

Protocol Number: M1095-HS-201

**Official Title: Phase 2, Randomized, Parallel-group, Double-blind, Placebo-controlled Study of
Sonelokimab in Patients With Active Moderate to Severe Hidradenitis Suppurativa**

NCT Number: NCT05322473

Document Date: 23 May 2023

Protocol Title: Phase 2, randomized, parallel-group, double-blind, placebo-controlled study of sonelokimab in patients with active moderate to severe hidradenitis suppurativa

Protocol Number: M1095-HS-201

Protocol Version, Date Final v2.1, 16 Aug 2022

Document Version, Date: FINAL 1.1, 23May2023

Prepared by:

On behalf of:

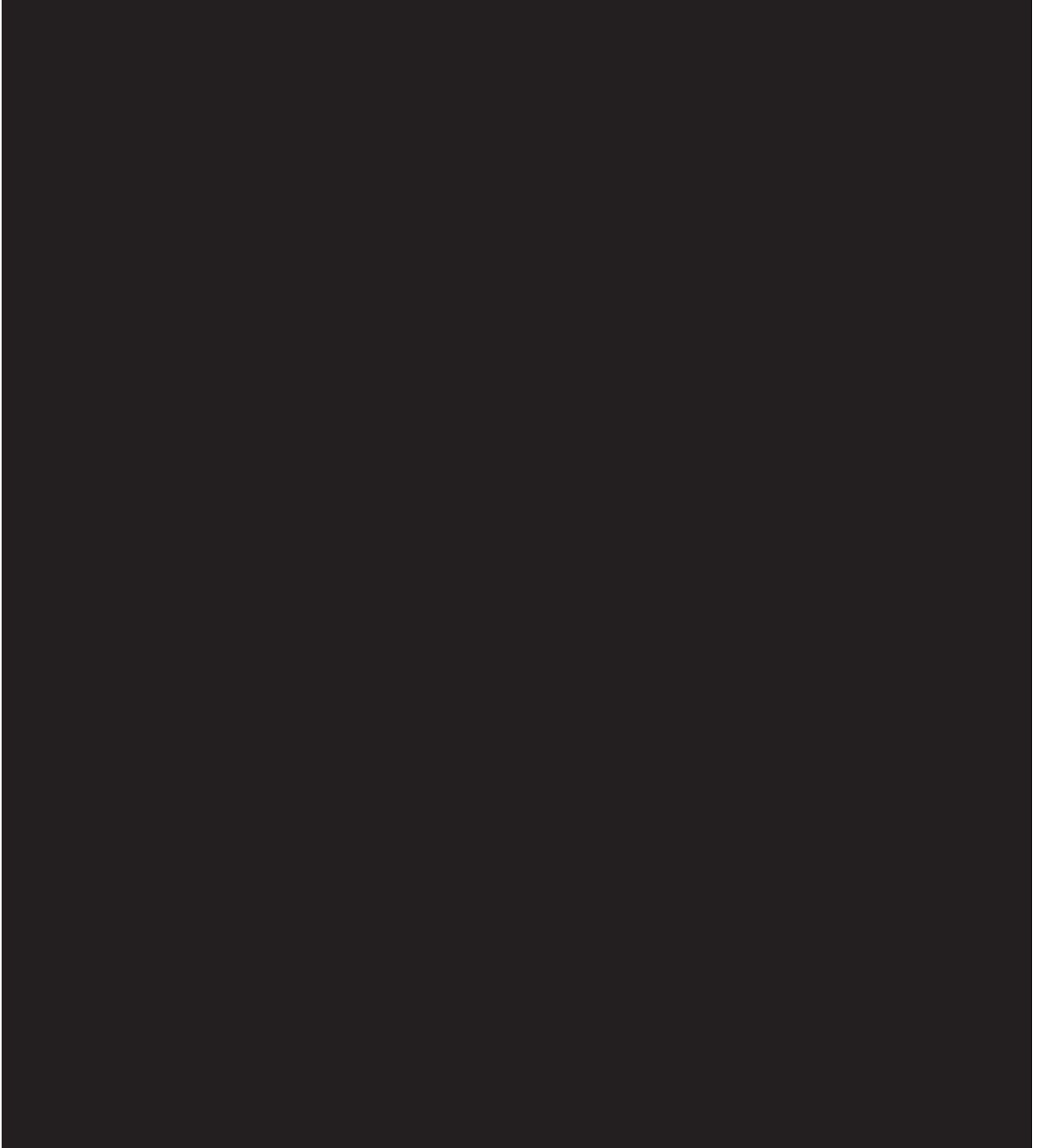
MoonLake Immunotherapeutics AG

Confidentiality statement:

- The information provided in this document is strictly confidential.
- The recipients of the SAP must not disclose the confidential information contained within this document or any related information to other persons without the permission of the sponsor.
- In addition, the recipients of the SAP must keep this confidential document in a controlled environment which prevents unauthorized access to the document.



SIGNATURE PAGE







Statistical Analysis Plan (SAP)



TABLE OF CONTENTS

SIGNATURE PAGE	2
REVISION HISTORY	4
TABLE OF CONTENTS	5
LIST OF ABBREVIATIONS	7
1 INTRODUCTION	9
2 STUDY OBJECTIVES	10
2.1 Primary objective.....	10
2.2 Secondary objectives	10
2.3 Exploratory objective	10
3 STUDY DESIGN	11
3.1 General study design	11
3.2 Randomization and blinding.....	13
3.3 Study treatments and assessments	14
4 STUDY ENDPOINTS.....	20
4.1 Primary efficacy endpoint	20
4.2 Secondary efficacy endpoints	20
4.3 Safety endpoints	22
4.4 Pharmacokinetic (PK) and Immunogenicity (IG) Endpoints	22
4.5 Exploratory endpoints.....	22
4.6 Ultrasound sub-study endpoints	22
5 STUDY ESTIMANDS	23
5.1 Estimand for the primary endpoint.....	23
5.2 Estimand for the key secondary endpoints	24
6 SAMPLE SIZE AND POWER	26
7 ANALYSIS POPULATIONS	27
7.1 Part A.....	27
7.2 Part B	28
7.3 Ultrasound Sub-study Set	28
7.4 Protocol deviations/violations and exclusions from analysis sets	28
8 STATISTICAL CONSIDERATIONS AND ANALYSIS.....	29
8.1 Derived Variables	29
8.2 Visit windows	30
8.3 Handling of missing data and randomization stratification errors	33
9 STATISTICAL METHODS.....	37
9.1 General statistical conventions	37
9.2 Participant disposition	38
9.3 Protocol deviations	39
9.4 Demographics and baseline characteristics	39
9.5 Extent of exposure	41
9.6 Efficacy analyses	41
9.7 PK and Immunological Analyses	63
9.8 Analysis of exploratory endpoint	64
9.9 Analysis of the ultrasound sub-study.....	64

9.10	Safety analyses	66
9.11	Subgroup analysis.....	73
9.12	Interim Analysis	73
9.13	Primary Analysis Timing.....	73
9.14	DSMB.....	73
9.15	DSUR.....	73
10	CHANGES TO PLANNED ANALYSIS FROM STUDY PROTOCOL.....	75
11	REFERENCES	76
12	APPENDICES	77
	Appendix A: Dermatology Life Quality Index.....	77
	Appendix B: Patient Global Assessment of Skin Pain	79
	Appendix C: Patient Global Impression of Severity (PGIS).....	80
	Appendix D: Hidradenitis Suppurativa-Physician's Global Assessment (HS-PGA)	
	81	
	Appendix E: Hidradenitis Suppurativa Quality of Life (HiSQOL) Score.....	82
	Appendix G: Work Productivity and Activity Impairment Questionnaire:	
	Hidradenitis Suppurativa, V2.0 (WPAI: Hidradenitis Suppurativa)	85
	Appendix H: FACIT-Fatigue Scale	87
	Appendix I: Clinical Laboratory Tests	89
	Appendix J: Exposure adjusted incidence rate	90
	Appendix K: Reasons for exclusion from the Per Protocol population	91
	Appendix L: CTCAE v5.0 laboratory grades	92

LIST OF ABBREVIATIONS

The following abbreviations will be used within this SAP.

Abbreviation	Explanation
ADA	Anti-Drug Antibodies
AE	Adverse Event
AESI	Adverse Events of Special Interest
AN	Abscess and inflammatory nodule
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
BP	Blood Pressure
CI	Confidence Interval
cm	Centimetre
COVID-19	Coronavirus disease 2019
CRF	Case Report Form
CRO	Contract Research Organization
CRP	C-Reactive Protein
CSR	Clinical Study Report
CV	Coefficient of Variation
DLQI	Dermatology Life Quality Index
EAIR	Exposure Adjusted Incidence Rates
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EOT	End of Treatment
EQ-5D	EuroQol five-dimension questionnaire
FACIT-F	Functional Assessment of Chronic Illness Therapy-Fatigue
FAS	Full Analysis Set
HiSCR	Hidradenitis Suppurativa Clinical Response
HiSQoL	Hidradenitis Suppurativa Quality of Life
HS	Hidradenitis Suppurativa
HS-PGA	Hidradenitis Suppurativa Physician's Global Assessment
ICH	International Conference on Harmonization
IG	Immunogenicity

IHS4	International Hidradenitis Suppurativa severity score system
IMP	Investigational Medicinal Product
IRT	Interactive Response Technology
kg	Kilogram
KM	Kaplan-Meier
LS	least squares
MAR	Missing at Random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
NLR	Neutrophil to Lymphocyte Ratio
NRS	Numerical Rating Scale
OR	Odds Ratio
PGA	Patient's Global Assessment
PGIS	Patient Global Impression of Severity

PK	Pharmacokinetic
PPS	Per-protocol Set
PT	Preferred Term
SAE	Serious Adverse Event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	Standard Deviation
SE	Standard Error
SOA	Scheduled of Assessments
SOC	System Organ Class
TE	Treatment-Emergent
TEAE	Treatment-Emergent Adverse Event
TFLs	Tables, Figures and Listings
VAS	Visual Analog Scale
WHODDE	World Health Organization Drug Dictionary Enhanced
WPAI	Work Productivity and Activity Impairment

1 INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to provide detailed descriptions of the statistical methods, data derivations and data displays for study protocol M1095-HS-201, Version Final v2.1, “Phase 2, randomized, parallel-group, double-blind, placebo-controlled study of sonelokimab in patients with active moderate to severe hidradenitis suppurativa” dated 16 Aug 2022 for final analysis in support of the Clinical Study Report (CSR), as well as describing the outputs that will be used for the sonelokimab Data Safety Monitoring Board (DSMB) and Development Safety Update Report (DSUR). The table of contents and templates for the Tables, Figures and Listings (TFLs) will be produced in a separate document.

Any deviations from this SAP will be described and justified in the CSR.

The preparation of this SAP has been based on International Conference on Harmonization (ICH) E9 guidelines.

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

2 STUDY OBJECTIVES

2.1 Primary objective

The primary objective of this study is to evaluate the efficacy of sonelokimab at 2 different dose levels (120 mg, 240 mg) compared with placebo in the treatment of participants with active moderate to severe hidradenitis suppurativa (HS).

2.2 Secondary objectives

The secondary objectives of this study are as follows:

- To evaluate the safety and tolerability of sonelokimab at 2 different dose levels (120 mg, 240 mg) compared with placebo in the treatment of participants with active moderate to severe HS.
 - To assess the pharmacokinetic (PK) and immunogenicity of sonelokimab at 2 different dose levels (120 mg, 240 mg) in the treatment of participants with active moderate to severe HS.
- 

3 STUDY DESIGN

3.1 General study design

This is a Phase 2, randomized, parallel-group, double-blind, placebo-controlled study of sonelokimab in participants with active moderate to severe HS. This study also has an active reference arm of adalimumab treatment; however, no formal comparison of sonelokimab vs adalimumab is planned.

The study includes a screening period of up to 28 days, a 24-week treatment period and a 4-week safety follow-up period. The treatment period will be divided into two parts: Part A (from baseline to Week 12) and Part B (from Week 12 to Week 24). The overall treatment period (Part A+B) will be from baseline to Week 24. The final dose of study treatment is scheduled to occur at Week 20. An End of Treatment (EOT) visit will be performed at Week 24. A safety follow-up visit will be performed 8 weeks after the administration of the final dose of study treatment.

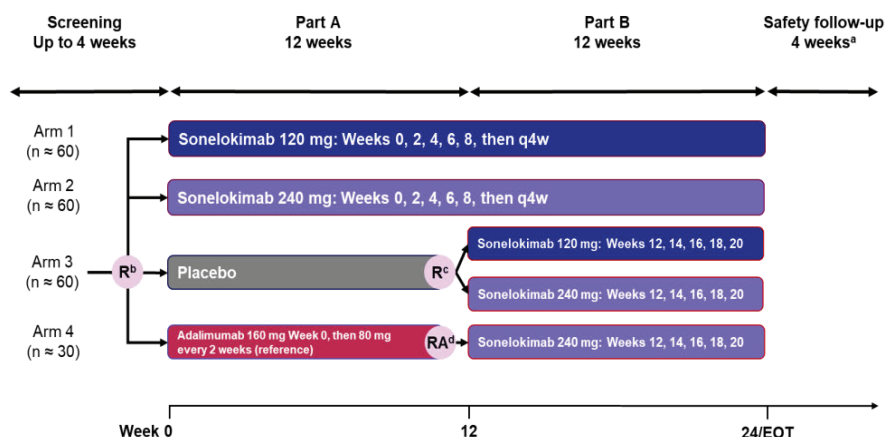
In Part A, approximately 210 participants will be assigned to randomized treatment (60 participants in each of the 2 sonelokimab study arms, 60 participants in the placebo arm, and 30 participants in the adalimumab reference arm). In Part B, participants who were initially randomized to sonelokimab 120 mg or 240 mg will continue treatment for the remainder of the study. Participants initially randomized to placebo will be re-randomized (1:1) to either sonelokimab 120 mg or 240 mg and will receive this treatment for the remainder of the study. Participants initially randomized to the adalimumab reference arm will be reallocated to treatment with sonelokimab 240 mg for the remainder of the study. Study will be conducted in approximately 60 centers in 9 countries.

A subset of participants at selected sites will undergo exploratory ultrasound assessment as part of a sub-study to evaluate the effect of sonelokimab at 2 different dose levels (120 mg, 240 mg) compared with placebo on target active draining tunnels and epidermal thickness.

The study flow chart is presented in [Figure 1](#).

Figure 1: Study Flow Chart

Statistical Analysis Plan (SAP)



Abbreviations: EOT = End of Treatment; q4w = once every 4 weeks; R = randomization; RA = reallocation

- The last dose is scheduled to occur at Week 20. The EOT Visit is scheduled to occur at Week 24. The Safety Follow-up Visit is to occur 8 weeks after the last dose is administered. The total duration of the Safety Follow-up period is 4 weeks.
- At the beginning of Part A at Day 1, all eligible participants will be randomized 2:2:2:1 to 1 of 4 arms and stratified by baseline Hurley Stage status (II and III) and prior biologic use (Yes/No).
- At the beginning of Part B at Week 12, participants initially randomized to placebo (Arm 3) will be re-randomized 1:1 to either sonelokimab 120 mg or 240 mg and stratified by Hurley Stage status (II and III) and prior biologic use (Yes/No) at screening.
- Participants initially randomized to adalimumab (Arm 4) will be reallocated to receive 240 mg sonelokimab and will continue this treatment through Week 20.

3.2 Randomization and blinding

3.2.1 Randomization – Part A

Initial randomization will take place on Day 1 after confirmation that the participant continues to meet the inclusion/exclusion criteria. Participants will be randomized 2:2:2:1 to 1 of 4 treatment arms by block randomization and stratified by Hurley Stage status (II and III) and prior biologic use (yes/no) at the Screening Visit, using a central randomization, as follows:

- Arm 1: sonelokimab 120 mg;
- Arm 2: sonelokimab 240 mg;
- Arm 3: placebo;
- Arm 4: adalimumab;
 - Participants in the adalimumab arm will receive 160 mg adalimumab at Day 1, and then 80 mg every 2 weeks from Week 2 until Week 10.

The Interactive Response Technology (IRT) will be responsible for the allocation of unique identification numbers to individual participants.

The blind will be maintained by the use of sonelokimab and matched placebo. The prefilled syringes used for all sonelokimab and placebo doses will be identical and outer packaging will be labeled in a blinded manner with no indication of the syringe contents (sonelokimab, placebo, or adalimumab). Participants will be asked to wear an eye-mask for all injections.

3.2.2 Randomization – Part B

At the beginning of Part B at Week 12, participants initially randomized to receive placebo (Arm 3) will be re-randomized 1:1 to receive either sonelokimab 120 mg or sonelokimab 240mg. Participants who received adalimumab (Arm 4) in Part A will be reallocated to sonelokimab 240 mg in Part B. Re-randomization of Arm 3 will also be stratified by Hurley Stage and prior biologic use as designated at the beginning of Part A. Participants who were randomized to receive sonelokimab in Part A will not be re-randomized and will continue on their originally assigned study treatment throughout Part B. Re-randomization will also be conducted via the IRT system and will not require any action from the investigator or site personnel.

