STATISTICAL ANALYSIS PLAN

Project Number:	GLPG3970
Study Number:	GLPG3970-CL-207
Study Title:	A randomized, double-blind, placebo-controlled, multicenter study to evaluate the efficacy, safety, tolerability, pharmacokinetics, and pharmacodynamics of orally administered GLPG3970 for 12 weeks in adult subjects with active primary Sjögren's Syndrome

Development Phase:	2a		
Status:	Final		
Version:	1.0	Date:	17-Sep-2021
EudraCT:	2020-003298-22	IND:	Not applicable
Sponsor:	Galapagos NV, Generaal De Wittelaan L11 A3, 2800 Mechelen, Belgium		
Biostatistician:	[redacted]		
Medical Leader:	[redacted]		

TABLE OF CONTENTS

VERSION HISTORY

SAP Amendment #	Date	Description of changes
SAP Version Final 1.0	17-Dec-2021	N/A

LIST OF ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical
AUC	area under the plasma concentration-time curve
BLOQ	below lower limit of quantification
BMI	body mass index
C3	complement component 3
C4	complement component 4
CI	confidence interval
CL	clearance
CLCr	creatinine clearance
CRF	case report form
CRO	contract research organization
CRP	C-reactive protein
CSP	clinical study protocol
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
C_{tau}	trough concentration
CTx	serum C terminal telopeptide type I collagen
ECG	electrocardiogram
eCRF	electronic case report form
ED	early discontinuation
eGFR	estimated glomerular filtration rate
ESSDAI	EULAR Sjögren's Syndrome Disease Activity Index
ESSPRI	EULAR Sjögren's Syndrome Patient Reported Index
EULAR	European League against Rheumatism
FAS	full analysis set
FU	follow-up
GLSM	geometric LS-mean
GM	geometric mean
GSD	geometric standard deviation
GSE	geometric standard error
H	high, above the upper limit of the normal range

HDL	high-density lipoprotein
High	high specificity test
HOMA-IR	homeostatic model assessment of insulin resistance
HR	heart rate
ICF	informed consent form
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IP	investigational product
INF	infinity
IWRS	interactive web response system
L	low, below the lower limit of the normal range
LDL	low-density lipoprotein
LLN	lower limit of the normal range
LLOQ	lower limit of quantification
LRV	lower reference value
LS	least squares
LSM	least square mean
MedDRA	medical dictionary for regulatory activities
[redacted]	[redacted]
MMRM	mixed models for repeated measures
n	number of non-missing data points
N	normal, with the limits of the normal range
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NRI	non-responder imputation
OC	observed cases
[redacted]	[redacted]
PK	pharmacokinetic(s)
pop-PK	population PK
PR	pulse rate
pSS	primary Sjögren's Syndrome
[redacted]	[redacted]
QTc	corrected QT interval
QTcF	QT interval corrected for the heart rate using Fridericia's formula
RR	respiratory rate
SAE	serious adverse event

SAP	statistical analysis plan
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
[redacted]	[redacted]
SI	standard international
[redacted]	[redacted]
SMC	Safety Monitoring Committee
SoC	standard-of-care
TE	Treatment-emergent
TEAE	treatment-emergent adverse event
[redacted]	[redacted]
TLF	tables, listings and figures
TV	target value
ULN	upper limit of the normal range
ULOQ	upper limit of quantification
[redacted]	[redacted]
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the final statistical analyses of study GLPG3970-CL-207. The results of the analysis will be described in the clinical study report (CSR).

Technical details on derivations and mock tables, listings and figures (TLFs) will be presented in a separate document.

The statistical analysis will process and present the results following the International Council for Harmonization (ICH) standards, particularly the ICH-E3, ICH-E6, ICH-E9, ICH-R9 (R1) and ICH-E14 guidelines.

On 26 November 2021, the Safety Monitoring Committee recommended to immediately terminate Study GLPG3970-CL-207 due to a safety signal of liver toxicity. Therefore, the GLPG3970-CL-207 study was prematurely stopped. In view of this, only the primary endpoint and safety analyses will be performed. Results will be presented in an abbreviated CSR. Any provisions no longer applicable following premature study termination are highlighted in grey.

2. STUDY DESIGN AND OBJECTIVES

2.1. Study Objectives

2.1.1. Primary Objectives

- To evaluate the efficacy of GLPG3970 compared to placebo on the signs and symptoms of primary Sjögren's Syndrome (pSS).
- To evaluate the safety and tolerability of GLPG3970 compared to placebo.

2.1.2. Secondary Objectives

- To further characterize the efficacy of GLPG3970 compared to placebo on the patient-reported signs and symptoms of pSS.
- To characterize the pharmacokinetics (PK) of GLPG3970.

2.1.3. Other Objectives

2.2. Study Endpoints

2.2.1. Primary Endpoints

- Change from baseline in European League against Rheumatism (EULAR) Sjögren's Syndrome Disease Activity Index (ESSDAI) score at Week 12.

- Number, incidence, and severity of treatment-emergent adverse events (TEAEs).

2.2.2. Secondary Endpoints

- Change from baseline in EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI) score at Week 4, 8, and 12.

- Change from baseline in ESSDAI score over time at Week 4, 8, and 12.

- Observed GLPG3970 plasma trough concentration (C_{tau}).

2.2.3. Other Endpoints

[redacted]

2.3. Study Design

This is a Phase 2a, randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the efficacy, safety, tolerability, PK, [redacted] of orally administered GLPG3970 for 12 weeks in adult subjects with active pSS.

One dose level of GLPG3970 will be evaluated in 30 subjects who will be randomized 2:1 to receive GLPG3970 (400 mg [2 x 200 mg] qd) or matching placebo, respectively.

The study will consist of 3 study periods:

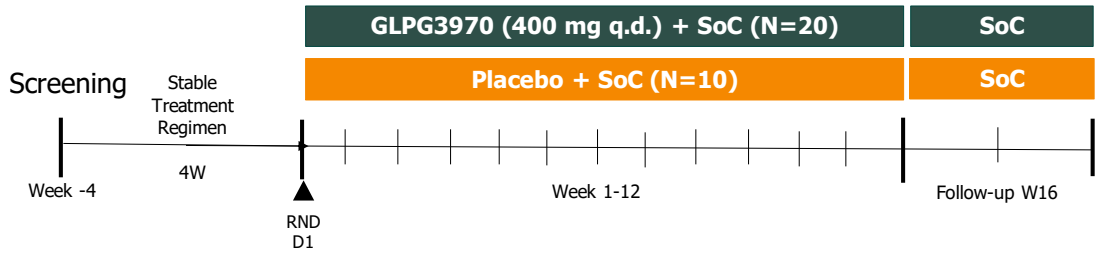
- Screening period: up to 4 weeks with 1 study visit. During the Screening visit, subjects will be selected based on the eligibility criteria. Subjects already on treatment will remain on a stable standard of care (SoC) treatment during the study.
- Treatment period: 12 weeks with up to 5 study visits (Days 1, 8, 29, 57, and 85). At any time during the course of the study, phone calls or unscheduled visits can be added if deemed necessary by the investigator. Subjects will receive GLPG3970 or placebo qd.
- Follow-up (FU) period: 4 weeks with 1 study visit. Subjects remain on stable SoC treatment, if applicable.

The subjects will be in the study for a duration of 20 weeks: up to 4 weeks of screening, 12 weeks of treatment, and 4 weeks of FU.

Refer to the Schedule of Assessments (Section 2.5) for further details.

A schematic diagram of clinical study design and periods is provided in Figure 1.

Figure 1 Schematic Diagram



2.4. Clinical Study Protocol (CSP) and CSP Amendments

This SAP was based on the protocol version 7.0, dated 12-May-2021.

2.5. Scheduled of Assessments

The study assessments will be undertaken at time points as specified in the Schedule of Assessments (Table 1). For detailed instructions on the clinical study procedures, please see the relevant CSP sections and CSP Section 8.1, “Timing of Assessments”.

Table 1 Schedule of Assessments

EVENT	SCREENING PERIOD	TREATMENT PERIOD[1]						FOLLOW-UP PERIOD
Study Visit	**S**	**1**	**2**	**3**	**4**	**5**	**ED[2]**	**Follow-up Visit**
Study Week	**-4 to -1**	**Baseline**	**1**	**4**	**8**	**12**		
Study Day (D) ± Days	**-28 to -1**	**1**	**8±1**	**29±2**	**57±2**	**85±2**	**Early treatment discontinuation**	**28 ± 3 after last IP dosing**
Informed consent	✓							
Inclusion/exclusion criteria	✓	✓						
Demographics	✓							
Medical history	✓							
Physical examination	✓	✓	✓	✓	✓	✓	✓	✓
Vital signs	✓	✓	✓	✓	✓	✓	✓	✓
Body weight and height[3]	✓	✓	✓	✓	✓	✓	✓	✓
12-lead ECG[4]	✓	Predose, 1.5h, 3h, 4.5h postdose[4]	Predose	Predose	Predose	Predose	✓	✓

[1] On dosing days, all assessments are to be performed predose, unless otherwise specified.
[2] Subjects who discontinue treatment early will be requested to return for an ED visit and to return for a FU visit 28 ± 3 days after last IP administration.
[3] Height only to be measured at screening.
[4] At Visit 1 (Day 1), triplicate ECGs will be recorded predose and at 4.5 hours postdose. At all other timepoints, single ECGs will be taken.

EVENT	SCREENING PERIOD	TREATMENT PERIOD[1]						FOLLOW-UP PERIOD
Study Visit	**S**	**1**	**2**	**3**	**4**	**5**	**ED[2]**	**Follow-up Visit**
Study Week	**-4 to -1**	**Baseline**	**1**	**4**	**8**	**12**		
Study Day (D) ± Days	**-28 to -1**	**1**	**8±1**	**29±2**	**57±2**	**85±2**	**Early treatment discontinuation**	**28 ± 3 after last IP dosing**
QuantiFERON-TB Gold test	✓							
SARS-CoV-2 RT-PCR test[5,6]	✓	✓	As needed					
SARS-CoV-2 serology test[5,6]		✓	As needed					
Randomization		✓						
Blood collection								
– Safety (hematology, coagulation, chemistry)	✓	✓[7]	✓	✓	✓[7]	✓[7]	✓	✓
– Safety (HBV, HCV, HIV serology)	✓							
– Pregnancy test serum (all females)	✓							
– FSH test (WOnonCBP, non-surgical postmenopausal women)	✓							

[5] RT-PCR from a nasal swab sample at screening and RT-PCR from a nasal swab sample and serology testing at Visit 1 (Day 1), and as needed when subject presents signs and symptoms of SARS-CoV-2 infection.
[6] SARS-CoV-2 RT-PCR and serology tests can (optionally) be sampled by a nurse at the subject's home 2-3 days before Visit 1 (Day 1).
[7] Fasted glucose, fasted insulin, and HOMA-IR only at Visit 1, 4, and 5.

EVENT	SCREENING PERIOD	TREATMENT PERIOD[1]						FOLLOW-UP PERIOD
Study Visit	**S**	**1**	**2**	**3**	**4**	**5**	**ED[2]**	**Follow-up Visit**
Study Week	**-4 to -1**	**Baseline**	**1**	**4**	**8**	**12**		
Study Day (D) ± Days	**-28 to -1**	**1**	**8±1**	**29±2**	**57±2**	**85±2**	**Early treatment discontinuation**	**28 ± 3 after last IP dosing**
– Blood samples for PK GLPG3970		Predose[8], 0.5-1.5h, 2-2.5h, 4.5h[9] postdose[10]	Predose[8]	Predose[8]	Predose[8], 0.5-1.5h, 2-2.5h postdose[10]	Predose[8]	✓	
[redacted]								
– [redacted]								
– [redacted]								
Urine sample								
– Safety (urine)	✓	Predose	Predose	Predose	Predose	Predose	✓	✓

[8] At Visit 1, 2, 3, 4, and 5, PK sample to be taken within 30 minutes predose.
[9] At Visit 1, the PK sample collected at 4.5 hours postdose must be taken immediately after the triplicate ECG.
[10] At Visit 1, 4, and 5 subjects must come for the study visits in a fasting state (no food intake for at least 8 hours). If the postdose samples could not be collected at Visit 4, they should be collected at Visit 5. It is important that the PK [redacted] samples are collected on the same day.
[redacted]

EVENT	SCREENING PERIOD	TREATMENT PERIOD[1]						FOLLOW-UP PERIOD
Study Visit	**S**	**1**	**2**	**3**	**4**	**5**	**ED[2]**	**Follow-up Visit**
Study Week	**-4 to -1**	**Baseline**	**1**	**4**	**8**	**12**		
Study Day (D) ± Days	**-28 to -1**	**1**	**8±1**	**29±2**	**57±2**	**85±2**	**Early treatment discontinuation**	**28 ± 3 after last IP dosing**
– Pregnancy testing (UPT, dipstick) (WOCBP only)		Predose	Predose	Predose	Predose	Predose		
Subject diary evaluation		✓	✓	✓	✓	✓	✓	✓
ESSDAI	✓	✓		✓	✓	✓	✓	
ESSPRI	✓	✓		✓	✓	✓	✓	
[redacted]								

[redacted]

EVENT	SCREENING PERIOD	TREATMENT PERIOD[1]						FOLLOW-UP PERIOD
Study Visit	**S**	**1**	**2**	**3**	**4**	**5**	**ED[2]**	**Follow-up Visit**
Study Week	**-4 to -1**	**Baseline**	**1**	**4**	**8**	**12**		
Study Day (D) ± Days	**-28 to -1**	**1**	**8±1**	**29±2**	**57±2**	**85±2**	**Early treatment discontinuation**	**28 ± 3 after last IP dosing**
[redacted]								
Dispense IP		✓	✓	✓	✓			
IP administration[14]		Once daily throughout the treatment period						
AE and concomitant medication	throughout the study							

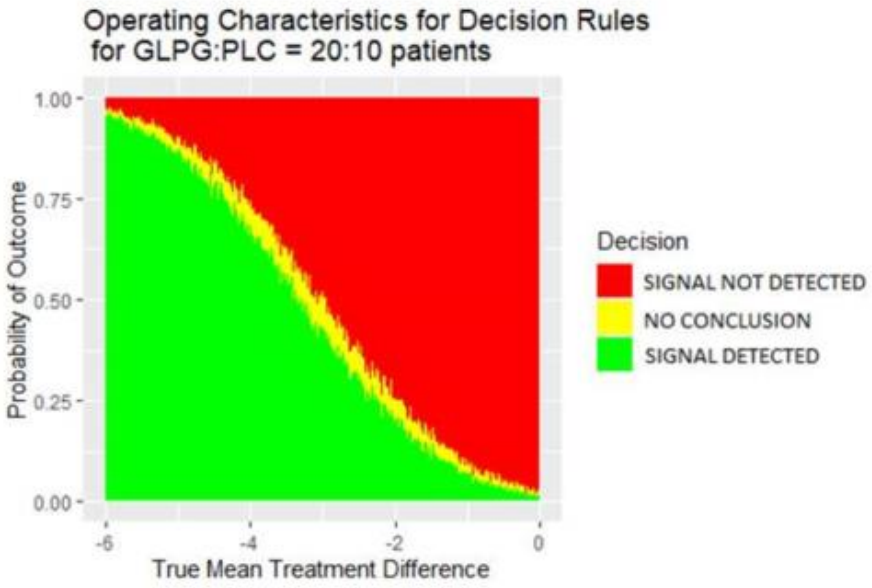
[14] At Visit 1 (Day 1), the IP will be administered at the study site in a fasted state. Breakfast will be provided to the subject 0.5-1 hour after the IP intake. Lunch should be eaten 1 hour before or after the ECG measurement at 4.5 hours postdose. At all other dosing days (including Visit 2, 3, 4, and 5), the IP should be taken in a fed state (together with a meal).

