STATISTICAL ANALYSIS PLAN

Project Number:	GLPG3970
Study Number:	GLPG3970-CL-210
Study Title:	A randomized, double-blind, placebo-controlled, multicenter study to evaluate the safety, tolerability, efficacy, and pharmacokinetics of GLPG3970, administered orally for 6 weeks in adult subjects with moderately to severely active ulcerative colitis

Development Phase:	2		
Status:	Final		
Version:	3.0	Date:	16-Jun-2021
EudraCT:	2020-000659-11	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 V1.0	13-Nov-2020	
SAP Version V 1.1	27-Jan-2021	Section 6.1.1: Updated derivation of Duration of UC (years)
SAP version 2.0	14-jun-2021	Several updates including analysis and derivation of the primary endpoint.
SAP version 3.0	16-jun-2021	Update typo to sensitivity analysis 1 for primary endpoint and a clarification to the SAS code in the appendix.

LIST OF ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical
AUC	area under the curve
BLOQ	below lower limit of quantification
BMI	body mass index
CI	confidence interval
CL	clearance
CMH	Cochran-Mantel-Haenszel
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_{trough}	trough concentration
CTx	serum C terminal telopeptide type I collagen
CV	coefficient of variation
[redacted]	[redacted]
ECG	electrocardiogram
ED	early discontinuation
ES	endoscopy subscore
FAS	full analysis set
FU	follow-up
H	high, above the upper limit of the normal range
HDL	high-density lipoprotein
High	high specificity test
[redacted]	[redacted]
HR	heart rate
hsCRP	high sensitivity C-reactive protein
[redacted]	[redacted]

ICF	informed consent form
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IP	investigational product
IXRS	interactive voice/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
MCII	minimum clinically important improvement
MedDRA	medical dictionary for regulatory activities
[redacted]	[redacted]
MMRM	mixed models for repeated measures
n	number of valid observations
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
PASS	patient acceptable symptom state
[redacted]	[redacted]
PGA	physician's global assessment score
PK	pharmacokinetic(s)
[redacted]	[redacted]
PR	pulse rate
q.d.	once daily
QTc	corrected QT interval
QTcF	QT interval corrected for the heart rate using Fridericia's formula
RA	rheumatoid arthritis
RB	rectal bleeding
Reg	regular test
[redacted]	[redacted]
RR	respiratory rate

SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SF	stool frequency
SI	standard international
SMC	Safety Monitoring Committee
TEAE	treatment-emergent adverse event
TLF	tables, listings and figures
TV	target value
UC	ulcerative colitis
[redacted]	[redacted]
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 analysis of study GLPG3970-CL-210. 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.

2. STUDY DESIGN AND OBJECTIVES

2.1. Study Objectives

2.1.1. Primary Objective

- To evaluate the effect of GLPG3970 compared to placebo on the signs and symptoms of ulcerative colitis (UC) in subjects with moderately to severely active UC.

2.1.2. Secondary Objectives

- To evaluate the safety and tolerability of GLPG3970 compared to placebo in subjects with moderately to severely active UC.
- To characterize the pharmacokinetic (PK) of GLPG3970 in subjects with moderately to severely active UC.

2.1.3. Other Objectives

2.2. Study Endpoints

2.2.1. Primary Endpoint

- Change from baseline in total Mayo Clinical Score (MCS) at Week 6.

2.2.2. Secondary Endpoints

- Number, incidence, and severity of treatment-emergent adverse events (TEAEs).
- Observed GLPG3970 plasma trough concentrations (C_{trough}).

2.2.3. Other Endpoints

2.3. Study Design

This is a Phase 2, randomized, double-blind, placebo-controlled, multi-center parallel-group study to evaluating GLPG3970 in adult subjects with moderately to severely active UC. In this study, subjects will receive 1 dose level of GLPG3970 (400 mg once daily [q.d.]).

The study is planned to randomize 30 subjects in a 2:1 ratio (20 subjects planned to receive GLPG3970 and 10 subjects planned to receive placebo).

The study will consist of 3 study periods:

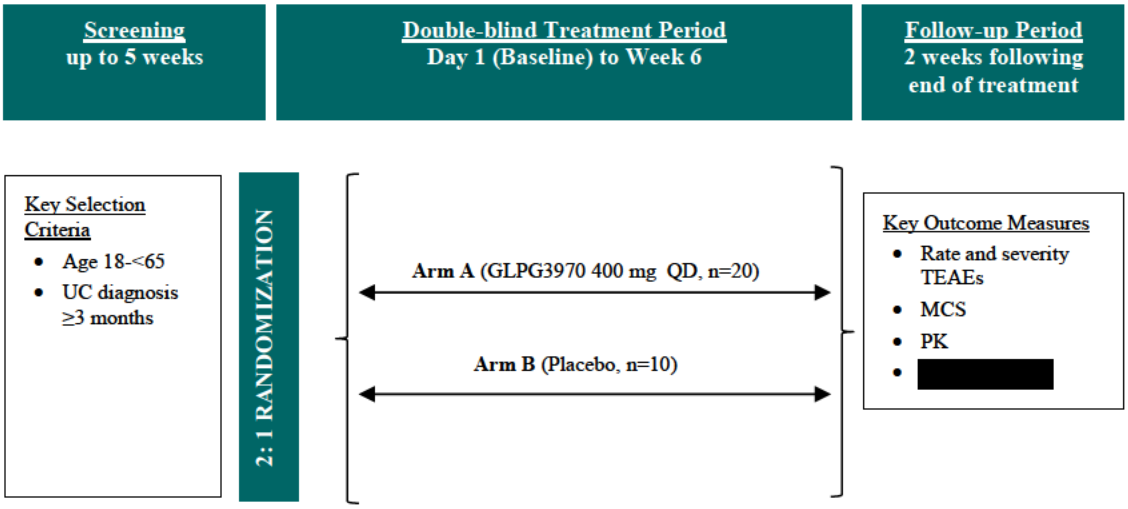
- Screening period: up to 5 weeks with 2 study visits, with a maximum of 2 weeks from screening visit 2 until randomization.
- Double-blind treatment period: 6 weeks with 4 study site visits (Days 1, 15, 29, and 43 or early discontinuation [ED]) and 1 telephone call (Day 8).
- Follow-up period: 2 weeks with 1 study visit.

The subjects will be in the study for a duration of a maximum of 13 weeks [screening to follow-up (FU)] and will receive oral GLPG3970 or placebo, as a reconstituted oral solution, once daily (q.d.) for 6 weeks.

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

A schematic diagram of clinical study design, procedures and stages is provided in Figure 1.

Figure 1 Schematic Diagram



2.4. Clinical Study Protocol (CSP) and CSP Amendments

This SAP is based on the protocol version 1.0, dated 02-Jul-2020.

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, 2						FOLLOW-UP PERIOD3
Study Visit	**S1**	**S2[4]**	**1**	**2**	**3**	**4**	**5**	**ED[3]**	**Follow-up Visit**
Study Day (D) ± Days	**-35 to 0**		**1**	**8 ± 2**	**15 ± 2**	**29 ± 2**	**43 ± 2**		**14± 3 after last IP dosing**
On-site visit	✓	✓	✓		✓	✓	✓	✓	✓
Telephone call				✓					
Informed consent (Section 8.3 of CSP)	✓								
Inclusion/exclusion criteria (Sections 6.1 and 6.2 of CSP)	✓	✓	✓						
Demographics (Section 8.3 of CSP)	✓								
Medical and smoking history including UC history (Section 8.3 of CSP)	✓								
Physical examination (Section 8.5.3 of CSP)	✓		✓				✓	✓	✓
Body weight and height[5] (Section 8.5.3 of CSP)	✓		✓		✓	✓	✓	✓	✓
Vital signs (Section 8.5.4 of CSP)	✓		✓		✓	✓	✓	✓	✓
12-Lead triplicate ECG (Section 8.5.5 of CSP)	✓		✓		✓	✓	✓	✓	✓

[1] On dosing days, all assessments are to be performed within 3 hours predose, unless otherwise specified.
[2] Subjects must come for the study visits at baseline (Visit 1), Day 15 (Visit 3), and Day 43 (Visit 5) and the early discontinuation (ED) visit, if applicable in the morning in a fasting state (no food intake for at least 8 hours).
[3] Subjects who discontinue treatment early will be requested to return for an ED visit to complete all Visit 5 assessments and to return for a follow-up (FU) visit 14±3 days after last IP administration.
[4] The endoscopy procedure will be done at Screening Visit 2 once the other in- and exclusion criteria have been confirmed at Screening Visit 1.
[5] Height only to be measured at Screening Visit 1.

EVENT	SCREENING PERIOD		TREATMENT PERIOD 1, 2						FOLLOW-UP PERIOD3
Study Visit	**S1**	**S2[4]**	**1**	**2**	**3**	**4**	**5**	**ED[3]**	**Follow-up Visit**
Study Day (D) ± Days	**-35 to 0**		**1**	**8 ± 2**	**15 ± 2**	**29 ± 2**	**43 ± 2**		**14± 3 after last IP dosing**
QuantiFERON TB Gold test (Section 8.5.2 of CSP)	✓								
SARS-CoV-2 RT-PCR test[6]	✓		✓	As needed					
SARS-CoV-2 serology test			✓	As needed					
Randomization (Section 6.5.1 of CSP)			✓						
Blood collection									
• Safety (hematology, coagulation, chemistry) (Section8.5.2 of CSP)	✓		✓[7]		✓[7]	✓	✓[7]	✓[7]	✓
• Serology (HBV, HCV, HIV) (Section 8.5.2 of CSP)	✓								
• Pregnancy test serum (all WOCBP females) (Section 8.5.2 of CSP)	✓		✓			✓	✓	✓	
• FSH test (WOnonCBP, non-surgical postmenopausal women) (Section 8.5.2 of CSP)	✓								
• PK (Section 8.6 of CSP)			✓[8]		✓[9]	✓[9]	✓[10]	✓	

[6] RT-PCR from a nasal swab sample at screening may be done at an earlier screening visit on a different day than the other screening assessments (after signature of the study informed consent form) and RT-PCR from a nasal swab sample and serology testing at a visit on Day -3, and throughout the study as needed when subject presents signs and symptoms of SARS-CoV-2 infection.
[7] Fasted glucose, fasted insulin, and HOMA-IR only at baseline (Visit 1), Day 15 (Visit 3), and Day 43 (Visit 5) and the ED visit, if applicable.
[8] Visit 1: predose (within 30 minutes prior to dosing), 1 sample within [0.5-1.5 hours postdose]; 1 sample within [2–2.5 hours postdose] and 1 sample within [3-4 hours postdose].
[9] Visit 3 and 4: predose (within 30 minutes prior to dosing).
[10] Visit 5: 1 sample predose (within 30 minutes prior to dosing) and 1 sample within [4-6 hours postdose].

EVENT	SCREENING PERIOD		TREATMENT PERIOD 1, 2						FOLLOW-UP PERIOD3
Study Visit	**S1**	**S2[4]**	**1**	**2**	**3**	**4**	**5**	**ED[3]**	**Follow-up Visit**
Study Day (D) ± Days	**-35 to 0**		**1**	**8 ± 2**	**15 ± 2**	**29 ± 2**	**43 ± 2**		**14± 3 after last IP dosing**
[redacted]									
Urine collection									
• Safety[12] (Section 8.5.2 of CSP)	✓		✓		✓	✓	✓	✓	✓
Stool sample collection[13]	✓		✓		✓	✓	✓	✓	
• Microbiology	✓								
• Microbiome (sampling before IP administration) (Section 8.7.3.2 of CSP)			✓				✓	✓	
[redacted]									

[11] [redacted]

[12] On dosing days, safety urine samples will be taken predose, on the same visit day. Visit 1 and 5: predose and 2 hours postdose. Calcium and phosphate should only be evaluated at Visits 1 and 5.

[13] At Screening Visit 1, a stool sample should be collected during the visit. If not possible, the subject can collect the sample at home within 24 hours after start of Screening. For Visits 1, 3, and 4, stool samples can be collected within 24 hours prior to the visit or during the visit. At Visit 5 (Day 43) and ED visit, if applicable, when the stool sample must be taken before the bowel preparation for endoscopy, stool can be collected more than 24 hours before the visit. One sample can be used for different assessments.

EVENT	SCREENING PERIOD		TREATMENT PERIOD 1, 2						FOLLOW-UP PERIOD3
Study Visit	**S1**	**S2[4]**	**1**	**2**	**3**	**4**	**5**	**ED[3]**	**Follow-up Visit**
Study Day (D) ± Days	**-35 to 0**		**1**	**8 ± 2**	**15 ± 2**	**29 ± 2**	**43 ± 2**		**14± 3 after last IP dosing**
Endoscopy [redacted]									
Mayo Clinical Score subscores (MCS-SF, MCS-RB, PGA) (Section 8.4.3 of CSP)		✓	✓		✓	✓	✓	✓	
[redacted]									
Subject diary evaluation (Section 8.8.1 of CSP)		✓	✓	✓	✓	✓	✓	✓	✓
Allocation and dispensing of IP supplies (Section 7.2 of CSP)			✓		✓	✓			
IP administration (Section 7.2 of CSP)			Once daily throughout the treatment period[16]						
Prior and concomitant medication (Section 6.3.2 of CSP)	Throughout the study								
Adverse events (Section 8.5.1 of CSP)	Throughout the study								

[14] Central endoscopy reading, [redacted].
[15] Screening endoscopic assessment (either surveillance colonoscopy or flexible sigmoidoscopy) should be completed prior to Day -8.
[16] On Visit 1, 3, 4, and 5, administration of IP must occur at the site.

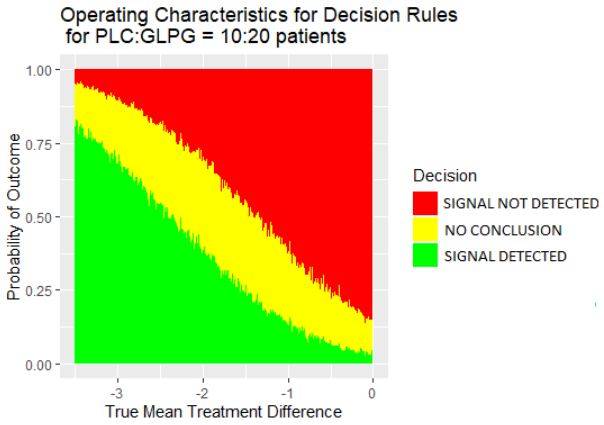
2.6. Sample Size Justification

A sufficient number of subjects with UC will be screened in order to have 30 subjects randomized (20 planned subjects on GLPG3970 and 10 planned subjects on placebo).