In Part B, the participants and investigators will remain blinded to each participant's treatment allocation. Blinding of participants and investigators will be maintained in the same manner as in Part A.

3.2.3 Maintenance of Blinding

The primary analysis of 12-week data will occur at the completion of Part A, while the study is still ongoing. All sponsor and study team personnel will remain blinded until Week 12. No unblinded

participant-level data will be supplied to the study sites until the final database lock after the conclusion of the study.

A Data and Safety Monitoring Board (DSMB) will be setup to review the benefit-risk profile of sonelokimab and to independently safeguard the interests of subjects in the studies and to enhance the integrity and credibility of the study. Members of the DSMB will be unblinded to study treatment as will the CRO unblinded statistician and programmer to support the generation of unblinded outputs for the DSMB review. To maintain the blind, the CRO unblinded statistician and programmer will work in a secure environment not accessible to the CRO study team.

Laboratory personnel performing the bioanalytical PK sample analysis may receive an open randomization list to enable analysis of relevant samples. In addition, a PK analyst and modeling/simulation scientist may receive an open randomization list to enable preparation of modeling/simulation activities. These individuals will not interact with site or contract research organization (CRO) personnel.

3.3 Study treatments and assessments

The planned study duration for individual participants will be up to 32 weeks, including a screening period of up to 4-weeks, 24-week treatment period, and a 4-week safety follow-up period.

Study treatments in this current study are:

- sonelokimab 120 mg,
- sonelokimab 240 mg,
- adalimumab,
- placebo.

3.3.1 Sonelokimab – Part A

Sonelokimab drug product is supplied as a fixed-dose, prefilled syringe to be administered via subcutaneous (SC) injection. Syringes and cartons will be supplied in a blinded manner.

Sonelokimab will be administered by a site-approved, study treatment administrator, who is unblinded. Blinded study personnel will not be in attendance while study treatment is being prepared or administered.

During Part A, participants in the sonelokimab treatment arms will receive either:

- Sonelokimab 120 mg at Weeks 0, 2, 4, 6, 8, followed by injections every 4 weeks through to Week 12; or
- Sonelokimab 240 mg at Weeks 0, 2, 4, 6, 8, followed by injections every 4 weeks through to Week 12;

3.3.2 Placebo – Part A

Placebo will be identical in appearance and packaging to sonelokimab in a fixed-dose, prefilled syringe to be administered via SC injection.

As with the administration of sonelokimab, placebo will be administered by a site-approved, study treatment administrator, who is unblinded. Blinded study personnel will not be in attendance while study treatment is being prepared or administered.

During Part A, participants in the placebo treatment arm will receive placebo via SC injection at Week 0, 2, 4, 6, 8, and 10.

Placebo will also be administered to participants in the sonelokimab arms at Week 10 to maintain study blinding.

3.3.3 Adalimumab – Part A

Adalimumab will be provided in prefilled syringes (40 mg) as sterile solution for SC injection. Adalimumab syringes will be relabelled and repackaged to facilitate maintenance of the study blinding.

Adalimumab will be administered by a site-approved, study treatment administrator, who is unblinded. Blinded study personnel will not be in attendance while study treatment is being prepared or administered.

During Part A, participants in the adalimumab treatment arm will receive adalimumab 160 mg SC at Week 0, followed by 80 mg once every 2 weeks, at Weeks 2, 4, 6, 8, and 10 according to the approved labeling.

3.3.4 Sonelokimab – Part B

At the beginning of Part B at Week 12, participants initially randomized to sonelokimab treatment arms will continue to receive their doses at Weeks 12, 16, and 20.

Participants initially randomized to placebo treatment arm will be re-randomized to one of the following 2 groups:

- Sonelokimab 120 mg at Weeks 12, 14, 16, 18, and 20;
- Sonelokimab 240 mg at Weeks 12, 14, 16, 18, and 20;

Participants initially randomized to adalimumab treatment arm will be reallocated to receive sonelokimab 240 mg treatment for the remainder of the study at Weeks 12, 14, 16, 18, and 20.

3.3.5 Placebo – Part B

Placebo will also be administered to participants initially randomized to sonelokimab arms at Weeks 14 and 18 to maintain study blinding.



Statistical Analysis Plan (SAP)

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

Table 1: Schedule of Study Assessments

Period	Screening ^a	Part A								Part B					Follow-up
		0	2	4	6	8	10	12*	14	16	18	20	24 EOT	28	
Scheduled Week	-														
Scheduled Day	-28 Day to -1 Day	1	15	29	43	57	71	85	99	113	127	141	169	8W after last dose ^b	
Visit Window (days)	-		±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±7	
Informed consent	X														
Eligibility criteria	X	X ^c													
Demography	X														
Relevant medical history / current medical condition	X														
HS Medical history and previous HS therapies	X														
Drug and alcohol history	X														
Smoking History	X														
AE/SAE evaluation ^d	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Concomitant medication review ^e	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
HiSQOL		X		X				X		X			X		
FACIT-F		X		X				X		X			X		
EQ-5-D		X		X				X		X			X		
WPAI		X						X					X		
DLQI		X		X		X		X		X		X	X		
PGI-S		X		X				X		X			X		

Statistical Analysis Plan
Protocol Number: M1095-HS-201

Statistical Analysis Plan (SAP)

Period	Screening ^a	Part A								Part B					Follow-up
Scheduled Week	-	0	2	4	6	8	10	12*	14	16	18	20	24 EOT	28	
Scheduled Day	-28 Day to -1 Day	1	15	29	43	57	71	85	99	113	127	141	169	8W after last dose ^b	
Visit Window (days)	-		±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±7	
Complete daily in eDiary															
PGA of Skin Pain NRS and analgesics use ^f															
Lesion Count (Inflammatory nodules, abscesses and draining fistulas/tunnels, Scars) including Hurley Stage ^g	X	X	X	X		X		X	X	X		X	X		
HS-Physician's Global Assessment	X	X		X		X		X		X		X	X		
12-lead ECG	X							X					X	X	
Height	X														
Weight	X	X						X					X	X	
Vital signs	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Complete Physical exam	X							X					X	X	
Serum/Urine pregnancy test ^h	X	X		X		X		X		X		X	X		
FSH ⁱ	X														
TB Screening ^j	X														
HIV Screening	X														
Hepatitis B and Hepatitis C Screening	X														
SARS-CoV-2 PCR virus testing ^k	X														
Safety laboratory testing (hematology, clinical chemistry including CRP, urinalysis)	X	X	X	X		X		X	X	X		X	X	X	

Statistical Analysis Plan
Protocol Number: M1095-HS-201

Version: FINAL 1.1, Date: 23 May 2023
Page 17 of 93

Statistical Analysis Plan (SAP)

Period	Screening ^a	Part A								Part B				Follow-up
Scheduled Week	-	0	2	4	6	8	10	12*	14	16	18	20	24 EOT	28
Scheduled Day	-28 Day to -1 Day	1	15	29	43	57	71	85	99	113	127	141	169	8W after last dose ^b
Visit Window (days)	-		±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±7
ESR	X	X	X	X		X		X	X	X		X	X	X
Fasting Lipid Panel and Glucose		X											X	
ADA sampling ^l		X	X	X		X		X		X		X	X	X
PK sampling ^l		X	X	X		X		X		X		X	X	X
Randomization ^m		X												
Sub-study ultrasound assessment ⁿ		X						X					X	
Study treatment administration ^o		X	X	X	X	X	X	X	X	X	X	X	X	X

Abbreviations: ADA= anti-drug antibody; AE= adverse event; DLQI= Dermatology Life Quality Index; ████████ ECG = electrocardiogram; EOT = End of Treatment; EQ-5D= EuroQol 5-dimension questionnaire visual analog scale; ESR = erythrocyte sedimentation rate; FACIT-F = mean Functional Assessment of Chronic Illness Therapy-Fatigue; FSH= follicle stimulating hormone; HS = hidradenitis suppurativa; HiSQoL= Hidradenitis Suppurativa Quality of Life; HIV= human immunodeficiency virus; hs-CRP = High-sensitivity C-reactive protein; IGRA = interferon-gamma release assay; MP = investigational medicinal product; NRS= numerical rating scale; ; PCR = polymerase chain reaction; PGA= Patient's Global Assessment of Skin Pain; PGI-S= Patient Global Impression of Severity; ████████ PK = pharmacokinetics; SAE = serious adverse event; TB= tuberculosis; W = week; WPAI = Work Productivity and Activity Impairment

*Assessments at Week 12 are final assessments of Part A. Treatment at Week 12 is the first treatment of Part B.

- The screening period is a minimum of 7 days to allow for pre-treatment patient diary completion. Screening is a maximum of 28 days, including any repeat assessments that may be required.
- Participants who discontinue treatment prematurely are required to complete the EOT Visit immediately and the Safety Follow-up Visit 8 weeks after the administration of the final dose.
- Recheck inclusion and exclusion criteria prior to randomization.
- AEs /SAEs occurring after the participants have provided informed consent will be reported until Safety Follow-up.
- While all concomitant medications are recorded on eCRF, participants will record their daily analgesics use on eDiary. Patients should be instructed to start collecting this information as soon as they receive the patient diary.

Statistical Analysis Plan (SAP)

- f. Each participant will receive a device upon which to complete the daily eDiary. From screening through to Week 24, participants will complete the eDiary daily for skin pain and analgesics use.
- g. Assessment of individual lesions e.g. abscesses, inflammatory nodules, number of active draining tunnels, including Hurley Stage.
- h. To be conducted in women of childbearing potential only (as defined in Appendix II in Section 21.2 of the protocol). Serum pregnancy test at screening to confirm study eligibility(central lab); local urine pregnancy tests at Week 0 and every 4 weeks thereafter.
- i. To be conducted only in female participants considered to be of non-childbearing potential due to being postmenopausal (as defined in Appendix II in Section 21.2 of the protocol), unless they are >55 years of age with amenorrhea for at least 12 months without an alternative medical cause, to verify that FSH levels are >40 IU/L (see Appendix II in Section 21.2 of the protocol). Women of childbearing potential (as defined in Appendix II in Section 21.2 of the protocol) and women >55 years of age with amenorrhea for at least 12 months without an alternative medical cause are not required to have FSH levels determined.
- j. QuantiFERON®-TB-Gold test (or IGRA equivalent) will be performed. A re-test is performed if the first is indeterminate; the result of the second test will then be used. If the results of both tests are indeterminate, the participant is considered a screen failure and to be excluded. A re-test can be performed if there is reasonable medical doubt about a potentially false-positive result after obtaining Medical Monitor approval; the result of the second test will prevail. If a chest X-ray or CT for tuberculosis (TB) screening is required per local guidance, the X-ray or CT must have been taken within 3 months prior to the Screening Visit.
- k. PCR test for SARS-CoV-2 infection at screening may be performed by the central laboratory, or if sample collection not possible at site due to local procedures, a test may be obtained at a suitable laboratory as long as evidence is provided and retained as source data. Local test must be between 2 days prior or 3 days post-screening visit. If a participant tests positive for SARS-CoV-2 infection at screening, another PCR or antigen test may be performed within the screening window. If the repeat test is negative and the participant is asymptomatic, they may enter the study.
- l. For visit where study treatment is administered the ADA, PK, [REDACTED] samples should be collected pre-dose.
- m. Participants in Arm 3 (placebo) will be re-randomized automatically at Week 12. The investigator does not need to perform any action for re-randomization.
- n. For participants taking part in the ultrasound sub-study (Section 21.3 of the protocol) only.
- o. Study treatment is to be administered after all other scheduled activities have been completed, in particular, after all blood samples have been collected.

4 STUDY ENDPOINTS

4.1 Primary efficacy endpoint

The primary efficacy endpoint of this study is the percentage of participants achieving HS Clinical Response 75 (HiSCR75) at Week 12.

4.2 Secondary efficacy endpoints

The key secondary efficacy endpoints of this study are:

- Proportion of participants achieving HiSCR50 at Week 12;
- Change from baseline in International HS Severity Score System (IHS4) at Week 12;
- Proportion of participants achieving a Dermatology Life Quality Index (DLQI) total score of ≤ 5 at Week 12;
- Proportion of participants achieving at least 30% reduction and at least 1-unit reduction from baseline in Numerical Rating Scale (NRS) 30 in Patient's Global Assessment (PGA) of Skin Pain at Week 12 among participants with Baseline NRS ≥ 3 .

The other secondary endpoints this study are:

- Proportion of participants who achieve HiSCR75 at time points other than Week 12;
- Proportion of participants who achieve HiSCR50 at time points other than Week 12;
- Proportion of participants who achieve HiSCR90;
- Proportion of participants who achieve HiSCR100;
- Proportion of participants achieving at least 30% reduction and at least 1-unit reduction from baseline in PGA of Skin Pain NRS among participants with baseline NRS ≥ 3 ;
- Proportion of participants achieving at least 30% reduction and at least 1-unit reduction from baseline in PGA of Skin Pain NRS at Week 12 among participants with Baseline NRS ≥ 3 , with participants using permitted analgesic therapy for ≥ 4 days the previous week counted as non-responders;
- Proportion of participants achieving Abscess and inflammatory nodule (AN) count of 0, 1, or 2;
- Proportion of participants achieving at least 50%, 75%, 90%, and 100% reduction in total AN count (AN50, AN75, AN90 and AN100);
- Proportion of participants achieving complete elimination of abscesses at Week 12 among those participants with abscesses at baseline;

- Proportion of participants achieving complete elimination of draining fistulas/tunnels at Week 12 among those participants with draining fistulas/tunnels at baseline;
- Proportion of participants achieving complete elimination of inflammatory nodules at Week 12 among participants with inflammatory nodules at baseline;
- Proportion of participants with change of ≥ 1 score points from baseline in Patient Global Impression of Severity (PGI-S);
- Proportion of participants reporting 'minimal symptoms' or 'absent symptoms' on the PGI-S;
- Change from baseline in IHS4 at time points other than Week 12;
- Change from baseline in HS-Physicians Global Assessment (HS-PGA) 6-point score;
- Change from baseline in the number of abscesses;
- Change from baseline in number of draining fistulas/tunnels;
- Change from baseline in number of inflammatory nodules;
- Change from baseline in HS Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, and activities-adaptations);

- Change from baseline in Work Productivity and Activity Impairment (WPAI) score;
- Change from baseline in mean Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) score;
- Change from baseline in EuroQol five-dimension questionnaire (EQ-5D) Visual Analog Scale (VAS) and in EQ-5D index scores by visit over time;
- Change from baseline in high-sensitivity C-reactive protein (hs-CRP);
- Change from baseline in Neutrophil to Lymphocyte Ratio (NLR);
- Proportion of participants experiencing one or more flares (defined as $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to baseline);
- Proportion of participants achieving DLQI Total Score ≤ 1 ;
- Proportion of participants achieving a DLQI Total Score of ≤ 5 ;
- Proportion of participants achieving a DLQI improvement (reduction) of ≥ 4 points among participants with DLQI ≥ 4 at baseline;
- Proportion of participants achieving a IHS4-55 improvement (reduction) of $\geq 55\%$.

4.3 Safety endpoints

The safety endpoints of this study are:

- Incidence, relatedness, severity, and seriousness of adverse events (AEs);
- Withdrawal due to AEs;
- Clinically relevant abnormalities in vital signs (blood pressure [BP] and heart rate) and body weight;
- Standard 12-lead electrocardiogram (ECG) intervals (RR, PR, QRS, QT, and QT intervals corrected for heart rate using Bazett's and Fridericia's formulas [QTcB and QTcF, respectively]), Clinically relevant abnormalities in ECG variables;
- Clinically relevant abnormalities in clinical laboratory variables (hematology, biochemistry, and urinalysis).

4.4 Pharmacokinetic (PK) and Immunogenicity (IG) Endpoints

The PK and Immunogenicity endpoints of this study are:

- PK of sonelokimab (trough levels);
- Anti-drug antibodies (ADAs).

4.6 Ultrasound sub-study endpoints

The ultrasound sub-study endpoints are:

- The effect of sonelokimab on the diameter of a target active draining tunnel at Weeks 12 and 24 compared to baseline;
- The effect of sonelokimab on power Doppler ultrasound in the area of a target active draining tunnel at Weeks 12 and 24 compared to baseline;
- The effect of sonelokimab on epidermal thickness at the target anatomical site as measured by the ultrasound entry signal at Weeks 12 and 24 compared to baseline;
- The effect of sonelokimab on the lumen opacity of a target active draining tunnel at Weeks 12 and 24 compared to baseline.