2.6. Sample Size Justification

Up to 30 subjects are planned to be randomized 2:1, i.e. 20 subjects will receive GLPG3970 (400 mg qd) and 10 subjects will receive placebo. In case of drop-outs due to SARS-CoV-2 infection, additional subjects may be randomized on top of the planned sample size. The number of additional subjects randomized will not exceed the number of subjects dropping out of the study in relation to SARS-CoV-2. Randomization of additional subjects will be decided by the sponsor before study lock or related unblinding has occurred.

This sample size provides favorable operating characteristics for the primary efficacy endpoint (i.e. ESSDAI change from baseline at Week 12) in terms of the modified signal detection methodology r. This methodology will be used to provide further insight into the treatment effect of GLPG3970 over placebo, and will support scenario analyses. The posterior distribution of this treatment effect will be estimated, and from this distribution probabilities of reaching at least a certain effect (delta) will be derived, e.g. a range of plausible effect size values going from as high as P (delta <=-5) to as low as P (delta <=-2). A common standard deviation (SD) of 4 was assumed for both active and placebo arm.

Probability	**Reference value**	**Outcome**
0.852	-5	Signal detected
0.041		No conclusion
0.107		Signal not detected
0.427	-3	Signal detected
0.088		No conclusion
0.485		Signal not detected
0.205	-2	Signal detected
0.050		No conclusion
0.745		Signal not detected



2.7. Randomization and Blinding

2.7.1. Randomization

At screening, subjects will be assigned a subject identification number. When a subject is confirmed to be eligible for the clinical study, the subject will be randomized. Allocation of each subject to a given treatment will be done using a centralized electronic system (interactive web response system [IWRS]). Subjects will be randomized in a 2:1 ratio to GLPG3970 or placebo.

2.7.2. Blinding and Unblinding

This is a randomized, double-blind clinical study. The subjects and the entire clinical study team, including the investigators, clinical study coordinators, and sponsor personnel, except for the Safety Monitoring Committee, are blinded to treatment assignment.

Blinded and packaged medication will be provided to the site. All IP formulations will be identical in appearance, shape, smell and taste, and will be packaged in the proper proportion to assure desired dosages and maintenance of the blinding.

The blind can be broken only if the investigator deems it necessary for the safety of a subject. The investigator is encouraged to discuss considerations to break the blind with the medical leader of the study, whenever possible and where the situation allows. However, the responsibility to break the treatment code in emergency situations resides solely with the investigator. The investigator is not required to discuss unblinding beforehand if he or she feels rapid emergency unblinding is necessary but is required to inform the sponsor in within 24 hours after unblinding has occurred.

The blind can be broken by the investigator via IWRS.

If the blind is broken for any reason during the course of the clinical study, the moment on which the blind was broken, and all other relevant information will be documented by the site. The reason for breaking the blind will be indicated and justified in the source documentation.

If an adverse event (AE) leads to unblinding, the AE will be given as the reason for unblinding. All subjects who are unblinded should, where possible, complete the FU visit assessments 28±3 days after last IP dosing. Any AEs will be followed until resolution.

The code-break information (via IWRS vendor) will be provided to the bioanalytical laboratory responsible for PK sample analysis, the sponsor person responsible for providing unblinded data to the Safety Monitoring Committee, the contract research organization (CRO) responsible for the population PK (pop-PK) [redacted], and to the pharmacovigilance vendor for serious adverse event (SAE) reporting purposes.

3. STUDY ESTIMANDS

Estimand for the primary endpoints:

Attribute	**Details**
Treatments	GLPG3970 (400 mg qd) vs. Placebo
Population	• Population target defined through the inclusion/exclusion criteria (see CSP section 6). • Analysis Set: Full Analysis Set (FAS), defined as all randomized subjects who have received at least one dose of IP.
Endpoint	Change from baseline in ESSDAI score at Week 12.
Population-level summary	Difference (GLPG3970-Placebo) in least square mean (LSM). For further details see study SAP section 6.2.2.2.
Intercurrent events and strategies to handle those	• Early treatment discontinuation (for any reason) is handled using the hypothetical strategy • Other intercurrent events (e.g. major protocol violations, intake of prohibited medication, lack of compliance) are handled using the

	treatment policy strategy (i.e. they are ignored)

Estimand for the safety endpoint

Attribute	**Details**
Treatments	GLPG3970 (400 mg qd) vs. Placebo
Population	• Population target defined through the inclusion/exclusion criteria (see CSP). • Analysis Set: Safety Analysis Set, defined as all randomized subjects who have received at least one dose of IP.
Endpoint	Presence of TEAE
Population-level summary	Percentage of subjects with TEAEs.
Intercurrent events and strategies to handle these	Intercurrent event: • Early treatment discontinuation (for any reason): handled using the while-on-treatment strategy. All collected data up to (and including) the last contact visit is regarded as possibly influenced by the drug exposure and thus will be used in the analysis. • Intercurrent event: major protocol deviations, intake of prohibited/rescue medication, lack of compliance): handled using the treatment policy strategy. All collected data will be used in the analysis.

4. GENERAL METHODOLOGY

4.1. Analysis Sets

The analysis set will always be indicated in a subtitle in the table, listing or figure.

4.1.1. All Screened Analysis Set

All subjects who signed and dated an informed consent form (ICF).

4.1.2. All Randomized Analysis Set

All screened subjects who were randomized into the clinical study.

4.1.3. Safety Analysis Set

All randomized subjects who received at least one dose of IP.

4.1.4. Full Analysis Set

All randomized subjects who have received at least one dose of IP.

4.1.5. Pharmacokinetic Analysis Set

Subset of the Safety Analysis Set for which plasma concentration data are available to facilitate development of the Population PK model as described in the pharmacometrics analysis plan and excluding CSP deviations which have an impact on the PK analysis.

For population PK modeling, besides the plasma concentration, the corresponding PK sampling time and the time of last drug intake prior to taking the sample should be available. The results from the population PK analysis will be presented in a separate document.

4.1.6. [redacted]

[redacted]

4.1.7. [redacted]

Intersection of the [redacted] Analysis Sets. The results from the [redacted] modeling will be presented in a separate document.

4.2. Randomized Versus Actual Treatment Group

For subject information and efficacy parameters, the treatment group as assigned by the randomization will be used in the analysis (i.e. planned analysis).

The actual treatment groups will be used for the analysis of safety, PK [redacted] parameters. The "actual" rather than "randomized" treatment will be used only in case the "wrong" treatment had been taken during the entire study duration.

Differences between the randomized and actual treatment group will be listed in the disposition/randomization section of the analysis.

4.3. Analysis Periods and Analysis Time Points

4.3.1. Relative Day

The timing of an assessment or an event relative to a reference date will be calculated as follows:

When the concerned date is before the reference date:

Relative day (days) = concerned date-reference date

When the concerned date is the equal or later than the reference date:

Relative day (days) = concerned date-reference date +1 day

Where:

- The *concerned date* could be the measurement date of the assessment, or the start or end date of the event.
- The *reference date* default is the date of the first dose of study drug administration, unless specified otherwise.
- *Date* implies a complete date having day, month and year available. Unless otherwise specified, the *relative day* will remain missing when it cannot be calculated due to absence or incompleteness of the concerned and/or reference dates.

The general terms of this formula also apply when similar relative timings are required in other time units, for example in minutes.

4.3.2. Analysis Periods

For treated subjects, AEs and assessments will be allocated to analysis periods according to Table 2.

Table 2 Analysis Periods

Analysis Period	Start Analysis Period	End Analysis Period
Screening	Date of signing the ICF, with 00:00 added as time part	Date (time) of first IP administration - 1 minute
Treatment	Date (time) of first IP administration.	Study termination date (i.e. date of last contact), with 23:59 added as time part

The last analysis period will always end on the date of last contact.

For every subject, all assessments performed and all events occurring during the study are expected to have a (start) date (/time) between the date of informed consent (IC) signature and the last contact date. Obvious exemptions are historical records like medical history or concomitant therapies having started before the study.

Assessments and events will be allocated to one of these analysis periods by incidence, meaning placing the record (start) date and time between the matching start and end dates and times of the subject's own analysis periods, assuming completeness of all dates and times, or further determined by the presence of timing indicators such as tick boxes flagging AEs starting before/after first IP administration, or by worst-case considerations for AEs if needed. For the parameters for which the time of assessment was not planned to be collected, their assessment on the day of first IP administration will be considered as baseline and will be reported under the 'Screening' analysis period with the exception of AEs.

4.3.3. Analysis Windows

For the efficacy, PK, [redacted] and safety assessments, all data (including data obtained at unscheduled visits) will be placed into time windows based on the relative day (ADY) of the assessment (relative to the first dose of IP), according to the following allocation tables:

For endpoints with multiple components, analysis windows will be derived separately for each component/parameter.

All assessments, including data collected on unscheduled visits, will be allocated to analysis windows based on the relative day of the assessment (see Section 4.3.1) and according to the algorithm in Table 3-6. TFLs will present the analysis windows.

Table 3: Analysis Visit Window for 12-Lead ECG Data

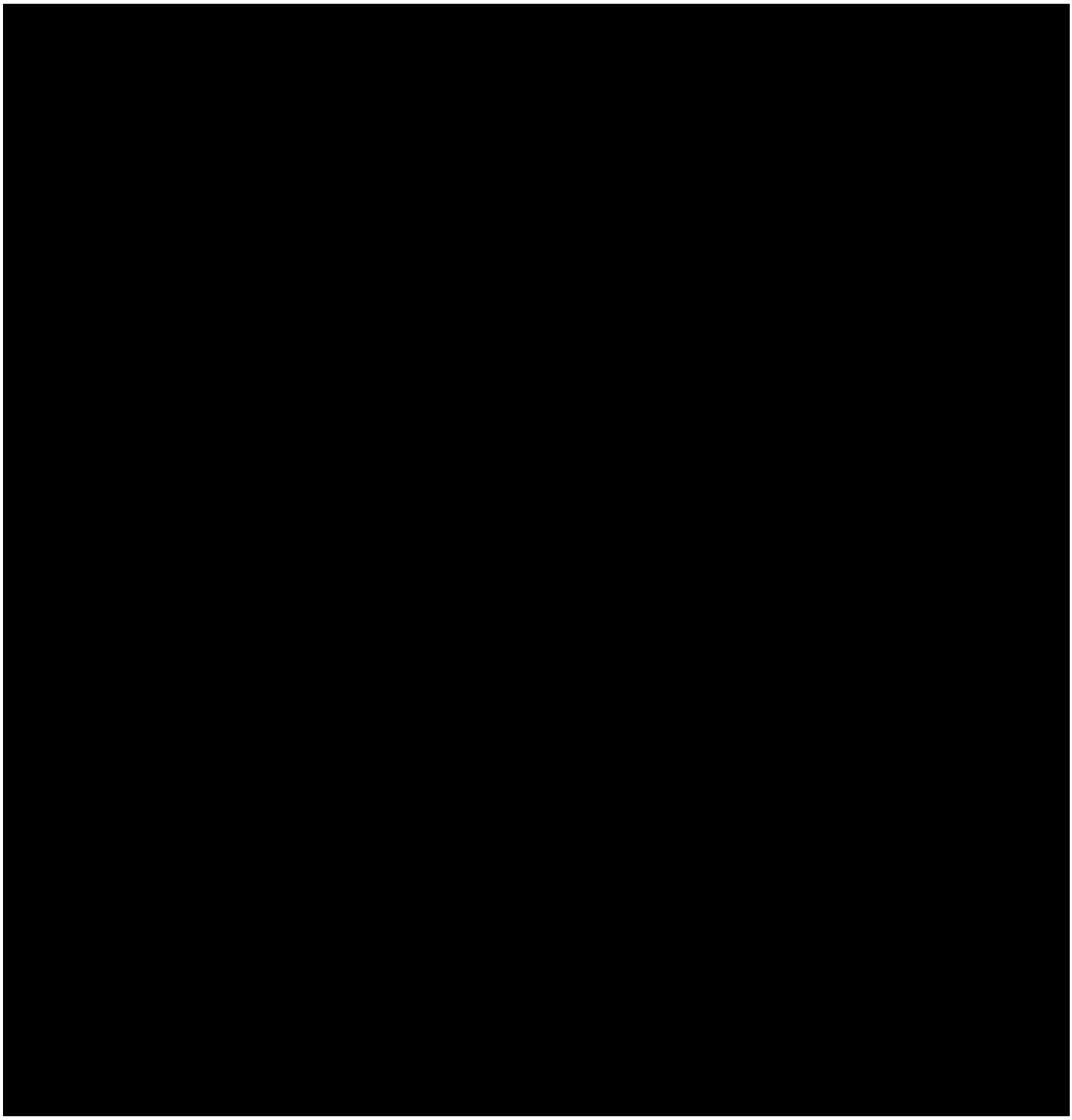
Visit date and time	Relative day (ADY)	Time window (in ADY)	Target day	Analysis visit (AVISIT)	Analysis time point (ATPT)
Before first study drug administration	<1 or predose at ADY=1	<1 or predose at ADY=1	1	Baseline	
Same as or after first study drug administration	≥ 1	1	1	Day 1	As indicated in the database: 1.5, 3, 4.5 hours postdose
		2 to 19	8	Week 1	Predose
		20 to 43	29	Week 4	Predose
		44 to 71	57	Week 8	Predose
		72 to 92	85	Week 12	Predose
		>7 days after last dose of IP	Last dose of IP + 28 days	Follow-up *	

** The Follow-up window has precendence over all other time windows.*
Note: At baseline and Day 1 predose and 4.5 hours postdose, triplicate ECGs will be recorded. At all other timepoints, single ECGs will be taken.

