					X				X				X				X	X
Vital signs	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
AEs/Concomitant medications[e]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Medical/Medication history[e]	X																		
ECG	X					X				X				X					X
IMP administration[f]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Hematology/chemistry/urinalysis[g]	X		X			X				X				X				X	X
Lipids[h]	X		X			X				X				X				X	X
Urine pregnancy test[i]	X	X	X	X	X	X	X	X	X	X		X		X		X		X	X
hsCRP and ESR[i]	X		X			X				X				X				X	X
Anti-drug antibodies	X		X			X				X				X				X	X[j]
PK sample[h]	X		X			X				X				X				X	X
C-SSRS[k]	X	X	X	X	X	X	X	X	X	X	X	X		X		X		X	X

[redacted]

Abbreviations: AE = adverse event; BL = Baseline; C-SSRS = Columbia-Suicide Severity Rating Scale; D = days; ECG = electrocardiogram; ESR = erythrocyte sedimentation rate; EoT = End of Treatment; FU = Follow-up; hsCRP = high sensitivity C-reactive protein; IMP = investigational medicinal product; M = months; PK = pharmacokinetic; PsA = psoriatic arthritis; W = weeks; Y = year(s).

a Visits will be scheduled every 12 weeks in the first two years, followed by every 24 weeks in the last two years. Collection of AEs and concomitant medications, as well as any unscheduled safety assessment deemed appropriate in the opinion of the Investigator to ensure subject safety, will be performed at each visit. In addition, following completion and database lock of the parent PsA Phase 2 study, the long-term extension study will have the option to become open-label for these subjects.

b The BL visit of the study Visit 1 (Week 0, Day 1) will be the same as Visit 16, Week 52 of parent PsA study (or later for subjects who entered the wash-out phase from the parent study due to the timing of study site activation of the long-term extension study).

c Subjects who withdraw from IMP at any time will complete the EoT/FU assessments approximately 4 weeks after administration of the last dose of IMP.

d To include body weight and height only. Other demographic data including date of birth, sex, and ethnicity will be obtained from the parent study.

e Report all AEs, SAEs, and concomitant medications that occur after signing of the Informed Consent Form for the long-term study. Any occurring prior to signing the Informed Consent Form for the long-term study should be recorded under Medical/Medication history.

f IMP administration will occur on Day 1 of Week 0; followed by subsequent dosing Q12 weeks.

g Blood samples are to be collected (pre-dose where applicable), after ECG and vital sign measurements.

h At visits where lipids are assessed, the blood samples are to be collected (pre-dose where applicable) after 12 hours fasting, following ECG and vital sign measurements.

i Urine pregnancy tests will be performed by using a urine dip stick. ESR will be performed at the site using materials supplied by the central laboratory. hsCRP samples will be sent to the central lab and analyzed by central lab.

j A final anti-drug antibodies sample needs to be collected after 20 weeks from last dose.

k Evaluations will continue from the parent study using the version with questions referring to 'since the last visit'. No lifetime version will be repeated at BL.

[redacted]

4 STUDY ENDPOINTS

4.1 Efficacy endpoint

The efficacy endpoints are:

- The proportion of subjects who achieve ACR20 at measured time points,
- The proportion of subjects who achieve ACR50 at measured time points,
- The proportion of subjects who achieve ACR70 at measured time points,
- Individual components of ACR response at measured time points:
  - tender joint counts (68),
  - swollen joint counts (66),
  - PGA of disease activity (VAS),
  - PtGA of disease activity (VAS),
  - patient's pain assessment (VAS),
  - patient's self-assessed disability (HAQ-DI),
  - acute-phase hsCRP,
  - ESR.
- SF-36 at measured time points.

The ACR20 is defined as at least a 20% improvement from Baseline in both tender joints (68) and swollen joints (66) along with at least a 20% improvements in 3 of 5 other items: 1) the Physician Global Assessment (PGA) of disease activity (as measured using a VAS), 2) the Patient Global Assessment (PtGA) of disease activity (as measured using a VAS), 3) patient pain assessment (as measured using a VAS), 4) patient self-assessed disability (as measured using the HAQ-DI), and 5) acute-phase CRP[1].

Similarly, ACR50 is defined as at least a 50% improvement from Baseline in both tender joints (68) and swollen joints (66) along with at least a 50% improvements in 3 of 5 other items listed above; ACR70 is defined as at least a 70% improvement from Baseline in both tender joints (68) and swollen joints (66) along with at least a 70% improvements in 3 of 5 other items listed above.

Baseline used to derive ACR is the baseline value for each component from parent study.

If the value in any of the components at a time point is missing, the component variables that are not missing will be used to determine the response status. As a general principle, if there are sufficient

[redacted]

non-missing components to determine whether the ACR endpoint is a response or non-response, then ACR endpoint is not missing, else if the available non-missing components are not sufficient to determine the response status of ACR endpoint then it is considered missing.

If the baseline value of any component is equal to 0, the following algorithm will be used in evaluating the percent change from baseline:

- If change from baseline is also equal to 0, then percent change from baseline is set to be 0%;
- If change from baseline is > 0, then percent change from baseline is set to be 9999999%.