5 STUDY ESTIMANDS

5.1 Estimand for the primary endpoint

Attribute	Details
Treatments	Sonelokimab at 2 different doses: • sonelokimab 120 mg or • sonelokimab 240 mg
Target Population	• Adults diagnosed with HS as determined by the investigator and with a history of signs and symptoms of HS dating back at least 6 months prior to the Screening Visit • Analysis Set: Full Analysis Set (FAS)
Endpoint	Percentage of participants achieving HiSCR75 at Week 12.
Population-level summary	The primary estimator will be the improvement compared to placebo over the 12-week Part A based on HiSCR75.
Intercurrent events and strategies to handle those	• Use of prohibited concomitant medication (Section 10.5.1 of protocol): participants will be analyzed as a treatment failure after taking prohibited medication, i.e., non-responder imputation will be performed after the first episode of prohibited medication for all future assessments (composite strategy). The prohibited concomitant medications will be confirmed during a data review meeting prior to database lock for part A by comparing the full list of concomitant medications to section 10.5.1 of the protocol; • Withdrawal from treatment due to any reason (Section 9.6 of protocol): participants will be considered non-responders to treatment if they discontinue the study prior to the Week 12 visit due to any reason (composite strategy); • Study drug non-adherence (missing doses) or drug administration error: participants will be analyzed in the same way as if the event had occurred or not (treatment policy strategy).

5.2 Estimand for the key secondary endpoints

Attribute	Details
Treatments	Sonelokimab at 2 different doses: • sonelokimab 120 mg or • sonelokimab 240 mg
Target Population	• Adults diagnosed with HS as determined by the investigator and with a history of signs and symptoms of HS dating back at least 6 months prior to the Screening Visit • Analysis Set: FAS
Endpoint	The key secondary endpoints are as follows: • Proportion of participants achieving HiSCR50 at Week 12; • Change from baseline in IHS4 at Week 12; • Proportion of participants achieving a DLQI total score of ≤ 5 at Week 12; • Proportion of participants achieving at least 30% reduction and at least 1-unit reduction from Baseline in NRS30 in PGA Skin Pain at Week 12 among participants with Baseline NRS ≥ 3.
Population-level summary	The primary estimator will be the improvement compared to placebo over the 12-week Part A based on the endpoint of interest.
Intercurrent events and strategies to handle those	• Use of prohibited concomitant medication (Section 10.5.1 of protocol): participants will be analyzed as a treatment failure after taking prohibited medication, i.e., non-responder imputation will be performed after the first episode of prohibited medication for all future assessments (composite strategy). The prohibited concomitant medications will be confirmed during a data review meeting prior to database lock for part A by comparing the full list of concomitant medications to section 10.5.1 of the protocol;

	• Withdrawal from treatment due to any reason (Section 9.6 of protocol): participants will be considered non-responders to treatment if they discontinue the study prior to the Week 12 visit due to any reason (composite strategy); • Study drug non-adherence (missing doses) or drug administration error: participants will be analyzed in the same way as if the event had occurred or not (treatment policy strategy).
--	---

6 SAMPLE SIZE AND POWER

The sample size calculation is based on the primary efficacy endpoint (i.e., HiSCR75 response rate at Week 12) with the primary comparison of interest being: sonelokimab 120 mg vs placebo group and sonelokimab 240 mg vs placebo group, assuming a HiSCR75

- 40% with the lower sonelokimab dose of 120 mg;
- 50% with the higher sonelokimab dose of 240 mg; and
- 15% with placebo.

A sample size of $n=60$ participants in each of the sonelokimab 120 mg, sonelokimab 240 mg, and placebo groups, therefore 180 participants in total, results in a power of more than 80%, while assuming an overall 2-sided alpha of 0.05 and a difference in HiSCR75 response rate of 25% between placebo and sonelokimab 120 mg and of 35% between placebo and sonelokimab 240 mg. An additional 30 participants will be randomized in the adalimumab group. The adalimumab treatment will be used as active reference arm, however, no formal direct comparison of sonelokimab or placebo vs adalimumab is planned.

Eligible participants will be randomized in a 2:2:2:1 ratio to sonelokimab 120 mg, sonelokimab 240 mg, placebo, or adalimumab, respectively. The number of participants in Hurley Stage III will be capped to approximately 40% and the number of participants currently using antibiotics will be capped at 30%.

Due to the use of the FAS with non-responder imputation, no accounting for dropout is required, therefore a total of approximately 210 participants will be randomized.

All calculations were performed in nQuery® Software version 8.5.1.0.

7 ANALYSIS POPULATIONS

7.1 Part A

7.1.1 Enrolled Set

The enrolled set is defined as all participants who signed an informed consent.

7.1.2 Safety Analysis Set (SAF)

The safety set is defined as all randomized participants who received at least one dose of study drug. Participants will be analyzed based on the treatment that they actually received, regardless of the treatment they are randomized to.

The SAF will be used for all safety analyses related to the Part A and Part A+B.

7.1.3 Full Analysis Set (FAS)

The full analysis set is defined as all randomized participants who are assigned treatment for Part A. Participants will be analyzed in accordance to the intention-to-treat principle, i.e., based on randomized treatment, irrespective of the treatment they actually received.

The FAS will be used for all efficacy analyses on efficacy endpoints related to the Part A and some Part A+B analyses and will serve as the primary analysis set.

7.1.4 Per-protocol Set (PPS)

The per-protocol set is defined as all participants in the FAS who had no major protocol deviations during the Part A, as confirmed during a data review meeting prior to database lock for part A.

The PPS will be used for the primary and key secondary efficacy endpoints related to the Part A and will serve as a secondary analysis set.

7.1.5 Immunogenicity Analysis Set (IG Analysis Set)

The IG analysis set is defined as all participants who received at least one dose of sonelokimab and who have at least 1 sample obtained for the detection of antibodies to sonelokimab.

In the immunogenicity analyses, participants will be analyzed according to the treatment they actually received, regardless of the treatment they are randomized to.

The IG analysis set will be used for all immunogenicity analyses related to the Part A and A+B.

7.1.6 Pharmacokinetic Analysis Set (PK Analysis Set)

The PK analysis set is defined as all participants who received at least one complete dose of sonelokimab and had at least 1 valid blood sample drawn for PK analysis after their first dose of sonelokimab.

In the PK analyses, participants will be analyzed according to the treatment they actually received, regardless of the treatments they are randomized to.

The PK analysis set will be used for all PK analyses related to the Part A and A+B.

7.2 Part B

7.2.1 Safety Analysis Set (SAF-B)

The SAF-B is defined as all randomized participants who received at least one dose of sonelokimab during Part B. Participants will be analyzed based on the treatment that they actually received, regardless of the treatments they are randomized to.

The SAF-B will be used for all safety analyses related to Part B.

7.2.2 Full Analysis Set (FAS-B)

The FAS-B is defined as all randomized participants who are assigned treatment for Part B. Participants will be analyzed in accordance to the intention-to-treat principle, i.e., based on randomized treatment, irrespective of the treatment they actually received.

The FAS-B will be used for some efficacy analyses on endpoints related to the Part B.

7.3 Ultrasound Sub-study Set

The ultrasound sub-study set is defined as all randomized participants who received at least one dose of study drug and consented for the sub-study.

The sub-study set will be used for all sub-study analyses.

7.4 Protocol deviations/violations and exclusions from analysis sets

All protocol deviations will be identified prior to their associated database lock (Part A and Part B) for any intentional or unintentional change or noncompliance with the approved protocol procedures or requirements as mentioned in protocol deviation criteria form (PDCF).

All protocol deviations will be classified as key or non-key and will be reviewed to determine their impact on the primary and key secondary endpoints during a data review meeting. Deviations that are identified as having an impact on the primary and key secondary endpoints may lead to a participant being excluded from the PP analysis set. A pre-defined set of reasons for exclusion from the PP population are given in Appendix K. All exclusions from an analysis population will be documented prior to unblinding, exceptions to this will be protocol deviations based upon unblinded data (e.g PK analysis set) which will be documented prior to release of the final outputs.

8 STATISTICAL CONSIDERATIONS AND ANALYSIS

8.1 Derived Variables

The below table provide the list of derived variables for demographic and baseline characteristics, various duration derivations, drug compliance, baseline derivations and other important derivations applicable for this study.

Variables	Formula
Demographic and Baseline characteristics	
Body mass index (BMI) (kg/m ²)	weight (kg)/[height (m) ²] Note: 1 lbs = 0.45359237 kg 1 in = 2.54 cm
Duration of HS (Year)	[(Date of first study drug administration – Date of Initial diagnosis)+1]/365.25 Details for imputing missing or partial start and/or stop dates of HS are described in Section 8.3.2
Derivation of Duration	
Study day at any visit	Date of interest – date of first dose of study drug. One day is added if this difference is ≥ 0
Duration of Exposure (Days) for Part A	Date of Week 12 or EOT visit, whichever occurs first – Date of first randomized study medication intake at baseline + 1
Duration of Exposure (Days) for Part B	Date of Week 24 or EOT visit, whichever occurs first – Date of first study medication intake for Part B + 1
Duration of Exposure (Days) to Sonelokimab for Part A+B	Date of Week 24 or EOT visit, whichever occurs first – Date of first sonelokimab intake + 1
Drug Compliance	
Compliance for Part A	Compliance (%) = 100 * (number of visits with treatment/ number of visits with treatment planned) during the subject's participation in the Part A study.
Compliance for Part B	Compliance (%) = 100 * (number of visits with treatment/ number of visits with treatment planned) during the subject's participation in the Part B study.
Compliance for Part A+B	Compliance (%) = 100 * (number of visits with treatment/ number of visits with treatment planned) during the subject's participation in the Part A+B study.
Baseline Derivations	

Baseline definition for Part A, Part B and Part A+B	Baseline will be defined as the last non-missing measurement prior to first dose of study medication.
Baseline definition for PGA skin pain	The average of the last 7 days, with at least 4 days completion, of non-missing assessments before the date of first dose of study drug injection. If less than 4 records exist between days -7 and -1 then use the period day -14 to -8 and if less than 4 records exist between days -14 and -8 then use the period day -21 to -15. If more than 3 out of the 7 days are missing the baseline will be set to missing
Baseline definition for ADA	The last sample prior to the first dose of sonelokimab
Change from baseline	The value at each time point minus the value at baseline
Actual Baseline Hurley	Maximum Hurley stage from all locations last non-missing measurements prior to first
Other Derivations	
Definition for the derived efficacy parameters are given in the respective efficacy analysis sections.	

8.2 Visit windows

Visit windows will be used for all scheduled assessments. To allow for unscheduled and discontinuation assessments, the data will be summarized using the assessment (whether reported as a scheduled or unscheduled visit) closest to the nominal visit date, calculated from the first day of dosing. The visit result will be missing if no assessment was reported within the specified visit window around the scheduled date. If two assessments are equidistant from a scheduled visit, the earlier of the two will be used.

The visit windows will be calculated by bisecting the scheduled visit assessments. The upper limit of each window will be the mean of the two adjacent planned study days, rounded down to the nearest integer. The lower limit of each window will be the upper limit of the preceding window plus one, except for the first post-treatment visit which will start at 2.

Visit windows for lesion count and safety laboratory testing

Nominal Week	Nominal study day	Visit window
0	1	All assessments prior to the start of study treatment
2	15	Day 2 – 22
4	29	Day 23 – 43
8	57	Day 44 – 71

Visit windows for lesion count and safety laboratory testing

Nominal Week	Nominal study day	Visit window
12	85	Day 72 – 92
14	99	Day 93 – 106
16	113	Day 107 – 127
20	141	Day 128 – 155
24	169	Day 156 +

Visit windows for ADA and PK Sampling

Nominal Week	Nominal study day	Visit window
0	1	All assessments prior to the start of study treatment
2	15	Day 2 – 22
4	29	Day 23 – 43
8	57	Day 44 – 71
12	85	Day 72 – 92
14	99	Day 93 – 106
16	113	Day 107 – 127
20	141	Day 128 – 155
24	169	Day 156 – 183
28	197	Day 183 +

Visit windows for HiSQOL, FACIT-F, EQ-5-D and PGI-S

Nominal Week	Nominal study day	Visit window
0	1	All assessments prior to the start of study treatment
4	29	Day 2 – 57
12	85	Day 58 – 99
16	113	Day 100 – 141

PAI, ECG, Weight and all ultrasound assessments		
Statistical Analysis Plan (SAP)		
	Nominal study day	
0	1	All assessments prior to the start of study treatment
12	85	Day 2 – 127
24	169	Day 128 +

Visit windows for HiSQOL, FACIT-F, EQ-5-D and PGI-S

Nominal Week	Nominal study day	Visit window
24	169	Day 152 +

Visit windows for DLQL, HS-Physician's Global Assessment

Nominal Week	Nominal study day	Visit window
0	1	All assessments prior to the start of study treatment
4	29	Day 2 – 43
8	57	Day 44 – 71
12	85	Day 72 – 99
16	113	Day 100– 127
20	141	Day 128 – 155
24	169	Day 156 +

Visit windows for vital signs

Nominal Week	Nominal study day	Visit window
0	1	All assessments prior to the start of study treatment
2	15	Day 2 – 22
4	29	Day 23 – 36
6	43	Day 37 – 50
8	57	Day 51 – 64
10	71	Day 65 – 78

Visit windows for vital signs

Nominal Week	Nominal study day	Visit window
12	85	Day 79 – 92
14	99	Day 92 – 106
16	113	Day 107– 120
18	127	Day 121– 134
20	141	Day 135 – 155
24	169	Day 156 +

8.3 Handling of missing data and randomization stratification errors**8.3.1 Missing data analysis methods**

Missing data in this study may result from participants discontinuing from the study prematurely or missing intermediate visits or selected assessments while remaining on study. Every reasonable effort will be made to obtain the protocol-required data for all study assessments that are scheduled for all participants who have been enrolled.

Dichotomous Efficacy Endpoints

For all dichotomous endpoints, a non-responder imputation method will be used for the primary analysis. Participants will be considered non-responders if:

- They discontinue the study, prior to the Week 12 visit for Part A or prior to the Week 24 visit for Part B and Part A+B (due to any reason). All future visits following the discontinuation will be imputed, or
- Have missing data for the endpoint of interest at baseline or at any intermediate time points for Part A, Part B and Part A+B. All visits with a missing assessment will be imputed, or
- Take prohibited protocol medication prior to the Week 12 visit for Part A or Week 24 visit for Part B and Part A+B. All future visits following the start date of the prohibited medication will be imputed.

In addition, two sensitivity approaches will be used for the dichotomous primary and key-secondary endpoints for analyses of Part A data:

1. The method for handling missing data will be dependent on the reason the data are missing. If a participant has discontinued due to lack of efficacy or starting prohibited medication the data will be imputed as a non-response. If the data are missing for any other reason the data will not be imputed. A Generalised Linear Mixed Model will be used including all available data up to Week 12 to enable the inclusion of participants with missing data at Week 12.
2. No data will be imputed and the data analysed by visit based on the observed data only.

Continuous Efficacy Endpoints

For continuous efficacy endpoints, missing observations will not be formally imputed. Mixed Model for Repeated Measures (MMRM) analysis will use with all available data, including participants with partial data (i.e., with missing data) in order to arrive at an estimate of the mean treatment effect. MMRM is based on the assumption of missing at random (MAR), i.e., dropout would behave similarly to other participants in the same treatment group, and possibly with similar covariate values, had they not dropped out.

Unless otherwise indicated, in other cases the observed data will be used and no imputation will be performed.

Scoring derivations

Several instruments will be used in this study to measure some underlying construct (e.g. the DLQI is a 10-item questionnaire that will be used to measure dermatology-specific quality of life). As such, missing data in this type of instruments can occur on the item scores. Missing data will be handled as described by the author of each instrument and detailed in the relevant endpoint section, or, if no information is stated by the author, no imputation will be done and such scores (and any resultant domain/total score) will remain missing.

8.3.2 Handling of missing or incomplete dates

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

If only Day of AE start date is missing:

If the AE start year and month are the same as that for the first dose date, then:

- If the full (or partial) AE end date is NOT before the first dose date or AE end date is missing, then impute the AE start day as the day of first dose date; otherwise, impute the AE start day as 1.

If the AE start year and month are not same as that for the first dose date, then, impute the AE start day as 1.

Compare the imputed AE start date with treatment-emergent (TE) period to determine whether the AE is pre-treatment AE, treatment-emergent AE (TEAE) or post-treatment AE.

If Day and Month of AE start date are missing:

If AE start year = first dose year, then:

- If the full (or partial) AE end date is NOT before the first dose date or AE end date is missing, then impute the AE start Month and Day as the Month and Day of first dose date; otherwise, impute the AE start Month as January and the Day as 1.

If AE start year not same as first dose year, then, impute the AE start MONTH as January and the DAY as 1.

Compare the imputed AE start date with TE period to determine whether the AE is pre-treatment AE, TEAE or post-treatment AE.

If Year of AE start date is missing:

If the year of AE start is missing or AE start date is completely missing, then query site with no imputation. Also compare the full (or partial) AE end date to the first dose date. If the AE end date is before the first dose date, then the AE should be considered as a pre-treatment AE. Otherwise, the AE will be considered as TEAE.