Table 4: Analysis Visit Window for PK Data

Visit date and time	Relative day (ADY)	Time window (in ADY)	Target day	Analysis visit (AVISIT)	Analysis time point (ATPT)
Before first study drug administration	<1 or predose at ADY=1	<1 or predose at ADY=1	1	Baseline	Predose (within 30 minutes prior to dosing)
Same as or after first study drug administration	≥ 1	1	1	Day 1	As indicated in the database: 0.5-1.5, 2-2.5, 4.5 hours postdose
		2 to 19	8	Week 1	As indicated in the database: Predose
		20 to 43	29	Week 4	As indicated in the database: Predose
		44 to 71	57	Week 8*	As indicated in the database: Predose, 0.5-1.5, 2-2.5 hours postdose
		72 to 92	85	Week 12	As indicated in the database: Predose
		>2 days after last dose of IP	Last dose of IP+28 days	Follow-up**	

**If postdose week 8 is missing and postdose week 12 are available, it means these postdose week 12 will be considered as postdose week 8, that records will be presented in week 8 in tables (with footnote) and with flagged in listings*
***The Follow-up window has precendence over all other time windows.*



4.3.4. Definition of Baseline

Baseline is defined as the last available assessment prior to the first intake of IP. For the parameters for which the time of assessment was planned to be collected, if for their assessment on the day of first intake the time is missing, their assessment at time point "PREDOSE" (when applicable) will be considered as baseline, otherwise, previous days last available assessment will be selected. For the parameters for which the time of assessment was not planned to be collected, their assessment on the day of first intake of IP will be considered as baseline. For efficacy endpoints, the baseline visit will be defined as Table 6. If multiple values on the same day qualify as last available assessment, the average of these values will be used in the analysis.

For ECG data assessed in triplicate, baseline is defined as the average (stored without rounding) of the combination of the most complete replicated ECG parameter before first intake of IP and at the same time the closest to it, selecting in the following order:

First consider Day 1, predose:

- Select the average of the last available triplicate ECG parameter on that day;
- If there is no triplicate, select the average of the last available duplicate ECG parameter;
- In the absence of triplicates and duplicates the last available single ECG parameter will become baseline.

If baseline is not yet determined, consider the previous day and repeat the above selection (average of last triplicate, average of last duplicate, last single ECG on that day). Repeat the whole process successively for each previous day(s) as needed until baseline is determined.

4.3.5. Selection and Use of Analysis Visit

Before first administration of IP, only Baseline will be used for the entire statistical analysis. Pre-baseline assessments will only be listed. After first administration of IP, if multiple valid, non-missing assessments exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- The record closest to the target day for that visit will be selected, or
- If there are 2 records that are equidistant from the nominal day, or more than 1 record (with time known) on the selected day, the latest record will be taken
- If chronological order cannot be determined (e.g., more than 1 record on the same day with time missing), the average of the records will be computed and reported in the analysis

4.4. Handling of Data

4.4.1. Handling of Missing Data

4.4.1.1. Handling of Missing Date-Time Data

No imputations will be done in case of missing date (time) fields, nor for the missing parts of partially known date (time) fields.

Assessments with missing date (time) will be omitted from the analysis.

Event-type data (e.g. adverse events, concomitant medications) with missing date (time) will be allocated to analysis periods using a worst-case approach as explained in the respective sections.

For the date of diagnosis of Sjogren's Syndrome disease to allow the computation of the disease duration, if missing month, impute as January (01), if missing day, impute as the first day of the month (01).

4.4.1.2. Handling of Missing Result Data

No imputation will be done of missing result data unless otherwise specified. That is, an observed case (OC) analysis will be performed. Endpoints analyzed via mixed model for repeated

measures (MMRM) handle missing data via the maximum-likelihood function. The MMRM will be applied only when the number of subjects with non-missing values is more than 30% of the sample size Following a hypothetical strategy estimand, non-PK data observed more than 7 days after the last dose of IP will not be used for the on-treatment related analysis visits. PK data observed more than 2 days after the last dose of IP will not be used for the on-treatment related analysis visits.

For analyses of binary endpoints, the non-responder imputation (NRI) method will be used for missing value imputation in addition to the observed case (OC) analysis.

4.4.1.3. Censoring of Time-to-Event Data

Not applicable.

4.4.2. Handling of Values Below or Above a Threshold

Results of continuous parameters, as well as normal limits of these reported as below (or above) the detection limit will be imputed by the value one precision unit smaller (or larger) than the detection limit itself. In listings, the original value will be presented.

Example: if the database contains the value "<0.04", then for the descriptive statistics the value "0.03" will be used. The value ">1000" will be imputed by "1001".

For PK data, values below the limit of quantification will be imputed by 0 for the calculation of descriptive statistics presentation; except for the geometric mean and the geometric coefficient of variation (CV%), where it will be imputed as lower limit of quantification (LLOQ)/2. These values will be listed as "below lower limit of quantification (BLOQ)".

[redacted]:

- [redacted]
- [redacted]

4.4.3. Handling of Outliers

There will be no exclusion of outliers, all measured values will be included in the analyses.

4.4.4. Stratification Factors

Not applicable.

4.5. Presentation of Results

4.5.1. Presentation of Treatment Groups

Results will be presented by treatment group:

- GLPG3970 400 mg qd
- Placebo

In the Subject Information, a grand total "All Subjects" will be added to summarize all subjects over all treatment groups in tables.

4.5.2. Calculation of Descriptive Statistics

For continuous parameters, descriptive statistics will be presented when $N \geq 2$. When $N = 1$, the observation will not be shown in tables nor in figures of summary statistics but will only be listed.

Descriptive statistics will include:

- the number of non-missing data points (n);
- the arithmetic mean;
- the standard error (SE) and standard deviation (SD);
- the median, minimum and maximum;
- 90% confidence interval (CI) of the mean (if indicated in the relevant section).
- Some parameters namely estimated glomerular filtration rate (eGFR), C-reactive protein (CRP), [redacted] and metabolic parameters (fasted insulin, homeostatic model assessment of insulin resistance (HOMA-IR) and HbA1c) are known to deviate heavily from the normal distribution but follow the log-normal distribution well. For these parameters, the geometric mean (GM), geometric standard deviation (GSD), and geometric standard error (GSE) will be also reported for the summaries by treatment group and visit. GM, GSD and GSE are not applicable for the summaries of the changes from baseline, however GM, GSD and GSE will be reported for the summaries of percent changes from baseline. This will be done by obtaining the mean, SD and SE of log(value) - log(baseline) and then transforming these statistics back to the 'percent change scale' as 100 X {exponential(statistic) - 1}.

For PK data, descriptive statistics will include:

- the number of non-missing observations (n);
- the number of data points above the LLOQ;
- the arithmetic mean;
- SE and SD;
- the median, minimum and maximum;
- the CV%;
- the geometric mean and geometric CV%.

If less than 50% of the subjects have quantifiable values, only the number of subjects with data, number of data points above the LLOQ, the arithmetic mean, median, minimum, and maximum will be presented with the original calculated value. The other descriptive statistics will be listed as “NC” (not calculated).

If the calculated descriptive statistic is BLOQ, then it will be presented as “BLOQ”.

Individual values and descriptive statistics of concentrations and PK parameters will be presented with 3 significant digits.

4.5.3. Calculation of Percentages

Frequencies and percentages will be generated for categorical parameters.

For event-type data (e.g. adverse events), the denominator will be all subjects in the analysis set and analysis period. For other data (e.g. worst-case analysis of assessments), the denominator will be all subjects with (post-baseline) data for the parameter, in the analysis set and analysis window/period.

5. INTERIM ANALYSES AND INTERNAL DATA SAFETY MONITORING COMMITTEE

No formal interim analysis is planned for this clinical study.

A Safety Monitoring Committee independent from the study will review unblinded data during the course of the study. This Committee will include a physician with experience in drug development, a biostatistician, and a safety physician. External medical experts (such as an expert in the field of pSS and an infectious diseases expert) may be included to support data interpretation. The Committee will review unblinded safety data and assess any potential safety issues arising during the conduct of the clinical study, including (but not limited to) any potential issues in the context of the SARS-CoV-2 pandemic. The process is described in a separate ‘Safety Monitoring Committee Charter’.

6. STATISTICAL ANALYSES

6.1. Subject Information

Subject information will be tabulated using the Safety Analysis Set, unless specified otherwise. No inferential testing will be performed, nor will p-values be provided.

Subject information will be tabulated with descriptive statistics per planned treatment group, as well as overall.

6.1.1. Demographic and Baseline Disease Characteristics

The following parameters will be summarized:

- date of ICF signature (listed);
- sex;
- age at signing the ICF (years);
- age, categorized (years): 18≤age<65; 65≤age≤74;
- race and ethnicity;
- height at baseline (cm);
- weight at baseline (kg);
- body mass index (BMI) at baseline (kg/m²) $= \frac{weight\ (kg)}{height^2\ (m^2)}$;
  (BMI will not be recalculated if already available in the database)
- duration of Sjogren's Syndrome disease (years)
  $= \frac{(date\ of\ first\ study\ drug\ administration) - (date\ of\ diagnosis) + 1\ day}{365.25}$

For the date of diagnosis, if missing month, impute as January (01), if missing day, impute as the first day of the month (01).

- duration of Sjogren's Syndrome disease, categorized (years): < 1; 1 ≤ duration < 5; 5 ≤ duration < 10;
- ESSDAI at baseline;
- ESSPRI at baseline;

[redacted]

Demographic and baseline disease characteristics will be listed.

6.1.2. Allocation and Randomization

The number of subjects (and percent) per treatment group and overall will be tabulated per country and site.

The following listings will be provided:

- Listing per subject of the IP kit numbers, IP lot numbers dispensed/returned.
- Randomization schemes and codes, with the treatment assigned to each subject. This listing includes a flag in case of discrepancies or errors between the assigned and the actual treatment taken, and should also include flags declaring any (potential) unblinding.

6.1.3. Disposition Information

The following tabulations will be provided, by treatment group and overall:

- The number of subjects screened, not-randomized (including reasons), randomized (treated and not treated), and treated with GLPG3970 and Placebo.
- Number (percent) of subjects randomized per country and site.
- The number (percent) of subjects in each analysis set as defined in Section 4.1.
- The number (percent) of subjects per analysis window as defined in Section 4.3.3.
- The number (percent) of subjects who completed/discontinued the IP administration schedule and the reasons for discontinuation.
- The number (percent) of subjects who completed/discontinued the study and the reasons for discontinuation.

Subjects who discontinued early from study or treatment will be listed. Discontinuation related to SARS-CoV-2 will be flagged in the listing.

Additionally, the following information will be provided in listings:

- Subject identification and randomization (country, site number, investigator, subject number).
- Randomization number and date planned and actual randomization group, with flags for any discrepancies.
- Subjects excluded from the safety and full analysis sets.
- First and Last Key Dates in the Study will be listed.

6.1.4. Protocol Deviations and Eligibility

Major protocol deviations are determined and recorded while the study is ongoing, and the list is finalized prior to database lock (and unblinding). For more details, please refer to the Protocol Deviations Plan.

The number (percent) of subjects with major protocol deviations will be tabulated, overall and per class of deviation, by treatment group and overall.

All available information concerning major protocol deviations, including COVID-19 related verbatim, violations on eligibility criteria and subjects not treated will be listed. Protocol deviations leading to the exclusion of subjects from any analysis set, as applicable, will be flagged.

Minor protocol deviations will not be tabulated and will be listed.

6.1.5. Medical History and Concomitant Diseases

Frequency tabulations, per treatment group and overall, per system organ class and preferred term will be provided for the medical history findings (i.e., conditions no longer present at the start of the study) as well as for the concomitant diseases (i.e. conditions present at the start of the study) using Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary.

All medical history findings and concurrent diseases will be listed separately.

6.1.6. Prior and Concomitant Medications

6.1.6.1. Coding of Reported Terms

All prior and concomitant therapy terms will be coded in the database using the World Health Organization (WHO) drug coding dictionary.

6.1.6.2. Classification of Therapies

All prior and concomitant medication records will be categorized as follows, considering their date and flags indicating the relative timing versus study (drug) start or end (before, after, ongoing):

- Prior only: when the record ended before the first IP administration date.
- Concomitant only: when the record started on or after the first IP administration date.
- Prior and concomitant: when the record started before the date of first IP administration, and ended on or after this point, or continued.

When the start or end date of the prior and concomitant medication records are incomplete (and no flags indicating relative timing are available), the date of first IP administration will be considered to the same level of information provided by these incomplete dates to categorize the timing of these records. This means that a record only having month and year will be categorized

comparing only to the month and the year of the date of first IP administration. The record only having year will be categorized comparing only to the year of the date of first IP administration.

6.1.6.3. Calculation of Relative Days

If both the start and the end dates of the concomitant medication records are completed, their day relative to the day of first IP administration will be calculated as described in Section 4.3.1.

6.1.6.4. Presentation of Results

A frequency tabulation per planned treatment group and overall of the Anatomical Therapeutic Chemical (ATC) classification classes level 4 by therapeutic subgroup (ATC level 2) and generic term of the prior medication (defined as 'prior only' and 'prior and concomitant') will be provided as well as of the concomitant medication (defined as 'concomitant only' and 'prior and concomitant').