The 30 subjects randomized in the study, are expected to be sufficient to provide preliminary insight in the safety, tolerability and PK of GLPG3970 in UC patients. 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 ultimately be decided by the sponsor before any study lock or related unblinding has occurred.

For the efficacy objective, the operating characteristics of the chosen decision framework for the most important endpoint (MCS change from baseline at Week 6) at a sample size of 30 subjects (10 on placebo and 20 on GLPG3970) are represented in the following graph and table, and were deemed acceptable. A common standard deviation of 3.3 was assumed in this small study. Lower Reference Value and Target Value for the decision framework were determined based on the published studies and the therapies in development or on the market. They were identified by the project team as the lowest delta of possible interest and the best case scenario for efficacy outcome respectively.

True Treatment Difference In MCS Change From Baseline At Week 6	**Signal Score**	**Probability Of Outcome**
-3.0	Signal detected	0.673
	No conclusion	0.220
	No signal detected	0.107
-1.3	Signal detected	0.205
	No conclusion	0.289
	No signal detected	0.506



PLC: placebo; GLPG: GLPG3970

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 at visit 1. Allocation of each subject to a given treatment will be done using a centralized electronic system (interactive voice/web response system [IXRS]). Investigational product (IP) 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 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 by the investigator for the safety of a subject. The investigator is encouraged to discuss considerations to break the blind with the medical leader, 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 unblinded has occurred.

The blind can be broken by the investigator via IXRS by the investigator.

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 study site. The reason for breaking the blind will be indicated and justified in the source documentation.

If an AE leads to unblinding, the AE will be given as the reason for unblinding. All subjects who are unblinded should, where possible, complete the follow-up visit assessments 14±3 days after unblinding. Any AEs will be followed until resolution.

The code-break information (via IXRS vendor) will be provided to the bioanalytical laboratory responsible for PK sample analysis, the person responsible for providing unblinded data to the internal safety monitoring committee (SMC), to the pharmacovigilance vendor for serious adverse event (SAE) reporting purposes, and, if applicable, to the contract research organization (CRO) performing population PK/PD analyses.

3. STUDY ESTIMANDS

Estimand for the primary study objective:

Attribute	Details
Treatments	GLPG3970 (400 mg q.d.) vs. Placebo
Population	• Population target defined through the inclusion/exclusion criteria (see CSP). • Analysis Set: Full Analysis Set (FAS) defined as all randomized subjects who have received at least 1 dose of IP.
Endpoint	Change from baseline in total MCS at Week 6.
Population-level summary	Difference (GLPG3970 minus placebo) in least square means (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 secondary study objective:

Attribute	Details
Treatments	GLPG3970 (400 mg q.d.) vs. Placebo
Population	• Population target defined through the inclusion/exclusion criteria (see CSP). • Analysis Set: Safety Analysis Set, defined as all subjects who were administered study drug at least once.
Endpoint	Presence of TEAE
Population-level summary	Percentage of subjects with TEAEs.
Intercurrent events and strategies to handle these	• Early treatment discontinuation (for any reason): handled using the while-on-treatment strategy. • Other intercurrent events (e.g. major protocol deviations, intake of prohibited/rescue medication, lock of compliance): handled using the treatment policy strategy (i.e. adverse events are counted regardless the occurrence of these other events).

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 subjects who used at least 1 dose of IP.

4.1.4. Full Analysis Set

All randomized subjects who have received at least 1 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 pharmacometric 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.

[redacted]

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. as-randomized analysis).

The actual treatment groups will be used for the analysis of safety, PK and PD parameters. The "actual" rather than "randomized" treatment will be used only in case the "wrong" drug 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 to 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, adverse events (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 study drug administration - 1 minute
Treatment	Date (time) of first study drug 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 study drug 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.

Table 3: Analysis visit window for non-PK data

Visit date and time	Relative day (ADY)	Assessment visit (VISIT)	Time window (in ADY)*	Target day	Analysis visit (AVISIT)
On or before first study drug administration	≤1	Last assessment at any point in time before first study drug administration	≤1	1	Baseline
After first study drug administration	>1	Any scheduled (non follow-up), unscheduled, or early discontinuation visit	2 to 12	8	Week 1
			13 to 22	15	Week 2
			23 to 36	29	Week 4
			37 to 49*	43	Week 6
		Follow-up	CRF collected follow-up visit		Follow-up

*For Week 1, Week 2, Week 3, Week 4, Week 5 and Week 6, time windows will be applied excluding assessments performed more than 2 days after the last dose of IP. Except for Endoscopy and [redacted] where assessments performed more than 7 days after last dose will be excluded and the Week 6 analysis window will be 37 to 52 days.

Table 4: Analysis visit window for the daily diary data

Visit date and time	**Relative day (ADY)**	**Assessment visit (VISIT)**	**Time window (in ADY or relative to last dose of IP for the follow-up)***	**Analysis visit (AVISIT)**
Before first study drug administration	<1	All days before the first IP administration.	<1	Baseline
Same as or after first study drug administration	≥ 1	Any scheduled (non follow-up), unscheduled, or early discontinuation visit	1 to 7	Week 1
			8 to 14	Week 2
			15 to 21	Week 3
			22 to 28	Week 4
			29 to 35	Week 5
			36 to 42	Week 6
		Follow-up	1 to 7, where day 1 = first day after the last dose of IP	Week 7
			8 to 14, where day 1 = first day after the last dose of IP	Week 8

*For Week 1, Week 2, Week 3, Week 4, Week 5 and Week 6, time windows will be applied excluding assessments performed more than 2 days after the last dose of IP.

Table 5: Analysis Visit Window for PK data

Visit date and time	Relative day (ADY)	Assessment visit (VISIT)	Time window (in ADY)*	Target day	Analysis visit (AVISIT)	Analysis time point (ATPT)
Before first study drug administration	<1 or pre-dose at ADY=1	Last assessment at any point in time before first study drug administration	<1 or pre-dose at ADY=1	1	Baseline	Pre-dose (within 30 minutes prior to dosing)
Same as or after first study drug administration	≥ 1	Any scheduled (non follow-up), unscheduled, or early discontinuation visit	1	1	Day 1	As indicated in the database: 0.5-1.5, 2-2.5, 3-4 hours post-dose
			13 to 22	15	Week 2	As indicated in the database: Pre-dose (within 30 minutes prior to dosing)
			23 to 36	29	Week 4	As indicated in the database: Pre-dose (within 30 minutes prior to dosing)
			37 to 49	43	Week 6	As indicated in the database:

Visit date and time	Relative day (ADY)	Assessment visit (VISIT)	Time window (in ADY)*	Target day	Analysis visit (AVISIT)	Analysis time point (ATPT)
						Pre-dose (within 30 minutes prior to dosing) and within [4-6 hours post dose].
		Follow-up	CRF collected follow-up visit		Follow-up	

*For Week 1, Week 2, Week 3, Week 4, Week 5 and Week 6, time windows have to be applied excluding assessments performed more than 2 days after the last dose of IP.

Table 6: Analysis Visit Window for urine phosphate and urine calcium

Visit date and time	Relative day (ADY)	Assessment visit (VISIT)	Time window (in ADY)*	Target day	Analysis visit (AVISIT)	Analysis time point (ATPT)
Before first study drug administration	<1 or pre-dose at ADY=1	Last assessment at any point in time before first study drug administration	<1 or pre-dose at ADY=1	1	Baseline	Pre-dose
Same as or after first study drug administration	≥ 1	Any scheduled (non follow-up), unscheduled, or early discontinuation visit	1	1	Day 1	2 hours post-dose
			13 to 22	15	Week 2	
			23 to 36	29	Week 4	
			37 to 49	43	Week 6	As indicated in the database: Pre-dose and 2 hours post-dose
		Follow-up	CRF collected follow-up visit		Follow-up	

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 "PRE-DOSE" (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. If multiple values qualify as last available assessment, the mean of these values will be used in the analysis.

For electrocardiogram (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, pre-dose:

- 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.

For diary-based efficacy parameters, baseline is defined as the average of the values before the first IP administration. The baseline value will be derived only when there are at least 3 days with data. Otherwise, the baseline value for that parameter will be treated as missing. The definition will exclude the day prior/of/after the colonoscopy.

4.3.5. Selection and Use of Analysis Visits

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 nominal 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

For the [redacted] endpoints including MCS, [redacted], the following steps will be used to derive each composite endpoint, unless otherwise specified:

- Step 1: Put [redacted] component value into an analysis visit window.
- Step 2: Within each visit window, select the data points for analysis.
- Step 3: Combine these selected data points within each analysis visit window across the components.

4.3.6. Handling of Missing Data

4.3.6.1. 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. AEs, concomitant medications) with missing date (time) will be allocated to analysis periods using a worst-case approach as explained in the respective sections.

4.3.6.2. 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 MMRM handle missing data via the maximum-likelihood function.

For MCS[redacted] a multiple imputation procedure will be used as described in section 6.2.2.2.[redacted]

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

4.3.6.3. Censoring of Time to Event Data

Not applicable.

4.3.7. 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 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]

4.3.8. Handling of Outliers

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

4.3.9. Stratification Factors

Not applicable.

4.4. Presentation of Results

4.4.1. Presentation of Treatment Groups

Results will be presented by treatment group:

- GLPG3970 400 mg q.d.
- Placebo

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

4.4.2. 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).
- [redacted]

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 coefficient of variation (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.

4.4.3. Calculation of Percentages

Frequencies and percentages will be generated for categorical parameters.

For event-type data (e.g. AEs), 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.

An internal SMC independent from the study will review unblinded safety data during the course of the study. This Committee may involve external medical experts (such as an expert in the field of UC and an infectious diseases expert) 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 clinical study team will remain blinded to the allocation of subjects to treatments during the course of the study. The process is described in a separate SMC 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 < 85; age ≥85;
- race and ethnicity;
- height at baseline (cm);
- weight at baseline (kg);
- body mass index (BMI) at baseline (kg/m^2) $= \frac{weight\ (kg)}{height^2\ (m^2)}$
  (BMI will not be recalculated if already available in the database);
- smoking status (Non-smoker/Ex- smoker/Current smoker);
- number pack-years;
- duration of UC (years)
  $= \frac{(date\ of\ first\ study\ drug\ administration) - (date\ of\ \ initial\ diagnosis) + 1}{365.25}$

If missing day: use 1st day of the month
If missing both day and month: use 1st of January;

- duration of UC, categorized: <6 months; 6 months ≤ duration<1 year; 1 ≤ duration<3 years; 3 ≤ duration<5 years; 5≤ duration<10 years; duration≥10 years;
- MCS at baseline;
- [redacted]
- MCS-SF at baseline
- MCS-RB at baseline;
- MCS-ES at baseline;
- PGA 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 study drug kit numbers, study drug batch 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 study drug administration schedule and the reasons for discontinuation.
– The number (percent) of subjects who completed/discontinued the study and the reasons for discontinuation.

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, including reasons.

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.

COVID-19 related protocol deviations will be listed together with the rest of major protocol violations. The fact that they are COVID-19 related will be present in that listing.

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. Ulcerative Colitis History

Frequency tabulations, per treatment group and overall, per system organ class and preferred term will be provided for the UC disease history.

In addition, frequency tabulations, per treatment group and overall, will be provided for the following parameters.

- extent of disease at time of most extensive disease activity per Montreal Classification;
- previous treatment failure and corresponding treatments;
- colonoscopy history (yes/no);
- stool history (number of stools per day under normal (healthy/remission) conditions)

All UC history will be listed.

6.1.7. Prior and Concomitant Medications

6.1.7.1. Coding of Reported Terms

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

6.1.7.2. Classification of Medications

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 study drug administration date.
– Concomitant only: when the record started on or after the first study drug administration date.
– Prior and concomitant: when the record started before the date of first study drug 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 study drug 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 study drug administration.

6.1.7.3. Calculation of Relative Days

For both the start and the end dates of the concomitant medication records, their day relative to the day of first study drug administration will be calculated as described in Section 4.3.1.

6.1.7.4. 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 listings 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).

6.1.8. Procedures

Procedures results will only be listed.

6.1.9. Exposure to Study Drug and Compliance

6.1.9.1. Derivation Rules

Derived Parameters: Extent of Exposure to Study Drug

- *Total treatment duration* (days) = last study drug administration date – first study drug administration date + 1 day.
- *Total treatment duration, excluding days off study drug*: Sum of the number of days with any study drug administration.
- *Total treatment duration, full compliance* (days): Number of days where the complete volume of IP has been administered.

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 (%) $= 100 \times \frac{total\ treatment\ duration,fully\ compliant}{total\ treatment\ duration}$

6.1.9.2. Presentation of Results

Summary statistics per planned treatment group and overall will be provided for each compliance and extent of exposure parameter for study drug 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 study drug administration records will be listed per subject. The listing will include all deviations from study drug schedules such as missed or reduced doses and GLPG3970/placebo switches and will also include food intake status at the time of study drug administration.

Study drug 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 endpoint for this study is the change from baseline in total MCS at Week 6. The total MCS score will be calculated only at baseline and week 6.

The MCS ranges from 0 to 12. A lower score means a better subject condition. It consists of 4 subscores. Each sub-score ranges from 0 to 3. The four components of the MCS are (see Appendix III):

- MCS-SF subscore
- MCS-RB subscore
- MCS-ES subscore
- PGA

For calculation of the MCS subscores performed by the investigator, MCS-RB and MCS-SF, subjects will be instructed to record symptoms of RB and SF daily in a diary.

The information on the UC symptomatology of the last 3 consecutive available days in the diary collected from the last week, will be used to score the MCS-RB and MCS-SF subscores, excluding the day before the endoscopy (with bowel preparation), the day of the endoscopy and the day after the endoscopy.

The PGA will be scored by the investigator or trained sub-investigator at visits specified in the Schedule of Activities.

6.2.2.1. Derivation Rules

The MCS will be calculated based on the following formula:

MCS = MCS-SF + MCS-RB + MCS-ES + PGA

If one of the component is missing at a specific analysis window, total score will not be calculated for that analysis visit and will be set to missing.

6.2.2.2. Analyses Methods

The MCS data will be analyzed descriptively as actual (observed) values and changes from baseline.