These imputed percentages of 0% and 9999999% will be used to derive the ACR endpoints only. Any ACR endpoint that utilizes these imputed values to achieve response will be flagged in by-subject listings.

In addition, any hsCRP value below the LLOQ will be reported as "<0.xxx" in database and will be set to the LLOQ and used in analyses.

HAQ-DI derivation is described in Section 12.1.1

SF-36 is derived per Section 12.1.2

4.2 Safety endpoints

The safety endpoints of this study are:

- Adverse events (AEs)
- Laboratory assessments
- Suicidal ideation and behavior (C-SSRS)
- Vital signs
- ECG
- Physical examination
- ADA to tildrakizumab

[redacted]

5 SAMPLE SIZE AND POWER

The sample size of this study is determined by the number of subjects who rollover from the parent PsA study. No power is calculated because 1) the study is populated with subjects consenting to extended treatment, and 2) the leading primary objective is to assess the long-term safety of tildrakizumab.

6 ANALYSIS POPULATIONS

6.1 All Subjects as Treated (ASaT) Population

The All Subjects as Treated (ASaT) population consists of all subjects who entered the extension study and received at least 1 dose of IMP. The ASaT is the primary population for safety/tolerability, efficacy, and PK analyses. The analysis will be based on the treatment subjects actually received.

7 STATISTICAL CONSIDERATIONS AND ANALYSIS

7.1 Derived variables

Details for deriving efficacy variables are discussed in Sections 4.1, 4.2, and 4.3. The below set of variables are the basic demographic variables and some general variables.

Variables	Formula
Age at informed consent (in years)	year of informed consent – year of birth
Body mass index (BMI;kg/m^2)	weight (kg)/[height (m)]^2
Extension Study day at any visit	(visit date – date of first IMP administration in the extension study) + 1 for visit dates on or after first IMP dosing date in the extension study; (visit date – date of first IMP administration in the extension study) for visit dates before first IMP dosing date in the extension study
Parent Study day at any visit	(visit date – date of first IMP administration in the parent study) +1 for visit dates on or after first IMP dosing date in the parent study; (visit date – date of first IMP administration in the parent study) for visit dates before first IMP dosing date in the parent study
Change from baseline	Post baseline value – Baseline
Percent change from baseline#	[(Post baseline value – Baseline)/Baseline]*100

#: If the baseline value of any component of ACR is equal to 0, the following algorithm will be used in evaluating the percent change from baseline: If change from baseline is also equal to 0, then percent change from baseline is set to be 0%; If change from baseline is > 0, then percent change from baseline is set to be 9999999%. Note: these conventions are only used when deriving ACR20/50/70 and will be described in the ACR listing. Such values will be flagged and explained in the footnote for presentation in the ACR listing.

7.2 Handling of missing data and outliers

7.2.1 Missing data analysis methods

In general, missing values for any of efficacy and safety endpoints will not be imputed when summarizing these endpoints using descriptive statistics.

7.2.2 Handling of missing or incomplete dates

Imputation rules for missing or partial AE start date are defined below:

If the start date has month and year but day is missing, the first day of the month will be imputed

- If this date is earlier than the first dose date, then the first dose date will be used instead.

[redacted]

- If this date is later than the AE stop date (possibly imputed), then the stop date will be used instead.

If the start date has year, but day and month are missing, the 1st of January will be imputed

- If this date is earlier than the first dose date, then the first dose date will be used instead.
- If this date is later than the AE stop date (possibly imputed), then the stop date will be used instead.

If the start date of an event is completely missing, then it is imputed with the first dose date.

Imputation rules for missing or partial AE stop date are defined below:

- If the stop date has month and year but day is missing, the last day of the month will be imputed
- If the stop date has year, but day and month are missing, the 31th of December will be imputed

After the imputation, the imputed dates will be compared against the date of death, if available. If the date is later than the date of death, the date of death will be used as the imputed date instead.

Imputation rules for missing or partial medication start/stop dates are defined below:

Missing or partial medication start date:

- If only Day is missing, use the first day of the month.
- If Day and Month are both missing, use the first day of the year.

Missing or partial medication stop date:

- If only Day is missing, use the last day of the month.
- If Day and Month are both missing, use the last day of the year.
- If Day, Month and year are all missing, assign 'continuing' status to stop date

[redacted]

8 STATISTICAL METHODS

8.1 General Statistical Conventions

All statistical procedures will be completed using SAS version 9.4 or higher.

Continuous variables will be summarized using descriptive statistics, including number of subjects (n), mean, median, standard deviation (SD), minimum and maximum. One additional decimal point for mean and median and 2 additional decimal points for SD will be used.

For categorical variables, summaries will include counts of subjects and percentages. Percentages will be rounded to one decimal place.

Two-sided 95% confidence intervals (CI) will be provided when relevant.

For summary purposes, baseline will be defined as the last available value prior to the date of first IMP administration in parent study.

For reporting purpose, summary tables for efficacy variables and safety variables will be displayed by nominal visit as appropriate. Unscheduled assessments will not be included in the summary tables, but will be in data listings.

All subject data, including those derived, will be presented in individual subject data listings. Unless otherwise stated, unscheduled visit results will be included in date/time chronological order, within patient listings only. All listings will be sorted by treatment group, subject number, date/time and visit. The treatment group as well as patient's sex and age will be stated on each listing. Unless otherwise stated, data listings will be based on ASaT population.

