KZR-616-208

A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PHASE 2A STUDY WITH OPEN-LABEL EXTENSION TO EVALUATE THE SAFETY AND EFFICACY OF ZETOMIPZOMIB (KZR-616) IN PATIENTS WITH AUTOIMMUNE HEPATITIS

Clinicaltrials.gov Identifier *NCT05569759*

Date of statistical analysis plan: *03 February 2025*

STATISTICAL ANALYSIS PLAN

A Randomized, Double-blind, Placebo-controlled, Phase 2a Study with Open-label Extension to Evaluate the Safety and Efficacy of Zetomipzomib (KZR-616) in Patients with Autoimmune Hepatitis

Sponsor: Kezar Life Sciences, Inc.
4000 Shoreline Court, Suite 300
South San Francisco, CA 94080

Protocol Number: KZR-616-208, V4.0

Blinding Management Plan: KZR-616-208, v2.0

Date: 03 February 2025

Version: 3.0

A Randomized, Double-blind, Placebo-controlled, Phase 2a Study with Open-label Extension to Evaluate the Safety and Efficacy of Zetomipzomib (KZR-616) in Patients with Autoimmune Hepatitis

Protocol Number: KZR-616-208, V4.0

SAP Version: 3.0

Statistical Analysis Plan Approval

Prepared by: [redacted] 06-Feb-2025 | 07:22 EST
[redacted] Date
Principal Biostatistician
[redacted]

Reviewed by: [redacted] 06-Feb-2025 | 11:32 MST
[redacted] Date
Sr. Director, Biostatistics
[redacted]

Reviewed by: [redacted] 06-Feb-2025 | 12:39 EST
[redacted] Date
QC Biostatistician
[redacted]

Approved by: *See electronic signature and date*
[redacted]
Kezar Life Sciences, Inc.
Date

Approved by: *See electronic signature and date*
[redacted]
Kezar Life Sciences, Inc.
Date

Approved by: *See electronic signature and date*
[redacted]
Kezar Life Sciences, Inc.
Date

TABLE OF CONTENTS

1. INTRODUCTION 12
2. STUDY OBJECTIVES 12
2.1. Objectives for Double-blind Treatment Period (DBTP) 12
2.1.1. Safety Objectives for DBTP 12
2.1.2. Efficacy Objective for DBTP 12
2.2. Objectives for Open-label Extension (OLE) Period 12
2.2.1. Safety Objectives for OLE Period 12
2.2.2. Efficacy Objective for OLE Period 12
3. ENDPOINTS 13
3.1. Definitions Related to Efficacy Endpoints 13
3.1.1. Successful glucocorticoid taper 13
3.1.2. Complete biochemical response (CR) 13
3.1.3. Partial biochemical response (PR) 13
3.1.4. Non-responder 13
3.1.5. Treatment failure 13
3.1.6. Disease flare 14
3.1.7. Baseline Value for ALT, AST, and IgG 14
3.1.8. Normal ALT, AST, and IgG 14
3.1.9. Baseline (DBTP-Baseline) 14
3.1.10. OLE-Baseline 14
3.1.11. Change from Baseline and Change from OLE-Baseline 15
3.1.12. EuroQol 5-dimension 5-level Index Value 15
3.1.13. Chronic Liver Disease Questionnaire Domain Scores and Overall Score 15
3.2. Endpoints for DBTP 15
3.2.1. Safety Endpoints for DBTP 15
3.2.2. Primary Efficacy Endpoint for DBTP 16
3.2.3. Secondary Efficacy Endpoints for DBTP 16
3.2.4. Exploratory Efficacy Endpoints for DBTP 16
3.3. Endpoints for the OLE Period 17
3.3.1. Safety Endpoints for the OLE Period 17
3.3.2. Primary Efficacy Endpoint for the OLE Period 17

3.3.3. Secondary Efficacy Endpoints for OLE Period 17
3.3.4. Exploratory Efficacy Endpoints for the OLE Period 17
3.3.5. Statistical Hypotheses 18
4. STUDY DESIGN AND PROCEDURES 19
4.1. General Study Design 19
4.1.1. DBTP 20
4.1.2. OLE Period 21
4.1.3. Methods of Assigning Patients to Treatment Groups 21
4.1.4. Study Treatment 22
4.1.5. Assessments Prior to Unblinding Data 22
4.1.5.1. Treatment Failure 22
4.1.5.2. Disease Flare/DILI 22
4.1.5.3. Protocol Deviations 23
4.1.5.4. Opportunistic Infection 24
4.1.6. Blinding/Unblinding/Database Freeze 24
4.1.7. Rescue Therapy 24
5. SAMPLE SIZE 25
6. ANALYSIS POPULATIONS 25
6.1. Full Analysis Set 25
6.2. Per Protocol Set 25
6.3. Safety Set 25
6.4. OLE Safety Set 25
6.5. Intent to Treat Population 25
7. GENERAL STATISTICAL CONSIDERATIONS 26
7.1. Overview 26
7.1.1. Timing of Analyses 26
7.1.2. Subgroups 26
7.1.3. Assessing Normality for Continuous Measures 26
7.2. Missing or Inconclusive Data Handling 26
7.2.1. Missing Efficacy Assessments 26
7.2.2. Missing Safety Assessments 26
7.2.3. QOL Missing Data 28
7.2.3.1. EQ-5D-5L Index Value 28