An ANCOVA will be used on the MCS change from baseline to compare treatment groups, with treatment as a fixed effect and the baseline MCS as a continuous covariate. Missing data will be imputed using Rubin's multiple imputation (Rubin, 1987). A multivariate normal model will be used as the imputation model. The Markov chain Monte Carlo (MCMC) method will be used to sample from the posterior distribution of the model parameters. The following variables will be used to impute missing data at week 6: MCS baseline value and [redacted] at baseline, week 2, 4 and week 6. The imputation model will be separately fitted for each treatment arm. In case the covariance matrix is singular, the imputation method will be changed as follows. The following variables will be used to impute missing data at week 6: MCS baseline value and [redacted] at baseline, week 2, 4 and week 6 and treatment arm. The imputation model will no longer be fitted separate for each treatment arm. If the imputation is still problematic after that, then treatment will not be used as a variable for imputation and imputation will not be done separately per treatment arm.

For each imputed dataset, the above ANCOVA model will be used to derive the least squares estimate and standard error of the treatment difference in mean change, and the LS means and standard error for the mean change within each treatment group. Then, these results will be combined using Rubin's rules (Rubin, 1987) to obtain the global least squares estimate and standard error of the treatment difference in mean change, its associated two-sided 90% confidence intervals and p-values for treatment group comparison. The global LS means and standard error and their associate two-sided 90% CIs for the mean change within each treatment group will be also obtained. Further details are provided in Appendix VIII with the SAS code.

The number of patients in the analysis population and number of patients in the analysis will be provided by treatment group.

Listings will be provided for MCS.

Sensitivity Analysis 1

As a sensitivity, the above analysis will be performed only allowing data points taken no more than 2 days after day of last dose.

Sensitivity Analysis 2

As sensitivity analysis the observed case analysis will be performed by using the same ANCOVA on the MCS change from baseline, with treatment as a fixed effect and the baseline MCS as a continuous covariate. The model will be used to derive least squares estimates of the treatment differences in mean change and two-sided 90% confidence intervals. Also, t-statistics corresponding to the type III sums of squares for the differences in the least squares means will be used to obtain p-values for treatment group comparisons. The LS means and associated two-sided 90% CIs for the mean change within each treatment group will be calculated. The p-values

for the effects will also be presented. The number of patients in the analysis population and number of patients in the analysis will be provided by treatment group.

6.2.2.3. Signal Detection Methodology

The signal detection methodology described by Frewer et al. (Frewer, Mitchell, Watkins, & Matcham, Decision-making in early clinical drug development., 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 $\leq$ -3.0) to as low as P (delta $\leq$ -1.3).

The chosen framework parameters are as follows:

- Target value (TV) = -3.0 and lower reference value (LRV) = -1.3
- False stop risk = 10% and false go risk = 20%
- Framework rules:
  - Signal detected if P10 < TV and P80 <LRV
  - No signal detected if P10 $\geq$ TV
  - No conclusion if P10 < TV

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.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

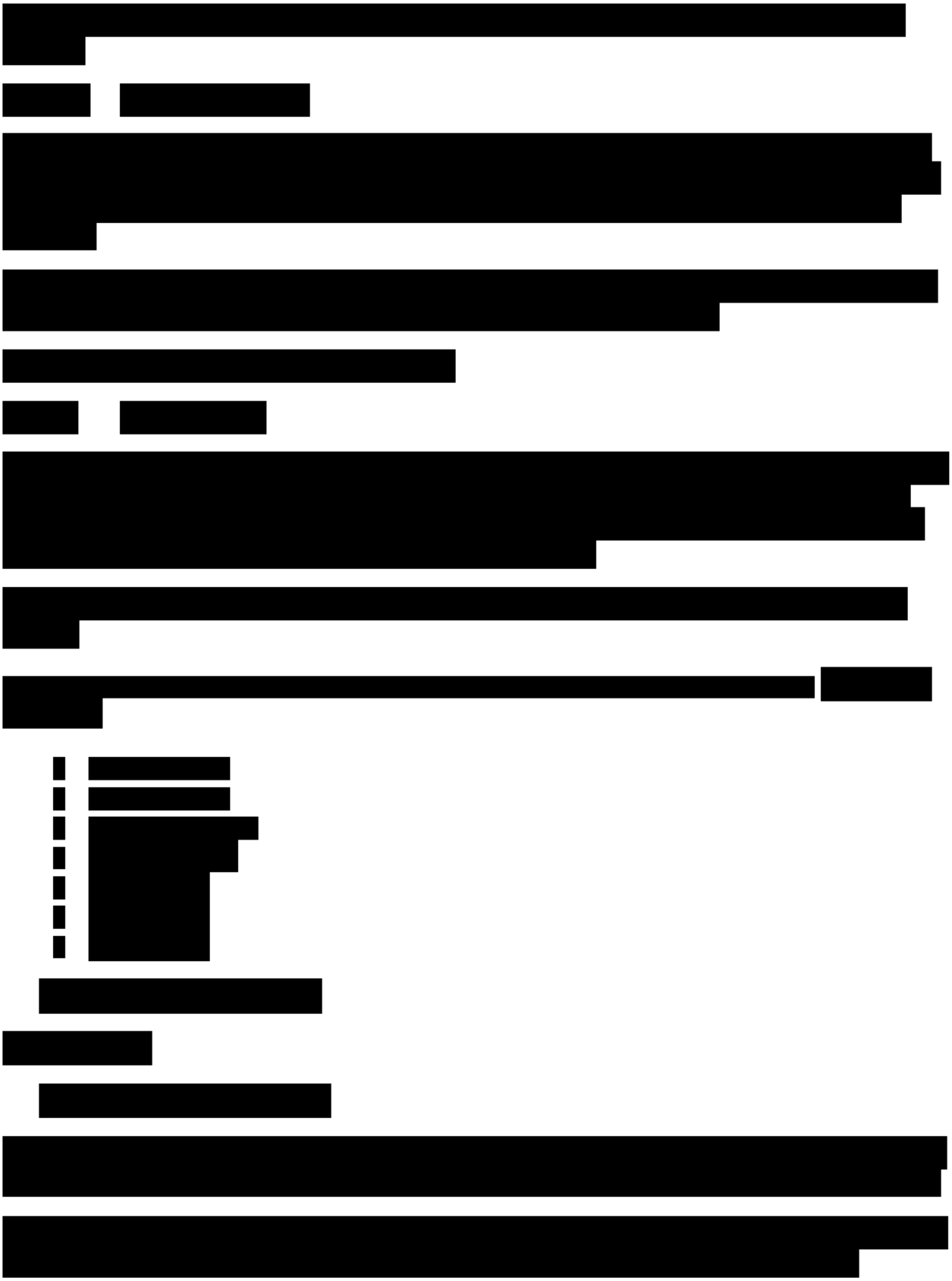
[redacted]

[redacted]

[redacted]

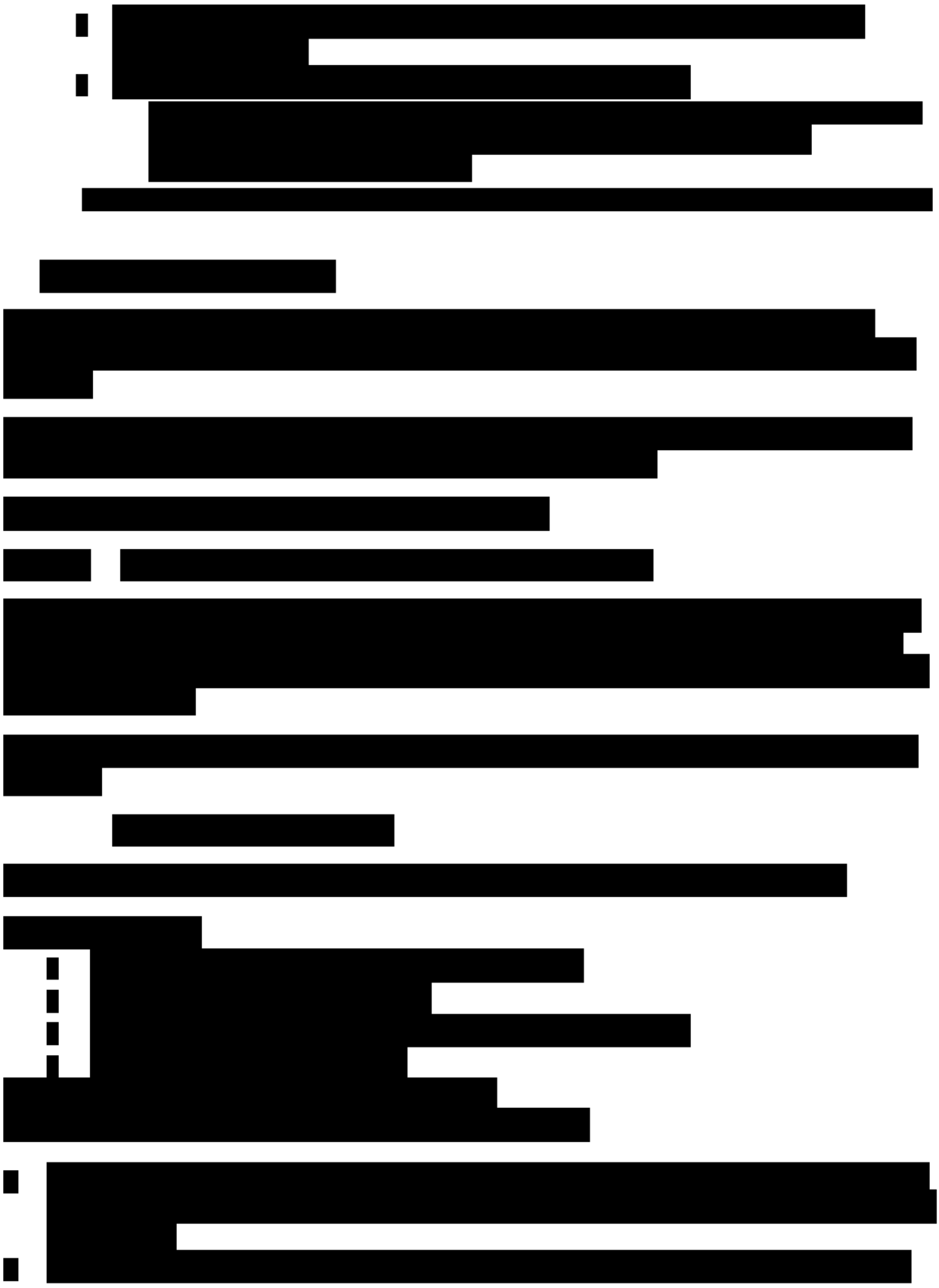
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

[redacted]



[redacted]

[redacted]



[redacted]

[redacted]

6.3. Safety Analyses

Safety analyses will be performed on the Safety Analysis Set.

Safety parameters will be analyzed descriptively (see Section 4.4.2 and Section 4.4.3) by treatment group and overall. No formal testing will be performed to compare the treatment groups.

6.3.1. Adverse Events

All 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 14 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 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.

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 study drug of an AE will be dichotomized as follows:

- *Not study drug related*: all non-missing weaker levels of relatedness than 'possibly drug related'.
- *Study drug 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 study drugs when relatedness has been collected separately per study drug; 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 study drug 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 study drug 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 study drug discontinuation of IP.

All AE data will also be listed, including separate lists for SAEs, AEs leading to death, and AEs leading to study drug 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 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 2 subjects in any treatment group.

6.3.2. Laboratory Safety

6.3.2.1. Available Data

Laboratory tests scheduled are described in the CSP section 8.5.2.

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.

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.

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 position indicator provided in the database as provided by the lab, 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 for these tests.

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

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

Continuous laboratory tests[17] 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, changes from baseline and percent change from baseline will be tabulated separately. Continuous values for metabolic markers will be summarized in the PD analysis section (section 6.5).

Profile line plots by subject of all actual observed values using relative day (ADY) will be presented by treatment group. Graphs of the mean (+- SE) actual values over time and of 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. Also, line plots for subjects will be provided by treatment group.

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

[17] With the exception of the metabolic markers [fasted insulin, homeostatic model assessment of insulin resistance (HOMA-IR) and HbA1c] and other markers hsCRP and fecal calprotectin .

grade. The table will be created per laboratory test, treatment group and analysis period. 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 the 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 $RR(ms) = 60 \times HR(bpm)$.

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, change from baseline and percent changes from baseline will be tabulated separately.

Profile line plots by subject of all actual observed values using relative day (ADY) will be presented by treatment group. Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline and percent change from baseline will be presented. A line plot by subject of the change from baseline will also be provided by treatment group.

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, 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), BMI (kg/m^2), diastolic and systolic blood pressure (mmHg), pulse rate (beats/min) and body temperature (C).

6.3.4.2. Derivation Rules

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, using all non-missing post-baseline records (including unscheduled and FU visits).

6.3.4.6. 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. 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, change from baseline and percent change from baseline will be tabulated separately.

Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline, and percent changes 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.

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. Thus, any subgroup of this would have small sample size to be relevant or even uninterpretable.

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. Pre-dose samples not collected within 22 to 26 hours after the last 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. The same will be done with post-dose samples taken outside of the defined time intervals.