The treatment groups will be displayed in the format below in tables:

- 200 mg Tildrakizumab q4wk
- 200 mg Tildrakizumab q12wk
- 100 mg Tildrakizumab q12wk

8.2 Protocol deviations

All major protocol deviations identified will be summarized by treatment group and overall.

A listing will include the inclusion/exclusion criteria violated at Screening and at Baseline Visits as well as other protocol deviations identified based on data recorded on the eCRF and/or protocol deviation Logs received from [redacted] Medical.

8.3 Subject disposition

The number of subjects in the following categories will be summarized overall and by treatment group:

[redacted]

- Subjects Screened
- Screen Failure(s)
- All Subjects as Treated (ASaT)

The number and percentage (based on ASaT) of subjects in each of the following disposition categories will be summarized overall and by treatment group:

- Completed treatment
- Prematurely discontinued the treatment and the reasons for discontinuation
- Completed study
- Prematurely discontinued the study and the reason for discontinuation

A listing will be provided for subjects who discontinued treatment or who discontinued study with reasons for discontinuation.

8.4 Demographics and baseline characteristics

The demographics and baseline characteristics, medical history, prior and concomitant medications will be summarized by treatment group and overall. Individual subject listings will be provided to support the summary tables.

8.4.1 Demographics

Age (years), height (cm), weight (kg) and other continuous demographic variables at Screening will be summarized descriptively. Gender, primary race, ethnicity and other categorical variables will be summarized.

Year of birth, age, gender, primary race, ethnicity, height, weight, and BMI will be listed as part of demographic listing

8.4.2 Baseline characteristics

Tender joint counts, swollen joint counts, PGA of disease activity, PtGA of disease activity, patient's pain assessment, HAQ-DI, hsCRP, ESR, and SF-36 at baseline will be summarized. Note baseline is the Baseline values from parent study.

8.4.3 Medical history

A summary of medical history will be presented by system organ class (SOC) and preferred term (PT) using the most recently available version of the Medical Dictionary for Regulatory Affairs® (MedDRA). A listing of medical history will be provided as well.

8.4.4 Prior and concomitant medications

Medications used in this study will be coded by using the World Health Organization Drug Dictionary and categorized as follows:

[redacted]

Prior medications and concomitant medications will be summarized descriptively using frequency tables by ATC class and preferred term by treatment group on the Safety Analysis Set. Prior medications are those with a stop date before the first dose of study drug, and concomitant medications are those with a stop date on or after the first dose of study drug or ongoing.

Details for imputing missing or partial start and/or stop dates of medication are described in Section 7.2.2.

8.5 Extent of exposure

8.5.1 Treatment Duration

Duration of study drug exposure (in days) will be calculated as: last dose date – first dose date + 1 day, regardless of study drug interruption.

Study drug exposure will be summarised by treatment group using descriptive statistics.

Exposure to randomized study drug will also be categorised in intervals <=3 months (<91 days), 3-6 months (91-180 days), 6-12 months (181-365 days), 1-2 years (366-730 days), 2-3 years (731-1095 days), >3 years (>1095 days) and summarized by treatment group.

8.5.2 Treatment Compliance

Study drug compliance will be calculated as: 100 × total number of doses administered /total number of doses planned to be administered during the study drug exposure period. If only one syringe among the two is administered at a visit, the number of dose will be counted as 0.5. If both syringes are administered at a visit, the number of dose will be counted as 1. The total number of doses planned will count Weeks 0, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, and all subsequent 4-weekly time points to Week 200. If a subject discontinues treatment, the total number of doses will be counted from Week 0 until the last dose before discontinuation.

Study drug compliance will be summarized by treatment group by the number of subjects (n), mean, SD, median, min, and max. They will also be summarized in categories "<80% compliant" and "≥80% compliant" using frequency tables.

8.6 Efficacy analyses

Responder-type endpoints (ACR20, ACR50, ACR70) will be summarized with response rates and 95% confident intervals (CI) by treatment group and visit.

Observed and change from baseline in continuous endpoints (tender joint counts, swollen joint counts, PGA of disease activity, PtGA of disease activity, patient's pain assessment, HAQ-DI, acute-phase hsCRP, ESR, SF-36) will be summarized by treatment group and visit with n, mean, SD, median, and range. No statistical tests will be performed.

At the start of the long-term extension study, each subject will continue to receive the same dose

[redacted]

regimen received at the end of the parent study. A decision to remove one or more treatment arms from the long-term extension study may occur as information from parent studies becomes available, and subjects in this extension study may be switched to selected dose regimens at that time. The subjects will be summarized based on the actual treatment received at the beginning of the extension study.

All analyses will be performed using ASaT population.

8.7 Safety analyses

Safety analyses will be conducted on the ASaT population and will be performed for all safety variables specified below.

The subjects will be summarized based on the actual treatment received at the beginning of the extension study.

The safety analyses of changes from baseline to a specific time point in safety variables (e.g., laboratory parameters, vital signs, and ECG) will only include subjects from ASaT who have data available for both the baseline and the time point under consideration unless otherwise specified. Where relevant, temporal relationships to parent study baseline will be described.

No statistical tests will be performed.