A listing of prior and concomitant medications will be provided. Also, prior and concomitant corticosterioid listing will be provided separately (the list of the ATC codes to be selected in this listing will be provided by the Galapagos medical services).

6.1.7. Procedures

Procedures results will only be listed.

6.1.8. Exposure to Study Drug and Compliance

6.1.8.1. Derivation Rules

Derived Parameters: Extent of Exposure to IP

- *Total treatment duration* (days) = last IP administration date – first IP administration date + 1 day, inclusive, regardless of temporary interruptions in IP administration.
- *Total treatment duration, excluding days off IP*: Sum of the number of days with any IP administration.
- *Total treatment duration, full compliance* (days): Number of days where IP has been taken.

Derived Parameters: Compliance

- Overall compliance (%) $= 100 \times \frac{number\ of\ doses\ actually\ taken}{number\ of\ doses\ that\ should\ have\ been\ taken}$
- Percent days with any intake (%) $= 100 \times \frac{total\ treatment\ duration,\ \ excluding\ days\ off\ drug}{total\ treatment\ duration}$
- Percent days full compliance (%) $= 100 \times \frac{total\ treatment\ duration, fully\ compliant}{total\ treatment\ duration}$

6.1.8.2. Presentation of Results

Summary statistics per planned treatment group and overall will be provided for each compliance and extent of exposure parameter for IP administration. Frequency tables will be provided for the compliance parameters, using the following categories: <80%; 80% ≤ x < 100%; 100%; 100% < x ≤ 120%; > 120%.

All original IP administration records will be listed per subject. The listing will include all deviations from IP schedules such as missed or reduced doses and GLPG3970/placebo switches. IP exposure and compliance data will be listed.

6.2. Efficacy Analyses

Efficacy analyses will be performed on the FAS.

Tabulations will be shown per planned treatment group.

6.2.1. Level of Significance

Statistical tests for efficacy analysis will be done at a 2-sided significance level of 10%.

6.2.2. Primary Efficacy Parameter

The primary efficacy endpoints for this study is the change from baseline in ESSDAI score at Week 12.

The ESSDAI is a systemic disease activity index to assess 12 domains (i.e. organ systems: constitutional, lymphadenopathy, glandular, articular, cutaneous, pulmonary, renal, muscular, peripheral nervous system, CNS, hematological, biological) in patients with pSS, where each of the domains is divided into 3-4 levels of activity. These scores are summed across the 12 domains in a weighted manner to provide the total score (range 0-123). A clinically meaningful reduction from baseline (>=3 points) indicates the improvement of symptoms.

The ESSDAI will be assessed at baseline, week 4, week 8 and week 12.

6.2.2.1. Derivation Rules

The binary categorizations of the ESSDAI actual values calculated per visit is defined as follows:

- ESSDAI Responder: Subjects having reduction from baseline ≥ 3 points by analysis window

6.2.2.2. Analyses Methods

Analysis will be performed on total score as well as the 12 domain scores for each treatment group and analysis visit. Observed and change from baseline will be summarized using descriptive statistics by treatment group and visit.

An MMRM will be used on the change from baseline to week 12 to compare treatment groups, with a 90% CI of the treatment difference.

The following covariates will be used: treatment-by-visit interaction and baseline-by-visit interaction. An unstructured variance-covariance matrix will be used to model the residuals. If the default Newton–Raphson algorithm used by SAS PROC MIXED fails to converge, the following will be tried to avoid lack of convergence while maintaining an unstructured variance:

A. The Fisher scoring algorithm (via the SCORING option of the PROC MIXED statement) will be used to obtain the initial values of covariance parameters.

B. If the above fails, the no-diagonal factor analytic structure will be used which effectively performs the Cholesky decomposition via the TYPE=FA0(V) option of the REPEATED statement, where V is the total number of visits in the response vector (V = 3).

C. If all the above fails, the variance-and-correlations parameterization will be attempted using TYPE =UNR

In the rare case where all the above fails, the Toeplitz structure with the sandwich variance estimator (EMPIRICAL option) will be used. If this option also fails, the AR(1) + random intercept model will be tried with the sandwich variance estimator.

The LSM, LSM difference between GLPG3970 and the placebo group (reference category), standard error, two-sided 90% CIs for the LSM and difference between treatment groups and p-value for treatment differences will be presented. The number of subjects in the analysis population and number of subjects in the analysis will be provided by treatment group.

Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline will be presented. Also, graph of the LSM (+-SE) change from baseline will be provided. In addition, profile plot of the actual values and change from baseline will be provided by treatment group. All values, including unscheduled visits will be presented in profile plot of the actual values. If there are more than one visits on the same day, time will be considered to order them.

Analysis of Binary Endpoint

The ESSDAI response will be summarized for the OC and NRI using number and percentages by treatment group. Fisher's exact method will be used to compare the GLPG3970 and placebo based on NRI rates. Fisher's exact p-value and 90% modified Santner-Snell confidence interval of the difference in response rates will be presented (In particular, the modification of Santner-Snell methodology proposed by Chang and Zhang in 1999 will be used (score option)). In addition, 90% Clopper-Pearson confidence interval based on NRI rates for each treatment group will be provided.

A frequency plot over time will be provided for the subjects having reduction ESSDAI $\geq$ 3 points.

6.2.2.3. Decision Making Methodology

The decision making methodology described by Frewer et al (Frewer, Mitchell, Watkins, & Matcham, 2016) will be used to provide further insight into the treatment effect of GLPG3970 over placebo, and will support scenario analyses.

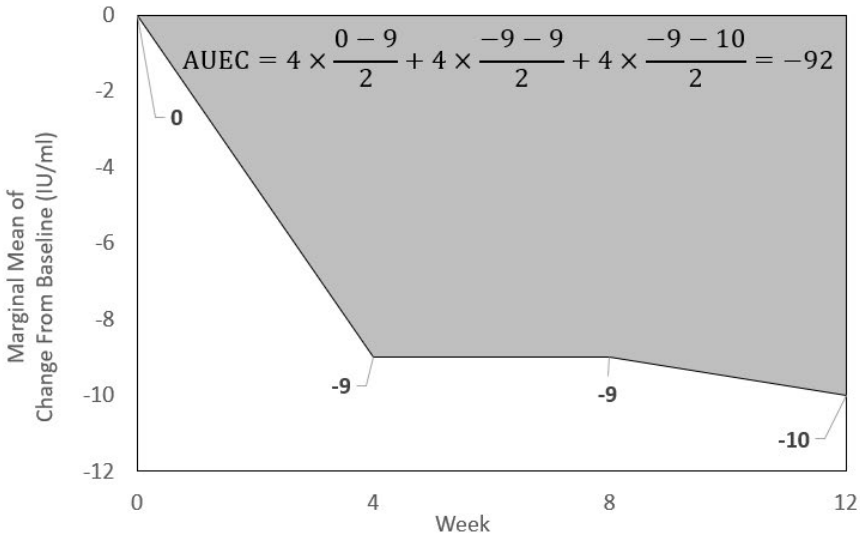
The posterior distribution of this treatment effect will be estimated, and from this distribution probabilities of reaching at least a certain effect (delta) will be derived, e.g., a range of plausible effect size values going from as high as P (delta <=-5) to as low as P (delta <=-2). A common SD of 4 was assumed for both active and placebo arm.

More details about this analysis will be provided in a separate document. This output will be generated for decision making purposes and will not be part of the CSR.

6.2.2.4. Area under the Efficacy Curve (AUEC)

To further explore the data in a slightly different fashion, the area under the efficacy curve (AUEC) based on the MMRM-derived LS-means will be derived for GLPG3970 and placebo. The AUEC will be calculated using the trapezoidal rule as illustrated in Figure 2. A 90% confidence interval and between-group p-value will be derived as well. All will be based on the MMRM model for change from baseline in total and 12 domain scores.

Figure 2. Graphical illustration, calculation of AUEC.



Week	Mean
0	0
4	-9
8	-9
12	-10

Period	Mean AUC
0-4	-18.0
4-8	-36.0
8-12	-38.0
0-12	**-92.0**

6.2.3. Secondary Efficacy Parameters

6.2.3.1. EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI) Score

The ESSPRI is a patient self-reported outcome measure to assess dryness, limb pain, fatigue, and mental fatigue, where each of the domains normally reported as 0 (not at all) to 10 (extremely severe). The final ESSPRI score is the average of 3: dryness, pain, and fatigue. A clinically significant reduction from baseline of the ESSPRI score (at least one point or 15% of the baseline value) indicates the improvement of symptoms.

The ESSPRI will be assessed at baseline, week 4, week 8 and week 12.

6.2.3.2. Derivation Rules

The binary categorizations of the ESSPRI actual values calculated at week 12 is defined as follows:

- ESSPRI Responder: Subjects having reduction ≥ 1 point from baseline or 15% change (i.e., decrease) from baseline

6.2.3.3. Analyses Methods

Analysis will be performed on average score as well as the 3 domain scores for each treatment group and analysis visit. Observed and change from baseline over time will be summarized using descriptive statistics by treatment group and visit.

The analysis of change from baseline up to week 12 in average and domain scores will use the same method as for analysis of the primary efficacy endpoint as detailed in section 6.2.2.2.

Graphs of the mean (+- SE) actual values over time, mean (+- SE) change from baseline and LSM (+-SE) change from baseline will be presented. In addition, profile plot of the actual values will be provided by treatment group. A frequency plot over time will be provided for the subjects having reduction ≥ 1 point from baseline or 15% change (i.e., decrease) from baseline

The analysis of the binary categorizations of ESSPRI at each time point will use same method as for the analysis of ESSDAI binary categorization as detailed in section 6.2.2.2.

To further explore the data in a slightly different fashion, an AUEC based on the MMRM-derived LS-means will be derived for GLPG3970 and placebo as detailed in section 6.2.2.4.

A listing will be provided for the ESSPRI domain scores and the total score along with actual values and change from baseline.

6.2.4. Other Efficacy Parameters

6.2.4.1. [redacted]

[redacted]

6.2.4.2. [redacted]

[redacted]

6.2.4.2.1 [redacted]

[redacted]

6.2.4.2.2 [redacted]

[redacted]

6.2.4.3. [redacted]

[redacted]

6.2.4.3.1 [redacted]

[redacted]

6.2.4.3.2 [redacted]

[redacted]

6.2.4.4. [redacted]

[redacted]

6.2.4.4.1 [redacted]

[redacted]

6.2.4.4.2 [redacted]

[redacted]

6.2.4.5. [redacted]

[redacted]

6.2.4.5.1 [redacted]

[redacted].

6.2.4.5.2 [redacted]

[redacted]

6.2.4.6. [redacted]

[redacted]

6.2.4.6.1 [redacted]

[redacted]

6.2.4.6.2 [redacted]

[redacted]

6.2.4.7. [redacted]

[redacted]

[redacted]

6.2.4.7.1 [redacted]

[redacted]

6.2.4.7.2 [redacted]

[redacted]

6.2.4.8. [redacted]

[redacted]

6.2.4.8.1 [redacted]

[redacted]

-
-
-

6.2.4.8.2

6.2.4.9.

6.2.4.9.1

6.2.4.9.2

6.2.5. Subgroup Analyses

No subgroup analysis is planned for this clinical study as the study sample size is too small. Thus, any subgroup of this would have small sample size to be relevant or even uninterpretable.

6.3. Safety Analyses

Safety analyses will be performed on the Safety Analysis Set.

Safety tables will be presented by treatment group.

Safety parameters will be analyzed descriptively (see Section 4.5.2 and Section 4.5.3). No formal testing will be performed to compare the treatment groups.

6.3.1. Adverse Events

All adverse events (AEs) and changes in attributes (worsening and improvement) of AEs are reported in the database. An identification number serves to link the records considered by the investigator as describing the evolution of one and the same event.

6.3.1.1. Definition of Treatment-Emergent Adverse Events

The analysis of AEs will be based on treatment-emergent events (TEAE).
TEAEs are defined as

- Any AE with an onset date on or after the IP start date and no later than 30 days after last dose of IP, or any worsening of any AE on or after the IP start date.
- Improvement or no change of any ongoing AEs on or after the IP start date are not considered treatment-emergent. If an AE is ongoing at the time of first IP intake, if there is no change or an improvement in its toxicity grade or its seriousness status (reported in the 'Change in Adverse event details Entry' CRF page), this AE will not be considered as treatment-emergent.

6.3.1.2. Coding of Reported Terms

All AE terms will be coded in the database using the latest version of MedDRA coding dictionary. AEs will be graded using the latest version of the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE).

All tables will show the AE preferred terms grouped into system organ class. Subject listings will also show the reported terms (verbatim). Any other coding levels will only be shown in a listing summarizing coding unless explicitly mentioned otherwise.

6.3.1.3. Allocation of Adverse Events to Analysis Periods

All AEs will be placed into analysis periods considering their start date, aiming to report the incidence of these events only in the analysis period during which they started.

The general rule for allocation of AEs to analysis periods follows:

$$\text{Analysis period start date} \leq \text{AE start date} \leq \text{analysis period end date}$$

If the start date of an AE is missing or incomplete to a level preventing a clear allocation of the AE to one single analysis period and no flag indicating timing relative to study medication is available, a worst-case consideration (see below) will be done aiming to allocate the AE to one single analysis period, if possible. When a worst-case consideration is needed, the end date of the AE, if and as available, should also be considered; if such AEs clearly end on a given point, this will exclude the possibility to allocate the AE to an analysis period after that point.

- An AE which according to the available information of its start date could belong to the screening as well as to the treatment analysis period will only be placed in the treatment period.
- An AE with a missing start date will be allocated to the treatment period unless the "Prior to First GLPG3970/Placebo Treatment" = "Yes".

6.3.1.4. Treatment Relatedness

Following the guideline ICH-E3 Structure and Content of Clinical Study Reports (Step 4 Version), the originally reported relatedness to IP of an AE will be dichotomized as follows:

- *Not IP related*: all non-missing weaker levels of relatedness than 'possibly drug related'.
- *IP related*: 'possibly drug related' and all stronger levels of relatedness (this class also includes any missing drug relatedness, as a worst-case consideration).