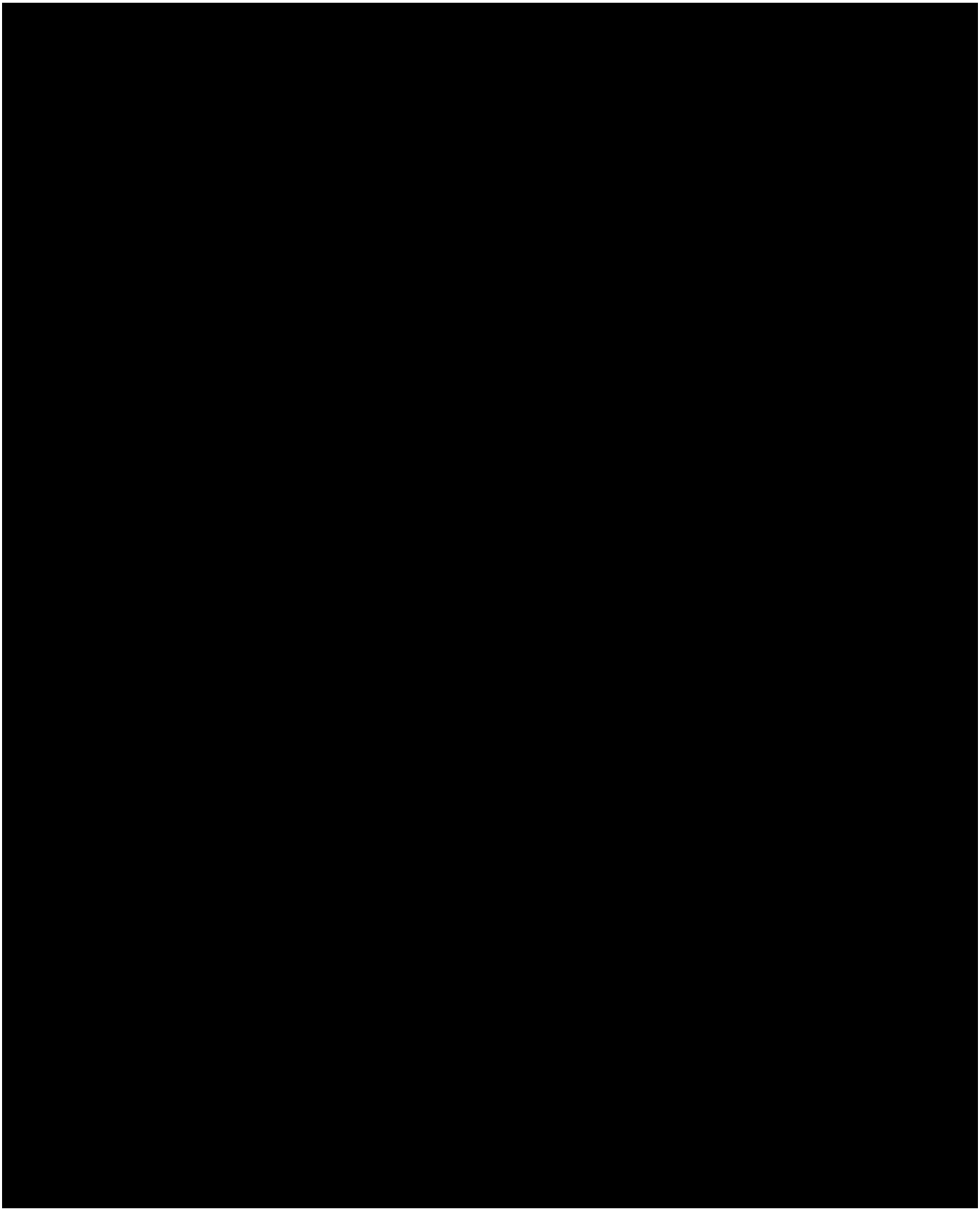
6.4.2. Presentation of Results

Descriptive statistics will be provided for GLPG3970, MTX and MTX-7-OH plasma levels at each assessment time point. Observed plasma C_{trough} for GLPG3970 will be reported in the CSR. Individual GLPG3970 plasma concentrations with descriptive statistics will be presented by time point/interval. Actual blood sampling times from the last drug intake will be listed.

For calculation of descriptive statistics of GLPG3970 C_{trough} (= pre-dose concentration), the following will apply:

- for pre-dose samples not collected within 21-27 hours after the last intake prior to the visit, or collected after the intake at the visit, the corresponding concentration will be excluded from descriptive statistics and flagged in the listing with appropriate footnoting.

- pre-dose samples collected within 3 days after a missed dose will also be excluded from descriptive statistics.



[redacted]

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

- As per protocol, MMRM should have been used for primary analysis [redacted]. However, as the data is collected only at two time points (baseline and week 6) for the primary endpoint, the MCS, an ANCOVA with a multiple imputation method to handle missing values will be used for treatment comparison as main analysis and an analysis on observed case will be performed as sensitivity analysis. [redacted]

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.

FDA CDER. (2016). Ulcerative Colitis: clinical trial endpoints. Guidance for Industry (Draft). Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Retrieved from Available at: https://www.fda.gov/regulatory-information/search-fdaguidance-documents/ulcerative-colitis-clinical-trial-endpoints-guidance-industry.

[redacted]

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

[redacted]

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).

[redacted]

Lu K, Mehrotra D. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. Statistics in Medicine. 2010; 29: 474-488. DOI: 10.1002/sim.3820

Lewis, J., Chuai, S., Nessel, L., Lichtenstein, G., Aberra F, & et al. (2008). Use of the non-invasive components of the Mayo score to assess clinical response in Ulcerative Colitis. *INflamm Bowel Dis*, 14(12):1660-6.

Mallinckrodt CH, Clark WS, Carroll RJ, Molenberghs G. Assessing response profiles from incomplete longitudinal clinical trial data under regulatory considerations. Journal of Biopharmaceutical Statistics 2003; 13: 179-190.1

[redacted]

Schroeder, K., Tremaine, W., & Ilstrup, D. (1987). Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. N Engl J Med, 317(26), 1625-1629.

APPENDICES

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 1 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 2 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

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
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
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

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
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	NA	≥160 mg/dL; ≥4.12 mmol/L	≥190 mg/dL; ≥4.90 mmol/L	NA
High-density lipoprotein (HDL)	NA	<40 mg/dL; <1 mmol/L	NA	NA	NA	NA	NA	NA
Triglycerides	NA	NA	NA	NA	>ULN	>300 mg/dL; >3.42 mmol/L	>500 mg/dL; >5.7 mmol/L	>1,000 mg/dL; >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

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
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
Blood urea nitrogen (BUN)	NA	NA	NA	NA	>ULN	>2.5 ULN	>5 ULN	>10 ULN
25-hydroxyvitamin D	< 20 ng/mL or < 50 nmol/mL	NA	NA	NA	> 60 ng/mL or > 125 nmol/mL	NA	NA	NA

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
PTH (0-20y)	<9 pg/mL or 1 pmol/L	NA	NA	NA	> 52 pg/mL or 5.5 pmol/L	NA	NA	NA
PTH (>20y)	<10 pg/mL or 1.1 pmol/L	NA	NA	NA	> 65 pg/mL or 6.9 pmol/L	NA	NA	NA
TRAP5b	<LLN	NA	NA	NA	>ULN	NA	NA	NA
P1NP	<LLN	NA	NA	NA	>ULN	NA	NA	NA
CTX Type I Collagen (20y - 51y), female	NA	NA	NA	NA	>574 pg/mL	NA	NA	NA
CTX Type I Collagen (20y - 30y) male	NA	NA	NA	NA	> 677 pg/mL	NA	NA	NA
CTX Type I Collagen (30y - 51y) male	NA	NA	NA	NA	> 685 pg/mL	NA	NA	NA
CTX Type I Collagen (> 51y) female	NA	NA	NA	NA	> 1009 pg/mL	NA	NA	NA
CTX Type I Collagen (51y - 71y) male	NA	NA	NA	NA	> 705 pg/mL	NA	NA	NA

Parameter	L1	L2	L3	L4	H1	H2	H3	H4
CTX Type I Collagen (>71y) male	NA	NA	NA	NA	> 855 pg/mL	NA	NA	NA

Table 3 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 sect ion of the CSR as per ICH E3 guidance, section 12.4.2.3.

APPENDIX II: Vital Signs and Electrocardiogram grading

Table 4 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

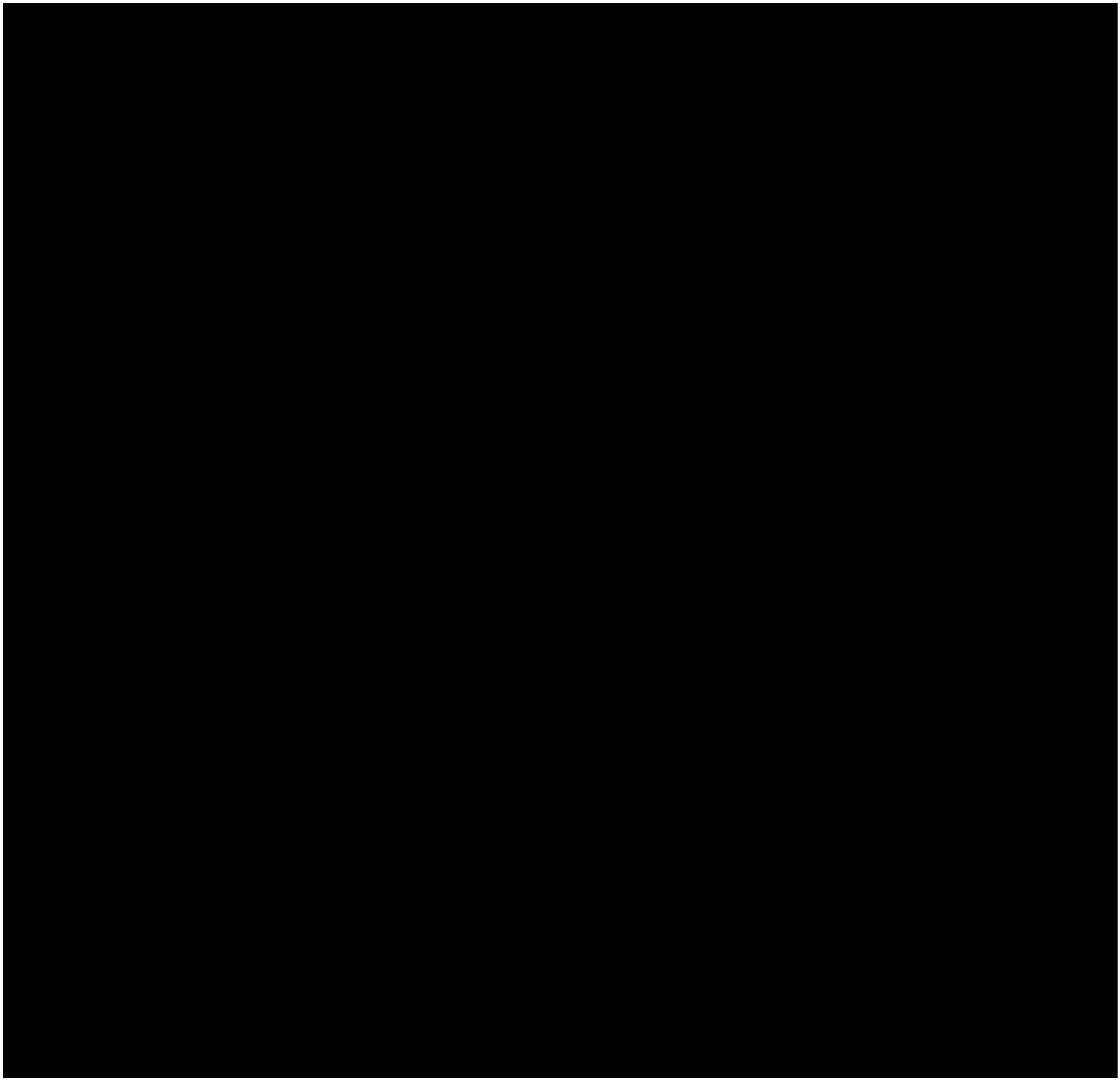
Parameter	L1	L2	L3	L4	H1	H2	H3	H4
QRS complex on ECG (ms)	<60	NA	NA	NA	>120	NA	NA	NA

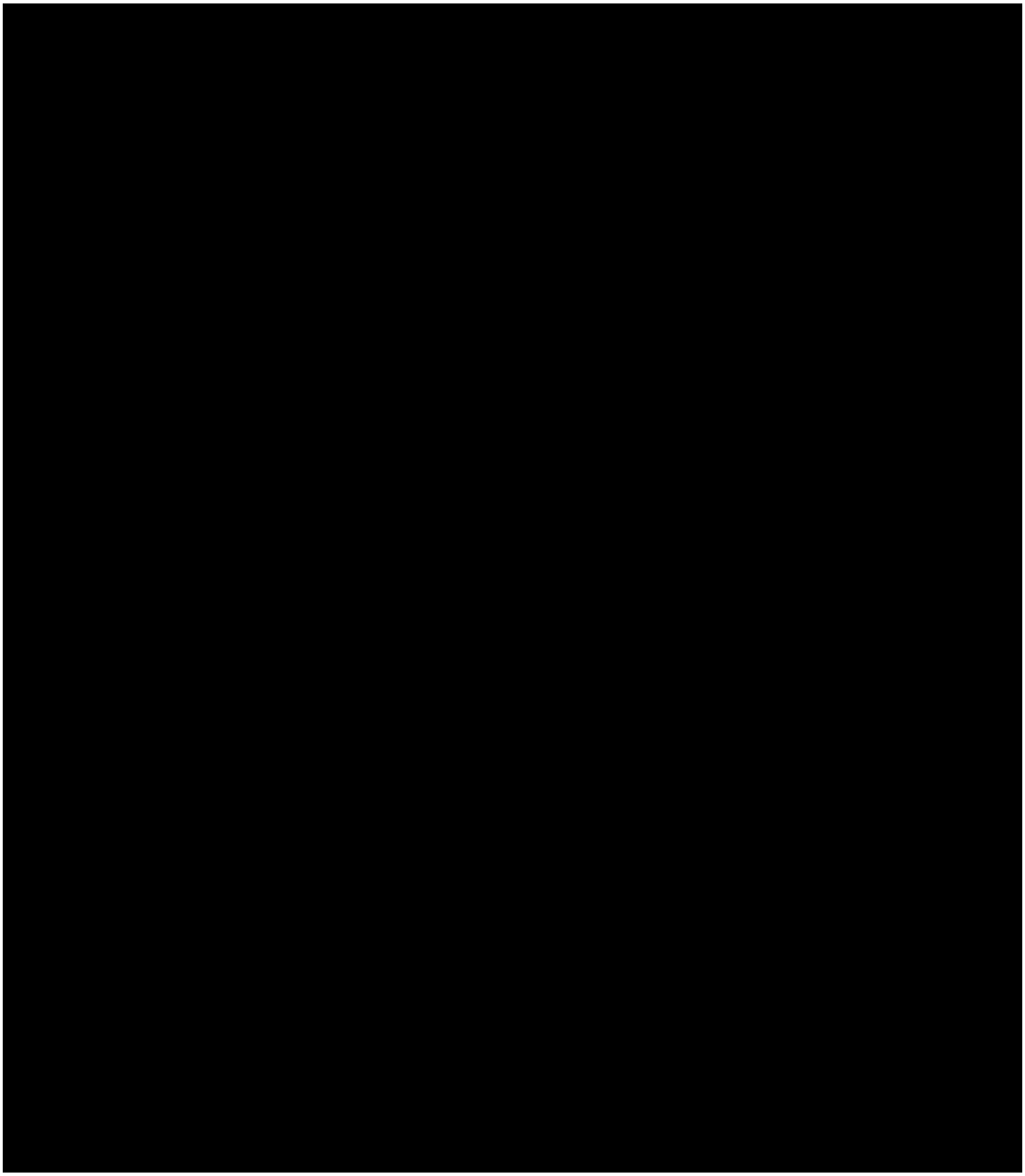
Source: FRM-MED-005/006, valid on 08JUL2020.