8.7.1 Adverse events

Any AE that started in the parent study must be reported on the medical history form in this extension study. Only those events that start on/after enrollment into the extension study or are ongoing at the start of the extension study and become worse after or concurrently to extension study enrollment must be reported as AEs.

All AEs will be classified by Primary System Organ Class (SOC) and Preferred Term (PT) according to the most recently available version of the MedDRA dictionary.

In summaries by SOC and PT, adverse events will be sorted within each SOC and PT in alphabetical order. In summaries by PT, AEs will be sorted within each PT in alphabetical order.

Details for imputing missing or partial start dates of adverse events are described in Section 7.2.2.

TEAEs are defined as any AEs occurring or worsening on or after the date of signing of the Informed Consent Form for the extension study. AE summary tables will be presented for TEAEs only and will include the following:

- All TEAEs
- Related TEAEs (AE will be defined as related if causality is possibly, probably, or certainly related)
- TEAEs by maximum severity

[redacted]

- TEAEs leading to study discontinuation
- Serious TEAEs
- TEAEs leading to death.

An overall summary for the categories above will be presented by treatment group and overall.

Due to the various durations of exposure to a treatment, the exposure-adjusted AE incidence rate (per 100 subject-years) will be summarized by treatment group. The incidence of AEs will be presented by system organ class and PT. In addition, all TEAEs will be summarized by SOC, PT and treatment group using frequency counts and percentages (i.e., number and percentage of subjects with an event).

Where a subject has the same adverse event, based on preferred terminology, reported multiple times in the study, the subject will only be counted once at the preferred terminology level in adverse event frequency tables.

Where a subject has multiple adverse events within the same system organ class in the study, the subject will only be counted once at the system organ class level in adverse event frequency tables.

When reporting adverse events by intensity, in addition to providing a summary table based on the event selection criteria detailed above, summary table will also be provided based on the most intense event - independent of relationship to study treatment.

Adverse Events of Special Interest (AESIs)

The events of severe infections, malignancies (including non-melanoma and melanoma skin cancer, excluding carcinoma in situ of the cervix), confirmed Major Adverse Cardiac Events (MACE), and drug-related hypersensitivity reactions will be identified a priori as AESIs for summarizing in this study. Severe infections are defined as any infection meeting the regulatory definition of a SAE, or any infection requiring IV antibiotics whether or not reported as a serious event, as per the regulatory definition. Major Adverse Cardiac Events include non-fatal stroke, non-fatal myocardial infarction and cardiovascular death. All MACE events will be evaluated and adjudicated by a Clinical Adjudication Committee. The AESIs will be summarized by SOC, PT and treatment group.

Adverse Events of Clinical Interest (AECIs)

An AECI is a non-serious AE or occurrence that is designated to be of special interest and must be reported to the Sponsor as though it were an SAE. The AECIs will be summarized by SOC, PT and treatment group.

The following events are considered AECIs for this study:

1. An overdose of the Sponsor's product, as defined as any dose greater than the intended protocol dose. An overdose that is not associated with clinical symptoms or abnormal

[redacted]

laboratory results is to be reported as a non-serious AECI, using the terminology "accidental or intentional overdose without adverse effect."

2. An elevated AST or ALT laboratory value that is ≥ 3 × the ULN and an elevated total bilirubin laboratory value that is ≥ 2 × ULN and, at the same time, an alkaline phosphatase laboratory value that is < 2 × ULN, as determined by way of protocol-specified laboratory testing or unscheduled laboratory testing, is to be reported as a non-serious AECI.
3. Infections that require IV antibiotics but do not meet the definition of an SAE will be designated a closely monitored AE for this study.
4. Depression and suicidal ideation and behavior events.

In addition, a listing containing individual subject adverse event data for TEAEs leading to discontinuation from study, Serious TEAEs, TEAEs leading to death, AESIs, and ECIs will be provided separately. All AEs, including pre- and post-treatment AEs, will be presented in an individual subject data listing.

8.7.2 Clinical laboratory evaluations

For the purposes of summarization in both the tables and listings, all laboratory values will be presented in SI units. If a lab value is reported using a nonnumeric qualifier e.g., less than (<) a certain value, or greater than (>) a certain value, the given numeric value will be used in the summary statistics, ignoring the nonnumeric qualifier.

Clinical laboratory parameters observed values and changes from Baseline will be summarized at each scheduled visit. Unscheduled visits will not be included in the summary table, but will be displayed in the listings.

Values outside the normal range will be categorized as H (above the normal range) or L (below the normal range) based on the laboratory's reference range and these will be flagged in the listings of individual subject data.

8.7.3 Vital signs

Vital signs observed values and changes from Baseline will be summarized at each scheduled visit.

A listing of vital signs by subject will be produced.

8.7.4 Physical examinations

Physical examination results will be summarized with incidence of "Normal" and "Abnormal" by body system at each scheduled visit. All abnormal physical examination results will be listed.

8.7.5 Electrocardiograms

The overall ECG interpretation will be summarized by presenting the number and percentage of subjects with "Normal", "Abnormal, not clinically significant", and "Abnormal, clinically

[redacted]

significant". In addition, shift tables from baseline to post-baseline in overall ECG interpretation will be displayed by visit and treatment group.