Only this dichotomized relatedness will be used in tables and can apply to different IPs when relatedness has been collected separately per IP; relatedness as originally reported will only be listed.

6.3.1.5. Worst-Case Selections

When cross-tabulating AE preferred terms versus an AE attribute (like intensity), only the worst-case within each same preferred term, same subject and same analysis period will be considered, i.e. when the same subject has more than once the same AE preferred term reported in the same treatment group, the subject will be counted only once and will be shown under the worst outcome (like the worst intensity for that AE in the concerned treatment period).

The severity grade of events for which the investigator did not record severity will be categorized as "missing" for tabular summaries and data listings. The missing category will be presented last in summary presentation.

6.3.1.6. Calculation of Relative Days and Duration

For each newly reported event, and reported worsening or improvement of an existing event, the start day in the study (the day of the AE start date relative to the date of first IP administration), the start day in the analysis period, and the duration (in days) will be calculated. In addition, the relative day and duration will be derived for the entire event; that is, the full evolution of the event, including the initial reporting and all subsequent worsening and/or improvement.

Relative days and durations will only be listed.

See Section 4.3.1 for the calculation of relative days.

6.3.1.7. Events of Special Interest

Not applicable.

6.3.1.8. Presentation of Results

The analysis will focus on AEs reported during the treatment period. AEs reported during the screening period will only be listed.

All AEs tables will show the number of subjects with TEAEs.

AEs which are not treatment-emergent will only be listed.

A summary table will be provided, showing the number (percent) of subjects with at least one:

- TEAE,
- IP-related TEAE,
- Serious TEAE,
- TEAE leading to death,
- TEAEs by worst intensity (CTCAE Toxicity Grade),
- TEAE leading to IP interruption,
- TEAE leading to IP discontinuation.

The AE terms will be presented sorted in descending order of frequency, prioritizing the GLPG3970 treatment group and then for placebo, first by system organ class, then by preferred term, and then alphabetically.

Frequency tabulations, by system organ class and preferred term, of the number (percent) of subjects with a TEAE will be presented. Similar tables will be provided by worst intensity, and for IP-related TEAEs, grade 3-4-5 TEAEs, serious TEAEs and TEAEs leading to IP discontinuation of IP.

All AE data will also be listed, including separate lists for SAEs, AEs leading to death, and AEs leading to IP discontinuation. Listings will clearly indicate AEs to be treatment-emergent or not.

In addition, AEs related to SARS-Cov-2 will be listed separately.

6.3.1.9. EudraCT Adverse Events Reporting

For the purpose of EudraCT reporting, the following tabulations will be created:

Frequency tabulations, by system organ class and preferred term, of the number (percent) of subjects and number of events with non-serious TEAE will be presented. A similar table will be

provided for all serious TEAEs, as well as a table for non-serious TEAEs reported in at least two subjects in any treatment group.

6.3.2. Laboratory Safety

6.3.2.1. Available Data

Laboratory tests scheduled are described in the protocol section 8.5.2. In addition, the derived laboratory tests described in Section 6.3.2.2 will be added.

The statistical analyses will only present results in Standard International (SI) standardized units. Other units will not be presented.

Only data provided by the central laboratory will be used in tables and figures. Results from local labs will be listed only.

[redacted]

6.3.2.2. Derivation Rules

Derived Laboratory Tests

Not applicable.

Fasted and Non-Fasted Results

Laboratory tests that are sensitive to the fasting status: glucose, triglycerides, and insulin.

For these laboratory tests, only results from blood samples drawn in a fasted state will be included in the analysis. Results from blood samples taken in a non-fasted (not fasting and fasting not declared) state will be listed only and no toxicities or abnormalities will be calculated. Laboratory results for which the fasting status is missing will be considered as taken non-fasted.

6.3.2.3. Definition of Toxicity Grades

Toxicity grades will only be derived for laboratory tests for which toxicity grades are available.

Toxicity grades will be determined as implemented in the attached table (Appendix I).

For elevations (H1 to H4), values under the threshold of H1 are reported as 'Grade 0'.

For elevations (L1 to L4), values above the threshold of L1 are reported as 'Grade 0'.

6.3.2.4. Definition of Non-Graded Abnormalities

For all laboratory tests provided by the laboratory, the position of the actual analysis values versus their normal ranges will be determined directly by using the derived position indicator, expressing the classes for these analysis values as low (L), normal (N) or high (H). L, N and H are further referred to as non-graded abnormalities.

6.3.2.5. Urinalysis Tests with Categorical Results

Results of urinalysis with qualitative results will be tabulated by time point. No toxicity grading or non-graded abnormalities will be derived.

6.3.2.6. Treatment-Emergent Principle

Toxicity Grades

A post-baseline toxicity grade 1, 2, 3 or 4 is defined as treatment-emergent when higher than the toxicity grade of the baseline result or when there is a change in direction (from H to L or from L to H). If the baseline result is missing, a post-baseline toxicity grade 1, 2, 3 or 4 will be considered as treatment-emergent.

Non-graded Abnormalities

A post-baseline non-graded abnormality class L or H is defined as treatment-emergent when it differs from the abnormality class of the baseline result. If the baseline result is missing, a post-baseline abnormality L or H will be considered as treatment-emergent.

6.3.2.7. Worst-Case Principle

Toxicity Grading

The worst-case post-baseline toxicity grade 0, 1, 2, 3 or 4 will be determined per subject, per laboratory test (and direction, increases and decreases) and for each analysis period, using all non-missing post-baseline records (including unscheduled and FU visits).

The worst-case toxicity grade is the highest toxicity grade corresponding to the highest laboratory test value (by direction, increases and decreases). In case of several equal highest post-baseline actual values, the earliest occurrence will be taken for the analyses and will be flagged in the listings.

- Grade 0: all post-baseline toxicity grades are classified as 0
- Grade 1: all post-baseline toxicity grades are classified as ≤1
- Grade 2: all post-baseline toxicity grades are classified as ≤2
- Grade 3: all post-baseline toxicity grades are classified as ≤3
- Grade 4: all post-baseline toxicity grades are classified as ≤4

If any record is missing, then the toxicity grade is considered as ‘missing’.

Non-graded Abnormalities

The following worst-case post-baseline abnormalities L, N or H will be determined per subject, per laboratory test and for each analysis period, using all non-missing post-baseline records (including unscheduled and FU visits):

- L = low: at least one post-baseline result is classified as L.
- N = normal: all post-baseline results are classified as N.
- H = high: at least one post-baseline result is classified as H.

If, for a subject, both L and H are reported, the subject will be counted twice in the table: once with a worst-case L and once with a worst-case H.

6.3.2.8. Hepatotoxicity

Hepatotoxicity will be investigated by tabulating the aspartate aminotransferase (AST) and alanine aminotransferase (ALT) values categorized as >3, >5, >10 and >20 times their upper limit of normal (per analyte and over both analytes combined), ALP categorized as > 1.5 times the upper limit and total bilirubin as > 2 times the upper limit of normal. Elevations of AT (AST or ALT) > 3 times their upper limits accompanied by elevated total bilirubin (>1.5×ULN, >2×ULN) on the same day will also be tabulated.

To assess the potential of the drug to cause severe liver damage, possible Hy's Law cases will be identified. These subjects are defined as having any elevated AT (AST or ALT) of >3×ULN, ALP <2×ULN, and associated with an increase in total bilirubin >2×ULN on the same day.

6.3.2.9. Presentation of Results

Continuous laboratory tests including, glucose, cholesterol, triglycerides, low-density lipoprotein (LDL), high-density lipoprotein (HDL) and fasted glucose, will be summarized by means of descriptive statistics (including 90% CI of the mean change) by laboratory test, treatment group and analysis window. Actual values and change from baseline will be tabulated separately.

[redacted]

Profile line plots by subject of all actual observed values using relative day (ADY) and profile plots on change from baseline by analysis window will be presented by treatment group. All values, including unscheduled visits will be presented in profile plot. If there are more than one visits on the same day, time will be considered to order them. Graphs of the mean (+- SE) actual values over time and the mean (+- SE) change from baseline, and percent changes from baseline will be presented for all continuous laboratory parameters, with the exception of the metabolic markers.

The analysis of abnormalities will focus on assessments reported during the treatment period. Results reported before baseline will only be listed.

Non-graded abnormalities and toxicities grades of the actual values will be presented as shift tables of the worst-case abnormality/toxicity grade versus the baseline abnormality/toxicity grade. The table will be created per laboratory test, treatment group and analysis window. The results of non-graded abnormalities and toxicities grades will be shown separately.

A frequency table of the number (percent) of subjects with treatment-emergent worst-case abnormalities/toxicity grade per laboratory test, treatment group and analysis period will be presented. The results of non-graded abnormalities and toxicities grades will be shown separately.

A frequency table of the number (percent) of subjects will be also provided for hepatotoxicity flags defined before. These abnormalities will be also flagged in the individual data laboratory listing.

Listing will be provided for the laboratory test along with flag for abnormality results. In addition, pregnancy result will be listed separately.

6.3.3. Electrocardiogram

6.3.3.1. Available Data

The following ECG parameters will be analysed: heart rate (HR), RR interval, PR interval, QRS interval, uncorrected QT interval, QTcF (derived).

6.3.3.2. Derivation Rules

Derived Parameters

The Fridericia's cube-root corrected QT (QTcF) will be calculated using the following formula.

Fridericia's cube-root corrected QT (Fridericia, 1920): $$QTcF\ (ms) = QT\ (ms) \times \sqrt[3]{\frac{1000}{RR\ (ms)}}$$

If RR is missing, then it will be derived from HR using the formula *RR(ms) = 60000 / HR(bpm)*.

Handling of ECGs Measured in Triplicate

If ECG is collected in triplicates, the following approach will be taken.

First, any derivation of ECG parameters will be done before handling ECG triplicates. Next, the mean of the triplicate ECG values will be calculated for each individual ECG parameter, without rounding the result. These calculated means will constitute the analysis values; any derivation (e.g. change from baseline, assignment of abnormalities) and statistic will be based on the mean value of the triplicates.

The values of the original members of a triplicate will be listed.

6.3.3.3. Abnormalities

The actual analysis values and changes from baseline of the QT and QTcF parameters will be categorized into the abnormality classes as defined in the SAP Appendix II.

Worst-Case Abnormality

The worst-case post-baseline categorized actual analysis value and the worst-case categorized change from baseline for QT and QTcF will be determined per subject, per parameter, and for each analysis period, using all non-missing post-baseline records (including unscheduled and FU visits).

The worst-case categorized actual analysis value is the category corresponding to the highest post-baseline actual value. In case of several equal highest post-baseline actual values, the earliest occurrence will be taken for the analyses and will be flagged in the listings.

The worst-case change from baseline is the category corresponding to the largest increase (positive change) from baseline. In case of several equal largest increase from baseline, the earliest occurrence will be taken for the analyses and will be flagged in the listings.

Treatment-Emergent Abnormalities

Actual value: An abnormal post-baseline abnormality is defined as treatment-emergent when the abnormality is worse compared to the abnormality at baseline or when there is a change in direction (from H to L or from L to H). When the baseline value is missing, post-baseline abnormalities are considered as treatment-emergent.

An abnormal category for change from baseline is always treatment-emergent.

6.3.3.4. Presentation of Results

No formal inferential statistics (p-values) will be derived.

Continuous parameters will be summarized by means of descriptive statistics (including 90% CI of the mean change) by parameter, treatment group and analysis window. Actual values and change from baseline will be tabulated separately.

Profile line plot by subject of all actual observed values using relative day (ADY) and profile plot on change from baseline by analysis window will be presented by treatment group. All values, including unscheduled visits will be presented in profile plot. If there is more than one visit on the same day, time will be considered to order them. Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline will be presented.

ECG will be further explored on the subset of subjects taking anti-malaria medications. Profile line plot by subject of actual observed values at baseline and at 1.5, 3 and 4.5 hours postdose will be presented by treatment group. In addition, profile line plot of the change from baseline to 1.5, 3 and 4.5 hours postdose will be presented.

The analysis of abnormalities will focus on assessments reported during the treatment period. Results reported during the screening period will only be listed.

Abnormalities of the actual values will be presented as shift tables of the worst-case abnormality versus the baseline abnormality. The table will be created per parameter, treatment group and analysis period.

A frequency table of the number (percent) of subjects with treatment-emergent worst-case abnormalities and the worst change per parameter, treatment group and analysis period will be presented.

A frequency table per treatment group and time point of the ECG interpretations as recorded in the case report form (CRF) will be provided.

A listing will be provided along with abnormality.

6.3.4. Vital Signs

6.3.4.1. Available Data

The following vital signs parameters will be analyzed: weight (kg), diastolic and systolic blood pressure (mmHG), heart rate (beats/min) and body temperature (C).

6.3.4.2. Derivation Rules

Derived Parameters

Not applicable.

6.3.4.3. Abnormalities

Vital signs data will be categorized according to the cutoffs provided in the SAP Appendix II.

6.3.4.4. Treatment-Emergent Principle

A post-baseline abnormality class L or H is defined as treatment-emergent when it differs from the abnormality class at baseline. If the baseline result is missing, a post-baseline abnormality L or H will be considered as treatment-emergent.

An abnormal category for change from baseline is always treatment-emergent.

6.3.4.5. Worst-Case Abnormality

The worst-case post-baseline categorized actual analysis value will be determined per subject, per parameter and for each analysis period, using all non-missing post-baseline records (including unscheduled and FU visits).

The worst-case categorized actual analysis value is the category corresponding to the highest post-baseline actual value. In case of several equal highest post-baseline actual values, the earliest occurrence will be taken for the analyses and will be flagged in the listings.

6.3.4.6. Presentation of Results

No formal inferential statistics (p-values) will be derived.

Continuous parameters will be summarized by means of descriptive statistics (including 90% CI of the mean change) by parameter, treatment group and analysis window. Actual values and change from baseline will be tabulated separately.

Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline will be presented.

The analysis of abnormalities will focus on assessments reported during the treatment period. Results reported during the screening or post-treatment (if defined) period will only be listed.

Abnormalities of the actual values will be presented as shift tables of the worst-case abnormality versus the baseline abnormality. The table will be created per parameter (and position if applicable), treatment group and analysis period.