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.
- If DAY, Month and Year are all missing, use a date before the first dose date.

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, then consider the medication is still ‘continuing’ and date will not be imputed.

Imputation rules for missing or partial HS start dates are defined below:

Missing or partial HS 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.
- If DAY, Month and Year are all missing, use a date before the first dose date.

8.3.3 Randomization stratification errors

Randomization stratification errors can occur if the incorrect Hurley Stage or Prior Biologic Use is

entered in the IRT system at the time of randomization. The FAS analyses will follow the ITT principle and analyse these subjects according to the randomized stratum. The PP analyses will use the correct baseline Hurley Stage and Prior Biologic Use values for the stratification factors.

9 STATISTICAL METHODS

9.1 General statistical conventions

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

For continuous variables, descriptive statistics (number of observations, mean, standard deviation (SD), minimum, first quartile, median, third quartile, and maximum and 95% confidence intervals (CIs) as appropriate) will be presented. Summaries of hs-CRP and PK concentration will include the geometric mean and coefficient of variation. The count of missing observations will be provided in all descriptive tables.

Categorical variables will be presented as counts and percentages with non-missing values in each category. The number of missing values will be presented as a separate category, if there is at least one missing value. Unless otherwise stated, all statistical testing will be two-sided and will be performed using a significance (alpha) level of 0.05. For all statistical analyses, p-values and 95% CIs will be provided. All sensitivity analyses will be regarded as exploratory and as such no formal adjustment for multiplicity will be made.

Number of decimal place will be presented as follow:

- continuous variables: for min and max, the decimals will be used as captured in the database; for arithmetic mean, median and quartiles will be depicted with one more decimal than captured in the database; and for standard deviation will be depicted with two more decimals than captured in the database unless stated otherwise
- categorical variables: percentages will be rounded to one decimal place and a percentage of 100% will be reported as 100%
- The p-values will be rounded and displayed to 3 decimal places, values less than 0.001 will be displayed as 'p<0.001' and values higher than 0.999 will be displayed as 'p>0.999'.

This study has two parts Part A and Part B:

Part A Analysis: For Part A, summaries and analyses will be based on the data collected from the visits Week 0 to Week 12. Unless otherwise stated, the summary statistics and analyses will be presented for the following treatment groups:

- Sonelokimab 120 mg
- Sonelokimab 240 mg
- Placebo
- Adalimumab

Part B and Part A+B Analysis: For Part B and Part A+B (when applicable), no statistical testing will be performed. Summary statistics will be based on the data collected from the visits Week 12 to Week 24 for Part B and from the visits Week 0 to Week 24 for Part A+B.

Unless otherwise stated, the summary statistics will be presented for the following treatment groups:

- Sonelokimab 120 mg
- Sonelokimab 240 mg
- Placebo-sonelokimab 120 mg
- Placebo-sonelokimab 240 mg
- Adalimumab-sonelokimab 240 mg

All participant data, including those derived, will be presented in individual participant data listings. Unless otherwise stated, unscheduled visit results will be included in date/time chronological order, within participant listings only. All listings will be sorted by treatment group, investigational site, participant number, date/time and visit. The treatment group as well as participant's sex, age, and race will be stated on each listing. Unless otherwise stated, safety data listings will be based on SAF and efficacy data listings will be based on FAS.

9.2 Participant disposition

The number and percentage of participants screened, randomized, randomized and treated, randomized and not treated, and the number of participants in each of the analysis sets (including reasons for exclusion) separately for Part A and Part A+B (see [section 7](#)) will be summarized by treatment group on the enrolled set for disposition and the randomized subjects for summary of analysis sets.

The number and percentage of participants will also be summarized for the following participant disposition categories:

- Completed study drug for Part A
- Discontinued study drug in Part A
- Received study drug in Part B
- Completed study drug treatment for both part A and part B per-protocol;
- Discontinued study drug earlier during Part B (including the primary reason for discontinuation);
- Completed the study at week 24;
- Discontinued from study early (including the reason for discontinuation);

Given data from Part A and Part B will be combined together (including participants not entering Part B), all distinct treatment groups detailed in section 9.1 from Part A and B will be presented.

The time in days to study drug discontinuation due to any reason will be displayed with a Kaplan-Meier (KM) plot by initial randomized treatment for the FAS during Part A. The time to discontinuation will be defined as the discontinuation date minus the first dose date plus one. Subjects who complete the treatment at Week 12 will be censored at the completion day (Week 12).

The KM plot will be repeated for Part A+B to show the time to discontinuation of sonelokimab. This will only include subjects exposed to sonelokimab in either Part A or B. The time to discontinuation will be defined as the discontinuation date minus the first sonelokimab dose date plus one. Subjects who complete the treatment at Week 24 will be censored at the completion day (Week 24).

9.3 Protocol deviations

A summary will be provided for all key protocol deviations by deviation sub-category and treatment group based on the FAS will be presented for Part A and Part A + B separately.

In addition, all protocol deviations will be listed and protocol deviations occurring due to the COVID-19 pandemic will be listed separately.

9.4 Demographics and baseline characteristics

All demographic and baseline characteristics data will be summarized for Part A by treatment group using the FAS. Inferential statistical analysis of demographic or baseline data will not be performed.

9.4.1 Demographics

Continuous variables such as age (years) at screening collected on the eCRF, height (cm), weight (kg) and BMI (kg/m²) will be summarized using descriptive statistics.

Categorical variables such as age categories (18-64, 65-84, >84), sex, ethnicity and race will be summarized using counts and percentages.

Childbearing potential status, contraception methods used and reason for non-child-bearing potential status for female participants will be provided in the listing only.

9.4.2 Baseline and disease characteristics

The following variables will be summarized:

- Drug and alcohol history
- Smoking history
- HS medical history
 - Duration of HS (Year)
 - Family history of HS
 - If yes, the relationship
- Geographic region: North America, Western Europe and Eastern Europe;
- Prior exposure to non-biologic systemic agents (yes/no);
- Prior exposure to biologic agents (yes/no);

- Prior exposure to systemic antibiotics (yes/no)
- Concomitant antibiotic use at baseline (yes/no);
- Baseline Abscesses (A) count
- Baseline inflammatory nodules (N) count
- Baseline Abscess and inflammatory nodule (AN) count
- Baseline draining fistulas (DF) count
- Baseline IHS4 severity score
- Baseline IHS4 Category (mild, moderate, severe)
- Baseline DLQI total score as well as 6 domain scores
- Baseline PGA of Skin Pain NRS
- Baseline analgesic use (0 days, 1-3 days, 4-7 days)
- Baseline hs-CRP;
- Baseline Hurley Stage status: Stage II – Moderate; Stage III – Severe
- Baseline combined strata: Hurley stage II and prior biologic use; Hurley stage II and no prior biologic use; Hurley stage III and prior biologic use; Hurley stage III and no prior biologic use;

Continuous variables will be summarized using descriptive statistics and categorical variables will be summarized using counts and percentages. Percentage calculations will be based on the number of participants in FAS including participants with missing value in each treatment group.

9.4.3 Medical history

Medical history will be summarized by system organ class (SOC) and preferred term (PT) coded using Medical Dictionary for Regulatory Affairs® (MedDRA) Version 25.0 or higher. Counts and percentages will be presented by treatment group using the SAF. Medical history will only be summarized for Part A.

9.4.4 HS Previous therapies

HS previous therapies will be summarized separately by category of previous therapy and therapy name. . Counts and percentages will be presented by treatment group using the SAF. HS previous therapies will only be summarized for Part A.

The categories summarised will be

- Previous Exposure for systemic biologics
- Previous Exposure to Non-biologics and non-antibiotics systemic therapy
- Topical Therapy
- Other Previous Therapies
- Previous Exposure to Systemic Antibiotics

9.4.5 Prior and concomitant medications

Prior and concomitant medications will be summarized separately by WHODDE ATC level 2 and

preferred name. Counts and percentages will be presented by treatment group using the SAF for Part A and Part A+B and using SAF-B for Part B. Prior medications will be summarized for Part A and concomitant medication will be summarized separately for Part A, Part B and Part A+B.

Prior medications: are defined as those medications with a start date and stop date prior to the first dose of study drug.

Concomitant medications:

- for part A, are defined as those medications with a start date on or after the first dose of study drug of part A but not after the first dose of Part B. A medication which started prior to first dose of part A and continued after first dose will also be considered as concomitant medications for part A.
- for part B, are defined as those medications with a start date on or after the first dose of study drug of part B. A medication which started prior to first dose of part B and continued after first dose of part B will also be considered as concomitant medications for part B.
- for part A+B, are defined as those medications with a start date on or after the first dose of study drug of part A. A medication which started prior to first dose of part A and continued after first dose will also be considered as concomitant medications for part A+B.

Details for imputing missing or partial start and/or stop dates of medication are described in [section 8.2.2](#).

9.5 Extent of exposure

Extent of exposure will be summarized using the SAF separately for Part A and Part A+B and SAF-B for Part B.

For Part A+B, only two groups will be presented:

- Sonelokimab 120 mg,
- Sonelokimab 240 mg,

9.5.1 Treatment duration

The duration (days) ([Section 8.1](#) for derivations) of exposure to study drug will be summarized using descriptive statistics by treatment group. In addition, the number of injections received will be summarized with count and percentages.

9.5.2 Treatment compliance

Treatment compliance (See section 8.1 for derivations) will be summarized by treatment group using descriptive statistics. Counts and percentages will be provided using the following categories: <80%; $80\% \leq x < 100\%$; 100%; $100\% < x \leq 120\%$ and $> 120\%$.

9.6 Efficacy analyses

This section addresses the analyses to be conducted on the primary, secondary and exploratory

efficacy variables.

9.6.1 Analysis methods

9.6.1.1 Cochran-Mantel-Haenszel tests (CMH)

Dichotomous efficacy endpoints (e.g. HiSCR75 responder) will be analyzed using Cochran-Mantel-Haenszel (CMH) test stratified by Hurley Stage status (II and III) and by prior biologic use (yes and no) at baseline.

The CMH test will be used to assess the association between the event of interest (e.g. HiSCR75 responder) and the treatment group of interest (e.g. sonelokimab 120 mg or sonelokimab 240 mg) while providing a summary measure of the association, the odds ratio (OR) (active treatment group/placebo), after taking into account strata effects. This procedure estimates the OR within each stratum and then estimates a weighted average of OR across the different strata of the confounding factor as a combined measure.

The OR is a ratio of two odds, the odds of the event of interest in individuals exposed to active treatment and the odds of the event of interest in individuals exposed to placebo. The OR can be interpreted as:

- An $OR=1$ indicates that the odds of having the event of interest (e.g. HiSCR75 responder) is the same in the exposed and unexposed groups,
- An $OR>1$ indicates that the odds of having the event of interest (e.g. HiSCR75 responder) is greater in exposed group than in the unexposed group,
- An $OR<1$ indicates that the odds of having the event of interest (e.g. HiSCR75 responder) is lower in exposed group than in the unexposed group,

Using the procedure, the null hypothesis to be tested is that the OR is equal to one versus the OR is not equal to one.

The number and percentage of responders, the OR and corresponding 95% confidence interval (CI) will be presented along with the associated p-value. The 95% CI will be used to estimate the precision of the OR. In the case of having no responders in one group, the estimated p-value will be presented, but the odds ratio and 95% CI will not be displayed.

In the case of the primary analysis, the Breslow Day test will be utilized to test the null hypothesis that the odds ratios for the strata are all equal, as such the resultant Breslow Day p-value will also be presented. The response rates, odds ratios and 95% CIs will also be presented for each of the 4 strata to allow investigation of any differences between strata if the Breslow Day test indicates that the odds ratio is not consistent across the strata.

The common risk difference, across the strata, and associated 95% CI will also be presented for dichotomous efficacy endpoints. The common risk difference and 95% CI will be calculated using the Mantel-Haenszel Estimate of the common risk difference.

9.6.1.2 Mixed Model Repeated Measures (MMRM)

Continuous efficacy endpoints (e.g. IHS4 score) assessed over time will be analyzed using an MMRM model including treatment group, visit, Hurley stage status, prior biologic use, associated baseline measurements as well as the visit by treatment interaction as fixed effects, and participant as random effect. An unstructured covariance matrix will be assumed, unless the model does not converge, in which case the covariance matrix (e.g. variance components, autoregressive, compound symmetry, and Toeplitz) will be decided upon model convergence status and the Akaike information criterion. The degrees of freedom will be approximated using Kenward Rogers approximation.

The MMRM will be used to test for differences in the mean of the endpoint of interest (e.g. change from baseline in IHS4 at week 12) between groups of interest (e.g. each active treatment group versus placebo).

From the model, the adjusted least-squares (LS) means in each group and the difference in LS means for the group comparisons will be estimated together with 95% CIs for the differences and p-values for the pairwise comparisons.

9.6.1.3 Analysis of Covariance (ANCOVA)

Continuous efficacy endpoints (e.g. WPAI score) assessed only at two times points (e.g. baseline and week 12) will be analyzed using ANCOVA including treatment group, associated baseline measurement, Hurley stage status and prior biologic use as fixed effects.

The ANCOVA will be used to test for differences in the mean of the endpoint of interest (e.g. change from baseline in WPAI at week 12) between groups of interest (e.g. each active treatment group versus placebo).

From the model, the adjusted least-squares (LS) means in each group and the difference in LS means for the pairwise comparisons will be estimated together with 95% CIs for the differences and p-values for the pairwise comparisons.

9.6.1.4 Logistic Regression Model

As a sensitivity analysis, the primary will also be analyzed using a logistic regression model including treatment group, Hurley Stage status, prior biologic and baseline AN count use as fixed effects.

The logistic regression will be used, as a sensitivity analysis, to test association between the event of interest (e.g. HiSCR75 responder) and the treatment group of interest (e.g. sonelokimab 120 mg or sonelokimab 240 mg) while providing a summary measure of the association, the odds ratio (OR) (active treatment/placebo), after adjusting for or taking into account confounding. The OR can be interpreted as described in section 9.6.1.1.

The number and percentage of responders, the OR and corresponding 95% confidence interval (CI) will be presented along with the p-value.

9.6.1.5 Multiple Imputation

As a sensitivity analysis of the primary and key secondary dichotomous efficacy endpoint HiSCR75 and HiSCR50, multiple imputation of missing data will be performed and analyzed using logistic regression. Multiple imputation uses a simulation based approach, replacing missing values from multiple draws from conditional distributions given the observed data and covariates. Assuming an arbitrary missing data pattern the missing values for each component will be imputed separately based on a fully conditional specification (FCS) approach, this method uses all variables in the imputation model for any other variable.

The components of the endpoints; number of inflammatory nodules, number of abscesses and number of draining fistulae will be imputed separately. HiSCR50/75 will be derived for each subject and imputation using the imputed values of number of inflammatory nodules, number of abscesses and number of draining fistulae.

The imputations will be done using the corresponding baseline value, Hurley stage and prior biologic use at baseline as covariates. If there are any missing values in those variables they will also be imputed, in the model a single component at a specific visit, will consider all previous visits as covariates. The full conditional specification regression model will be used for continuous variables and the number of imputations will be 100 with 50 burn in imputations and the seed for all imputations will be 422754. For all visit results a minimum value of 0 will be added to the imputation. If the imputation does not input values within the specified range then the number of maximum iterations by step will be increased accordingly.

For each completed imputed dataset, the necessary derived variables will be computed using logistic regression as described in section 9.6.1.4 and the results obtained will be combined into one multiple imputation inference using proc MIANALYZE.



9.6.1.6 Multiplicity

The following multiplicity adjustments will be used to control the overall type I error rate at 0.05 for the primary endpoint and key secondary endpoints in Part A. All sensitivity analyses will be regarded as exploratory and no formal adjustment for multiplicity will be made.

Primary endpoint

To adjust for multiple testing and to control the family-wise error rate, the Bonferroni-Holm method will be applied. The Bonferroni-Holm procedure will work as follows:

- 1) Let p_{SLK120mg} and p_{SLK240mg} be the p-value resulting from the comparison between sonelokimab 120 mg vs placebo and sonelokimab 240 mg vs placebo groups respectively.
- 2) If the smallest p-value is greater than or equal to $\alpha/2 = 0.025$, the procedure is stopped and no p-values are significant.