ECG parameter (e.g., QTcF) observed values and changes from Baseline will be summarized at each scheduled visit.

An accompanying listing of subjects will be produced and it will display all ECG findings during the study in subjects with abnormal ECGs, as determined by the investigator.

8.7.6 Assessment of Suicidal Ideation and Behavior

The C-SSRS consists of 2 major aspects: Suicidal Ideation and Suicidal Behavior. Based on outcomes and data analyses suggested by the C-SSRS website, the following endpoints will be used to analyze C-SSRS data:

- Presence of Suicidal Ideation: Set to 1 if any ideation is present and 0 otherwise.
- Presence of Suicidal Behavior: Set to 1 if any type of suicidal behavior is present and 0 otherwise.
- Suicidal ideation score: Defined as the maximum suicidal ideation category (1-5) on the C-SSRS present at the assessment. A score of 0 is assigned if no ideation is present.

The number (%) of subjects with presence of suicidal ideation/behavior will be summarized at each assessment time. The C-SSRS data will be summarized using worst-case shift tables. Worst-case shift tables will be the cross tabulation of the Baseline result score with the worst-case result score during the treatment period.

8.8 Other analysis

8.8.1 Analysis of Pharmacokinetic Endpoints

Plasma tildrakizumab concentration data will be listed by individual subject and summarized by time and tildrakizumab dose group. The subjects will be summarized based on the ASaT population with the actual treatment received at the beginning of the extension study.

8.8.2 ADA to Tildrakizumab

For each subject, tildrakizumab serum concentrations and ADA sample results are matched to actual sampling times and treatment. Subjects are grouped based on the actual treatment received. Subjects have baseline samples taken prior to dosing to assess for any preexisting immune response that may be detected by the ADA assays. Subjects are considered positive if at least one pre-treatment or post-dose sample is positive at any time. Positive subjects are subsequently categorized into treatment-emergent positive if the positive sample occurs following treatment with tildrakizumab or non-treatment emergent positive if the subject has an immune response present at baseline and the response is not boosted following treatment.

[redacted]

The presence of tildrakizumab can interfere with the ADA assay at concentrations above the drug tolerance level (DTL) of 6 µg/mL. Therefore, samples with a negative test result in the ADA assay can only be described as negative in the case of a tildrakizumab concentration below 6 µg/mL. The immunogenicity status of a subject is considered to be negative if all pre-treatment and post-dose samples tested negative in the ADA assay and if the concentration of tildrakizumab in the last post-dose sample is below the DTL. Therefore, an integrated evaluation of ADA results and drug serum concentrations is required for interpretation of immunogenicity results.

To summarize, subjects are categorized in one of four immunogenicity categories as described in Table 2. The overall immunogenicity incidence is defined as the proportion of treatment-emergent positive subjects to the number of evaluable subjects. The proportion of non-treatment emergent positive subjects is similarly reported. For ADA positive subjects (based on the confirmatory assay), the immune response is further characterized for antibody titer and neutralizing capacity. When titer is measured to be less than 1, it is considered to be 0.5 for the purposes of deriving ADA subject status.

Table 2 Immunogenicity Subject Status Definitions

Subject Status	Definition
Negative	All pre-treatment and post-dose samples were negative in the ADA assay and the drug concentration in the last post-dose sample was below the respective DTL (6 µg/mL) for the ADA assay.
Inconclusive	All pre-treatment and post-dose samples were negative in the ADA assay AND the drug concentration in the last post-dose sample was equal or above the respective DTL for the ADA assay.
Treatment-emergent Positive	Pre-treatment sample was negative and at least 1 post-dose sample was positive in the ADA assay (treatment-induced positive).
	Pre-treatment and post-dose samples were both positive in the ADA assay and the titer increased post-dose by ≥2-fold (treatment boosted positive).
Non-treatment-emergent Positive	Pre-treatment sample was positive and post-dose samples were negative in the ADA assay.
	Pre-treatment and post-dose samples were positive in the ADA assay with a < 2-fold increase in titer post-dose
ADA = anti-drug antibodies; DTL = drug tolerance level.	

The anti-tildrakizumab immunogenicity status of evaluable subjects, along with titer and neutralizing antibody, will be summarized by dose level. The anti-tildrakizumab immunogenicity status of evaluable subjects will be further summarized by treatment group and by dose level. ADA data will be presented in a listing.

[redacted]

[redacted]

8.9 Interim analysis

An IA to review safety and efficacy data will be performed following the last subject's Week 52 visit or early termination prior to Week 52. Additional IAs (for safety and PK/ADA assessments only) may be performed as needed to support reporting requirements or as needed for internal decision-making. Details of the statistical methodology and operational processes for later IAs will be described in a separate document if different from those described here.

9 REPORTING CONVENTION

This section details the format and layout of all TFLs and statistical output that will be produced in conjunction with the Clinical Study Report. The table of contents and templates for the TFLs will be produced in a separate document.

All data analyses and generation of TLFs will be performed using SAS 9.4® or higher.