7.2.3.2. CLDQ Domain Scores and Overall Score 28
7.3. Scheduled/Unscheduled Visits 28
7.4. Data Analysis Conventions....29
7.5. Adjustments for Multiplicity 30
8. DISPOSITION OF PATIENTS....30
9. PROTOCOL DEVIATIONS 31
10. DEMOGRAPHICS AND BASELINE CHARACTERISTICS 31
11. MEDICAL HISTORY 32
12. PRIOR AND CONCOMITANT MEDICATIONS....32
13. DOSING COMPLIANCE AND TREATMENT EXPOSURE....32
13.1. Dosing Compliance 32
13.2. Treatment Exposure....33
14. EFFICACY ANALYSES 34
14.1. General Considerations....34
14.2. Summary of Efficacy Endpoints....35
14.3. Analysis of Primary Endpoint 44
14.3.1. DBTP 44
14.3.2. OLE Period 45
14.3.3. Sensitivity Analyses....46
14.3.3.1. Worst Observation Carried Forward 46
14.3.3.2. Tipping Point Analysis 46
14.4. Analyses of Secondary Efficacy Endpoints....47
14.4.1. Proportion of Patients Who Achieve PR, Disease Flare, or Treatment Failure 47
14.4.2. Change from Baseline in ALT, AST, IgG, Ishak, LSM, CAP scores 47
14.4.3. Time to [CR] or [PR with a glucocorticoid taper]....47
14.4.4. Duration of [CR] or [PR with successful glucocorticoid taper] 48
14.4.5. Time to patient experiencing first Systemic Injection Reaction....49
14.4.6. Sensitivity Analysis of Secondary Endpoints....49
14.5. Analyses of Exploratory Efficacy Endpoints 49
14.5.1. DBTP Exploratory Efficacy Endpoints 49
14.5.2. OLE Exploratory Efficacy Descriptive Endpoints 51
14.5.3. OLE Exploratory Efficacy Endpoints....52
15. SAFETY ANALYSIS 52

15.1. Adverse Events 52
15.2. TEAE by category 54
15.3. DILI 55
15.4. Systemic Injection Reactions 55
15.5. Clinical Laboratory Data 55
15.6. Vital Signs 56
15.7. ECG 12-Lead 56
15.8. Physical Examination 57
16. INTERIM ANALYSES (PRIMARY ANALYSES) 57
17. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL 57
18. REFERENCES 63
19. APPENDICES 64
APPENDIX A. SAS® CODES FOR EQ-5D-5L INDEX VALUES 64
APPENDIX B. SCHEDULE OF VISITS AND ASSESSMENTS 66
APPENDIX C. WEIGHTING, SCORING, AND ANALYSIS OF THE GTI 75
APPENDIX D. COMPOSITE GLUCOCORTICOID TOXICITY INDEX 75

LIST OF TABLES

Table 1: The Domains of Chronic Liver Disease Questionnaire 15
Table 2: Analysis Visit Windows for DBTP 28
Table 3: Analysis Visit Windows for Open Label Extension Period 29
Table 4: Efficacy Endpoints and Analysis Methods 36
Table 5: Sensitivity Analyses for Primary Analysis 46
Table 6: Stages of Liver Fibrosis Based on FibroScan Liver Stiffness Measurement (LSM) 51
Table 7: Changes to Analysis Planned in the KZR-616-208 Protocol 58
Table 8: Schedule of Assessments (Visits with Study Assessments) – Double Blind Treatment Period 66
Table 9: Schedule of IMP Administration (Visits with no Study Assessments) – Double Blind Treatment Period 70
Table 10: Schedule of Assessments (Visits with Study Assessments) – Open Label Extension Period 71

Table 11: Schedule of IMP Administration (Visits with no Study Assessments) – Open Label Extension Period....74

LIST OF FIGURES

Figure 1: Study Design Schema....19

Figure 2: Flowchart Showing Paths that a Patient Can Take to Achieve CR, PR, Non-Responder, or Treatment Failure at Week 12, 16, 20, or 24....45