A frequency table of the number (percent) of subjects with treatment-emergent worst-case abnormalities per parameter, treatment group and analysis period will be presented.

A listing will be provided along with abnormality.

6.3.5. Physical Examinations

Physical examination results will only be listed.

6.3.6. SARS-CoV-2 Infection

Results of the SARS-CoV-2 infection test will be listed for the subjects who have at least one positive test.

6.3.7. Subgroup Analyses

No subgroup analysis is planned for this clinical study as the study sample size is too small.

6.4. Pharmacokinetic Analysis

6.4.1. Available Data

Blood samples for the PK assessment of GLPG3970 should be collected on the visits specified in the Schedule of Assessments in Section 2.5.

6.4.2. Presentation of Results

Descriptive statistics will be provided for GLPG3970 plasma level at each assessment time point. Observed predose concentrations (C_{tau}) for GLPG3970 will be reported in the CSR. Individual GLPG3970 plasma concentrations with descriptive statistics will be presented by time point/interval.

Predose samples not collected within 21 to 27 hours after the previous drug intake, or that were taken after the drug administration at the site will be excluded from descriptive statistics and flagged in the table with appropriate footnoting. In case there was a missed dose within 3 days prior to the sample, it will also be excluded from descriptive statistics and flagged in the table with appropriate footnoting. The same will be done with postdose samples taken outside of the defined time intervals. Actual blood sampling times from the last drug intake will be listed.

6.5. [redacted]

6.5.1. Available Data

[redacted]

6.5.2. Presentation of Results

[redacted]

6.6. Changes to the Planned Analyses, not Covered by Protocol Amendments

- Section 2.2.2: As per the protocol secondary endpoint is "Observed GLPG3970 plasma trough concentrations (Ctrough)". Replaced Ctrough with Ctau.
- The prednisolone conversion table in Appendix XII will be used instead of the version in the CSP.

7. REFERENCES

Chan, I. S. F., and Zhang, Z. (1999). “Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions.” Biometrics 55:1202–1209.

Fridericia, L. (1920). Die Systolendauer im Elektrokardiogramm bei normalen Menschen und bei Herzkranken. Acta Medica Scandinavia 53:469-486.

ICH E3 Structure and Content of Clinical Study Reports (Step 4 Version).

ICH E9 Statistical Principles for Clinical Trials (Step 4 Version).

ICH E14 The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs (Step 4 Version).

ICH E6 Intergrated Addendum to ICH E6 (R1): Guideline for good clinical practice (Step 4 Version).

ICH E9(R1) Estimands and Sensitivity Analysis in Clinical Trials (Step 2 Version).

APPENDIX

Appendix I: Laboratory: Toxicity Grading

The following tables contain a list of safety tests with associated gradings. There may be tests in this table that are not measured in this particular study.

Table 6 Gradings for Hematology Parameters

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
Hemoglobin	<LLN	<10 g/dL; <6.2 mmol/L	<8 g/dL; <4.9 mmol/L	NA	>ULN	Increase in >2 g/dL above ULN	Increase in >4 g/dL above ULN	NA
Hematocrit	<LLN	NA	NA	NA	>ULN	NA	NA	NA
Mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC)	<LLN	NA	NA	NA	>ULN	NA	NA	NA
Platelet count - *(assuming no platelet cluster)*	<LLN	<75 x 10^9/L	<50 x 10^9/L	<25 x 10^9/L	>ULN	>600 x 10^9/L	>1000 x 10^9/L	NA
Leukocytes	<LLN	<3.0 x 10^9/L	<2.0 x 10^9/L	<1.0 x 10^9/L	>ULN	>20.0 x 10^9/L	>100.0 x 10^9/L	NA
% Polymorphonuclear Leukocytes + Band Cells	NA	NA	NA	NA	>ULN	≥90%	>95%	NA

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
Neutrophils	<LLN	<1.5 x 10^9/L	<1.0 x 10^9/L	<0.5 x 10^9/L	NA	NA	NA	NA
Eosinophils	NA	NA	NA	NA	>ULN	>5.0 x 10^9 or eosinophils >5%	NA	NA
Lymphocytes	<LLN	<0.8 x 10^9/L	<0.5 x 10^9/L	<0.2 x 10^9/L	>ULN	>4.0 x 10^9/L	>20.0 x 10^9/L	NA
Red blood cells	<LLN	NA	NA	NA	>ULN	NA	NA	NA

Table 7 Gradings for Blood Chemistry Parameters

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
Alanine aminotransferase (ALT)	NA	NA	NA	NA	>ULN	>3.0 x ULN	>5.0 x ULN	>8.0 x ULN
Aspartate aminotransferase (AST)	NA	NA	NA	NA	>ULN	>3.0 x ULN	>5.0 x ULN	>8.0 x ULN
Gamma-glutamyl transferase (GGT)	NA	NA	NA	NA	>ULN	>2.5 x ULN	>5.0 x ULN	>20.0 x ULN
Alkaline Phosphatase (total)	NA	NA	NA	NA	>ULN	>2.5 x ULN	>5.0x ULN	>20.0 x ULN
Lactate dehydrogenase (LDH)	NA	NA	NA	NA	>ULN	NA	NA	NA
Total bilirubin	NA	NA	NA	NA	>ULN	>1.5 x ULN	>3.0 x ULN	>10.0 x ULN
Amylase	NA	NA	NA	NA	>ULN	>1.5x ULN	>2.0 x ULN	>5.0 x ULN
Lipase	NA	NA	NA	NA	>ULN	>1.5x ULN	>2.0 x ULN	>5.0 x ULN
Total protein	<LLN	<5.5 g/dL	<5.0 g/dL	NA	NA	NA	NA	NA
C-reactive protein (CRP)	NA	NA	NA	NA	>5mg/L	>10mg/L	>20mg/L	>30mg/L
Activated partial thromboplastin time (APTT)	NA	NA	NA	NA	>ULN	>1.5 x ULN	>2.5 x ULN	NA
Prothrombin time (PT)	NA	NA	NA	NA	≥1.10 x ULN	≥1.25 x ULN	≥1.50 x ULN	≥3.00 x ULN

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
International normalized ratio (INR)	NA	NA	NA	NA	>1.2 x Baseline	>1.5 x Baseline	>2.5 x Baseline	NA
Creatinine	NA	NA	NA	NA	>ULN	>1.5 x ULN	>3.0 x ULN	>6.0 x ULN
Glucose *(fasting)*	<LLN	<55 mg/dL; <3.0 mmol/L	<40 mg/dL; <2.2 mmol/L	<30 mg/dL; <1.7 mmol/L	>ULN	>125 mg/dL ; ≥6.95 mmol/ L	>250 mg/dL ; ≥13.89 mmol /L	≥500 mg/dL; ≥27.75 mmol/L
Glucose *(non-fasting)*	<LLN	<55 mg/dL; <3.0 mmol/L	<40 mg/dL; <2.2 mmol/L	<30 mg/dL; <1.7 mmol/L	>ULN	>160 mg/dL; ≥8.89 mmol/ L	>250 mg/dL; ≥13.89 mmol /L	≥500 mg/dL; ≥27.75 mmol/L
HbA1c	NA	NA	NA	NA	> 6.0 %	> 6.5 %	NA	NA
Fasting insulin	<2.6 µU/mL or 18.1 pmol/L	NA	NA	NA	>24.9 µU/m or 172.9 pmol/L	NA	NA	NA
Cholesterol	NA	NA	NA	NA	>ULN	>300 mg/dL; >7.75 mmol/L	>400 mg/dL ; >10.34 mmol/L	>500 mg/dL; >12.92 mmol/L
Low-density lipoprotein (LDL)	NA	NA	NA	NA	>ULN	≥160 mg/dL; ≥4.12 mmol/ L	≥190 mg/dL; ≥4.90 mmol/ L	NA
High-density lipoprotein (HDL)	<LLN	<40 mg/dL; <1 mmol/L	NA	NA	NA	NA	NA	NA

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
Triglycerides	NA	NA	NA	NA	>ULN	>300 mg/dL; >3.42 mmol/ L	>500 mg/dL; >5.7 mmol/ L	>1,000 mg/d L; >11.4 mmol/L
Calcium *(corrected for albumin)*	<LLN	<8.0 mg/dL; <2.0 mmol/L	<7.0 mg/dL; <1.75 mmol/L	<6.0 mg/dL; <1.5 mmol/L	>ULN	>11.5 mg/dL ; >2.9 mmol/L	>12.5 mg/dL ; >3.1 mmol/L	>13.5 mg/dL ; >3.4 mmol/L
Ionized calcium	<LLN	<1.0 mmol/L	<0.9 mmol/L	<0.8 mmol/L	>ULN	>1.5 mmol/L	>1.6 mmol/L	>1.8 mmol/L
Sodium	<LLN	<130 mEq/L; <130 mmol/L	<125 mEq/L; <125 mmol/L	<120 mEq/L; <120 mmol/L	>ULN	>150 mEq/L; >150 mmol/L	>155 mEq/L; >155 mmol/L	>160 mEq/L; >160 mmol/L
Chloride	<LLN	NA	NA	NA	>ULN	NA	NA	NA
Potassium	<LLN	NA	<3.0 mmol/L	<2.5 mmol/L	>ULN	>5.5 mmol/L	>6.0 mmol/L	>7.0 mmol/L
Phosphate	<LLN	<2.0 mg/dL; <0.65 mmol/L	<1.4 mg/dL; <0.45 mmol/L	<1.0 mg/dL; <0.32 mmol/L	>ULN	NA	NA	NA
Creatine phosphokinase (CPK)	NA	NA	NA	NA	> ULN	>2.5 x ULN	>5 x ULN	>10 x ULN
Uric acid	NA	NA	NA	NA	>ULN	≥10 mg/dL; ≥0.59 mmol/L	≥12 mg/dL; ≥0.71 mmol/L	≥15 mg/dL; ≥0.89 mmol/L
Albumin	<LLN	<30 g/L	<20 g/L	NA	NA	NA	NA	NA
eGFR (or Cr/Cl)	<LLN	<60 ml/min/ 1.73 m^2	<30 ml/min/ 1.73 m^2	<15 ml/min/ 1.73 m^2	NA	NA	NA	NA

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
Blood urea nitrogen (BUN)	NA	NA	NA	NA	>ULN	>2.5 ULN	>5 ULN	>10 ULN

[redacted]

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
[redacted]								

Table 8 Gradings for Urine Analysis Parameters

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
Urine erythrocytes	NA	NA	NA	NA	≥ULN	≥10 cells/HPF	NA	NA
Urine protein	NA	NA	NA	NA	1+ proteinuria	2+ and 3+ proteinuria; urinary protein <3.5 g/24 hrs	urinary protein ≥3.5 g/24 hrs; 4+ proteinuria	NA
Urine glucose	NA	NA	NA	NA	>ULN (presence of glucose)	NA	NA	NA

LLN: lower limit of normal, ULN: upper limit of normal, NA: not applicable.

Any laboratory parameter with treatment-emergent (i.e., worsening from baseline) abnormalities of grade 2 or above (ie, H2/L2 or higher/lower) should be considered a 'marked laboratory abnormality'. Marked laboratory abnormalities should be described in a corresponding section of the CSR as per ICH E3 guidance, section 12.4.2.3.

Appendix II: Vital Signs and Electrocardiogram grading

Table 9 Gradings for Vital Signs and Electrocardiogram Parameters

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
Systolic blood pressure (mmHg)	<90	<60	NA	NA	NA	≥140	≥160	≥180
Diastolic blood pressure (mmHg)	<60	<45	NA	NA	NA	≥90	≥100	≥120
Heart rate (bpm)	<60	<50	<40	NA	>100	>115	>130	NA
Body temperature (°C)	NA	<35	NA	NA	>38	NA	NA	NA
Respiratory rate (breaths per minute)	<12	NA	NA	NA	>20	NA	NA	NA
O_2 saturation (%)	<95	<90	NA	NA	NA	NA	NA	NA
Weight	≥7% decrease from baseline	NA	NA	NA	≥7% increase from baseline	NA	NA	NA
QTc interval on ECG (ms)	NA	NA	NA	NA	>450	>480	>500	NA
QTc change from baseline (ms)	NA	NA	NA	NA	NA	>30	>60	NA
PR interval on ECG (ms)	<110	NA	NA	NA	>200	NA	NA	NA
QRS complex on ECG (ms)	<60	NA	NA	NA	>120	NA	NA	NA

Source: FRM-MED-005/006, valid on 08JUL2020.