The following reporting conventions will be adopted for the presentation of study data:

- The tables and listings will be provided in a Word document in landscape format, Courier New 8 with the following margins: top: 1.50 in, bottom: 1.00 in, left: 1.00 in, right: 1.00 in, header: 1.50 in, footer: 0.50 in. The output alignment will be centered on the page and the titles and footnotes will be center aligned.
- The ICH[2,3] numbering convention will be used for all TLFs.
- The analysis population represented on the tables will be clearly identified in the title of the table.
- Dates will appear as DDMMMYY format; times as HH:MM format (24 hour clock).
- For the presentation of summary data, results will be aligned on the decimal point and be centered within the column. Unless otherwise stated, tables will summarize the results per treatment group and overall.
- All listings will be ordered by investigational site, patient number, date/time and visit. The treatment group as well as patient's sex and age will be stated on each listing. Line break spacing will be added (for example between different patients, visits) to facilitate review of the listings.
- All TLFs will have the SAS program path and name, output filename and date/time of production in the footnote, and will include the following hierarchy of titles and footnotes (as an example):

[redacted]

[redacted]

[redacted]

10 CHANGES TO PLANNED ANALYSIS FROM STUDY PROTOCOL

Not applicable.

11 REFERENCES

1. Felson DT, Anderson JA, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology Preliminary Definition of Improvement in Rheumatoid Arthritis. Arthritis Rheum. 1995; 38(6):727-35.
2. ICH Topic E3: Structure and Content of Clinical Study Reports (CPMP/ICH/137/95-adopted December 1995).
3. ICH Topic E9: Statistical Principles for Clinical Trials (CPMP/ICH/363/96 – adopted March 1998).
4. Bruce, B., Fries, J.F. The health assessment questionnaire (HAQ). Clinical and Experimental Rheumatology; 2005, 23 (Supp. 39), S14-S18.
5. Ware JE. User's Manual for the SF-36v2 Health Survey; 2007.

12 APPENDICES

12.1 Endpoint definition and derivation details

12.1.1 Health Assessment Questionnaire – Disability Index (HAQ-DI)

There are 8 categories assessed by the HAQ-DI[4] 1) dressing and grooming, 2) arising, 3) eating, 4) walking, 5) hygiene, 6) reach, 7) grip, and 8) common daily activities. For each of these categories, patients report the amount of difficulty they have in performing 2 or 3 specific activities. The time frame for the disability questions is the PAST WEEK and each question can be scored as 0 (without any difficulty), 1 (with some difficulty), 2 (with much difficulty) or 3 (unable to do). The use of aids and devices for these activities is also recorded. Use of any device or aid will result in a minimum score of 2 for that category as described below.

Domain	If domain score is either 0 or 1, adjust to 2 when the following is satisfied.
Dressing and grooming	"Devices used for dressing (button hook, zipper pull, long-handled shoe horn, etc.)" or help from others on "Dressing and Grooming" is checked
Arising	"Special or built up chair" or help from others on "Arising" is checked.
Eating	"Built up or special utensils" or help from others on "Eating" is checked.
Walking	"Cane", "Walker", "Crutches", "Wheelchair", or help form others on "Walking" is checked
Hygiene	"Raised toilet seat", "Bathtub bar", "Long-handled appliances in bathroom", "Bathtub seat" or help from others on "Hygiene" is checked.
Reach	"Long-handled appliances for reach" or help from others on "Reach" is checked
Grip	"Jar opener (for jars previously opened)" or help from others on "Gripping or opening things" is checked.
Activities	Help from others on "Errands and chores" is checked.
Note: For "Other, (specify)", whether checked or unchecked or specifying the other "aids or devices" in this category, is not to be used in the adjustment of the domain score.	

For each question in the questionnaire, the level of difficulty is scored from 0 to 3 with

0 =no difficulty
1 = some difficulty
2 = much difficulty
3 = unable to do.

In order to compute score patient must complete 6 of the 8 categories

1. A category score is determined from the highest score of the sub categories, components in that category. If there are 3 sub categories and the patient responds with a score of 2, 0, 1 respectively then the score of the category will be 2.
2. When there are NO aids or devices or help indicated for a category the category score is not modified. But when aids or devices or help indicated score of the category item is raised from a 0 or a 1 to a 2. If patient highest score for a sub category is 3 it remains at 3
3. Sum eight category score obtained as result of step 1 and 2

Divide the sum by number of categories answered (range should be 6-8). This yields a single disability score.

If more than 2 of the categories, or 25%, are missing, the scale is not scored. A higher score indicates greater disability.

12.1.2 36-item Short Form Health Status Questionnaire

The SF-36 v2[5] is a multi-purpose survey that measures 8 domains of health: physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health. It yields scale scores for each of these 8 domains and 2 summary measures of physical and mental health. All domains and summary components are scored such that a higher score indicates a higher functioning or health level.

These 8 domains are as follows:

a. Physical Functioning (PF). This score is based on the responses to the 10 items that compose Question 3 and reflects the degree to which various physical activities have been limited in the previous week by the subject's health.

b. Role-Physical (RP). This score is based on the responses to the four items that compose Question 4 and reflects the relative amount of time that the subject has had problems with work or other regular daily activities as a result of their physical health during the previous week.