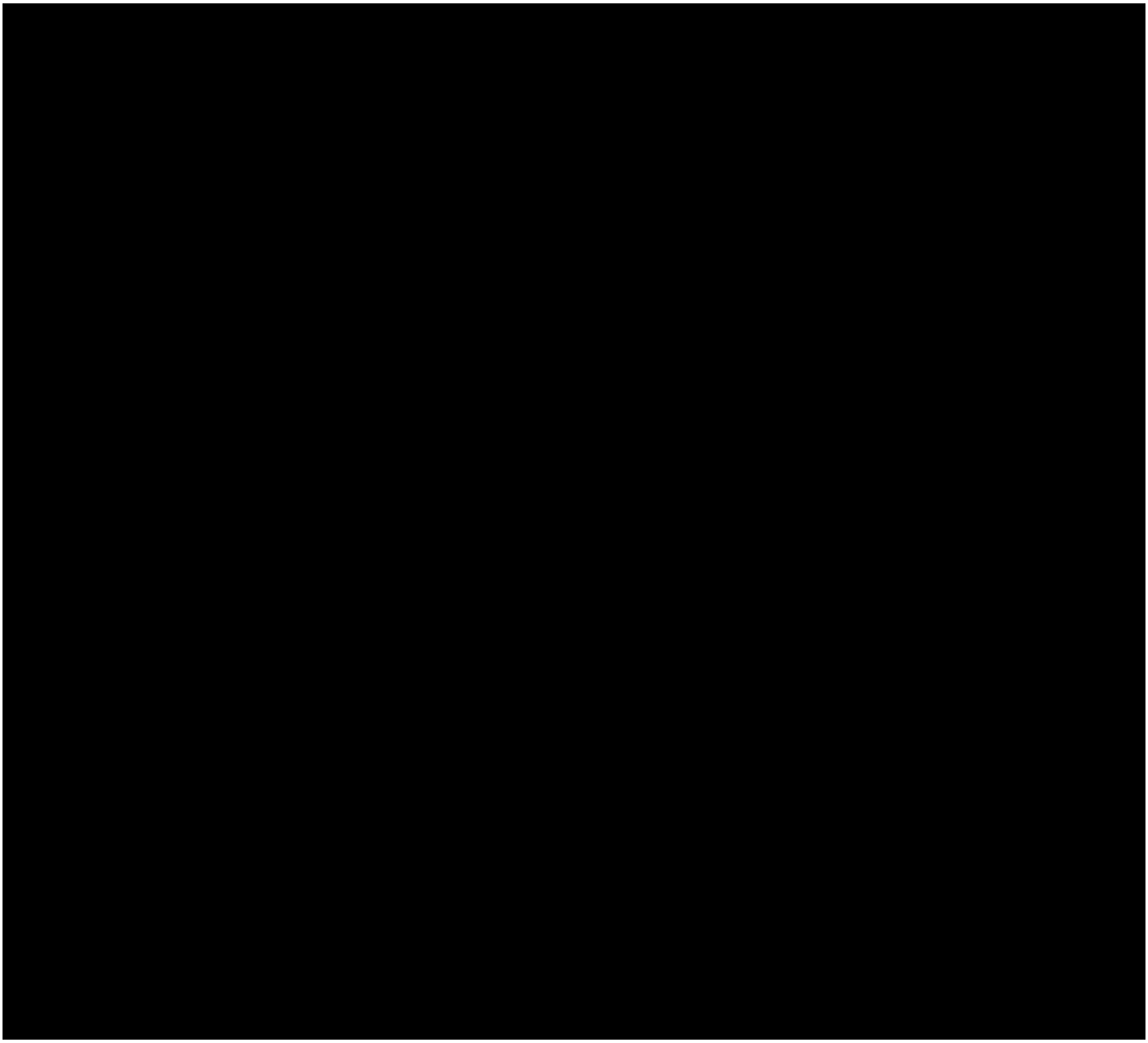
APPENDIX III: Mayo Scoring System

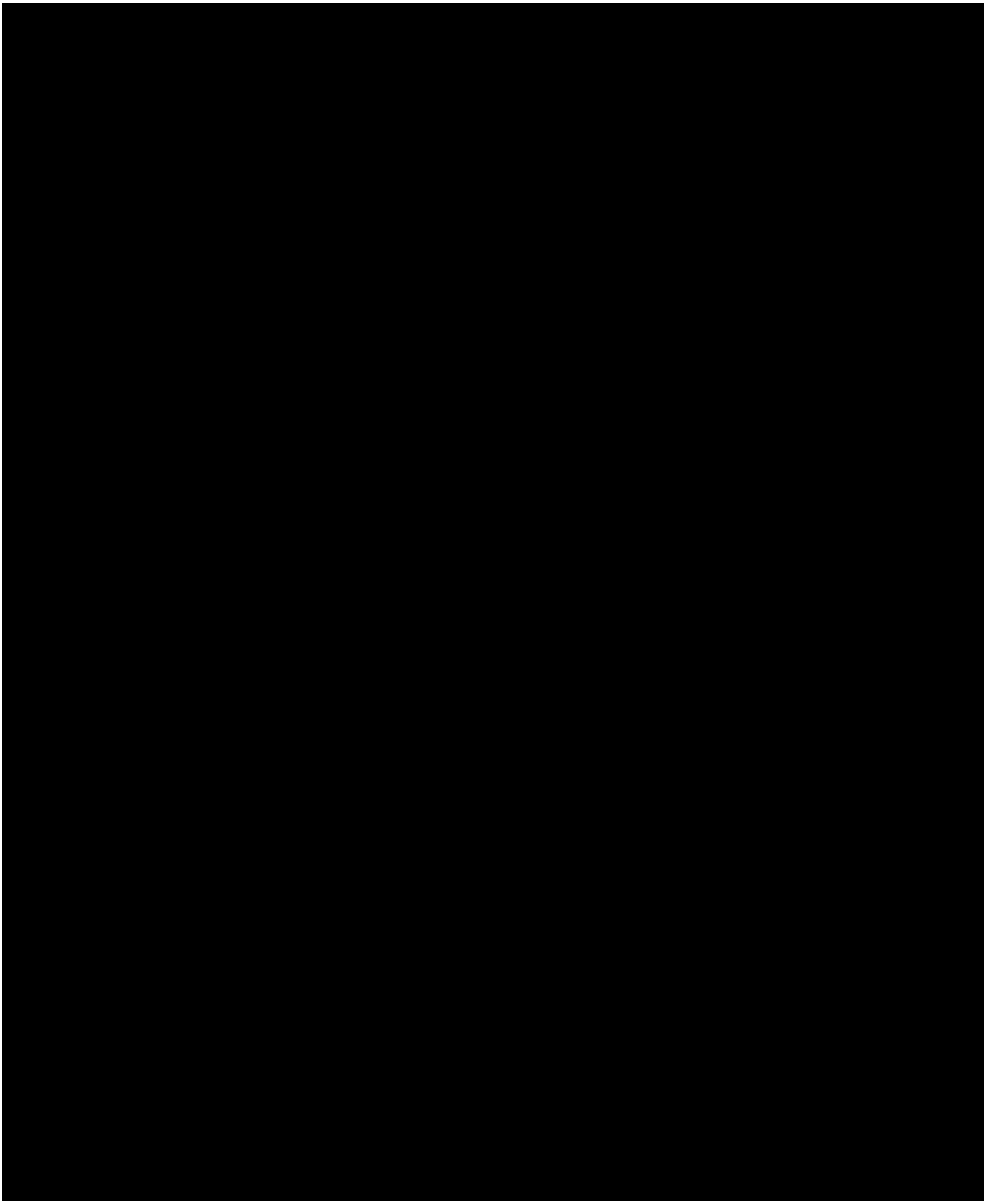
	0	1	2	3
Stool frequency [MCS-SF] [a]	Normal	1 -2 stools above normal	3 -4 stools above normal	>4 stools above normal
Rectal bleeding [MCS-RB] [a]	No blood seen	Streaks of blood less than half the time	Obvious blood with stool most of the time	Blood alone passed
Flexible sigmoidoscopy [MCS-ES] [b]	Normal or inactive disease	Mild disease (Erythema, decreased vascular pattern, no friability)	Moderate disease (Marked erythema, absent vascular pattern, friability, erosions)	Severe disease (Spontaneous bleeding, ulceration)
Physician's global assessment [PGA] [a]	Normal	Mild disease	Moderate disease	Severe disease

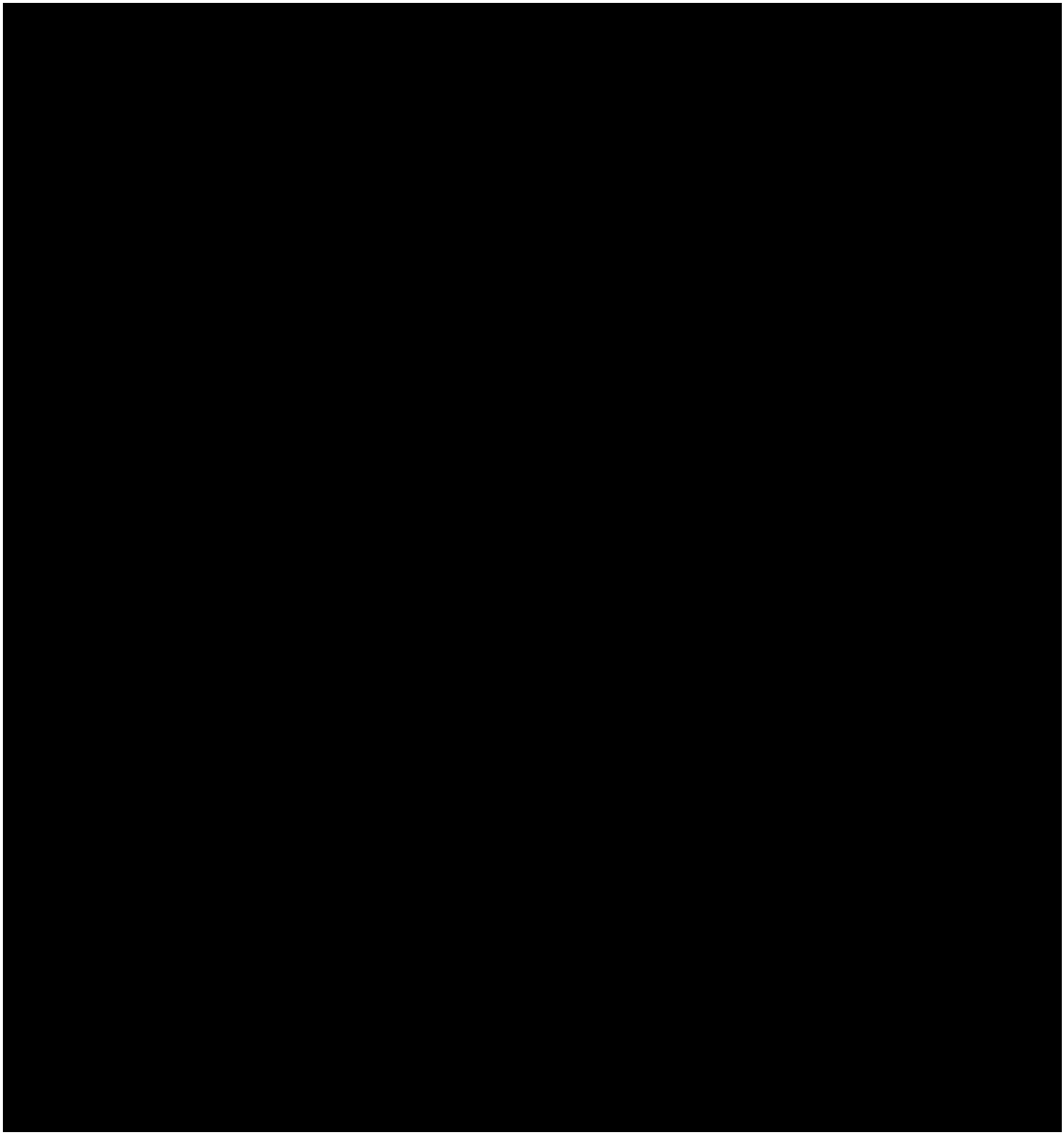
a (Lewis, et al., 2008)
b (Schroeder, Tremaine, & Ilstrup, 1987)