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse Event
AESI	Adverse Events of Special Interest
AIH	Autoimmune Hepatitis
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
AZA	Azathioprine
CAP	Controlled attenuation parameter
CI	Confidence Interval
CLDQ	Chronic Liver Disease Questionnaire
CMH	Cochran-Mantel-Haenszel
CR	Complete Biochemical Response
CRF	Case Report Form
DBTP	Double-blind Treatment Period
DILI	Drug-induced liver injury
ECG	Electrocardiogram
EOS	End of Study
EOT	End of Treatment
EQ-5D-5L	EuroQol 5-dimension 5-level
EQ-VAS	EuroQol Visual Analog Scale
ETV	Early Termination Visit
FAS	Full Analysis Set
GC	Glucocorticoid
GTI	Glucocorticoid Toxicity Index
GTI-AIS	GTI Aggregate Improvement Score
GTI-CWS	GTI Cumulative Worsening Score
HAI	Histological Activity Index
HgbA1c	Hemoglobin A1c
ICF	Informed Consent Form

Abbreviation	Definition
ICH	International Council for Harmonisation
IDMC	Independent Data Monitoring Committee
IgG	Immunoglobulin G
IMP	Investigational Medicinal Product
IRT	Interactive Response Technology
ISR	Injection Site Reaction
ITT	Intent to Treat Population
KZR-616	Zetomipzomib
LDL	Low-density Lipoprotein
LLN	Lower Limit of Normal Range
LS	Least-square
LSM	Liver Stiffness Measurement
MedDRA	Medical Dictionary for Regulatory Activities
MMF	Mycophenolate Mofetil
MMRM	Mixed Model for Repeated Measures
OLE	Open-label Extension
PP	Per Protocol
PR	Partial Biochemical Response
PT	Preferred Term
QTcF	QT interval Corrected for Pulse Rate by Fridericia's Formula
RTSM	Randomization and Trial Supply Management
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	Standard Deviation
[redacted]	[redacted]
SIR	Systemic Injection Reaction
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TESAE	Treatment-Emergent Serious Adverse Event
ULN	Upper Limit of Normal Range
VCTE	Vibration-controlled Transient Elastography

Abbreviation	Definition
WHODrug	World Health Organization Drug Dictionary

1. INTRODUCTION

The purpose of this amended statistical analysis plan (SAP) is to reflect the changes made in the Addendum 1 to SAP version 2.0, dated 14 January 2025.

This SAP is being written with due consideration of the recommendations outlined in the most recent International Conference on Harmonisation (ICH) E9 Guideline entitled, Guidance for Industry: Statistical Principles for Clinical Trials, and the most recent ICH E3 Guideline entitled, Guidance for Industry: Structure and Content of Clinical Study Reports.

This SAP provides details and scope of the planned analyses and methodology which will be used to assess the KZR-616-208 data. The statistical analysis methods presented in this document will supersede the statistical analysis methods described in the clinical protocol. If additional analyses are required to supplement the planned analyses described in this SAP, they may be completed and will be identified in the Clinical Study Report.

PK data analyses will be provided in the PK Analysis Plan and results will be provided in a stand-alone PK report. Description of these analyses is outside the scope of this SAP.

2. STUDY OBJECTIVES

2.1. Objectives for Double-blind Treatment Period (DBTP)

2.1.1. Safety Objectives for DBTP

To evaluate the safety and tolerability of zetomipzomib (KZR-616) in patients with autoimmune hepatitis (AIH) during the double-blind treatment period (DBTP).

2.1.2. Efficacy Objective for DBTP

The efficacy objective for the DBTP is to evaluate the efficacy of zetomipzomib in addition to standard-of-care in patients with AIH who have failed standard-of-care treatment, had an incomplete biochemical response to ≥3 months of standard-of-care treatment, or had a disease flare after experiencing complete remission induced by standard-of-care treatment.

2.2. Objectives for Open-label Extension (OLE) Period

2.2.1. Safety Objectives for OLE Period

To evaluate the long-term safety and tolerability of zetomipzomib in patients with AIH who continue in the open-label extension (OLE) period.

2.2.2. Efficacy Objective for OLE Period

The efficacy objective for the OLE period is to evaluate the long-term efficacy of zetomipzomib in patients with AIH.

3. ENDPOINTS

3.1. Definitions Related to Efficacy Endpoints

3.1.1. Successful glucocorticoid taper

Prednisone/prednisone equivalent dose tapered from the starting dose (at Baseline) to a prednisone/prednisone equivalent dose of ≤10 mg/day by Week 24 or earlier.

3.1.2. Complete biochemical response (CR)

Normal alanine aminotransferase (ALT), aspartate aminotransferase (AST), and immunoglobulin G (IgG) values (if IgG level is elevated at Baseline) with glucocorticoid dose not higher than starting dose (at Baseline).

Note: Once a patient has a treatment failure they are not eligible for a CR during the treatment period in which they failed treatment (ie, the DBTP or the OLE Period).

3.1.3. Partial biochemical response (PR)

ALT, AST, and IgG values (if IgG level is elevated at Baseline) with glucocorticoid dose not higher than starting dose (at Baseline) meeting one of the following criteria:

- <2 × ULN (values must be lower than Baseline)
- >80% improvement from Baseline.