[redacted]

c. Bodily Pain (BP). This score is based on the responses to Questions 7 and 8 and reflects bodily pain and its effects on normal work during the previous week.

d. General Health (GH). This score is based on responses to Question 1 and the four items in Question 11 and reflects the subject's perception of their general health during the previous week.

e. Vitality (VT). This score is based on responses to Question 9 items a, e and g, and reflects the subject's physical energy level relative to time during the previous week.

f. Social Functioning (SF). This score is based on responses to Questions 6 and reflects how physical health or emotional problems have interfered with social activities during the previous week.

g. Role-Emotional (RE). This score is based on responses to the three items in Question 5 and reflects the amount of time during the previous week that emotional problems have interfered with work or regular daily activities.

h. Mental Health (MH). This score is based on responses to Question 9 items b, c, d, f, and h and reflects various mental/emotional states relative to time during the previous week.

The summary component scores are:

a. Physical Component Summary (PCS).

b. Mental Component Summary (MCS).

Data Derivation Details to Obtain Scale Scores for SF-36

VARIABLE	DERIVATION
SF-36 PF scale score	raw score = sum (items 3A, 3B, 3C, 3D, 3E, 3F, 3G, 3H, 3I, 3J) PF = (raw score -10) * 5 PF_Z = (PF – 82.62455) / 24.43176 PF scale score = (PF_Z*10) + 50 When calculating the raw score, if 5 or more of the items are non-missing then replace any missing values as follows: Calculate the mean of the answered questions, using the number of answered questions as the denominator. Substitute the mean for any missing scores. Otherwise, if less than 5 of the items are non-missing then PF scale score is missing. The response scale for each activity ranges from 1 to 3 where 1=limited a lot, 2=limited a little, and 3=not limited at all. A higher PF scale score indicates better physical functioning.
SF-36 RP scale score	raw score = sum (items 4A, 4B, 4C, and 4D) RP = [(raw score -4)/16] * 100 RP_Z = (RP – 82.65109) / 26.19282 RP scale score = (RP_Z * 10) + 50 When calculating the raw score, if 2 or more of the items are non-missing then replace any missing values as follows: Calculate the mean of the answered questions, using the number of answered questions as the denominator. Substitute the mean for any missing scores. Otherwise, if less than 2 of the items are non-missing then RP scale score is missing. The response scale for each item ranges from 1 to 5 where 1=All of the time, 2=Most of the time, 3=Some of the time, 4=A little of the time, and 5=None of the time. A higher RP scale score indicates better role-physical functioning.

[redacted]

VARIABLE	DERIVATION
SF-36 BP scale score	raw score = sum (reversed item 7 and reversed item 8) BP = (raw score -2) * 10 BP_Z = (BP – 73.86999) / 24.00884 BP scale score = (BP_Z * 10) + 50 Reverse direction of Item 7 as follows if =1, set to 6 if =2, set to 5.4 if =3, set to 4.2 if=4, set to 3.1 if=5, set to 2.2 if=6, set to 1 Reverse direction of item 8 as follows: if =1 and original value of item 7=1, set to 6 if =1 and original value of item 7>=2, set to 5 if=2, set to 4 if=3, set to 3 if=4, set to 2 if=5, set to 1 If item 7 is answered and item 8 is missing, set 8 = reversed 7 as defined above. If 8 is answered and 7 is missing, set 7 as reverse item 8 as follows if =1, set to 6 if =2, set to 4.75 if=3, set to 3.5 if=4, set to 2.25 if=5, set to 1 If 1 or more questions were answered, calculate BP scale score as defined above. If neither question was answered then BP scale score is missing. The scale for Question 7, amount of bodily pain, ranges from 1 to 6 where 1=None, 2=Very mild, 3=mild, 4=Moderate, 5=Severe, and 6=Very severe. The scale for Question 8, the degree to which pain interfered with normal work, ranges from 1 to 5 where 1=Not at all, 2=A little bit, 3=Moderately, 4=Quite a bit, and 5=Extremely. A higher BP scale score indicates lack of bodily pain.

VARIABLE	DERIVATION
SF-36 GH scale score	raw score = sum (reversed item 1, item 11A, reversed 11B, 11C and reversed 11D) GH = (raw score -5) * 5 GH_Z = (GH – 70.78372) / 21.28902 GH scale score = (GH_Z * 10) + 50 Reverse direction of Item 1 as follows: if =1, set to 5 if =2, set to 4.4 if =3, set to 3.4 if =4, set to 2 if =5, set to 1 Reverse direction of item 11B and 11D by subtracting score from 6. When calculating the raw score, if 3 or more of the items are non-missing then replace any missing values as follows: Calculate the mean of the answered questions, using the number of answered questions as the denominator. Substitute the mean for any missing scores. Otherwise, if less than 3 of the items are non-missing then GH scale score is missing. Responses for Question 1, an assessment of self-perceived health status, range from 1 to 5 where 1=Excellent, 2=Very good, 3=Good, 4=Fair, and 5=Poor. Responses for the items in Question 11 range from 1 to 5 where 1=Definitely true, 2=Mostly true, 3=Don't know, 4=Mostly false, and 5=Definitely false and reflect the subject's perception of their relative health and expectations of their future health status. A higher GH scale score indicates better general health perceptions.