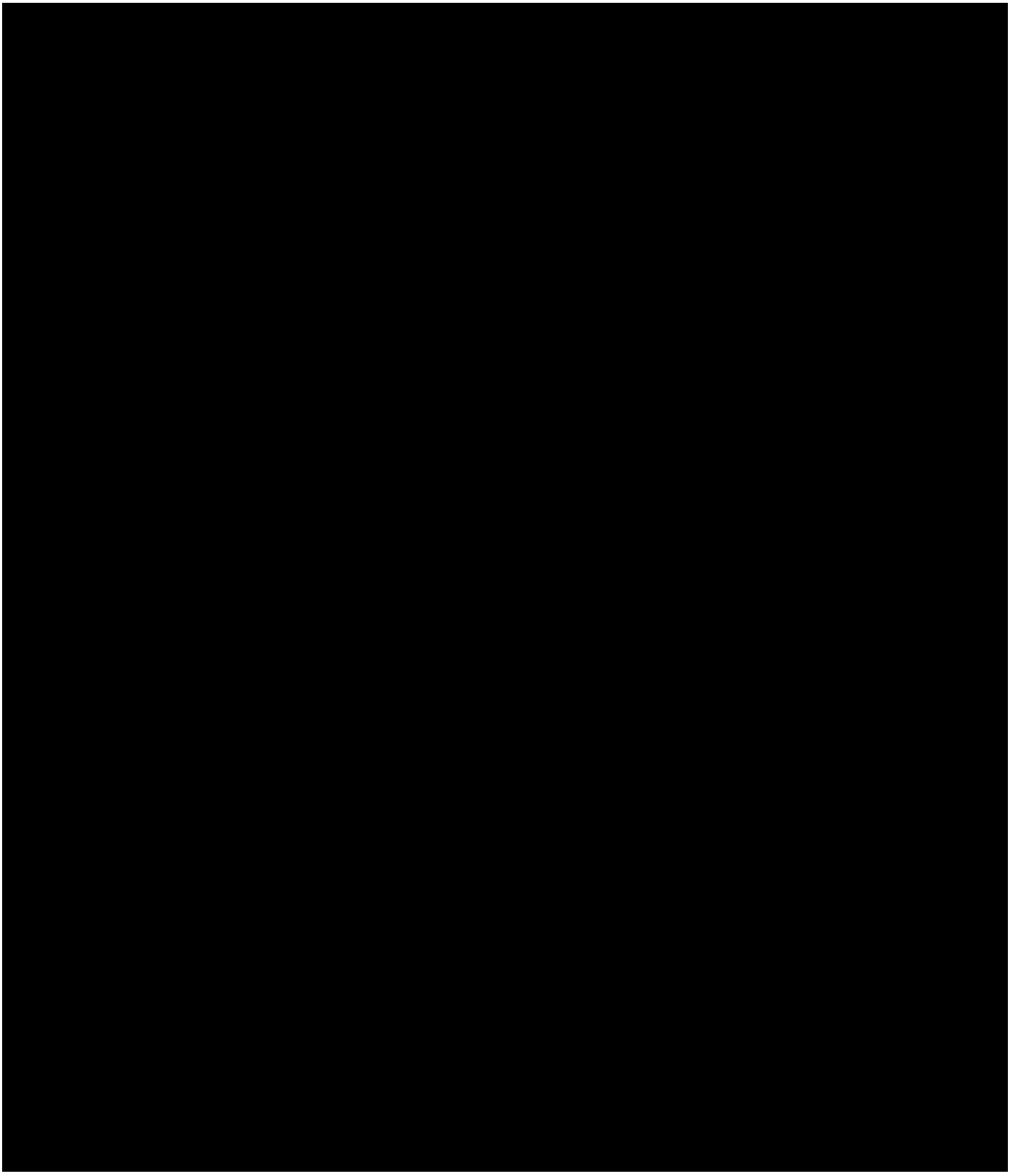


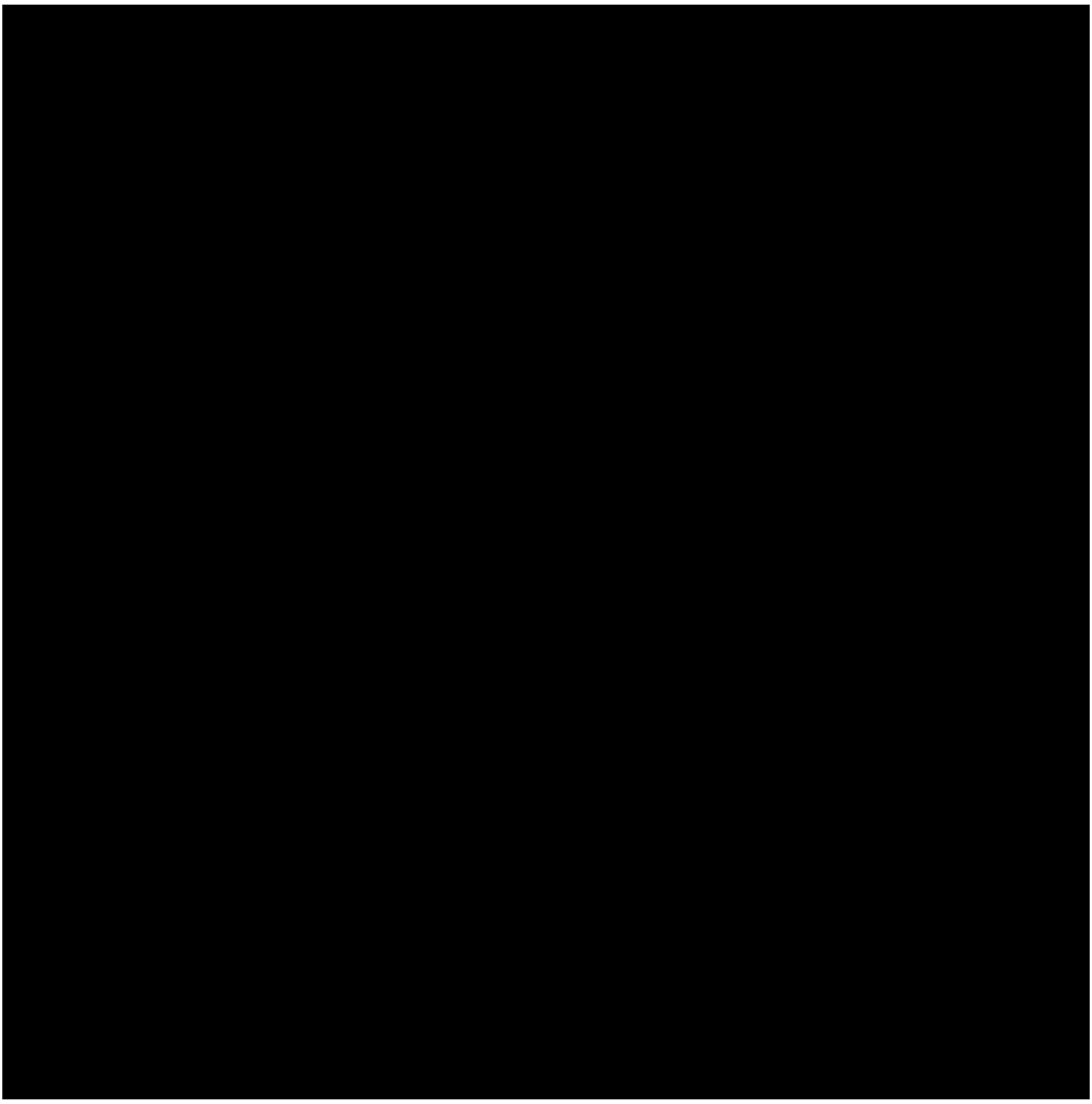


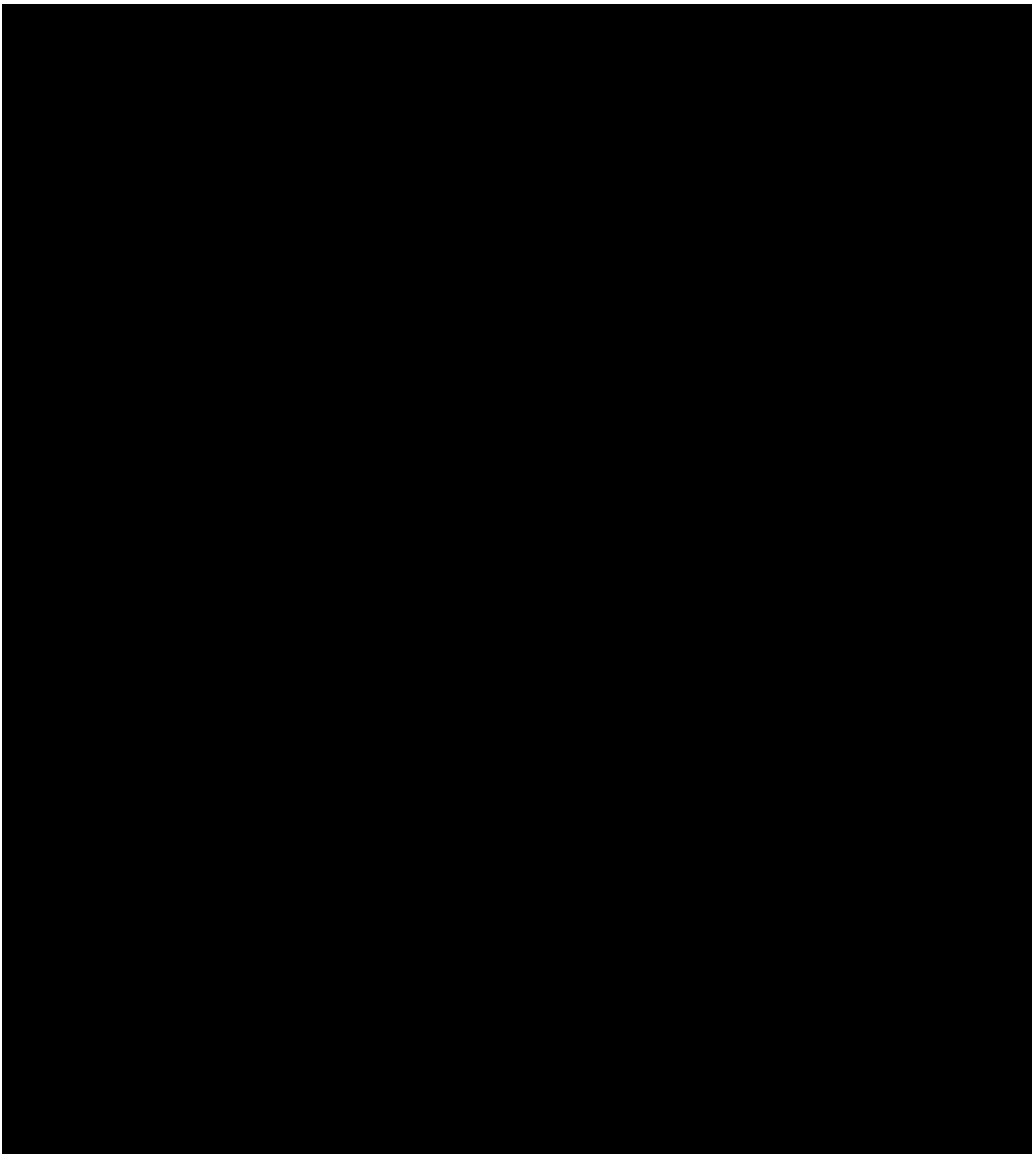


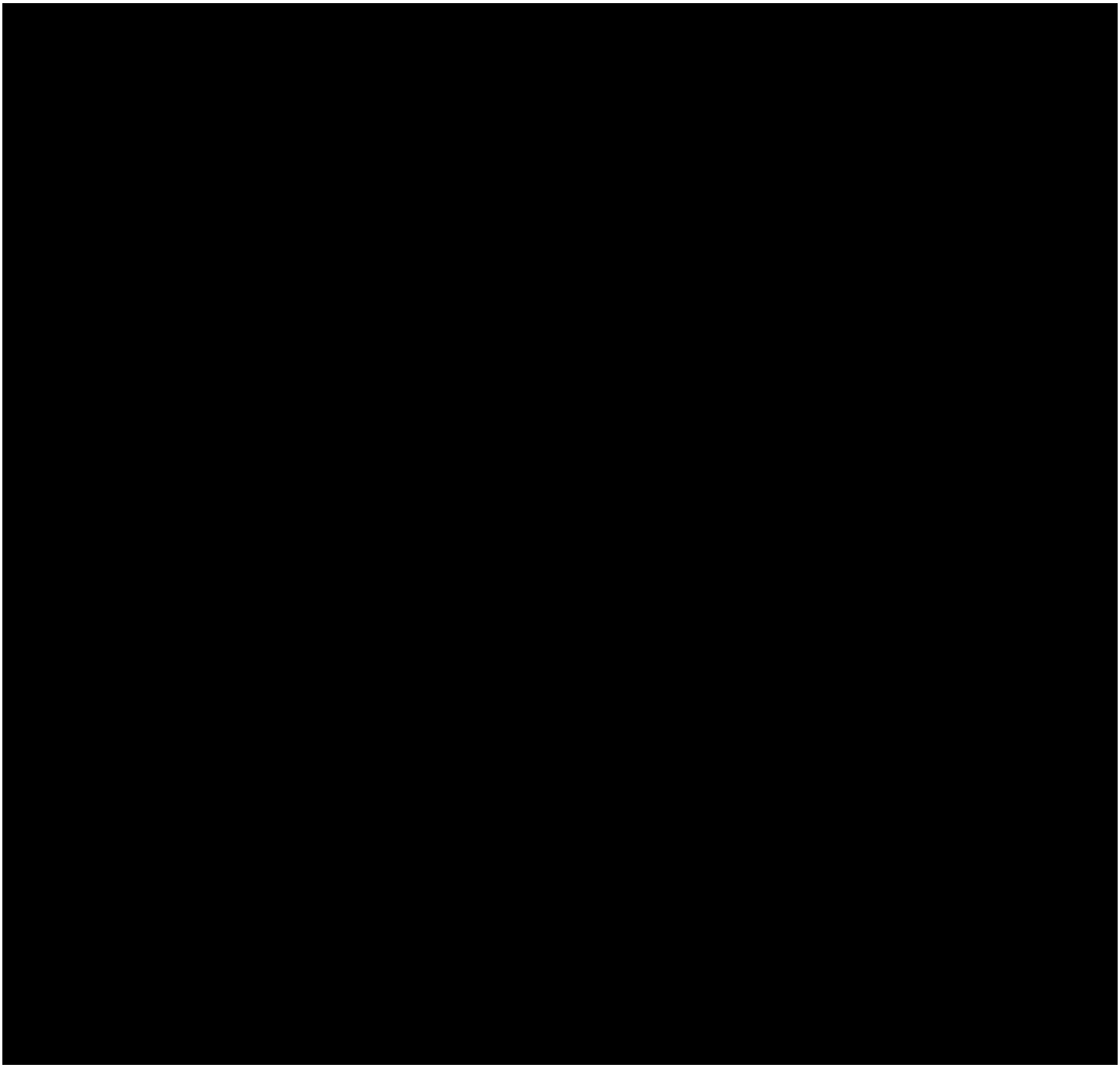


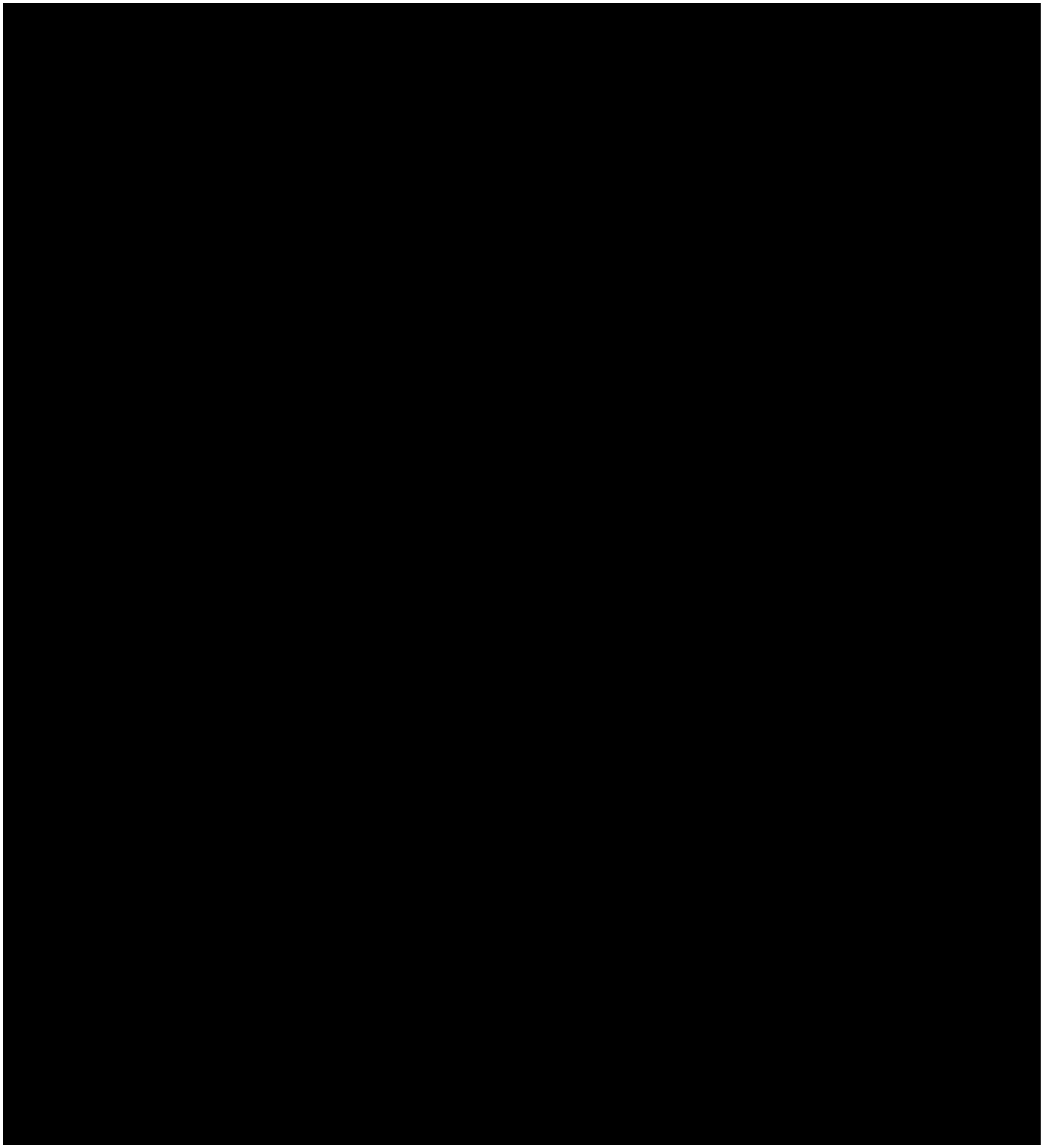


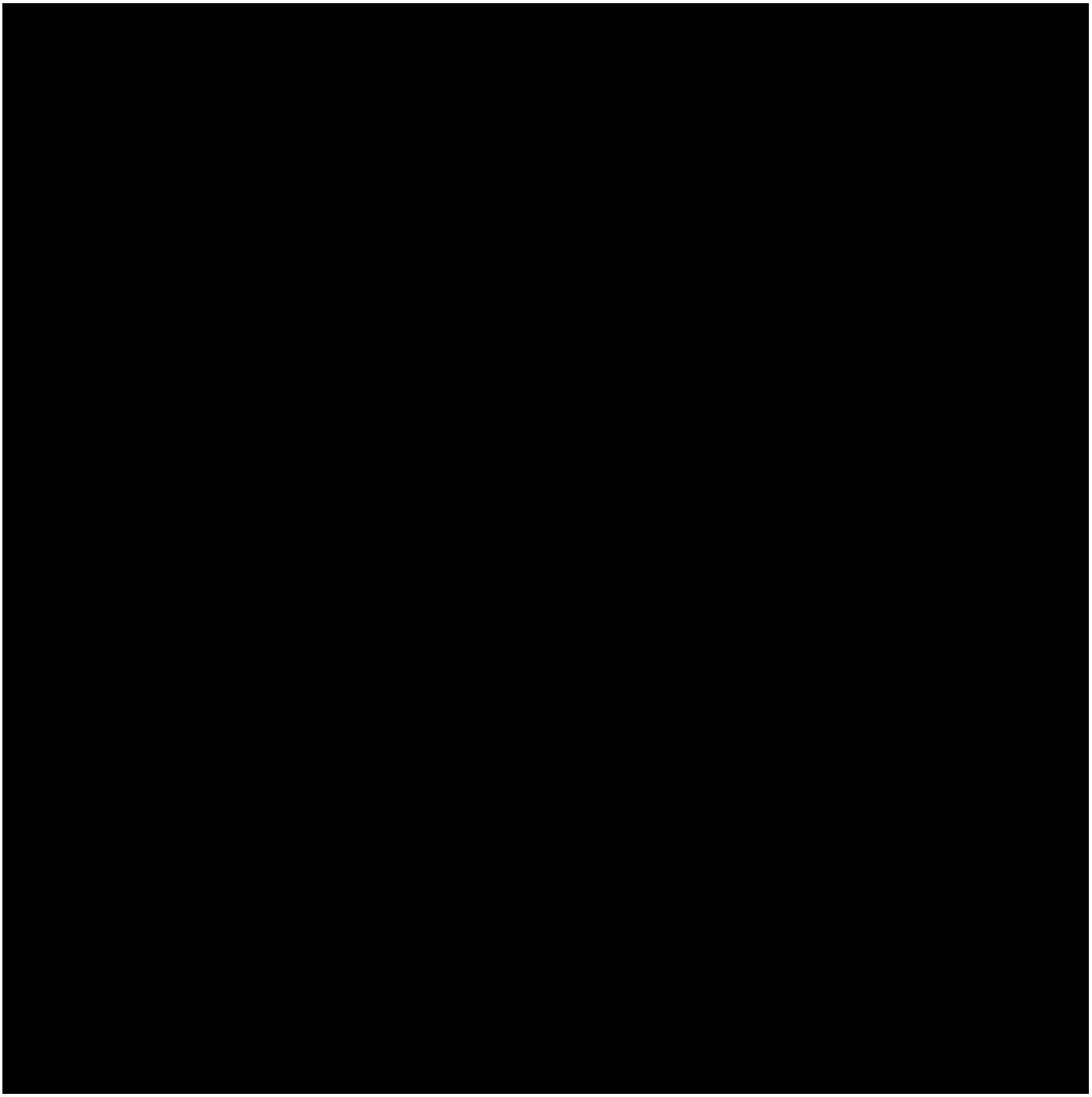