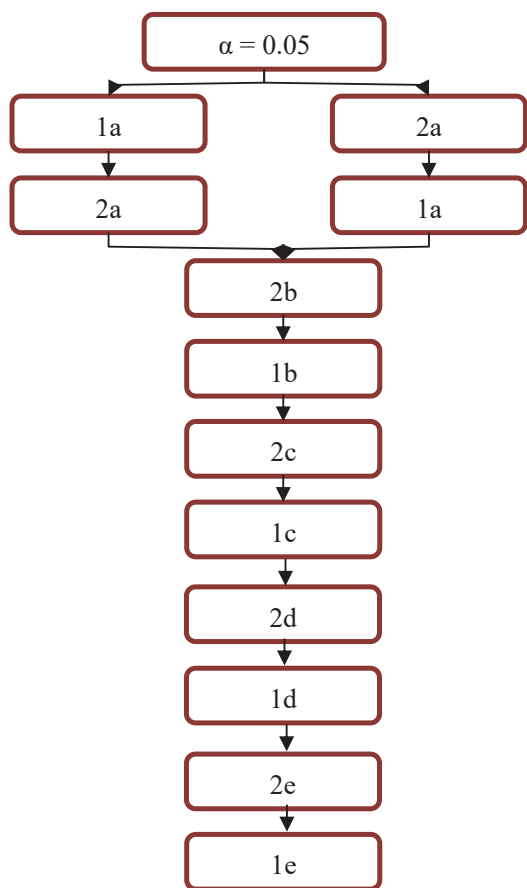
- 3) If the smallest p-value is declared significant (i.e p-value less than $\alpha/2 = 0.025$), then compare the largest p-value with $\alpha = 0.05$. If the p-value is less than 0.05, then declared as significant. Otherwise not significant.

Key Secondary Endpoints

To adjust for multiple testing and to control the family-wise error rate, a serial gatekeeping testing strategy will be used. It works as follows:

- 1) The key secondary endpoints family will be analyzed as confirmatory only if both previous tests in the primary endpoint family have been successful, i.e., the comparisons of both sonelokimab 120 mg and sonelokimab 240 mg groups vs placebo on the primary endpoint demonstrated a significant difference. Using this strategy, an alpha level of 0.05 will be used for the confirmatory analysis of key secondary endpoints.
- 2) The family of the key secondary endpoints will be analyzed in a fixed sequence testing procedure to control for multiplicity, within the ordered key secondary endpoints as listed:
 - Proportion of participants achieving HiSCR50 at Week 12;
 - Change from baseline in International HS Severity Score System (IHS4) at Week 12;
 - Proportion of participants achieving a Dermatology Life Quality Index (DLQI) total score of ≤ 5 at Week 12;
 - Proportion of participants achieving at least 30% reduction and at least 1-unit reduction from Baseline in Numerical Rating Scale (NRS30) in Patient's Global Assessment of Skin Pain (PGA skin pain) at Week 12 among participants with Baseline NRS ≥ 3 .
- 3) Sonelokimab 240 mg will be compared to placebo first, sonelokimab 120 mg to placebo second. Each test will be considered confirmatory only if the previous steps were successful. An alpha level of 0.05 will be used for the confirmatory analysis for all steps until the sequence breaks.

The following diagram demonstrates the control of the overall type I error rate, where 1 = sonelokimab 120mg, 2 = sonelokimab 240mg, a = primary endpoint and b-e the key secondary endpoints.



9.6.1.7 Treatment-by-center interaction analysis (multi-center study)

Unless otherwise specified, data from all investigational centers/sites will be pooled for analyses.

9.6.2 Analysis of primary efficacy endpoint

The primary analysis will be performed using the FAS. The PPS will be used for the primary endpoint and key secondary endpoints only. The FAS will be considered primary and PPS will be considered as supportive.

All sensitivity analyses will be performed using FAS population.

9.6.2.1 Hidradenitis Suppurativa Clinical Response 75 (HiSCR75)

The HiSCR75 will be assessed at Week 2, Week 4, Week 8 and Week 12, Week 14, Week 16, Week 20 and Week 24, however the primary endpoint is the proportion of HiSCR75 responders at Week 12.

HiSCR75 response criteria is defined as:

- HiSCR75 Responder: *participant who had a reduction from baseline in AN count $\geq 75\%$ and an absolute change in abscess ≤ 0 and an absolute change in draining fistula count ≤ 0 ;*
- HiSCR75 Non-responder: *Otherwise*

A Non-responder imputation method will be used for all missing data as detailed in section 8.3.1.

The null hypotheses to be tested are that the proportion of HiSCR75 responders at Week 12 are no different for the following comparisons:

- sonelokimab 120 mg versus placebo
- sonelokimab 240 mg versus placebo

To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1. This analysis will also be conducted for Weeks 2, 4 and 8.

The data for Part A+B (Week 2-24) will be presented by FAS. The HiSCR75 response rates by initial randomized treatment will be presented with the placebo and adalimumab arms not presented after Week 12. After Week 12 the placebo and adalimumab arms will be split into the Part B treatment for Week 14-24. Non-responder imputation will be used for all missing data, including participants who did not take part in Part B. No statistical analysis will be presented.

The data for Part B only will be presented by FAS-Part B: the participants will be split by Part A treatment and whether they were HiSCR75 response responders at Week 12 or not. The response rates for Weeks 12 to 24 by Part B treatment will be presented separately for Week 12 non-responders and responders. Non-responder imputation will be used for all missing data. No statistical analysis will be presented. These results will be also be presented in a line plot.

The following sensitivity analyses of the Part A data will be presented:

- an exact logistic regression model as detailed in section 9.6.1.4 with a non-response imputation for all missing data
- A GLMM as detailed in section 9.6.1.5 with missing data due to lack of efficacy or taking prohibited medication is imputed as a non-response
- An as Observed analysis using a pairwise stratified Cochran-Mantel-Haenszel test but with no imputation of missing data

Line plots for HiSCR75 response rate by week will also be provided, with week on the x-axis and HiSCR75 response rate on the y-axis will. Graphs for Part A (FAS), Part A+B (FAS) will be presented.

A HiSCR definition plot will be also presented for the Week 12 data. The x-axis will be the minimum required percentage reduction in AN count to be a response, the y-axis will show the percentage of responders and the plot will include one stepped line per treatment group. Subjects with missing data at Week 12 will be included as non-responders with an AN count percentage change of 0 for all

panels to ensure the plot matches the primary analysis. The plot will show four panels to illustrate the components of the HiSCR definition. The four panels will show the percentage of responders:

1. By AN count percentage change only
2. By AN count percentage change & no increase in abscesses
3. By AN count percentage change & no increase in draining fistulae
4. By AN count percentage change & no increase in abscesses and draining fistulae

9.6.3 Analysis of key secondary efficacy endpoints

For the key secondary endpoints, summaries and analyses will be performed using both the FAS and PPS. The FAS will be considered primary and PPS will be considered supportive.

In addition, for the key secondary dichotomous endpoints, a sensitivity analysis will be performed using FAS population.

9.6.3.1 Hidradenitis Suppurativa Clinical Response 50 (HiSCR50)

A key secondary endpoint is the proportion of HiSCR50 responders at Week 12.

HiSCR50 response criteria is defined as:

- HiSCR50 Responder: *participant who had a reduction from baseline in AN count $\geq 50\%$ and an absolute change in abscess ≤ 0 and an absolute change in draining fistula count ≤ 0 ;*
- HiSCR50 Non-responder: *Otherwise*

The same analysis as per the primary efficacy endpoint will be conducted.

9.6.3.2 International Hidradenitis Suppurativa Severity Scoring System (IHS4)

A key secondary endpoint is change from baseline in IHS4 at Week 12.

The IHS4 is a validated tool to dynamically assess HS severity. Determination of IHS4 requires counting of nodules, abscesses and draining fistulas/sinus tracts (tunnels). The IHS4 severity score will be calculated using the following formula:

$$\text{IHS4 Severity Score} = (\text{number of inflammatory nodules} \times 1) + (\text{number of abscesses} \times 2) + (\text{number of draining tunnels} \times 4);$$

A total score of 3 or less signifies mild, 4 to 10 signifies moderate and 11 or higher signifies severe disease. All three parameters need to be present otherwise the IHS4 Severity Score will be set to missing.

The IHS4 will be assessed at Baseline, Week 2, Week 4, Week 8, Week 12, Week 14, Week 16, Week 20 and Week 24.

The null hypotheses to be tested are that the mean change from baseline in IHS4 at Week 12 are not

different for the following comparisons:

- sonelokimab 120 mg versus placebo
- sonelokimab 240 mg versus placebo

To test these hypotheses, a MMRM will be used, as detailed in section 9.6.1.2, including the Baseline, Week 2, Week 4, Week 8 and Week 12 visits.

IHS4-55 is defined as:

- IHS4-55 Responder: *participant who had a reduction from baseline in IHS4 $\geq 55\%$;*
- IHS4-55 Non-responder: *Otherwise*

A non-responder imputation will be applied to IHS4-55 for all missing data as detailed in section 8.3.1.

IHS4-55 will be analysed using the same method as the primary endpoint, a pairwise stratified Cochran-Mantel-Haenszel test as detailed in section 9.6.1.1.

Two plots of the results will be presented:

1. The lsmean change from baseline for each arm until Week 12. This plot will match the primary analysis of IHS4.
2. The raw mean change from baseline for each arm until Week 24.

An IHS4 definition plot will be presented for the Part A data. The x-axis will be the Part A visit, the y-axis will show the mean change from baseline and the plot will include one line per treatment group. The plot will show four panels to illustrate the components of the IHS4 definition. The four panels will show the mean change from baseline for:

1. Inflammatory nodules
2. Abscesses
3. Draining fistulae
4. IHS4

9.6.3.3 Dermatology Life Quality Index (DLQI)

A key secondary endpoint is the proportion of $DLQI \leq 5$ responders at Week 12.

The DLQI is a dermatology-specific quality of life instrument designed to assess the impact of the disease on a participant's quality of life. It is a 10-item questionnaire that, in addition to evaluating overall quality of life, can be used to assess 6 different aspects that may affect quality of life: symptoms and feelings, daily activities, leisure, work or school performance, personal relationships, and treatment (See [Appendix A](#)).

The scoring of each question is as follows:

Very much = 3

A lot = 2

A little = 1

Not at all = 0

Not relevant = 0

Question 7: 'prevented work or studying' = 3

The DLQI is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. If one question is missing this will be scored 0 and the scores will be summed and expressed as usual out of a maximum of 30. If two or more questions are left unanswered the questionnaire will not be scored.

A higher score indicates more severe disease. A score of ≤ 1 indicates no effect at all of disease on participant's health related quality of life, a score of 2-5 indicates a small effect on the participant's life and a reduction of 5 points or more in total DLQI score is considered clinically meaningful improvement.

The 6 domain scores are calculated by summing the score of each question as specifies below:

- Symptoms, feelings: 1, 2
- Daily activities: 3, 4
- Leisure: 5, 6
- Work/School: 7
- Personal relationships: 8, 9
- Treatment: 10

If any question is missing within a domain, the domains score will also be missing.

DLQI ≤ 5 response criteria is defined as:

- DLQI ≤ 5 Responder: *participant who had a DLQI total score of ≤ 5 ;*
- DLQI > 5 Non-responder: *Otherwise*

A Non-responder imputation method will be used for all missing data as detailed in section 8.3.1.

The DLQI ≤ 5 response criteria will be assessed at Baseline, Week 4, Week 8, Week 12, Week 16, Week 20 and Week 24. Any DLQI ≤ 5 responders at baseline will be excluded from the analysis.

The null hypotheses to be tested are that the proportion of DLQI ≤ 5 responders at Week 12 are not different for the following comparisons:

- sonelokimab 120 mg versus placebo
- sonelokimab 240 mg versus placebo

To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1. This analysis will also be conducted for Weeks 2, 4 and 8.

The same sensitivity analyses as for the primary endpoint will be presented.

9.6.3.4 Patient's Global Assessment (PGA) of Skin Pain

A key secondary endpoint is the proportion of PGA Skin Pain responders at Week 12 among participants with Baseline NRS ≥ 3 .

The PGA skin pain will be self-administered by the participant on a daily basis (See [Appendix B](#)).

The PGA Skin Pain - NRS in the past 24 hours will be used to assess pain “at its worst” due to HS in the last 24 hours. The NRS is a segmented numeric version of the VAS in which a respondent selects a whole number (0–10 integers) that best reflects the intensity of their pain ranging from 0 (no skin pain) to 10 (skin pain as bad as you can imagine). The PGA of Skin Pain - NRS will be completed by the participant using the eDiary device.

For the visit-wise analysis at Baseline, Week 1, Week 2, Week 3, ..., up to Week 24, the weekly averages score will be calculated based on the available daily scores from the 7 days on or prior to the nominal day of the visit. The average of the 7 days on or before the date of first dose of study drug injection will be used as Baseline. The weekly average will be considered missing if there were fewer than 4 diary entries.

The PGA Skin Pain response criteria are defined as:

- PGA Skin Pain Responder: *participant who achieves NRS score of 30 (NRS30) defined as had a reduction from baseline $\geq 30\%$ and an absolute reduction from baseline ≥ 1 in NRS in PGA Skin Pain;*
- PGA Skin Pain Non-responder: *Otherwise*

A Non-responder imputation method will be used for all missing data as detailed in section 8.3.1.

For analysis proposes, only the following times point will be considered: Baseline, Week 1 - Week 12, all other timepoints will be summarised only. Any participants with Baseline NRS ≥ 3 will be excluded from the analysis.

The null hypotheses to be tested are that the proportion of PGA Skin Pain responders at Week 12 are not different for the following comparisons:

- sonelokimab 120 mg versus placebo
- sonelokimab 240 mg versus placebo

To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1.

The same sensitivity analyses as for the primary endpoint will be presented.

9.6.4 Analysis of other secondary efficacy endpoints

For the other secondary endpoints, no confirmatory statistical testing will be performed, but statistical analyses may be used for exploratory proposes for Part A.

FAS for Part A summarises and analyses and FAS-B for Part B and Part A+B summaries will be used for all other secondary endpoints analysis.

9.6.4.1 Hidradenitis Suppurativa Clinical Response (HiSCR50/75/90/100)

Other secondary endpoints are:

- Proportions of HiSCR50 responders over time.
- Proportions of HiSCR75 responders over time.
- Proportions of HiSCR90 responders over time.
- Propotions of HiSCR100 responders over time.

HiSCR50 and HiSCR75 response criteria are defined in section 9.6.3.1 and in section 9.6.2.1 respectively.

HiSCR90 response criteria is defined as:

- HiSCR90 Responder: *participant who had a reduction from baseline in AN count $\geq 90\%$ and an absolute change in abscess ≤ 0 and an absolute change in draining fistula count ≤ 0 ;*
- HiSCR90 Non-responder: *Otherwise*

HiSCR100 response criteria is defined as:

- HiSCR100 Responder: *participant who had a reduction from baseline in AN count $\geq 100\%$ and an absolute change in abscess ≤ 0 and an absolute change in draining fistula count ≤ 0 ;*
- HiSCR100 Non-responder: *Otherwise*
-

Missing data imputation method will be used as detailed in section 8.2.1.

The HiSCR50/75/90/100 will be assessed at Week 2, Week 4, Week 8, Week 12, Week 14, Week 16, Week 20 and Week 24.

The exploratory null hypotheses to be tested are that the proportion of HiSCR50/75/90/100 responders at each time point in Part A are not different for the comparisons of interest for part A, analysis as described in section 9.1. For exploratory proposes, the comparison of interest will be:

- sonelokimab 120 mg versus placebo
- sonelokimab 240 mg versus placebo

To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1.

HiSCR50/75/90/100 responders will be also summarized descriptively by treatment group and visit for Part A, Part B and Part A+B as described in section 9.1.

9.6.4.1 Dermatology Life Quality Index (DLQI)

Other secondary endpoints include:

- Proportion of $DLQI \leq 1$ responders over time.
- Proportion of $DLQI \leq 5$ responders over time.
- Proportion of DLQI improvement responders over time among participants with $DLQI \geq 4$ at baseline.

DLQI Total score calculation and $DLQI \leq 5$ response criteria is defined in section 9.6.3.3.

$DLQI \leq 1$ response criteria is defined as:

- $DLQI \leq 1$ Responder: *participant who had a DLQI total score of ≤ 1 ;*
- $DLQI \leq 1$ Non-responder: *Otherwise*

DLQI improvement response criteria is defined as:

- DLQI improvement Responder: *participant who achieving a DLQI improvement (reduction) of ≥ 4 points;*
- DLQI improvement Non-responder: *Otherwise*

A Non-responder imputation method will be used for all missing data as detailed in section 8.3.1.

The $DLQI \leq 1/\leq 5$ /improvement will be assessed at Baseline, Week 2, Week 4, Week 8, Week 12, Week 14, Week 16, Week 20 and Week 24. Any subjects with $DLQI \leq 1/\leq 5$ at baseline will be excluded from the $\leq 1/\leq 5$ analyses, respectively.

The exploratory null hypotheses to be tested are that the proportion of $DLQI \leq 1/\leq 5$ /improvement responders at each time point in Part A are not different for the comparisons of interest described in section 9.6.4.1. To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1.

The $DLQI \leq 1/\leq 5$ /improvement responders, and the DLQI total score as well as 6 domain scores and corresponding change from baseline will also be summarized descriptively by treatment group and visit under Part A, Part B and Part A+B as described in section 9.1.