Note: Each laboratory value (ALT, AST, IgG) must be considered separately for PR criteria above. For example, if AST is in the normal range, IgG is in the normal range, but ALT is <2 x ULN, then this qualifies as a PR. If all three laboratory values (ALT, AST, IgG) are less than the ULN, then this is considered a CR. A patient cannot be a PR and a CR at the same time.
A patient's response can improve from a PR to a CR. Once a patient has a treatment failure, they are not eligible for a PR during the treatment period in which they failed treatment (ie, the DBTP or the OLE Period).

3.1.4. Non-responder

Patients who do not achieve a [CR], [PR with a successful glucocorticoid taper], or TF (Treatment Failure). Patients who have insufficient data for response determination at a time point will be considered non-responders for that time point.

3.1.5. Treatment failure

Treatment failure is defined by meeting one of the following criteria:

- Patient's ALT, AST, or IgG (if IgG level is elevated at Baseline) level worsened ≥2 times that of the Baseline value that is sustained for ≥1 week as verified via repeat laboratory assessments, despite compliance with standard of care (ie, with regard to inclusion criteria) or protocol-defined therapy.

  OR

- If the glucocorticoid or immunosuppressant dose is increased above the Baseline dose or if a new immunosuppressant is added during the study due to an elevation in liver enzymes, it may be considered a treatment failure unless it is attributed to an adverse event (AE) not relating to AIH. Patients falling under this criterium will be classified through a blinded evaluation prior to the DBTP database freeze (see Section 4.1.5.1).

Note: Treatment failure can only occur if a patient has not already achieved a [CR] or [PR with a successful glucocorticoid taper].

3.1.6. Disease flare

Elevation of ALT after achieving [CR] or [PR with a successful glucocorticoid taper] that meets all the following criteria:

- is ≥25% above the [CR] or [PR with a successful glucocorticoid taper] value
- is ≥1.25 ULN
- is sustained for ≥1 week as verified via repeat laboratory assessments
- requires the patient to re-start or escalate glucocorticoid therapy or immunosuppressants
- is not due to other identifiable causes [eg, viral hepatitis, drug-induced liver injury (DILI), concomitant medications, alcoholic liver injury, etc.]

The last 2 bullets of the Disease Flare definition (above) will be adjudicated through a blinded evaluation prior to the DBTP database freeze (see Section 4.1.5.2).

If a patient experiences a [PR with a successful glucocorticoid taper] first and then experiences a [CR], the laboratory values will be based upon the [CR].

3.1.7. Baseline Value for ALT, AST, and IgG

Mean of all collected Screening and Day 1 values prior to first dose of investigational medicinal product (IMP). When only one ALT, AST, or IgG result prior to first dose of IMP is available, that single value will be used for the Baseline (DBTP-Baseline).

3.1.8. Normal ALT, AST, and IgG

Normal ranges will be defined according to the definitions provided by the central laboratory ULN. Minor dips in the ALT/AST/IgG below LLN are not considered clinically meaningful in this patient population. Thus, normal ALT, AST, and IgG will be defined by the ULN.

3.1.9. Baseline (DBTP-Baseline)

Last non-missing assessment prior to first dose of IMP in the DBTP, except for those specified in Section 3.1.10 .

3.1.10. OLE-Baseline

Last non-missing assessment prior to first dose of IMP in the OLE period.

3.1.11. Change from Baseline and Change from OLE-Baseline

Change from Baseline = Post-Baseline value – Baseline (ie, DBTP-Baseline) value.

Change from OLE-Baseline = Post OLE-Baseline value – OLE-Baseline value.

3.1.12. EuroQol 5-dimension 5-level Index Value

By using the individual responses to the EuroQol 5-dimension 5-level (EQ-5D-5L) descriptive system, index values for the EQ-5D-5L will be calculated using the SAS codes in Appendix A from https://euroqol.org/wp-content/uploads/2024/01/US_valueset_SAS.txt.

3.1.13. Chronic Liver Disease Questionnaire Domain Scores and Overall Score

The Chronic Liver Disease Questionnaire (CLDQ) contains 29 items divided into 6 domains as described below in Table 1. Overall CLDQ score is an average of the 6 domain scores.

Table 1: The Domains of Chronic Liver Disease Questionnaire

Domain	No of Items	Question #	Domain Score
Abdominal Systems	3	1, 5, 17	Average of item scores
Fatigue	5	2, 4, 8, 11, 13	Average of item scores
Systemic Symptoms	5	3, 6, 21, 23, 27	Average of item scores
Activity	3	7, 9, 14	Average of item scores
Emotional Functions	8	10, 12, 15, 16, 19, 20, 24, 26	Average of item scores
Worry	5	18, 22, 25, 28, 29	Average of item scores

3.2. Endpoints for DBTP

The DBTP safety and efficacy endpoints are described below in subsections. The flowchart shown in Figure 2 shows the methodology for determining the status of a patient at Week 12, 16, 20, and 24, where the status can be determined as one of the following: complete biochemical response (CR), partial biochemical response (PR), non-responder, or treatment failure.