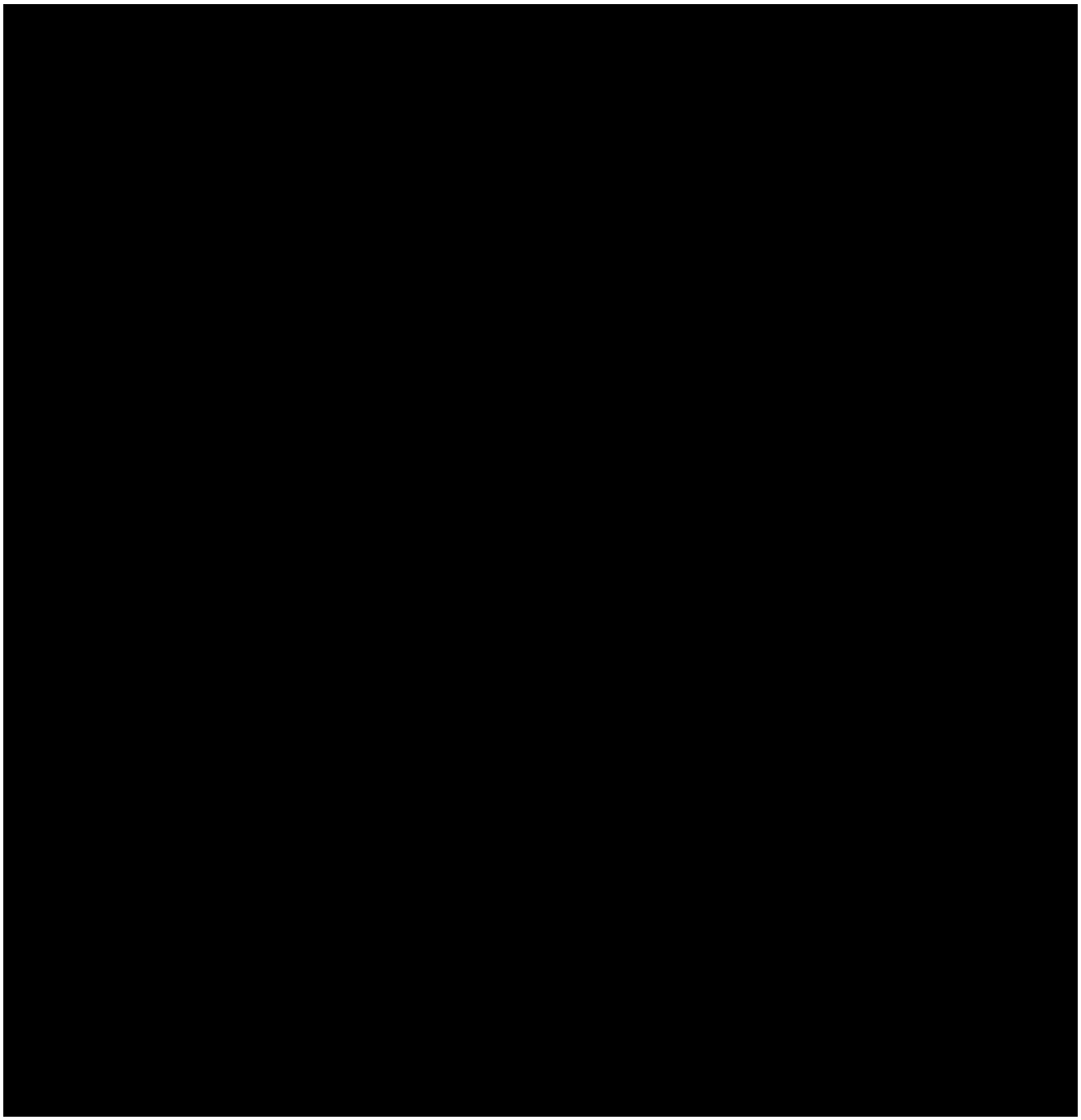




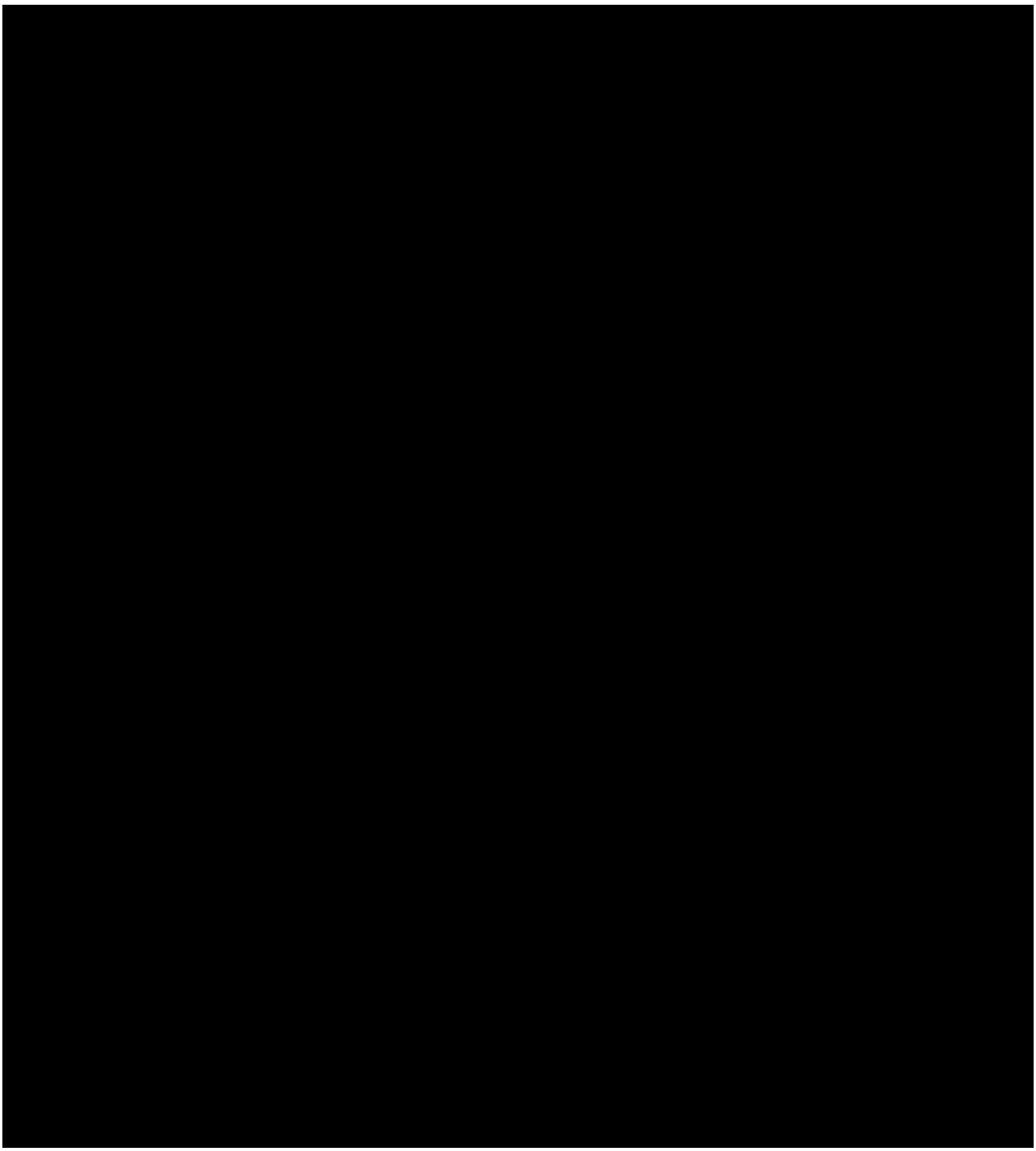


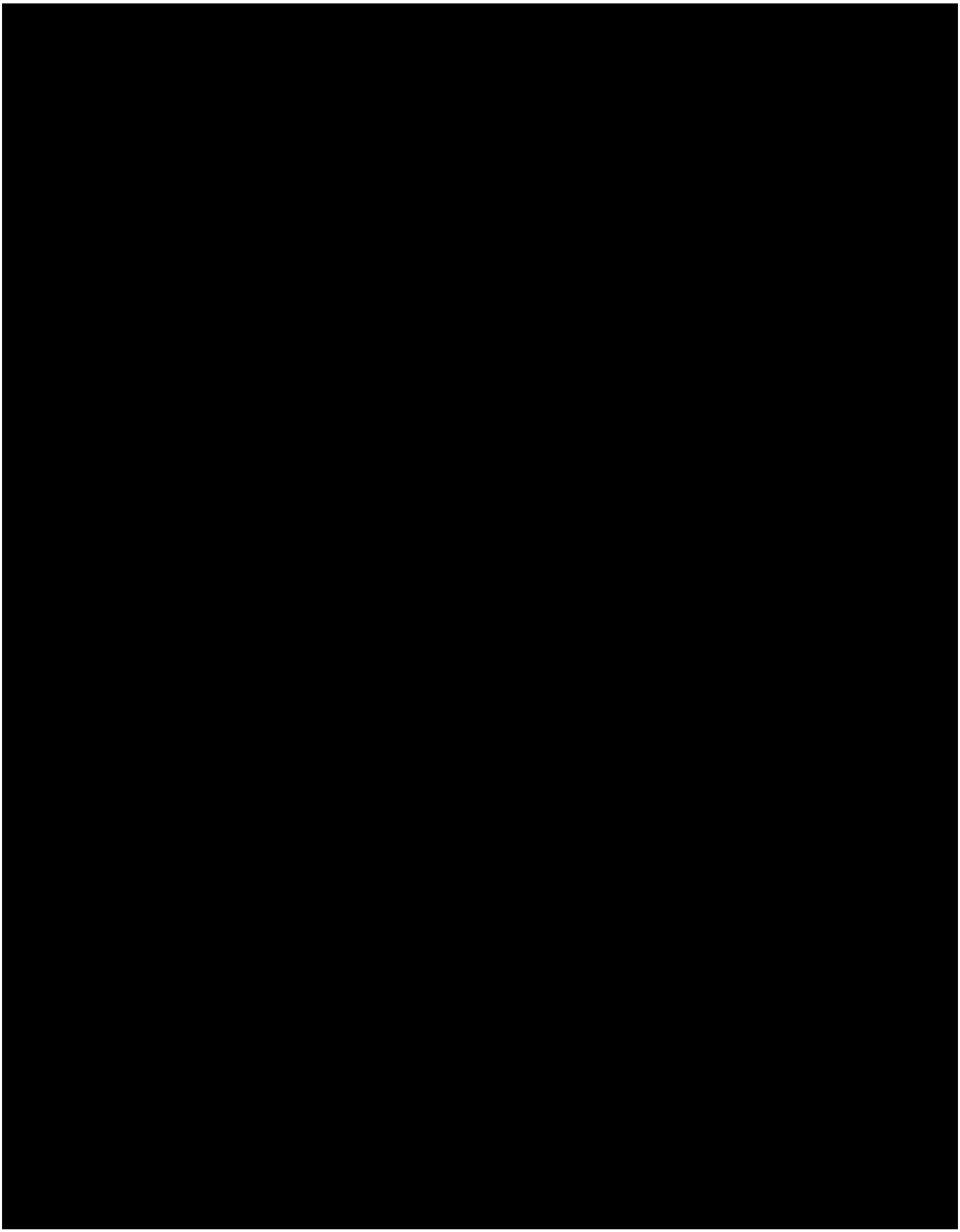


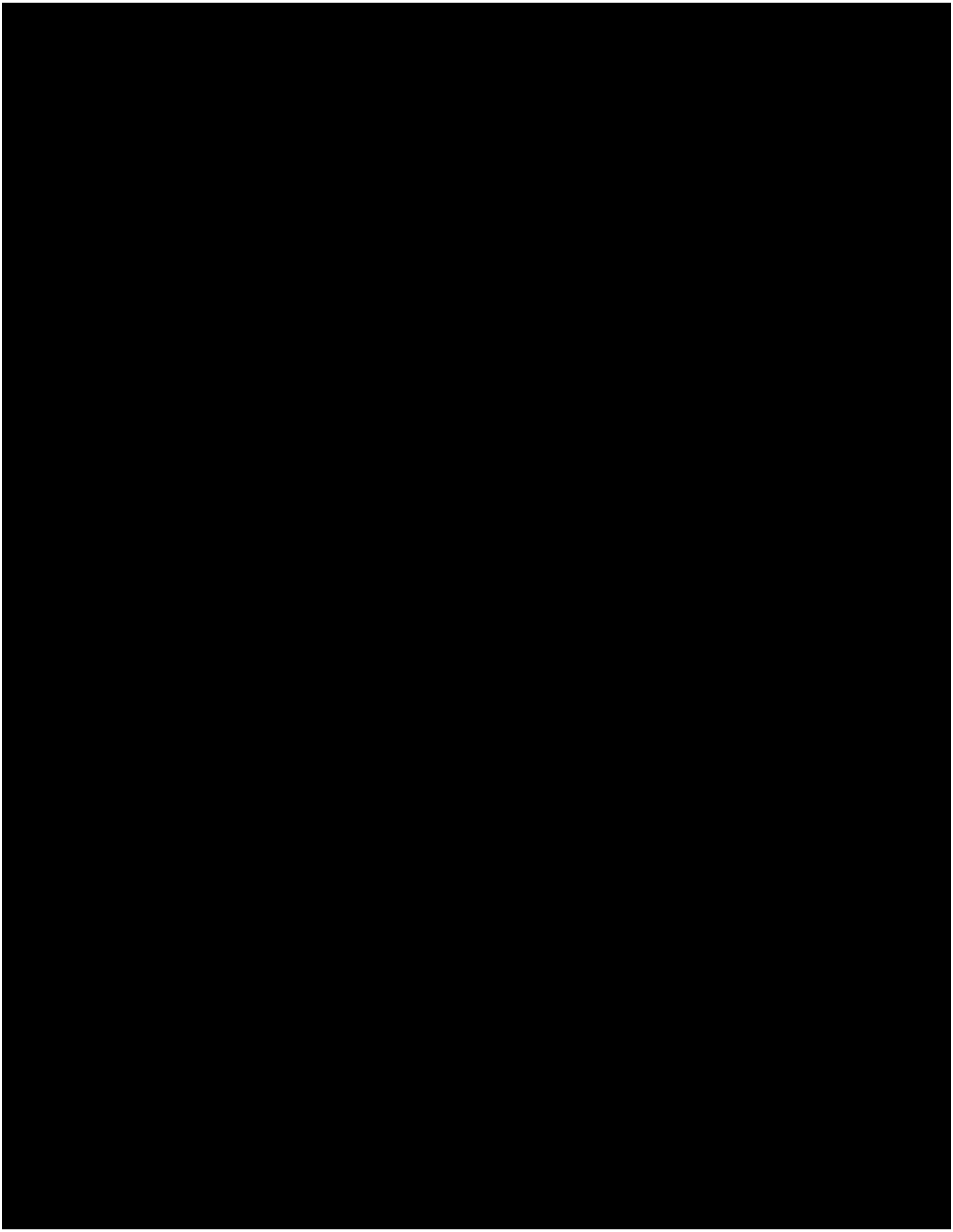


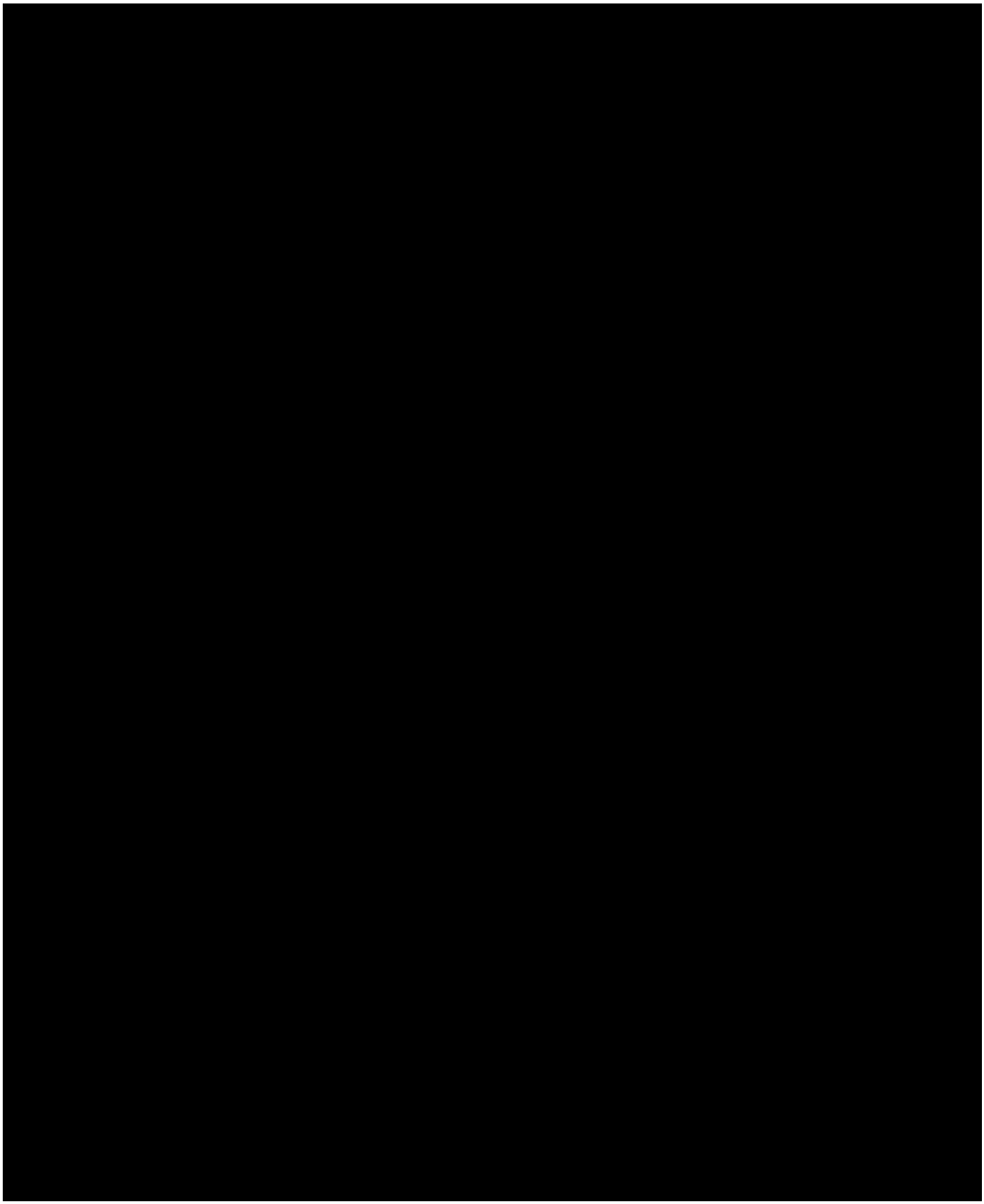