9.6.4.2 Patient's Global Assessment of Skin Pain NRS (PGA Skin Pain)

Other secondary endpoints include:

- Proportion of PGA Skin Pain responders over time
- Proportions of PGA Skin Pain responders over time, with participants using permitted analgesic therapy for $\geq 4/7$ days assigned as non-responders

The PGA Skin Pain NRS and response criteria are defined in section 9.6.3.4. In the second analysis, participants who have taken permitted analgesic therapy for at least 4 days of a given week will be set to non-responders for that week. This is intended to avoid assigning subjects with an improvement in pain due to analgesic therapy as responders.

A Non-responder imputation method will be used for all missing data as detailed in section 8.3.1.

For analysis proposes, only the following times point will be considered: Baseline, Week 1-Week 12, all other timepoints will be summarised only.

The exploratory null hypotheses to be tested are that the proportion of PGA Skin Pain responders at each time point are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for part A. To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1.

The PGA Skin Pain response rate, the PGS Skin Pain NRS and change from baseline will also be summarized descriptively by treatment group and visit as described in section 9.1.

9.6.4.3 Lesion Count Endpoints

Other secondary endpoints include:

- Proportion of participants achieving AN count of 0, 1, or 2, over time;
- Proportion of AN50/75/90/100 responders over time;
- Proportion of abscesses responders over time among those participants with abscesses at baseline;
- Proportion of draining fistulas/tunnels responders over time among those participants with draining fistulas/tunnels at baseline
- Proportion of inflammatory nodules responders over time among those participants with inflammatory nodules at baseline
- Absolute change from baseline in the number of abscesses over time;
- Absolute change from baseline in number of draining fistulas/tunnels over time;
- Absolute change from baseline in number of inflammatory nodules over time;
- Proportion of participants with at least one flare (flare is defined as $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to baseline)

Lesion counts will be assessed by the investigator and will consist in count of the number of abscesses (A), inflammatory nodules (N), non-inflammatory nodules, draining fistulas (DF), non-draining fistulas, HS scars, and other lesions (papules and pustules) by anatomical area.

AN count is defined as the total abscess and inflammatory nodule count. Assessments on both parameters need to be available otherwise the AN count will be set to missing.

AN count of 0, 1 or 2 response criteria is defined as:

- Responder: Participants with an AN count ≤ 2
- Non-responder: AN count > 2

AN50/75/90/100 response criteria is defined as:

- AN50/75/90/100 Responder: *Participant who had a reduction of $\geq 50\%/75\%/90\%/100\%$ respectively in total AN count from baseline*
- AN50/75/90/100 Non-responder: *Otherwise*

Abscesses response criteria is defined as:

- Abscesses Responder: *participant who had abscesses count = 0*
- Abscesses Non-responder: *Otherwise*

Draining fistulas/tunnels response criteria is defined as:

- Draining fistulas/tunnels Responder: *participant who had draining fistulas/tunnels count = 0*
- Draining fistulas/tunnels Non-responder: *Otherwise*

Inflammatory nodules response criteria is defined as:

- Inflammatory nodules Responder: *participant who had inflammatory nodules count = 0*
- Inflammatory nodules Non-responder: *Otherwise*

A Non-responder imputation method will be used for all missing data for response criteria as detailed in section 8.3.1.

The lesion count endpoints will be assessed at each visit and will be analysed at Week 2, Week 4, Week 8, Week 12. All abscesses, draining fistulas/tunnels and inflammatory nodules responders at baseline will be excluded from the analysis.

For dichotomous endpoints (AN50/75/90/100, abscesses, draining fistulas/tunnels, inflammatory nodules response, flares), the exploratory null hypotheses to be tested are that the proportion of responders at each time point are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for part A. To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1.

For continuous endpoints (change from baseline in the number of abscesses, draining fistulas/tunnels and inflammatory nodules), the exploratory null hypotheses to be tested are that the mean change from baseline over time are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for Part A. To test these hypotheses, a MMRM will be used as detailed in section 9.6.1.2.

Additionally, the number of abscesses, inflammatory nodules, non-inflammatory nodules, draining fistulas (DF) and AN Count will be also summarized descriptively by treatment group and visit for Part A, Part B and Part A+B as described in section 9.1.

9.6.4.4 Patient Global Impression - Severity of Illness (PGI-S)

Other secondary endpoints include:

- Proportion of PGI-S improvement Responder over time
- Proportions of PGI-S ‘minimal symptoms’ or ‘absent symptoms’ Responder over time

The PGI-S is a self-reported single item to evaluate all aspects of participants’ HS symptoms (See Appendix C). Participants are asked to choose the response that best describes the severity of their HS symptoms over the past week. The question is rated on a 7-point scale ranging from 0 to 6 (0 = absent; 1 = minimal; 2 = mild; 3 = moderate; 4 = moderately severe; 5 = severe; 6 = very severe). PGI-S improvement response criteria is defined as:

- PGI-S improvement Responder: *participant who had a change of ≥ 1 score points from baseline in PGI-S;*
- PGI-S improvement Non-responder: *Otherwise*

PGI-S ‘minimal symptoms’ or ‘absent symptoms’ response criteria is defined as:

- PGI-S ‘minimal symptoms’ or ‘absent symptoms’ Responder: *participant who had a PGI-S score = 0 or 1;*
- PGI-S ‘minimal symptoms’ or ‘absent symptoms’ Non-responder: *Otherwise*

A non-responder imputation method will be used as detailed in section 8.2.1.

The PGI-S will be self-administered by the participant at Baseline, Week 4, Week 12, Week 16 and Week 24. All PGI-S ‘minimal symptoms’ or ‘absent symptoms’ responder at baseline will be excluded from the analysis.

The exploratory null hypotheses to be tested are that the proportion of PGI-S responders at each time point in Part A are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for Part A. To test these hypotheses, a pairwise stratified Cochran-Mantel-Haenszel test will be used as detailed in section 9.6.1.1.

The derived PGI-S responders, the PGI-S total score and change from baseline will also be summarized descriptively by treatment group and visit as described in section 9.1.

9.6.4.5 Hidradenitis Suppurativa Physician’s Global Assessment (HS-PGA)

Other secondary endpoints include change from baseline in HS-PGA.

The HS-PGA is a validated 6-point scale that is used to measure improvement in inflammatory nodules, abscesses, and draining fistulas (See [Appendix D](#)). The HS-PGA scale ranges from “Clear” (no inflammatory or non-inflammatory nodules) to “Very severe” (more than 5 abscesses or draining

fistulas). The scores will be derived on a participant-level basis (i.e., across all anatomical regions) based on the following rules in below table.

Score	Rating	Description
0	Clear	0 abscesses, 0 draining fistulas, 0 inflammatory nodules, and 0 non-inflammatory nodule
1	Minimal	0 abscesses, 0 draining fistulas, 0 inflammatory nodules, and presence of non-inflammatory nodules
2	Mild	0 abscesses, 0 draining fistulas, and 1-4 inflammatory nodules or 1 abscess or draining fistula and 0 inflammatory nodules
3	Moderate	0 abscesses, 0 draining fistulas, and ≥ 5 inflammatory nodules or 1 abscess or draining fistula and ≥ 1 inflammatory nodules or 2 to 5 abscesses or draining fistulas and < 10 inflammatory nodules
4	Severe	2–5 abscesses or draining fistulas and ≥ 10 inflammatory nodules
5	Very Severe	> 5 abscesses or draining fistulas

The HS-PGA will be assessed by the investigator or a qualified designee at Baseline, Week 4, Week 12, Week 16, Week 20 and Week 24.

The exploratory null hypotheses to be tested are that the mean change from baseline over time are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for Part A. To test these hypotheses, a MMRM will be used as detailed in section 9.6.1.2.

The HS-PGA score as well as the corresponding change from baseline will be also summarized descriptively by treatment group and visit for Part A, Part B and Part A+B as described in section 9.1.

9.6.4.6 Hidradenitis Suppurativa Quality of Life (HiSQOL)

Other secondary endpoints include change from baseline in HiSQOL

The HiSQOL is a 17-item HS specific HRQoL questionnaire with a 7-day recall period (See [Appendix E](#)). The 17 items use a 5-point Likert-type scale (0 = “Not at all”, 1 = “Slightly”, 2 = “moderately”, 3 = “very much”, 4 = “Extremely”) and included four symptom items (pain, itch, drainage, odor), eight activities-adaptation items (walking, exercising, sleeping, washing yourself, getting dressed, what you wear, ability to work/study and sexual activity difficult), and five psychosocial items (down or depressed, embarrassed, anxious or nervous, concentration, sexual

desire). The item scores are summed to create a total ranging from 0 to 68, with higher scores indicating more severe impact on HRQoL. The sub-scale score for symptoms range from 0 to 16. The symptoms domain includes questions 8-11 (Pain, Itch, Drainage, Odor). The psychosocial sub-scale ranges from 0 to 20 and includes questions 6, 12-14, 16 (Concentrating, Depressed, Embarrassed, Anxious, Sexual desire). The activities-adaptations sub-scale ranges from 0 to 32 and includes questions 1-5, 7, 15, 17 (walking, exercising, sleeping, washing, getting dressed, what you wear, sexual activity, ability to work).

The HiSQOL will be self-administered by the participant at Baseline, Week 4, Week 12, Week 16 and Week 24.

The exploratory null hypotheses to be tested are that the mean change from baseline over time are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for Part A. To test these hypotheses, a MMRM will be used for both the total score change from baseline and the 3 domain scores' change from baseline as detailed in section 9.6.1.2.

The HiSQOL total score as well as the 3 domain scores and corresponding change from baseline will also be summarized descriptively by treatment group and visit for Part A, Part B and Part A+B as described in section 9.1.

9.6.4.1 Work Productivity and Activity Impairment Questionnaire (WPAI)

Other secondary endpoints include:

- Change from baseline of Overall Work Impairment (OWI) score from WPAI
- Change from baseline of absenteeism score from WPAI
- Change from baseline of presenteeism score from WPAI
- Change from baseline of activity impairment score from WPAI

The Work Productivity and Activity Impairment (WPAI) is a questionnaire used to evaluate lost productivity due to any physical or emotional problem or a symptom related with the HS. (See [Appendix G](#)). The questionnaire is self-administered and includes 6 questions which need to be filled by the participant: Q1 = currently employed; Q2 = hours missed due to health problems; Q3 = hours missed other reasons; Q4 = hours actually worked; Q5 = degree health affected productivity while working (using a 0 to 10 Visual Analogue Scale (VAS)); Q6 = degree health affected productivity in regular unpaid activities (VAS). The recall period for the questions 2 to 6 is seven days.

Four main scores can be generated from the WPAI and expressed in percentages by multiplying the following scores by 100 (any missing items will result in a missing score):

- 1) Absenteeism – percent work time missed due to health = $Q2/(Q2 + Q4)$ for those who were currently employed;
- 2) Presenteeism – percent impairment while working due to health = $Q5/10$ for those who were currently employed and actually worked in the past seven days;
- 3) Overall work impairment (OWI) – percent overall work impairment due to health $Q2/(Q2 + Q4) + ((1 - Q2/(Q2 + Q4)) \times (Q5/10))$ for those who were currently employed;
- 4) Activity impairment – percent activity impairment due to health $Q6/10$ for all respondents

The WPAI will be self-administered by the participant at Baseline, Week 12 and Week 24.

The exploratory null hypothesis to be tested is that the mean change from baseline at Week 12 is not different for the comparisons of interest described in section 9.6.4.1. To test this hypothesis, a ANCOVA will be used as detailed in section 9.6.1.3.

The four main WPAI scores as well as the corresponding change from baseline will be also summarized descriptively by treatment group and visit as described in section 9.1.

9.6.4.2 Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue)

Other secondary endpoints include:

- Change from baseline in Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) score

FACIT-Fatigue is a 13-item self-assessment of fatigue and the effect on daily activities and function over the previous 7 days (See [Appendix H](#)). The 13 items use a 5-point Likert-type scale (0 = “Not at all”, 1 = “A little”, 2 = “Somewhat”, 3 = “Quite a bit”, 4 = “Very much”). The FACIT-Fatigue assessment score range between 0-52 with 0 corresponding to the highest level of fatigue and 52 corresponding to the lowest level of fatigue. The scores for the 11 items where a higher score indicates more fatigue are reversed (4-score) before being summed together with the other 2 items to give the total. If any items are missed, then the total score is adjusted by multiplying by 13 and dividing by the number of items answered.

The FACIT-Fatigue will be self-administered by the participant at Baseline, Week 4, Week 12, Week 16 and Week 24.

The exploratory null hypotheses to be tested are that the mean change from baseline over time are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for Part A. To test these hypotheses, a MMRM will be used including the Part A visits as detailed in section 9.6.1.2.

The FACIT-F score as well as the corresponding change from baseline will also be summarized descriptively by treatment group and visit for Part A, Part B and Part A+B as described in section 9.1.

9.6.4.3 European Quality of Life 5 Dimensions Questionnaire-3-Level (EQ-5D-3L)

Other secondary endpoints include:

- Change from baseline in EuroQol five-dimension questionnaire 3-Level (EQ-5D-3L) Visual Analog Scale (VAS) over time
- Change from baseline in EQ-5D-3L index scores over time

The EQ-5D comprises 5 questions on mobility, self-care, pain, usual activities, and psychological status with 3 possible answers for each item (1=no problem, 2=moderate problem, 3=severe problem). A summary index with a maximum score of 1 can be derived from these 5 dimensions by conversion with a table of scores. The UK⁴ index value calculator will be used. The score from each domain will be converted to a deduction from full health and aggregated to form the index score. The maximum score of 1 indicates the best health state, by contrast with the scores of individual questions, where higher scores indicate more severe or frequent problems:

Full health (11111) =	1.000
Mobility level 1 (No problems):	0
Mobility level 2 (Moderate):	-0.069
Mobility level 3 (Severe):	-0.314

Self-care level 1 (No problems):	0
Self-care level 2 (Moderate):	-0.104
Self-care level 3 (Severe):	-0.214
Usual activities 1 (No problems):	0
Usual activities 2 (Moderate):	-0.036
Usual activities 3 (Severe):	-0.094
Pain/discomfort 1 (No pain):	0
Pain/discomfort 2 (Moderate):	-0.123
Pain/discomfort 3 (Severe):	-0.386
Anxiety/depression 1 (No Anxiety)	0
Anxiety/depression 2 (Moderate):	-0.071
Anxiety depression 3 (Severe):	-0.236
For any dysfunctional (score >1 in any domain):	- 0.081
For any severe domain (score = 3 in any domain):	-0.269

In addition, there is a VAS to indicate the general health status with 100 indicating the best health status. The EQ-5D VAS is a measure of overall self-rated health status, used and analysed separately from the index score. The VAS ranges from 0 to 100, with higher scores indicative of better overall health.

The EQ-5D-3L will be assessed at Baseline, Week 4, Week 12, Week 16 and Week 24.

The exploratory null hypotheses to be tested are that the mean change from baseline over time are not different for the comparisons of interest described in section 9.6.4.1 and to be performed under Part A. To test these hypotheses, a MMRM will be used including the Part A visits as detailed in section 9.6.1.2.

The EQ-5D-3L index score and VAS score as well as the corresponding change from baseline will also be summarized descriptively by treatment group and visit for Part A, Part B and Part A+B analysis as described in section 9.1.

9.6.4.4 Clinical laboratory evaluations

Other secondary endpoints include:

- hs C-reactive protein (hs-CRP)
- Neutrophil to Lymphocyte Ratio (NLR);

CRP and NLR values are assessed at Baseline, Week 2, Week 4, Week 8, Week 12, Week 14, Week 16, Week 20 and Week 24.

The exploratory null hypotheses to be tested are that the geometric mean ratio to baseline over time are not different for the comparisons of interest described in section 9.6.4.1 and to be performed for

Part A. To test these hypotheses, the data will be log transformed prior to using a MMRM including as detailed in section 9.6.1.2. The results will be back-transformed to be presented in terms of adjusted least-squares (LS) geometric means, and ratios in LS geometric means together with 95% CIs and p-values.

If the data suggest that either endpoint is normally distribution rather than log-normally distributed then the analysis will be performed on the change from baseline without log-transformation.

The CRP and NLR assessments as well as the corresponding ratio to baseline will also be summarized descriptively by treatment group and visit for Part A, Part B and Part A+B as described in section 9.1.

9.7 PK and Immunological Analyses

For PK and Immunological Analyses, no confirmatory statistical testing will be performed, but statistical tests may be used for exploratory proposes only. IG and PK analysis sets for Part A will be used.

9.7.1 Immunological Analyses

Antibodies to sonelokimab will be evaluated in serum samples collected for Arms 1 & 2 (120mg & 240mg sonelokimab) at Baseline, Week 2, Week 4, Week 8, Week 12, Week 16, Week 20, Week 24 and Week 28 and for Arms 3 & 4 (placebo and adalimumab) at the last visit prior to Sonelokimab (Week 12), Week 16, Week 20, Week 24 and Week 28. At dosing visits the samples will be collected prior to dose administration.