Appendix III: EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI) Score

Domain	Activity level	Description
Constitutional *Exclusion of fever of infectious origin and voluntary weight loss*	No=0	Absence of the following symptoms
	Low=3	Mild or intermittent fever (37.5–38.5°C)/night sweats and/or involuntary weight loss of 5–10% of body weight
	Moderate=6	Severe fever (>38.5°C)/ night sweats and/or involuntary weight loss of >10% of body weight

Domain	Activity level	Description
Articular *Exclusion of osteoarthritis*	No=0	Absence of currently active articular involvement
	Low=2	Arthralgias in hands, wrists, ankles and feet accompanied by morning stiffness (>30 min)
	Moderate=4	1–5 (of 28 total count) synovitis
	High=6	≥6 (of 28 total count) synovitis

Domain	Activity level	Description
Lymphadenopathy and lymphoma *Exclusion of infection*	No=0	Absence of the following features
	Low=4	Lymphadenopathy ≥1 cm in any nodal region or ≥2 cm in inguinal region
	Moderate=8	Lymphadenopathy ≥2 cm in any nodal region or ≥3 cm in inguinal region, and/ or splenomegaly (clinically palpable or assessed by imaging)
	High=12	Current malignant B-cell proliferative disorder

Domain	Activity level	Description
Cutaneous *Rate as 'No activity' stable long-lasting features related to damage*	No=0	Absence of currently active cutaneous involvement
	Low=3	Erythema multiforma
	Moderate=6	Limited cutaneous vasculitis, including urticarial vasculitis, or purpura limited to feet and ankle, or subacute cutaneous lupus
	High=9	Diffuse cutaneous vasculitis, including urticarial vasculitis, or diffuse purpura, or ulcers related to vasculitis

Domain	Activity level	Description
CNS *Rate as 'No activity' stable long-lasting features related to damage or CNS involvement not related to the disease*	No=0	Absence of currently active CNS involvement
	Moderate=10	Moderately active CNS features, such as cranial nerve involvement of central origin, optic neuritis or multiple sclerosis-like syndrome with symptoms restricted to pure sensory impairment or proven cognitive impairment
	High=15	Highly active CNS features, such as cerebral vasculitis with cerebrovascular accident or transient ischaemic attack, seizures, transverse myelitis, lymphocytic meningitis, multiple sclerosis-like syndrome with motor deficit

Domain	Activity level	Description
Biological	No=0	Absence of any of the following biological feature
	Low=1	Clonal component and/or hypocomplementemia (low C4 or C3 or CH50) and/or hypergammaglobulinemia or high IgG level between 16 and 20 g/L
	Moderate=2	Presence of cryoglobulinemia and/ or hypergammaglobulinemia or high IgG level >20 g/L, and/or recent onset hypogammaglobulinemia or recent decrease of IgG level (<5 g/L)

Domain	Activity level	Description
Pulmonary *Rate as 'No activity' stable long-lasting features related to damage, or respiratory involvement not related to the disease (tobacco use, etc)*	No=0	Absence of currently active pulmonary involvement
	Low=5	Persistent cough **due to** bronchial involvement with no radiographic abnormalities on radiography Or radiological or HRCT evidence of interstitial lung disease with: no breathlessness and normal lung function test
	Moderate=10	Moderately active pulmonary involvement, such as interstitial lung disease shown by HRCT with shortness of breath on exercise (NHYA II) or abnormal lung function tests restricted to: 70% >DL_{CO} ≥40% or 80% >FVC≥60%
	High=15	Highly active pulmonary involvement, such as interstitial lung disease shown by HRCT with shortness of breath at rest (NHYA III, IV) or with abnormal lung function tests: DL_{CO} <40% or FVC <60%

FVC, forced vital capacity; HRCT, high-resolution CT; NYHA, New York Heart Association.

Domain	Activity level	Description
Haematological *For anaemia, neutropenia, and thrombopenia, only auto-immune cytopenia must be considered Exclusion of vitamin or iron deficiency, drug-induced cytopenia*	No=0	Absence of autoimmune cytopenia
	Low=2	Cytopenia of autoimmune origin with neutropenia (1000<neutrophils<1500/mm^3), and/or anaemia (10<haemoglobin<12 g/dL), and/or thrombocytopenia (100 000<platelets<150 000/mm^3) Or lymphopenia (500<lymphocytes<1000/mm^3)
	Moderate=4	Cytopenia of autoimmune origin with neutropenia (500≤ neutrophils ≤1000/mm^3), and/or anaemia (8≤ haemoglobin ≤10 g/dL), and/or thrombocytopenia (50 000 ≤ platelets ≤100 000/mm^3) Or lymphopenia (≤500/mm^3)
	High=6	Cytopenia of autoimmune origin with neutropenia (neutrophils<500/mm^3), and/or or anaemia (haemoglobin<8 g/dL) and/or thrombocytopenia (platelets<50 000/mm^3)

Domain	Activity level	Description
Glandular *Exclusion of stone or infection*	No=0	Absence of glandular swelling
	Low=2	Small glandular swelling with enlarged parotid (≤3 cm), or limited submandibular (≤2 cm) or lachrymal swelling (≤1 cm)
	Moderate=4	Major glandular swelling with enlarged parotid (>3 cm), or important submandibular (>2 cm) or lachrymal swelling (>1 cm)

Domain	Activity level	Description
Muscular *Exclusion of weakness due to corticosteroids*	No=0	Absence of currently active muscular involvement
	Low=6	Mild active myositis shown by abnormal EMG, MRI* or biopsy with no weakness and creatine kinase (N≤CK≤2N)
	Moderate=12	Moderately active myositis proven by abnormal EMG, MRI* or biopsy with weakness (maximal deficit of 4/5), or elevated creatine kinase (2N<CK≤4N),
	High=18	Highly active myositis shown by abnormal EMG, MRI* or biopsy with weakness (deficit ≤3/5) or elevated creatine kinase (>4N)

*We decided to add this item not included in the initial version since the value of this examination for the diagnosis of myositis was not clear until recently.
EMG, electromyogram.

Domain	Activity level	Description
Renal *Rate as 'No activity' stable long-lasting features related to damage and renal involvement not related to the disease.* *If biopsy has been performed, please rate activity based on histological features first*	No=0	Absence of currently active renal involvement with proteinuria <0.5 g/day, no haematuria, no leucocyturia, no acidosis or long-lasting stable proteinuria due to damage
	Low=5	Evidence of mild active renal involvement, limited to tubular acidosis without renal failure or glomerular involvement with proteinuria (between 0.5 and 1 g/day) and without haematuria or renal failure (GFR ≥60 mL/min)
	Moderate=10	Moderately active renal involvement, such as tubular acidosis with renal failure (GFR <60 mL/min) or glomerular involvement with proteinuria between 1 and 1.5 g/day and without haematuria or renal failure (GFR ≥60 mL/min) or histological evidence of extra-membranous glomerulonephritis or important interstitial lymphoid infiltrate
	High=15	Highly active renal involvement, such as glomerular involvement with proteinuria >1.5 g/day, or haematuria or renal failure (GFR <60 mL/min), or histological evidence of proliferative glomerulonephritis or cryoglobulinemia related renal involvement

GFR, glomerular filtration rate.

Domain	Activity level	Description
PNS *Rate as 'No activity' stable long-lasting features related to damage or PNS involvement not related to the disease*	No=0	Absence of currently active PNS involvement
	Low=5	Mild active PNS involvement, such as pure sensory axonal polyneuropathy shown by NCS or trigeminal (V) neuralgia **Proven small fibre neuropathy*
	Moderate=10	Moderately active PNS involvement shown by NCS, such as axonal sensory–motor neuropathy with maximal motor deficit of 4/5, pure sensory neuropathy with presence of cryoglobulinemic vasculitis, ganglionopathy with symptoms restricted to mild/moderate ataxia, inflammatory demyelinating polyneuropathy (CIDP) with mild functional impairment (maximal motor deficit of 4/5 or mild ataxia) Or cranial nerve involvement of peripheral origin (except trigeminal (V) neuralgia)
	High=15	Highly active PNS involvement shown by NCS, such as axonal sensory–motor neuropathy with motor deficit ≤3/5, peripheral nerve involvement due to vasculitis (mononeuritis multiplex, etc), severe ataxia due to ganglionopathy, inflammatory demyelinating polyneuropathy (CIDP) with severe functional impairment: motor deficit ≤3/5 or severe ataxia

*We decided to add this item not included in the initial version since the link between this entity and SS was not clear until recently.
CIPD, chronic inflammatory demyelinating polyneuropathy; NCS, nerve conduction study.

Appendix IV: EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI) Score

1) How severe has your **dryness** been during the last 2 weeks?

No dryness	0	1	2	3	4	5	6	7	8	9	10	Maximal imaginable dryness

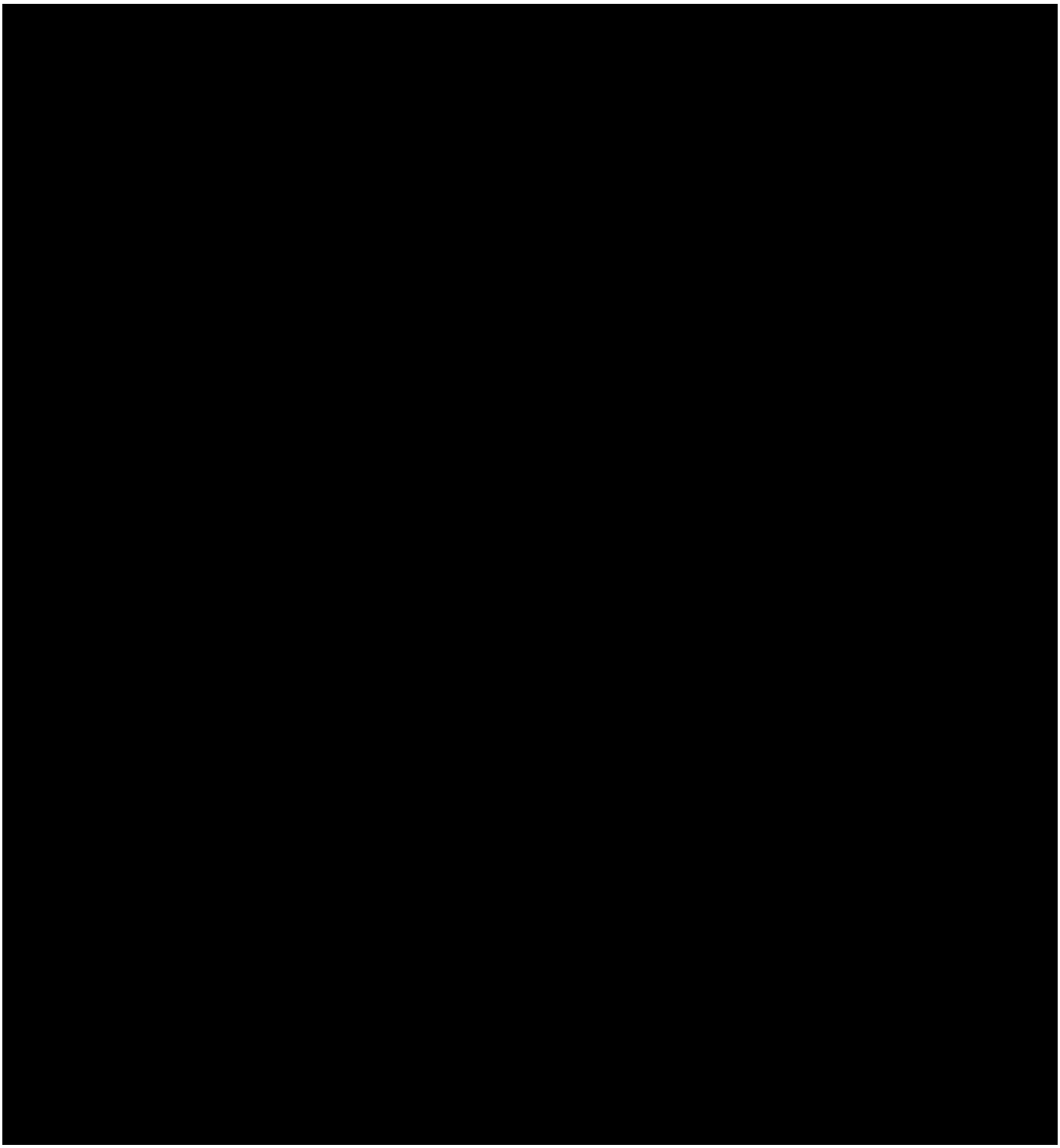
2) How severe has your **fatigue** been during the last 2 weeks?

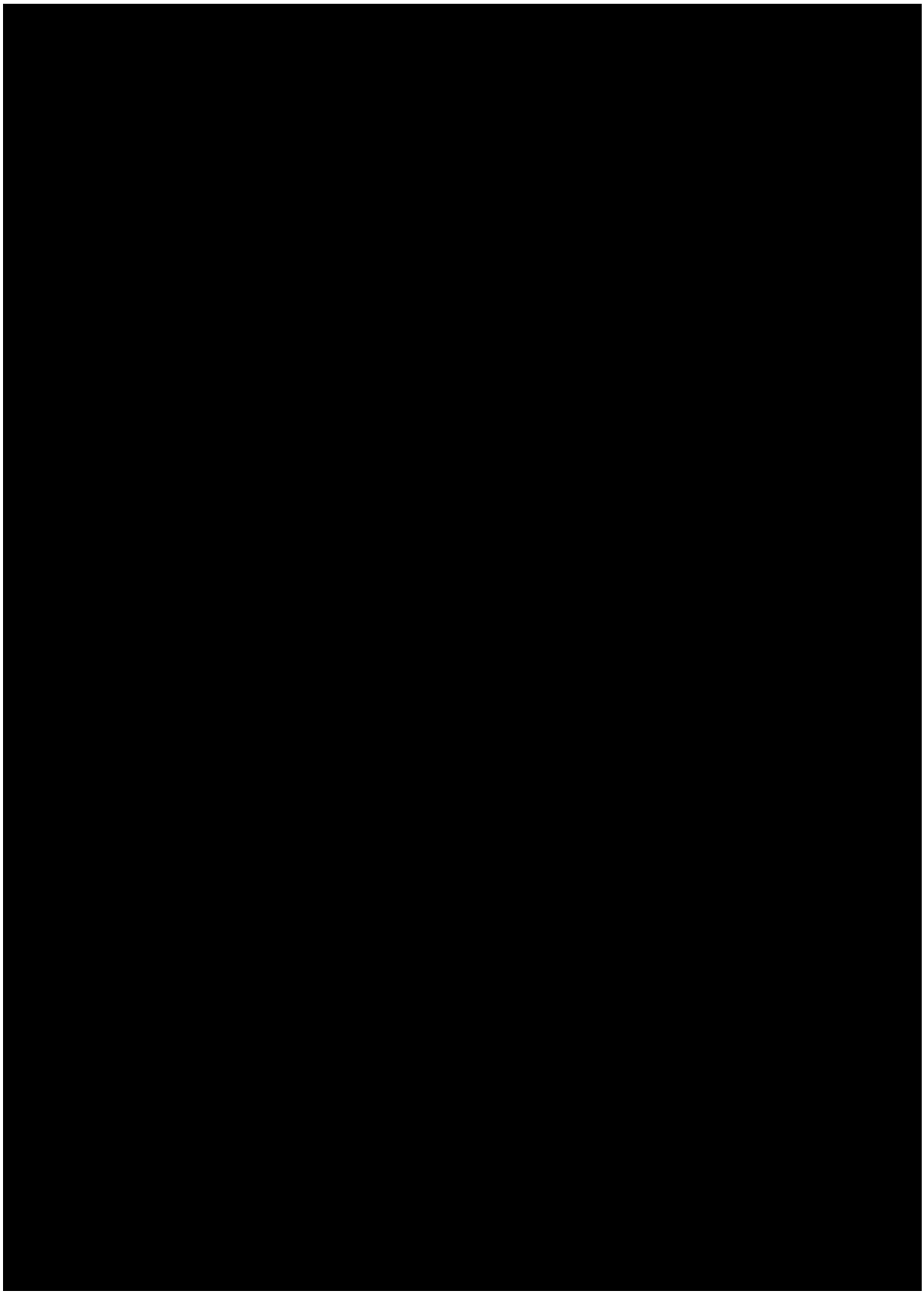
No fatigue	0	1	2	3	4	5	6	7	8	9	10	Maximal imaginable fatigue