3.2.1. Safety Endpoints for DBTP

The primary safety endpoint is the proportion of patients who experience AEs and serious adverse events (SAEs). Other safety endpoints include:

- Incidence, nature, and severity of AEs and SAEs
- Incidence of AEs leading to IMP discontinuation, dose reduction, and dose interruptions
- Changes in standard laboratory parameters, vital signs, autoantibodies, and electrocardiogram (ECG) parameters
- Drug-induced liver injury (Hy's law)

3.2.2. Primary Efficacy Endpoint for DBTP

Proportion of patients who achieve CR (Section 3.1.2) by Week 24; analyses will also be performed on data collected at Weeks 12, 16, and 20.

3.2.3. Secondary Efficacy Endpoints for DBTP

- Change from Baseline in ALT at Weeks 12, 16, 20, and 24
- Proportion of patients who achieve PR (Section 3.1.3) at Weeks 12, 16, 20, and 24
- Time to CR
- Time to [CR] or [PR with a successful glucocorticoid taper]
- Proportion of patients experiencing a disease flare (Section 3.1.9) after CR or PR
- Duration of CR during the DBTP
- Duration of [CR] or [PR with a successful GC taper] during the DBTP
- Proportion of patients who are treatment failures (Section 3.1.8)
- Proportion of patients who achieve CR with successful glucocorticoid taper (Section 3.1.1) by Week 24; analyses will also be performed on data collected at Weeks 12, 16, and 20
- Proportion of patients who achieve CR with glucocorticoid taper to ≤5 mg by Week 24
- Proportion of patients who achieve CR with glucocorticoid taper to 0 mg by Week 24
- Proportion of patients who achieve PR with successful glucocorticoid taper (Section 3.1.1) by Week 24
- Proportion of patients who achieve PR with glucocorticoid taper to ≤5 mg by Week 24
- Proportion of patients who achieve PR with glucocorticoid taper to 0 mg by Week 24

3.2.4. Exploratory Efficacy Endpoints for DBTP

- Proportion of patients who achieve ALT normalization based on Prati criteria (males, 30 U/L; females, 19 U/L) (Prati et al., 2002), with successful glucocorticoid taper, by Week 24
- Change from Baseline in AST at Weeks 12, 16, 20, and 24
- Change from Baseline in IgG at Weeks 12, 16, 20, and 24
- Change from Baseline in glucocorticoid dose at Weeks 12, 16, 20, and 24
- Proportion of patients who achieve CR that had a baseline ALT >2x ULN
- Glucocorticoid Toxicity Index (GTI) Cumulative Worsening Score (GTI-CWS) and the Aggregate Improvement Score (GTI-AIS) at Week 24. Please See Appendix C and Appendix D for information on calculating the scores for GTI.

- Change from Baseline in liver stiffness measurement (LSM) and controlled attenuation parameter (CAP) at Week 24, assessed by vibration-controlled transient elastography (VCTE) utilizing Fibroscan®
- Descriptive Summary Liver Fibrosis stages shown in Table 6 at timepoints that LSM data was collected during DBTP
- Change from Baseline in the EuroQol Visual Analog Scale (EQ-VAS) and EQ-5D-5L index value at Weeks 16 and 24
- Change from Baseline in the CLDQ domain and overall scores at Weeks 16 and 24
- Change from Baseline in liver histopathology at Week 24, based on Ishak (modified Histological Activity Index [HAI]) score (Ishak et al., 1995)
- Descriptive Summary of Baseline and 24-week liver histopathology characteristics.

3.3. Endpoints for the OLE Period

3.3.1. Safety Endpoints for the OLE Period

The primary safety endpoint is the proportion of patients who experience AEs and SAEs. Other safety endpoints include:

- Incidence, nature, and severity of AEs and SAEs
- Incidence of AEs leading to IMP discontinuation, dose reduction, and dose interruptions
- Changes in standard laboratory parameters, vital signs, and ECGs
- Drug-induced liver injury (Hy's law)

3.3.2. Primary Efficacy Endpoint for the OLE Period

Proportion of patients experiencing a disease flare among the patients who achieved a CR during the DBTP.

3.3.3. Secondary Efficacy Endpoints for OLE Period

There are no secondary efficacy endpoints in the OLE Period.