Up to three separate ADA types may be assessed: ADAs that bind to sonelokimab ('ADA'); ADA to a modified version of sonelokimab ('mADA'), and ADA that neutralise the activity of sonelokimab ('NAb ADA'). Results for mADA will only be available for ADA baseline positive participants. NAb ADA results will only be available for the ADA positive samples.

Summary tables for each ADA type will be prepared to record where applicable the number and percentage of participants who are:

- ADA positive at baseline
- ADA positive at any visit post-baseline
- ADA positive at baseline and any visit post-baseline
- ADA positive post-baseline but not at baseline
- persistent positive (positive at more than one post-baseline visit where the first and last ADA-positive visits are separated by a period of 16 weeks or longer; or positive at last post-baseline assessment and not at baseline; or positive at a visit with less than 16 weeks before an ADA-negative last visit)
- transient positive (positive at only 1 post-baseline visit and not at baseline; or positive at two or more visits where the first and last ADA-positive visits are separated by a period less than 16 weeks and the last visit is negative)

The log10 titer values for the ADA and mADA (not NAb ADA) will be summarised by visit and treatment arm.

Additional analyses may be performed at the sponsor's discretion but they are beyond the scope of this SAP and will be specified in a separate document, if needed.

9.7.2 Pharmacokinetic Analyses

Blood samples will be collected for the measurement of trough serum sonelokimab levels at Baseline, Week 2, Week 4, Week 8, Week 12, Week 16, Week 20, Week 24 and Week 28. At dosing visits the samples will be collected prior to dose administration.

Serum concentrations of sonelokimab will be summarized by visit including n, number of concentrations below the quantifiable limit (nBLQ), mean, SD, geometric mean (GM), coefficient of variation (CV%), median, minimum, and maximum. All concentration values reported as no results (not collected or ND) will be treated as missing. For the calculation of concentration summaries, all concentrations below the quantifiable limit (BLQ) will be set to the BLQ value/2.

The PK measurements for sonelokimab will be listed by-participant, by visit and by treatment group as described in section 9.7.1. Individual values of concentrations and PK parameters will be presented with 3 significant digits.

Additional analyses may be performed at the sponsor's discretion but they are beyond the scope of this SAP and will be specified in a separate document, if needed.

9.9 Analysis of the ultrasound sub-study

At each visit during the sub-study, measurements will be made of maximal epidermal thickness at the target site and the maximum transverse dimension (in mm) at the widest part of the target tunnel. Three separate images of the target anatomical site will be captured at each session, and one measurement of each parameter from each image performed and recorded. The mean of the 3 measurements will be used in the data analysis.

Opacity within the lumen of the tunnel (representing accumulated matter) will be measured as present or absent (Yes/No).

Maximal Power Doppler intensity (as a representation of hyperemia) will be assessed in the region of the tunnel using a semi-quantitative scoring method from 1 to 3, with "1" indicating minimal focal or no intensity, "2" indicating focal high intensity or diffuse low intensity, and "3" indicating diffuse high intensity.

At each applicable study visit, including baseline (Week 0/Day 1), ultrasound measurements will always be performed before any clinical assessments of HS lesions are performed. All measurements will be made by site personnel experienced in the use of ultrasound. Whenever possible, the same operator will perform all measurements for an individual participant.

All analysis of the ultrasound sub-study will be conducted using the ultrasounds sub-study set and deemed exploratory.

In addition to the 4 treatment arms a combined Sonelokimab dose group will be presented.

Descriptive statistics will be presented for each visit and for the change from baseline for the 4 endpoints below:

- Epidermal thickness at the target lesion (mm);
- Maximum transverse dimension of the target tunnel (mm);
- Power Doppler intensity score;
- Proportion of subjects with a luminal opacity score of “No” among participants with a score of “Yes” at baseline.

The mean difference from placebo will be analyzed using an ANCOVA model as per section 9.6.1.3 for each active treatment group for each of the continuous endpoints at Week 12. Binary variables will be analysed using an exact logistic regression model as per section 9.6.1.4. Stratification factors will not be included in any models of the sub-study. These comparisons may be omitted if there are too few participants in each group.

Additional these analyses will be repeated using a combined sonelokimab group for the change from baseline for each of the endpoints at Week 12.

Additional exploratory analyses may be defined to explore the correlation of these endpoints with the endpoints for the main study.

9.10 Safety analyses

For all the safety endpoints, no formal statistical testing will be performed. Safety summarises will be presented separately for the Part A (SAF), Part B (SAF-Part B), and Part A+B (SAF)

For Part A, 5 groups will be considered:

- Sonelokimab 120 mg,
- Sonelokimab 240 mg,
- Placebo,
- Adalimumab
- Total sonelokimab

For Part B and follow-up period, 6 groups will be considered:

- Sonelokimab 120 mg,
- Sonelokimab 240 mg,
- Placebo-sonelokimab 120 mg,
- Placebo-sonelokimab 240 mg,
- Adalimumab-sonelokimab 240 mg groups,
- Total sonelokimab

For Part A+B, 10 groups will be considered:

- Placebo
- Adalimumab
- Placebo/sonelokimab 120mg
- Placebo/sonelokimab 240mg
- Adalimumab/sonelokimab 240mg
- Sonelokimab 120 mg,
- Sonelokimab 240 mg,
- Total Sonelokimab 120mg
- Total Sonelokimab 240mg
- Total Sonelokimab

9.10.1 Adverse events

All AEs will be classified by Primary System Organ Class (SOC) and Preferred Term (PT) according to MedDRA Version 25.0 or higher and will be graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 5.0.

AEs will be summarized by SOC and PT using counts and percentage (i.e., number and percentage of subjects with an AE). AEs will be sorted by decreasing frequency within each SOC and PT as per the overall column.

Where a subject has the same AE, based on PT reported multiple times, the subject will only be counted once at that PT level. Where a subject has multiple AEs within the same system organ class, the subject will only be counted once at that SOC level.

If a subject experiences the same AE more than once with different severity grades, then the event with the highest severity will be tabulated by maximum severity. In addition, AEs with a missing severity will be presented in the summary table under the category of “Missing” and will not be imputed.

TEAEs will be assigned to Part A, Part B and Part A+B using the event start date and time and will be defined as:

- Part A: events started on or after the first dose of study medication for Part A but not after Part B or events present prior to the first dose of study medication for Part A but increased in severity on or after dosing on Part A but not after Part B
- Part B: events started on or after the first dose of study medication for Part B or events present prior to the first dose of study medication for Part B but increased in severity on or after dosing on Part B
- Part A+B: events started on or after the first dose of study medication for Part A or events present prior to the first dose of study medication for Part A but increased in severity on or after dosing on Part A

Treatment-related AEs are those events with relationship to study drug “Related” as recorded on the case report form (CRF). If the relationship to study drug is missing, the AE will be considered as treatment related.

Separate summaries tables by treatment will be presented with the total number of events and number and percentage of participants with:

- All TEAEs;
- Grade 3/4/5 TEAEs
- TEAEs by maximum intensity
- Related TEAEs;
- TEAE leading to treatment discontinuation;
- Serious AEs (SAEs);
- Related SAEs;
- Fatal AEs;
- AESI.
- AEs by DSUR safety topics (Hypersensitivity reactions SMQ; Opportunistic Infections SMQ; Gastrointestinal nonspecific inflammation and dysfunctional conditions SMQ)

AEs by DSUR safety topics will be summarized by SMQ and PT using counts and percentage (i.e., number and percentage of subjects with an AE). AEs will be sorted by decreasing frequency within each SMQ and PT as per the overall column.

Where a subject has the same AE, based on PT reported multiple times, the subject will only be counted once at that PT level. Where a subject has multiple AEs within the same SMQ, the subject will only be counted once at that SMQ level.

An overall summary presenting the categories above will also be prepared by treatment group for Part A, Part B, and Part A+B. In addition, for all Part A+B summarizes, the total number of participant treatment years and exposure adjusted incidence rates [EAIR] will be provided as described in [Appendix J](#).

9.10.2 Clinical laboratory evaluations

All central laboratory (hematology, biochemistry, urinalysis) parameters will be collected as described in [Appendix I](#).

Observed and change from baseline values for clinical laboratory parameters will be summarized using descriptive statistics by treatment group and visit separately for Part A and Part B.. CTCAEv5.0 criteria are described in Appendix L. Subjects meeting any grade 3 criteria will also be listed.

In addition, the number and percentage of participants meeting the criteria for normal ranges in laboratory parameters will be summarized by treatment group separately for Part A and Part B.

Laboratory results together with the normal ranges will be listed, flagging values that are below or above the normal range. Serum/urine pregnancy tests assessments will be listed only.

Boxplots of the observed and change from baseline values for all clinical laboratory parameters will be presented by treatment group and visit for Part A and Part B separately. Reference lines for the normal ranges, that consistent across all participants being presented, will be displayed.

For subjects with an AE of grade 3 or higher the hematology and chemistry lab parameters for that subject will be presented in a panel plot with a panel for each lab parameter. The x-axis will be the study week and the y-axis will be the absolute lab value.

9.10.3 Vital signs

Vital sign parameters collected are weight (kg), systolic and diastolic blood pressure (SBP and DBP) (mmHg), heart rate (beats/min), respiratory rate (breaths/min) and temperature (°C).

Observed values and the corresponding change from baseline values will be summarized using descriptive statistics by treatment group and visit separately for Part A and Part B.

In addition, the number and percentage of participants meeting criteria for potentially clinically significant abnormalities in vital signs (SBP, DBP and heart rate) and body weight will be summarized by treatment group separately for Part A and Part B. The following table presents the Criteria for clinically relevant abnormalities.

Table: Criteria for clinically relevant abnormalities in Vital sign parameters

Parameter	Criteria
SBP (mmHg)	Value ≤ 90 or decrease ≥ 20 mmHg from baseline Value ≥ 180 or increase ≥ 20 mmHg from baseline
DBP (mmHg)	Value ≤ 50 or decrease ≥ 15 mmHg Value ≥ 105 or increase ≥ 15 mmHg
Hear Rate (beats/min)	Value ≤ 50 bpm or decrease from baseline ≥ 20 bpm Value ≥ 120 bpm or increase from baseline ≥ 20 bpm
Weight (kg)	$\geq 7\%$ decrease from baseline $\geq 7\%$ increase from baseline

9.10.4 Physical examinations

The following parameters and body systems will be examined and any abnormalities described general appearance, skin (presence of rash), HEENT (head, ears, eyes, nose, throat), lungs (auscultation), heart (auscultation for presence of murmurs, gallops, rubs), lower extremity exam,

abdomen (palpation and auscultation), neurologic (mental status, station, gait, reflexes, motor and sensory function, coordination) and lymph nodes.

All physical examination findings will be listed only.

9.10.5 Electrocardiograms

The standard 12-lead ECG parameters collected are: intervals (RR, PR, QRS, QT, and QT intervals corrected for heart rate using Bazett's and Fridericia's formulas [QTcB and QTcF, respectively]).

Observed values and the corresponding change from baseline values will be summarized using descriptive statistics at each scheduled assessment by treatment group separately for Part A and Part B. If ECG was done in triplicate at a visit then the average result will be taken for numerical scores, and the worst interpretation will be analyzed.

The overall ECG interpretation will be summarized by presenting the number and percentage of participants with "normal" and "Abnormal" by visit and treatment group for Part A and Part B. Shift tables from baseline (Normal/Abnormal) to each visit and to the worst post-baseline ECG interpretation will be summarized.

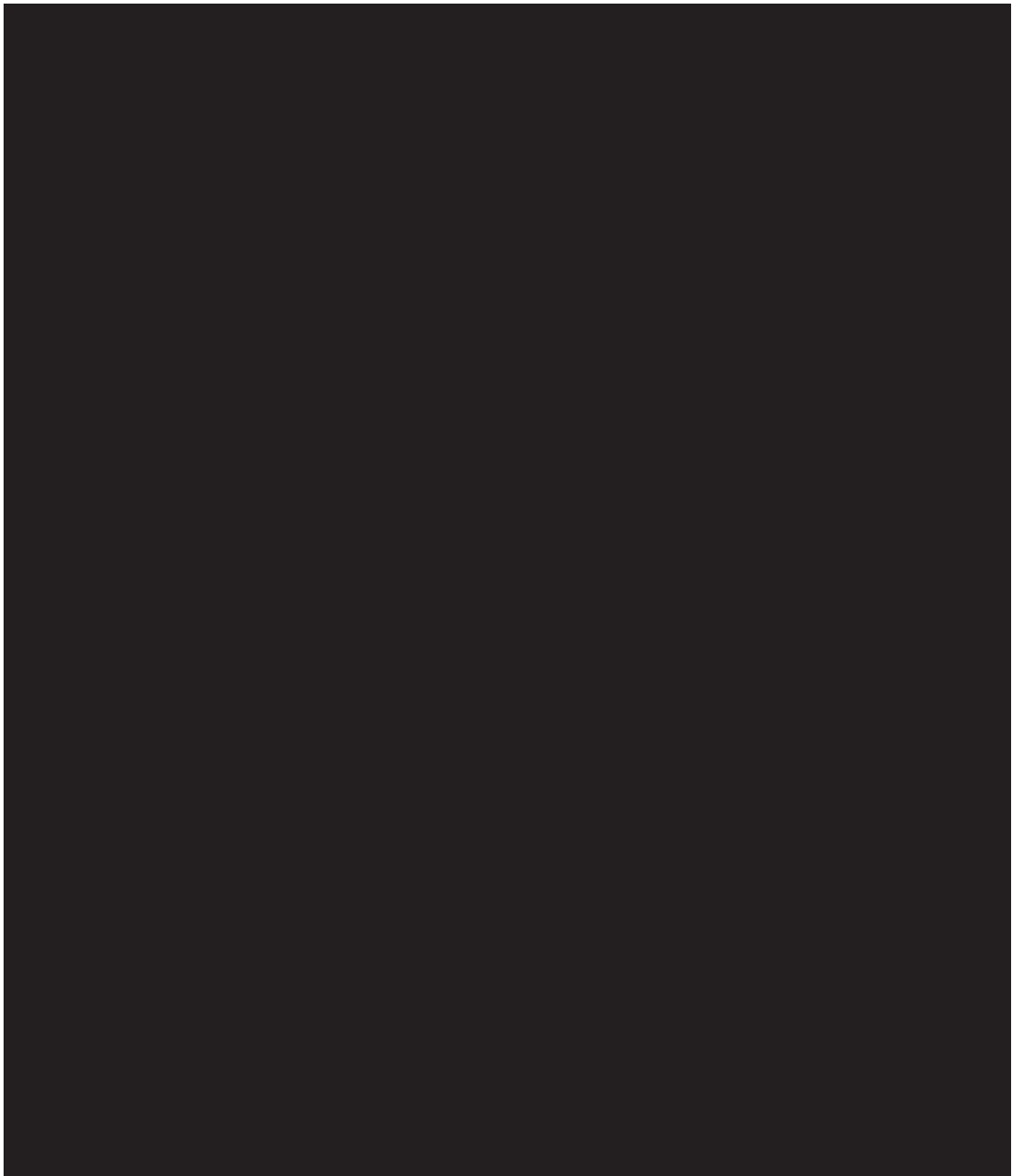
In addition, the number and percentage of participants meeting the criteria for potentially clinically significant abnormalities in ECG parameters will be summarized by treatment group separately for Part A and Part B. The following table presents the Criteria for clinically relevant abnormalities.

Table: Criteria for clinically relevant abnormalities in ECG parameters

Parameter	Criteria
PR Interval	>200 ms >200 ms and increase from baseline $\geq 25\%$ > 220 ms >220 ms and increase from baseline $\geq 25\%$ > 240 ms > 240 ms and increase from baseline $\geq 25\%$
QRS Interval	>110 ms >110 msec and increase from baseline $\geq 25\%$ >120 ms >120 ms and increase from baseline $\geq 25\%$
QT Interval	>500 ms
QTcB and QTcF	Absolute Values >450 to ≤ 480 msec



	>480 to ≤500 msec >500 msec Change from baseline: Change from baseline >30 to ≤60 msec Change from baseline >60 msec
--	--





9.11 Subgroup analysis

Subgroup analyses for the primary endpoint and key secondary endpoints may be considered for the following baseline variables:

- Weight ($< 100\text{kg}$, $\geq 100\text{kg}$);
- BMI ($< 30\text{kg/m}^2$, $\geq 30\text{kg/m}^2$)
- Prior exposure to biologic agents (yes/no);
- Concomitant antibiotic use (yes/no);
- Hurley Stage status (II, III).

These subgroup analyses will be carried out using the participants from FAS for Part A only.