VARIABLE	DERIVATION
SF-36 VT scale score	raw score = sum (reversed item 9a, reversed 9e, 9g and 9i) VT = [(raw score -4)/16] * 100 VT_Z = (VT – 58.41968) / 20.87823 VT scale score = (VT_Z * 10) + 50 Reverse direction of Items 9a and 9e by subtracting score from 6. When calculating the raw score, if 2 or more of the items are non-missing then replace any missing values as follows: Calculate the mean of the answered questions, using the number of answered questions as the denominator. Substitute the mean for any missing scores. Otherwise, if less than 2 of the items are non-missing then VT scale score is missing. The scale for these items ranges from 1 to 5 where 1=All of the time, 2=Most of the time, 3=Some of the time, 4=A little of the time, and 5=None of the time. A higher VT scale score indicates more vitality.
SF-36 SF scale score	raw score = sum (reversed 6 and 10) SF = [(raw score -2) / 8] * 100 SF_Z = (SF – 85.11568) / 23.24464 SF scale score = (SF_Z * 10) + 50 Reverse direction of score for item 6 by subtracting score from 6. When calculating the raw score, if 1 of the items is missing then substitute the missing score with the score on the non-missing item. If both items are missing then SF scale score is missing. Responses to Question 6, an assessment of the extent to which health/emotional problems interfered with social activities, range from 1 to 5 where 1=Not at all, 2=Slightly, 3=Moderately, 4=Quite a bit, and 5=Extremely. Responses to Question 10 reflect the amount of time that health/emotional problems interfered with social activities and range from 1 to 5 where 1=All of the time, 2=Most of the time, 3=Some of the time, 4=A little of the time, and 5=None of the time. A higher SF scale score indicates better social functioning.

[redacted]

VARIABLE	DERIVATION
SF-36 RE scale score	raw score = sum (items 5A, 5B, and 5C) RE = [(raw score -3) / 12] * 100 RE_Z = (RE – 87.50009) / 22.01216 RE scale score = (RE_Z * 10) + 50 When calculating the raw score, if 2 or more of the items are non-missing then replace any missing values as follows: Calculate the mean of the answered questions, using the number of answered questions as the denominator. Substitute the mean for any missing scores. Otherwise, if less than 2 of the items are non-missing then RE scale score is missing. Responses to the items in Question 5 range from 1 to 5 where 1=All of the time, 2=Most of the time, 3=Some of the time, 4=A little of the time, and 5=None of the time. A higher RE scale score indicates better role-emotional functioning.
SF-36 MH scale score	raw score = sum (items 9B, 9C, reversed 9D, 9F and reversed 9H) MH = (raw score - 5) * 5 MH_Z = (MH – 75.76034) / 18.04746 MH scale score = (MH_Z * 10) + 50 Reverse direction of scores for 9D and 9H, by subtracting score from 6. If 3 or more of the items are non-missing then replace any missing values as follows: Calculate the mean of the answered questions, using the number of answered questions as the denominator. Substitute the mean for any missing scores. Otherwise, if less than 3 of the items are non-missing then MH scale score is missing. The scale for these items ranges from 1 to 5 where 1=All of the time, 2=Most of the time, 3=Some of the time, 4=A little of the time, and 5=None of the time. A higher MH scale score indicates better mental health.

[redacted]

VARIABLE	DERIVATION
SF-36 TR scale score	raw score = item 2 TR scale score = raw score The scale for this item ranges from 1 to 5 where 1=Much better now than one week ago, 2=Somewhat better now than one week ago, 3=About the same as one week ago, 4=Somewhat worse now than one week ago, and 5=Much worse now than one week ago. A higher TR scale score indicates worse general health currently relative to one week previous.
SF-36 PCS score	PCS score includes the 8 scales for GH, PF, RP, RE, SF, MH, BP, and VT. PF1= (PF-82.62455)/24.43176; RP1=(RP-82.65109)/26.19282; BP1=(BP-73.86999)/24.00884; GH1 = (GH-70.78372)/21.28902; VT1= (VT-58.41968)/20.87823; SF1=(SF-85.11568)/23.24464; RE1= (RE-87.50009)/22.01216; MH1=(MH-75.76034)/18.04746; Raw Score = ((GH1*.24954)+(PF1* .42402)+(RP1*.35119)+ (RE1*-.19206)+(SF1*-.00753)+(MH1*-.22069)+(BP1*.31754)+ (VT1*.02877)) PCS Summary Scale Score = (raw score *10) + 50 Raw Score is missing if one of the component scale scores is missing.

[redacted]

VARIABLE	DERIVATION
SF-36 MCS score	MCS score includes the 8 scales for GH, PF, RP, RE, SF, MH, BP, and VT. PF1= (PF-82.62455)/24.43176; RP1=(RP-82.65109)/26.19282; BP1=(BP-73.86999)/24.00884; GH1 = (GH-70.78372)/21.28902; VT1= (VT-58.41968)/20.87823; SF1=(SF-85.11568)/23.24464; RE1= (RE-87.50009)/22.01216; MH1=(MH-75.76034)/18.04746; Raw Score =((GH1*-.01571)+(PF1*-.22999)+(RP1*-.12329)+(RE1*.43407)+(SF1*.26876)+(MH1*.48581)+(BP1*-0.09731)+(VT1*.23534)) MCS Summary Concept Score = (raw score *10) + 50 Raw Score is missing if one of the component scale scores is missing.

[redacted]