3.3.4. Exploratory Efficacy Endpoints for the OLE Period

- Proportion of patients who have CR during the OLE period of the study among the patients who did not achieve a CR during the DBTP
- Proportion of patients who have PR with a successful glucocorticoid taper during the OLE period of the study among the patients who did not achieve a PR with a glucocorticoid taper during the DBTP
- Proportion of patients who have a CR during the OLE period of the study with successful glucocorticoid taper who did not achieve a CR during the DBTP

- Proportion of patients who have a PR with a successful glucocorticoid taper during the OLE period of the study who did not achieve a PR with a successful glucocorticoid taper during the DBTP
- Proportion of patients who achieve CR with glucocorticoid taper to ≤5 mg by OLE Week 25
- Proportion of patients who achieve CR with glucocorticoid taper to 0 mg by OLE Week 25
- Proportion of patients who achieve PR with glucocorticoid taper to ≤5 mg by OLE Week 25
- Proportion of patients who achieve PR with glucocorticoid taper to 0 mg by OLE Week 25
- Duration of CR during the entire study (DBTP + OLE)
- Duration of PR with a successful glucocorticoid taper during the entire study (DBTP + OLE)
- Change from Baseline and change from OLE-Baseline in ALT
- Change from Baseline and change from OLE-Baseline in AST
- Change from Baseline and change from OLE-Baseline in IgG
- Change from Baseline and change from OLE-Baseline in the EQ-5D-5L Index Value and Visual Analog Scale (EQ-VAS)
- Change from Baseline and change from OLE-Baseline in CLDQ domain and overall scores
- GTI-CWS and the GTI-AIS at OLE-Week-25.
- Change from Baseline and change from OLE-Baseline in liver stiffness and controlled attenuation parameter score, assessed by VCTE utilizing Fibroscan® at OLE-Week-25
- Descriptive Summary Liver Fibrosis stages shown in Table 6 at timepoints that LSM data was collected during study
- Change from Baseline and OLE-Baseline in Ishak score at OLE Week 25 (modified Histological Activity Index [HAI]) scores (Ishak et al., 1995)

3.3.5. Statistical Hypotheses

The statistical analyses of the study data will be descriptive in nature, and no formal hypothesis testing will be performed. Reported p-values are for hypothesis-generating and should be interpreted with caution. These analyses are intended to explore potential relationships or patterns within the data, rather than to confirm pre-specified hypotheses. As such, any significant findings should be considered as suggestive rather than conclusive evidence. Additionally, this study does not apply corrections for multiple comparisons which may elevate the risk of Type I errors (false positives), and thus, all significant findings should be interpreted cautiously.