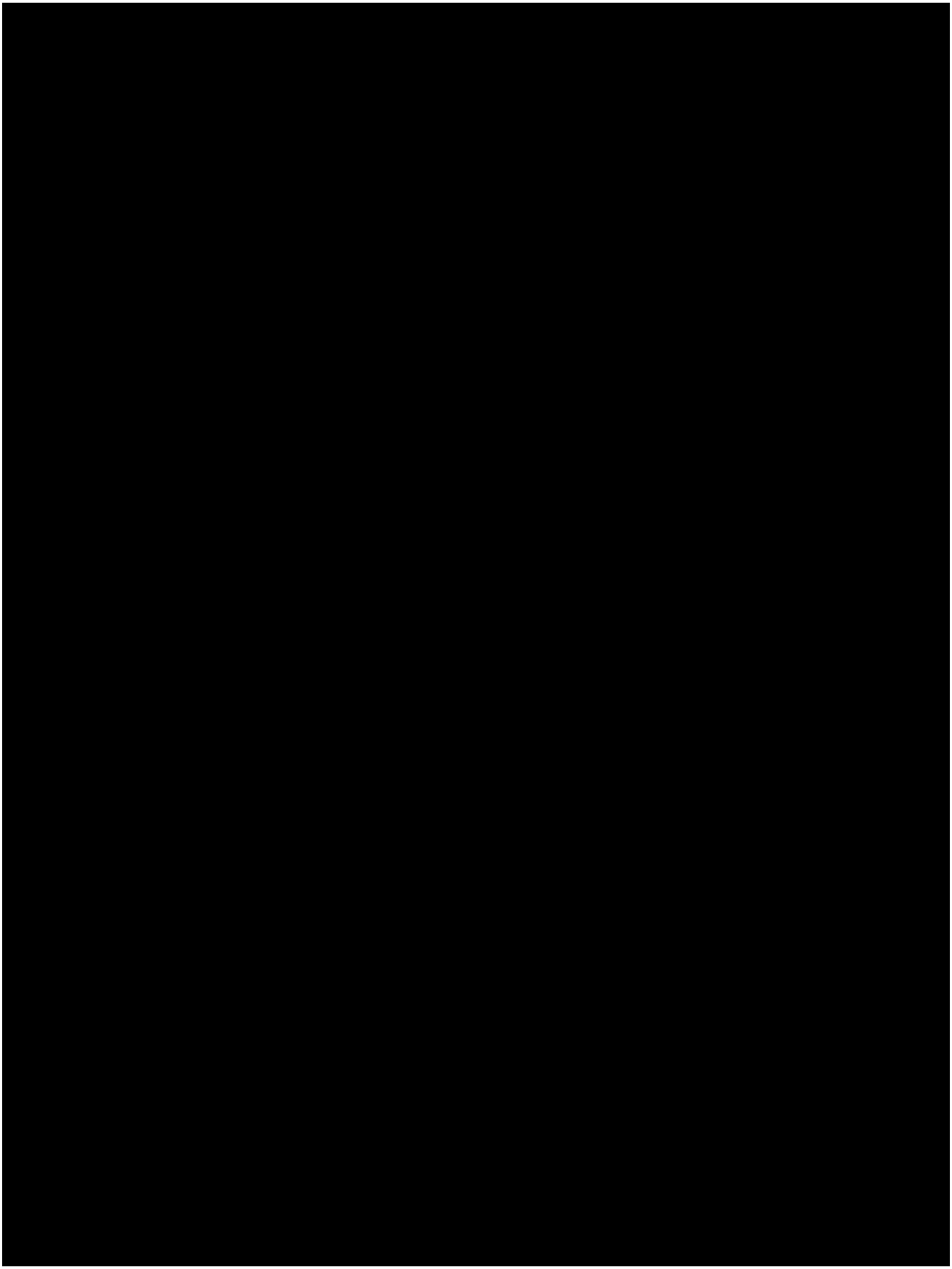
3) How severe has your **pain** (joint or muscular pains in your arms or legs) been during the last 2 weeks?

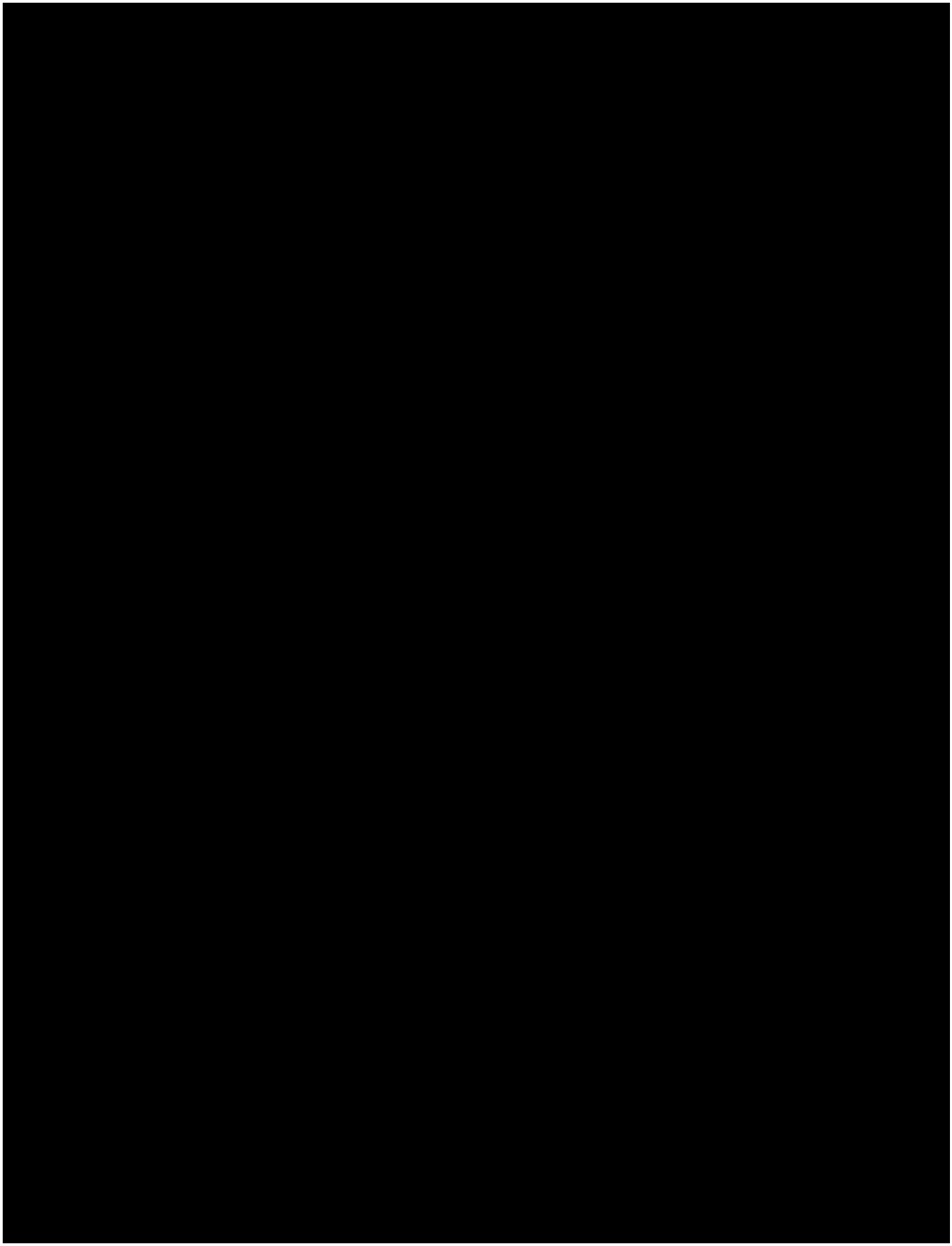
No pain	0	1	2	3	4	5	6	7	8	9	10	Maximal imaginable pain

Fig. 1. The EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI). The ESSPRI score is the mean of the 3 scales.









Appendix XI: SAS Code

Mixed Model for Repeated Measures

The SAS code for MMRM model with unstructured variance-covariance matrix for the primary efficacy endpoint is:

```
PROC SQL data = dataset NOPRINT;
      SELECT mean(AVAL)   INTO : b   FROM ADXX where paramcd = 'XXX' and fasfl = 'Y' and avisit = 'Baseline';
QUIT;
```

For parameters planned to be collected/measured at Week 4, Week 8 and Week 12 the following SAS code will be used:

```
PROC MIXED data = dataset;
     CLASS subjid trt visit ;
     MODEL change =  visit*trt bsl*visit/ NOINT DDFM=KenwardRoger;
     REPEATED visit / SUBJECT=subjid TYPE=un;
    LSMEANS visit*trt / cl alpha=0.1 at (bsl) = (&meanbsl.);
    ESTIMATE 'GLPG3970 vs. Placebo: Week 4'  avisit*trtp 1 0 0 -1  0  0 /CL ALPHA = 0.1;
    ESTIMATE 'GLPG3970 vs. Placebo: Week 8'  avisit*trtp 0 1 0  0 -1  0 /CL ALPHA = 0.1;
    ESTIMATE 'GLPG3970 vs. Placebo: Week 12' avisit*trtp 0 0 1  0 -0 -1 /CL ALPHA = 0.1;
    ESTIMATE 'AUEC GLPG3970'  avisit*trtp 4 4 2 0 0 0  bsl*avisit %sysevalf(4*&meanbsl.)
%sysevalf(4*&meanbsl.)
 %sysevalf(2*&meanbsl.)  / divisor = 12  CL ALPHA = 0.1;
    ESTIMATE 'AUEC Placebo'   avisit*trtp 0 0 0 4 4 2     bsl*avisit %sysevalf(4*&meanbsl.)
%sysevalf(4*&meanbsl.)
  %sysevalf(2*&meanbsl.)  / divisor = 12  CL ALPHA = 0.1;
     ESTIMATE 'AUEC GLPG3970 vs. Placebo'  avisit*trtp 4 4 2 -4 -4 -2 / divisor = 12  CL
ALPHA = 0.1;
RUN;
```

Code for endpoints measured at Baseline, Week 8 and Week 12:

```
     LSMESTIMATE trt*visit

          ' GLPG3970 vs Placebo: Week 8'   1 0      -1 0,

          ' GLPG3970 vs Placebo: Week 12'   0 1    0 -1;

RUN;
```

For the log transformed parameters the following will be applied:

- In the *PROC SQL*: *AVAL* will be replaced by *log(AVAL)*
- In the *PROC MIXED*:
  - *Change* will be replaced by *geopchg=log(aval)-log(bsl)*
  - And *bsl* will be replaced by *log(bsl)*

If the default Newton–Raphson algorithm used by SAS PROC MIXED fails to converge, the following will be tried to avoid lack of convergence while maintaining an unstructured variance:

A. The Fisher scoring algorithm (via the SCORING option of the PROC MIXED statement) will be used to obtain the initial values of covariance parameters:

PROC MIXED data = dataset ***SCORING=5****;*

…
The rest of the code is the same as the first full proc mixed code provided.

B. If the above fails, the no-diagonal factor analytic structure will be used which effectively performs the Cholesky decomposition via the TYPE=FA0(V) option of the REPEATED statement, where V is the total number of visits in the response vector (V = 3):

PROC MIXED data = dataset;

…
*REPEATED visit / subject=subjid type=****FA0(3)****;*
…
The rest of the code is the same as the first full proc mixed code provided.

C. If all the above fails, the variance-and-correlations parameterization will be attempted using TYPE =UNR:

PROC MIXED data = dataset;
…
*REPEATED visit / subject=subjid type=****UNR****;*
…
The rest of the code is the same as the first full proc mixed code provided.

In the rare case where all the above fails, the Toeplitz structure with the sandwich variance estimator (EMPIRICAL option) will be used:

PROC MIXED data = dataset ***EMPIRICAL****;*
…
*REPEATED visit / subject=subjid type=****TOEP****;*
…
The rest of the code is the same as the first full proc mixed code provided.

If this option also fails, the AR(1) + random intercept model will be tried with the sandwich variance estimator:

PROC MIXED data = dataset ***EMPIRICAL****;*

…

*REPEATED visit / subject=subjid type=**AR(1)**;*
RANDOM intercept / subject=subjid;

…

The rest of the code is the same as the first full proc mixed code provided.

Appendix XII: Prednisolone Conversion Table

Glucocorticoid Name	Dose Equivalent to 5 mg Prednisolone
Prednisone oral	5 mg
Betamethasone oral, topical, IM, IV	0.6 mg
Cortisol (hydrocortisone) oral, topical, IM, IV	20 mg
Cortisone oral, IV, intra-articular, transcutaneous	25 mg
Deflazacort oral, IM, IV	7.5 mg
Dexamethasone oral, IM, IV	0.75 mg
Methylprednisolone Oral IM, IV	4 mg
Methylprednisolone Acetate Oral IM, IV	4 mg
Methylprednisolone Sodium Succinate	4 mg
Paramethasone oral, IM, intra-articular	2 mg
Prednisolone oral, topical, eyes drops, IV	5 mg
Triamcinolone oral, IM, inhalation	4 mg
Beclometasone Dipropionate oral, topical, nasal, inhalation	1.25 mg
Budesonide oral, nasal, inhalation, rectal	1.5 mg
Hydrocortisone Sodium Succinate IM, IV	20 mg

(Steroid Conversion Calculator)

Signature Page for glpg3970-cl-207-sap 19634

Approval	[redacted] Senior Medical Director Clinical Development 17-Dec-2021 09:19:08 GMT+0000

Approval	[redacted] Consultant Biostatistics Biometrics 17-Dec-2021 09:48:52 GMT+0000

Signature Page for glpg3970-cl-207-sap43952_33739_62673