Binary variables will be analysed using an exact logistic regression model as per section 9.6.1.4 but with the addition of the subgroup and subgroup by treatment interaction as covariates. Similarly, continuous endpoints will be analyzed using MMRM model as per section 9.6.1.3 with the addition of the subgroup and subgroup by treatment interaction as a covariate.

A forest plot will be used to display the results of the subgroup analyses.

If an individual subgroup for a variable includes fewer than 5% of the total planned study size, the subgroup analyses for that variable may be omitted.

Additional subgroup analyses may be performed at the sponsor's discretion. These additional analyses are beyond the scope of this SAP and will be specified in a separate document, if needed.

9.12 Interim Analysis

Not applicable.

9.13 Primary Analysis Timing

The primary analysis and unblinding of the data collected during Part A will be performed when all data up to and including Week 12 has been collected and database lock declared.

The study will be continued regardless of the results of the primary analysis.

9.14 DSMB

As required, a select number of safety outputs will be generated for the review of the DSMB. For full details of the scope of the DSMB, please see the DSMB charter. In addition a summary of the observed HiSCR50/75/90 responders will be presented for Part A+B by the FAS.

9.15 DSUR

As required, the below DSUR summaries will be generated combining data from Part A and Part B (until unblinding only a blinded column will be presented):

- Number of participants randomized
- Number of participants exposed by age and gender
- Number of participants exposed by race
- Number of SAEs within the DSUR reporting period
- Number of SAEs
- Number of participants reporting a DSUR safety topics AE (see Section 9.10.1)
- Listing of SAEs

10 CHANGES TO PLANNED ANALYSIS FROM STUDY PROTOCOL

Logistic regression using multiple imputation will be used as a sensitivity analysis of HiSCR50/75/ instead of general linear mixed models.

HiSCR100 is added as an additional secondary endpoint.

11 REFERENCES

1. ICH Topic E3: Structure and Content of Clinical Study Reports (CPMP/ICH/137/95-adopted December 1995).
2. ICH Topic E9: Statistical Principles for Clinical Trials (CPMP/ICH/363/96 – adopted March 1998).
3. Zouboulis, C. C., Tzellos, T., Kyrgidis, A., Jemec, G. B. E., Bechara, F. G., Giamarellos-Bourboulis, E. J et.al., European Hidradenitis Suppurativa Foundation Investigator, G. Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity. Br J Dermatol. 2017;177(5), 1401-1409.
4. Modeling Valuations for EuroQol Health States (Medical Care , Nov., 1997, Vol. 35, No. 11 (Nov., 1997), pp. 1095-1108)

12 APPENDICES

Appendix A: Dermatology Life Quality Index

DLQI

Hospital No:

Date:

Score:

Name:

Diagnosis:

Address:

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

1.	Over the last week, how itchy, sore, painful or stinging has your skin been?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	
2.	Over the last week, how embarrassed or self conscious have you been because of your skin?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	
3.	Over the last week, how much has your skin interfered with you going shopping or looking after your home or yard ?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	Not relevant ☐
4.	Over the last week, how much has your skin influenced the clothes you wear?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	Not relevant ☐
5.	Over the last week, how much has your skin affected any social or leisure activities?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	Not relevant ☐
6.	Over the last week, how much has your skin made it difficult for you to do any sport ?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	Not relevant ☐
7.	Over the last week, has your skin prevented you from working or studying ?	yes no	☐ ☐	Not relevant ☐
	If "No", over the last week how much has your skin been a problem at work or studying ?	A lot A little Not at all	☐ ☐ ☐	

8.	Over the last week, how much has your skin created problems with your partner or any of your close friends or relatives ?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	Not relevant ☐
9.	Over the last week, how much has your skin caused any sexual difficulties ?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	Not relevant ☐
10.	Over the last week, how much of a problem has the treatment for your skin been, for example by making your home messy, or by taking up time?	Very much A lot A little Not at all	☐ ☐ ☐ ☐	Not relevant ☐

Please check you have answered EVERY question. Thank you.

Appendix B: Patient Global Assessment of Skin Pain

Patient Global Assessment of Skin Pain

Please answer the question before you go to bed. Please mark an "X" in the box (☒) which best describes the severity of your skin pain in the last 24 hours.

1. In the last 24 hours, which number best describes your <u>skin pain</u> at its worst due to your Hidradenitis Suppurativa (HS)?	No skin pain ▼ Skin pain as bad as you can imagine ▼ 0 1 2 3 4 5 6 7 8 9 10 ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
---	--



Appendix C: Patient Global Impression of Severity (PGIS)



Appendix D: Hidradenitis Suppurativa-Physician's Global Assessment (HS-PGA)

Score	Rating	Description
0	Clear	0 abscesses, 0 draining fistulas, 0 inflammatory nodules, and 0 non-inflammatory nodules
1	Minimal	0 abscesses, 0 draining fistulas, 0 inflammatory nodules, and presence of non-inflammatory nodules
2	Mild	0 abscesses, 0 draining fistulas, and 1-4 inflammatory nodules or 1 abscess or draining fistula and 0 inflammatory nodules
3	Moderate	0 abscesses, 0 draining fistulas, and ≥ 5 inflammatory nodules or 1 abscess or draining fistula and ≥ 1 inflammatory nodules or 2-5 abscesses or draining fistulas and < 10 inflammatory nodules
4	Severe	2-5 abscesses or draining fistulas and ≥ 10 inflammatory nodules
5	Very severe	> 5 abscesses or draining fistulas

Appendix E: Hidradenitis Suppurativa Quality of Life (HiSQOL) Score

This questionnaire is designed to measure the impact of
Hidradenitis Suppurativa (HS), also known as acne inversa, on you.

PLEASE READ THESE DIRECTIONS:

It is important to:

1. Think about your HS over the past 7 days.
2. Think about **your HS only**, not another condition.
3. For each item select the **single best option**.

Please select the single best option for each item

In the past 7 days, how much has your HS impacted:	UNABLE TO DO due to my HS	Extremely	Very much	Moderately	Slightly	Not at all
1. Walking (not for exercise)	[]					
2. Exercising (for example: swimming, jogging, biking, yoga, aerobics)	[]					
3. Sleeping						
4. Washing yourself						
5. Getting dressed						
6. Concentrating						

In the past 7 days, how has HS influenced:	Extremely	Very much	Moderately	Slightly	Not at all
7. What you wear to avoid discomfort					

In the past 7 days, due to HS, how impacted have you been by:	Extremely	Very much	Moderately	Slightly	Not at all
8. Pain					
9. Itch					
10. Drainage					
11. Odor					

In the past 7 days, how much has HS caused you to feel:	Extremely	Very much	Moderately	Slightly	Not at all
12. Down or depressed					
13. Embarrassed					
14. Anxious or nervous					

In the past 7 days, how much has HS:			Extremely	Very much	Moderately	Slightly	Not at all
15. Made sexual activities difficult	I am not sexually active []	UNABLE TO DO due to my HS []					
16. Affected your desire for sexual activities							

In the past 7 days, how much has HS influenced:			Extremely	Very much	Moderately	Slightly	Not at all
17. Your ability to work or study	I do not work or study []	UNABLE TO DO due to my HS []					



Appendix G: Work Productivity and Activity Impairment Questionnaire: Hidradenitis Suppurativa, V2.0 (WPAI: Hidradenitis Suppurativa)

The following questions ask about the effect of your hidradenitis suppurativa on your ability to work and perform regular activities. *Please fill in the blanks or circle a number, as indicated.*

1. Are you currently employed (working for pay)? _____ NO ____ YES
If NO, check "NO" and skip to question 6.

The next questions are about the **past seven days**, not including today.

2. During the past seven days, how many hours did you miss from work because of problems associated with your hidradenitis suppurativa? *Include hours you missed on sick days, times you went in late, left early, etc., because of hidradenitis suppurativa. Do not include time you missed to participate in this study.*

_____ HOURS

3. During the past seven days, how many hours did you miss from work because of any other reason, such as vacation, holidays, time off to participate in this study?

_____ HOURS

4. During the past seven days, how many hours did you actually work?

_____ HOURS *(If "0", skip to question 6.)*

5. During the past seven days, how much did your hidradenitis suppurativa affect your productivity while you were working?

Think about days you were limited in the amount or kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual. If hidradenitis suppurativa affected your work only a little, choose a low number. Choose a high number if hidradenitis suppurativa affected your work a great deal.

Consider only how much hidradenitis suppurativa affected productivity while you were working.

Hidradenitis
suppurativa
had no effect
on work

0 1 2 3 4 5 6 7 8 9 10

Hidradenitis
suppurativa
completely
prevented me
from working

CIRCLE A NUMBER



6. During the past seven days, how much did your hidradenitis suppurativa affect your ability to do your regular daily activities, other than work at a job?

By regular activities, we mean the usual activities you do, such as work around the house, shopping, childcare, exercising, studying, etc. Think about times you were limited in the amount or kind of activities you could do and times you accomplished less than you would like. If hidradenitis suppurativa affected your activities only a little, choose a low number. Choose a high number if hidradenitis suppurativa affected your activities a great deal.

Consider only how much hidradenitis suppurativa affected your ability to do your regular daily activities, other than work at a job.

Hidradenitis suppurativa had no effect on my daily activities		Hidradenitis suppurativa completely prevented me from doing my daily activities
---	-------------	--

CIRCLE A NUMBER

WPAI: Hidradenitis Suppurativa V2.0 (US English)

Appendix H: FACIT-Fatigue Scale

Below is a list of statements that other people with your illness have said are important. Please select one number per line to indicate your response as it applies to the past 7 days.

		Not at all	A little bit	Some- what	Quite a bit	Very much
HI7	I feel fatigued	0	1	2	3	4
HI12	I feel weak all over	0	1	2	3	4
An1	I feel listless (“washed out”)	0	1	2	3	4
An2	I feel tired	0	1	2	3	4
An3	I have trouble <u>starting</u> things because I am tired	0	1	2	3	4
An4	I have trouble <u>finishing</u> things because I am tired	0	1	2	3	4
An5	I have energy	0	1	2	3	4
An7	I am able to do my usual activities	0	1	2	3	4
An8	I need to sleep during the day	0	1	2	3	4
An12	I am too tired to eat	0	1	2	3	4
An14	I need help doing my usual activities	0	1	2	3	4
An15	I am frustrated by being too tired to do the things I want to do	0	1	2	3	4



An16	I have to limit my social activity because I am tired	0	1	2	3	4
------	---	---	---	---	---	---

Appendix I: Clinical Laboratory Tests

Laboratory Testing Profile	Tests Included
Laboratory Tests Required at screening Only	Hepatitis B serology (HBV PCR if required) including Hepatitis B surface antigen and antibody and Hepatitis B core antibody, Hepatitis C serology, HCV antibodies, confirmatory HCV PCR, HIV testing, QuantiFERON® TB-gold test (or equivalent interferon-gamma release assay), SARS-CoV-2 PCR, Serum pregnancy, FSH
Blood Chemistry	Sodium, potassium, chloride, bicarbonate, total protein, blood urea nitrogen, serum creatinine, albumin, alkaline phosphatase, AST, ALT, total bilirubin, direct bilirubin, indirect bilirubin, calcium, GGT, creatine kinase, CRP, hs-CRP
Hematology	Hemoglobin, hematocrit, WBC with differentials (monocytes, eosinophils, basophils, neutrophils, lymphocytes) as an absolute value, RBC count, platelet count, ESR
Urinalysis	Specific gravity, pH, semi-quantitative “dipstick” evaluation of glucose, protein, bilirubin, ketones, leukocytes, blood Microscopy and/or culture to be performed if clinically indicated or if urinalysis results positive (blood, protein or leukocyte esterase/WBC). Urine human chorionic gonadotropin pregnancy testing for women of childbearing potential; may be repeated more frequently than indicated if required by local practice, if a menstrual cycle is missed, or if potential pregnancy is suspected
Fasting Lipid Panel and Glucose to be performed at Day 1/Week 0 and EOT	Fasting: glucose, total cholesterol, low density lipoprotein, high density lipoprotein, triglycerides,

Abbreviations: ALT: Alanine aminotransferase, AST: Aspartate aminotransferase, EOT: End of Treatment; ESR: erythrocyte sedimentation rate; FSH: Follicle stimulating hormone; GGT: gamma-glutamyl transferase, HBV: hepatitis B virus; HCV: hepatitis C virus; HIV: Human immunodeficiency virus; hs-CRP: high-sensitivity C-reactive protein, PCR: Polymerase chain reaction; RBC: red blood cells; SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2; WBC: white blood cells

Appendix J: Exposure adjusted incidence rate

It will be assumed that for each of n participants in a clinical trial the time t_j ($j=1, \dots, n$) to the first occurrence of a certain event is observed, or if the event was not experienced, the (censored) time to the end of the observation period for each treatment group. The sequence of first occurrences of an event will be modeled to follow approximately a Poisson process with constant intensity θ . The rate

$$T = \sum_{j=1}^n t_j$$

parameter θ will be estimated as $\lambda=D/T$, where $\sum_{j=1}^n 1$ and D is the number of participants with at least one event. The end of the observation period for subjects randomized in Part A to sonelokimab is the last visit date in Part B or study discontinuation date. The end of the observation period for subjects randomized to placebo or Adalimumab will be the first dose of Part B study medication or study discontinuation if discontinued during Part A. The start of the observation period of sonelokimab for subjects who were randomized to placebo or Adalimumab in Part A will be the first dose of Part B study medication.

Appendix K: Reasons for exclusion from the Per Protocol population

Participants will be excluded from the PP population for the following reasons in Part A:

- Lack of compliance – miss ≥ 2 doses
- Incorrect treatment – subject receives the wrong study medication
- Prohibited medications – a subject receives a medication prohibited by the protocol prior to Week 12
- A subject is incorrectly included in the study when they do not meet the following inclusion/exclusion criteria:
 - Inclusion 2-5
 - Exclusion 3-5
 - Exclusion 8-10

This list is the minimum set of reasons for exclusion from the PP population. The review of the final list of protocol deviations prior to unblinding may identify other key deviations that will impact the primary or key secondary endpoints and will require exclusion from the PP population.

A participant who is randomised in the wrong strata will still be included in the PP population but will be analysed using the correct strata information.

Appendix L: CTCAE v5.0 laboratory grades

Lab parameter	Units	Grade 1	Grade 2	Grade 3	Grade 4
sodium	mmol/L	low 130 - < LLN	125-129	120-124	<120
		high >ULN - 150	>150 - 155	> 155-160	> 160
potassium	mmol/L	low 3.0 - < LLN	3.0 - < LLN	2.5 - < 3.0	< 2.5
		High > ULN - 5.5	> 5.5 - 6.0	> 6.0 - 7.0	> 7.0
Alanine aminotransferase	U/L	High >ULN - 3* ULN if baseline normal; >1.5 - 3*ULN if baseline abnormal	> 3 - 5*ULN	> 5 - 20*ULN	> 20*ULN
Alkaline phosphatase	U/L	High >ULN - 2.5* ULN if baseline normal; >2.0 - 2.5*ULN if baseline abnormal	> 2.5 - 5*ULN	> 5 - 20*ULN	> 20*ULN
Aspartate aminotransferase	U/L	High >ULN - 3* ULN if baseline normal; >1.5 - 3*ULN if baseline abnormal	> 3 - 5*ULN	> 5 - 20*ULN	> 20*ULN
Total bilirubin	umol/L	High >ULN - 1.5* ULN	> 1.5 - 3*ULN	> 3 - 10*ULN	> 10*ULN
Serum creatinine	umol/L	High > ULN - 1.5* ULN	> 1.5 - 3.0*ULN	> 3.0 - 6.0 ULN	> 6.0*ULN
Creatine kinase	U/L	High > ULN - 2.5* ULN	> 2.5 - 5.0*ULN	> 5.0 - 10.0 ULN	> 10.0*ULN
Albumin	g/L	low 30 - < LLN	20 - <30	< 20	
Calcium	mmol/L	Low 2.0 - < LLN	1.75 - < 2.0	1.5 - < 1.75	< 1.5
		High >ULN - 2.9	>2.9 - 3.1	>3.1 - 3.4	>3.4
GGT		High >ULN - 2.5*ULN if baseline normal; >2.0 -	>2.5 - 5.0*ULN	>5.0-20.0*ULN	>20.0*ULN

Statistical Analysis Plan

Protocol Number: M1095-HS-201

Version: FINAL 1.1, Date: 23 May 2023

Page 92 of 93



Statistical Analysis Plan (SAP)



			2.5*ULN if baseline abnormal		
Hemoglobin	g/L	low	100 - <LLN	80 - < 100	< 80
		High	Increase >0-20	Increase >20-40	Increase >40
Neutrophils	10 ⁹ /L	Low	1.5 - < LLN	1.0 - <1.5	0.5 - <1.0
Lymphocytes	10 ⁹ /L	Low	0.8 - <LLN	0.5 - <0.8	0.2 - <0.5
		High	-	>4-20	>20
Platelets	10 ⁹ /L	Low	75 - <LLN	50 - <75	25 - <50
					<25