4. STUDY DESIGN AND PROCEDURES

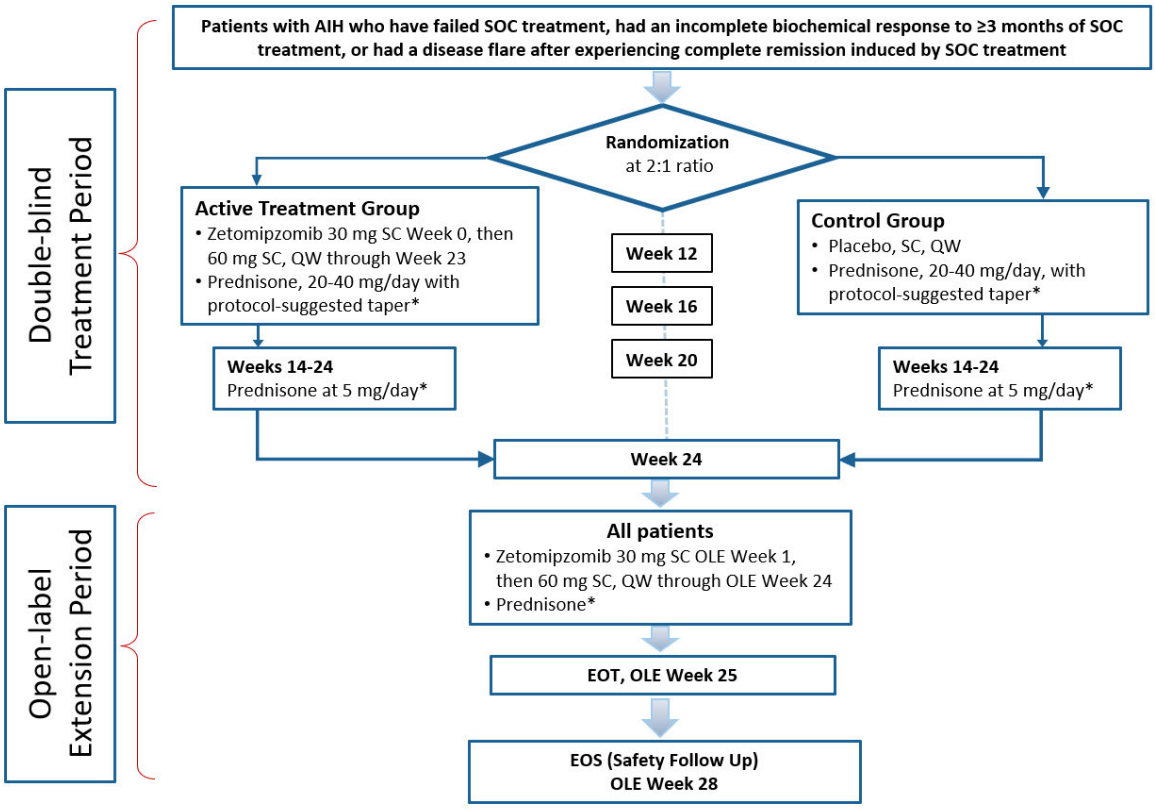
4.1. General Study Design

This is a Phase 2a, multi-center, randomized, double-blind, placebo-controlled study with an OLE Period to evaluate the safety, tolerability, and efficacy of zetomipzomib in patients with AIH who have failed standard-of-care treatment, had an incomplete biochemical response to ≥3 months of standard-of-care treatment, or had a disease flare (relapse) after experiencing complete remission induced by standard-of-care treatment.

An Independent Data Monitoring Committee (IDMC) will review safety findings as described in Protocol Section 9.11.

The study design schema can be visualized in Figure 1.

Figure 1: Study Design Schema



Abbreviations: AIH=autoimmune hepatitis; EOS=End-of-Study; EOT=End-of-Treatment; QW=once weekly; OLE=open-label extension; SC=subcutaneous; SoC=standard-of-care
Note: Patients who are not eligible or choose not to roll over to the OLE Period will have an EOS visit (safety follow up) 4 weeks after completion of the Double-blind Treatment Period. *For protocol-suggested taper schedule, refer to Protocol Table 6 and Protocol Table 7 for prednisone equivalents in Protocol Appendix B.

4.1.1. DBTP

In the DBTP, zetomipzomib or placebo will be administered weekly for a 24-week treatment period. Patients will be evaluated for eligibility according to the entry criteria (see Protocol Section 4) within 4 weeks prior to the first dose of IMP on Day 1 (Visit 2). Efficacy assessments will be performed for all patients at Weeks 12, 16, 20, and 24. Safety will be assessed throughout the study by monitoring of vital signs, clinical laboratory tests, and physical examinations, and by recording and analyzing all AEs and SAEs.

Eligible patients will be randomized in a 2:1 ratio on Day 1 (Visit 2) to receive either standard-of-care (glucocorticoids) with zetomipzomib (Active Treatment group), or standard-of-care with Placebo (Control group). Zetomipzomib and placebo (ie, sterile water for injection in an equivalent volume to the reconstituted zetomipzomib dose) will be administered subcutaneously once weekly by study personnel at the site or by a home trial service at the patient's home or alternate location (eg, place of employment). The first dose of IMP will be 30 mg zetomipzomib or placebo, followed by weekly doses of 60 mg zetomipzomib or placebo for the remaining 23 weeks of the DBTP.

Enrolled patients will be on standard-of-care with a prednisone/prednisone equivalent dose of 20 mg/day on Day 1 (Visit 2). Following consultation with the Medical Monitor, patients who are taking a prednisone/prednisone equivalent dose higher than 20 mg/day at Baseline may also be enrolled in the study provided the prednisone/prednisone equivalent dose does not exceed 40 mg/day. Patients are encouraged to undergo a glucocorticoid taper of prednisone/prednisone equivalent dose according to the protocol-suggested glucocorticoid taper schedule (Protocol Table 6) and prednisone equivalent doses (Protocol Table 7) as provided in Protocol Appendix B. Patients taking budesonide who are unwilling to switch to a prednisone/prednisone equivalent for the study may continue to take budesonide at the doses shown in Protocol Appendix B.

In addition to prednisone/prednisone equivalent, patients may continue to use up to two additional immunosuppressive agents if these agents are already used as Baseline therapy:

a. Oral azathioprine (AZA) or oral Mycophenolate mofetil (MMF) including the alternatives mycophenolate sodium or mycophenolic acid

   and/or

b. Oral tacrolimus, cyclosporine A, or 6-mercaptopurine

The dose of the immunosuppressants described above in a) and b), must be stable for at least 4 weeks prior to Screening and must be held constant unless tolerability/AEs require dose adjustment and glucocorticoid taper is completed. Following completion of the glucocorticoid taper, immunosuppressants described above in a) and b) may be decreased at the Investigator's discretion.

For patients who are confirmed treatment failures or who experience a disease flare during the treatment period, IMP may be discontinued upon consultation with the Medical Monitor, and additional care will be introduced as medically indicated. Patients who receive rescue therapy will temporarily discontinue IMP, complete the DBTP visits, and remain eligible for the OLE Period. Patients who do not receive rescue therapy but undergo a temporary interruption in IMP (≤2 consecutive doses or ≤4 total doses) should also remain in the study. If the patient cannot

continue with the planned assessments, the patient will complete the Early Termination Visit (ETV) followed by the End-of-Study (EOS) visit 4 weeks later.