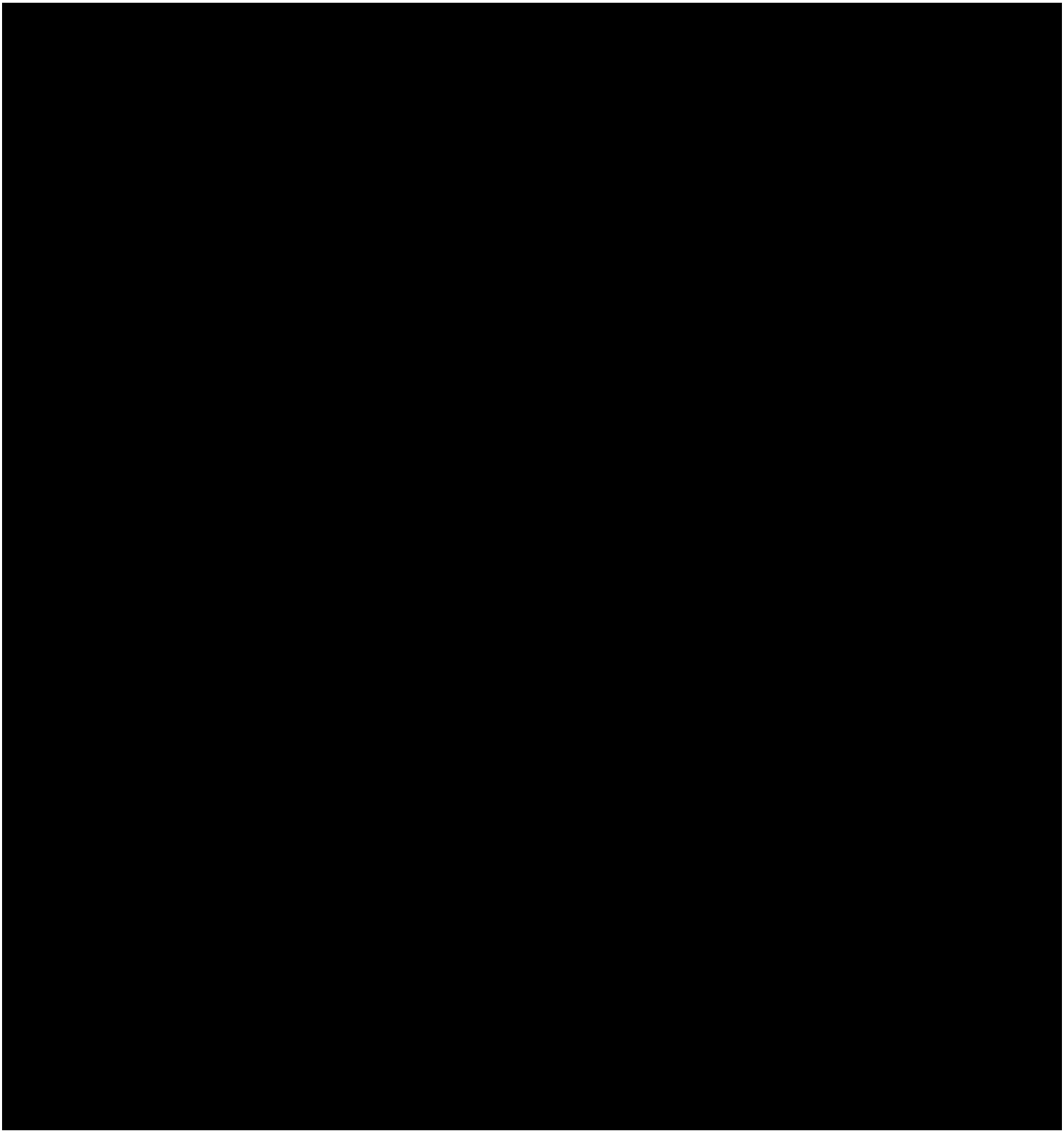
[redacted]











Signature Page for glpg3970-cl-210-sap 16159

Approval	[redacted] 16-Jun-2021 08:20:17 GMT+0000

Approval	[redacted] 16-Jun-2021 10:23:48 GMT+0000

Signature Page for glpg3970-cl-210-sap43952_17539_53034