A liver biopsy will be performed at the Week 24 visit to assess any changes from Baseline in liver histopathology.

Patients who are eligible to roll over to the OLE Period will receive their first subcutaneous (SC) injection of zetomipzomib at the dose and schedule described in the OLE Period section (Protocol Section 3.1.2). Patients who are not eligible or choose not to roll over to the OLE Period will have an EOS visit (assessments shown in Table 8) 4 weeks after completion of the DBTP.

4.1.2. OLE Period

Patients who complete the DBTP study visits through Week 24 are eligible to roll over to the OLE Period. If a patient does not meet eligibility criteria for the OLE Period, then the patient would proceed with the EOS visit.

All patients will receive a SC injection of 30 mg zetomipzomib at the OLE Week 1 visit, followed by weekly SC injections of 60 mg zetomipzomib, administered by study personnel at the site or by a home trial service at the patient's home or alternate location (eg, place of employment), for a total of 24 additional weeks of treatment. After the OLE Week 1 Visit, additional on-site study visits will occur every 4 weeks from OLE Week 4 through OLE Week 25 (EOT visit). At these visits, safety and efficacy assessments will be performed according to the OLE Schedule of Assessments (Table 10) patients will have a safety follow-up visit (EOS Visit) 4 weeks after their last dose of zetomipzomib, for a maximum potential length of participation in the OLE of 28 weeks.

During the course of the OLE Period, prednisone/prednisone equivalent can continue to be tapered and may be withdrawn (see Protocol Appendix B for suggested taper schedule). Patients with signs and symptoms of worsening disease (ie, worsening of ALT level ≥25% from OLE-Baseline for ≥1 week, as verified via repeat laboratory assessments) may increase glucocorticoid doses following consultation with the Medical Monitor. For patients who are confirmed treatment failures or who experience a disease flare (Section 3.1.9) during the OLE Period, IMP may be discontinued upon consultation with the Medical Monitor, and additional care will be introduced as medically indicated. Patients who receive rescue therapy (see Protocol Section 6.3) will temporarily discontinue IMP but will remain in the study to complete OLE visits. Patients who do not receive rescue therapy but undergo a temporary interruption in IMP (≤2 consecutive doses or ≤4 total doses) should also remain in the study. If the patient cannot continue with the planned assessments, the patient will complete the ETV followed by the EOS visit, 4 weeks later.

An optional liver biopsy may be performed at the OLE Week 25 visit to assess changes in liver histopathology.

4.1.3. Methods of Assigning Patients to Treatment Groups

Once patients have consented, undergone all screening procedures, and been determined to be eligible for the study, they will return for the Baseline Visit to be randomized in a 2:1 ratio to receive weekly SC administration of zetomipzomib plus standard-of-care or placebo control plus

standard-of-care during the DBTP. Patients will be stratified by steroid use at Screening (ie, steroid use, no steroid use at Screening).

The Randomization and Trial Supply Management (RTSM) System is developed and managed by [redacted] to help the sponsor with randomization, blinding, and dispensing of IMP. RTSM is an updated term for Interactive Response Technology (IRT).

4.1.4. Study Treatment

An unblinded pharmacist will prepare and blind the active and placebo doses. During the DBTP, IMP (active zetomipzomib or volume-matched placebo) will be administered by SC injection once weekly, according to randomization. The first dose will be 30 mg zetomipzomib or placebo, administered at the clinical trial site, followed by weekly doses of 60 mg zetomipzomib or placebo for the remaining 23 weeks of the DBTP. During the OLE Period, zetomipzomib will be administered to all patients by SC injection once weekly. The first open-label dose will be 30 mg zetomipzomib, followed by weekly doses of 60 mg zetomipzomib for the duration of the OLE Period.

After the first dose, IMP will be administered by study personnel at the site or by a home trial service at the patient's home or alternate location (eg, place of employment).

4.1.5. Assessments Prior to Unblinding Data

4.1.5.1. Treatment Failure

The definition of treatment failure can be seen in Section 3.1.5. Once a patient has a treatment failure they are not eligible for a [CR] or [PR with a successful glucocorticoid taper] during the treatment period in which they failed treatment (ie, the DBTP or the OLE Period).

Treatment failure assessment will occur in a blinded manner by the Medical Monitor and Safety Physician for patients who have a glucocorticoid increase above their glucocorticoid dose at Baseline prior to the database freeze by the following steps:

Step 1: The study programming team will identify all patients who had glucocorticoid and/or immunosuppressant increase or a new immunosuppressant added, study day that the change of GC or immunosuppressant falls within, baseline dose of Immunosuppressants, and produce an Excel spreadsheet with all necessary data recorded up until the database freeze.

Step 2: The study Medical Monitor and Safety Physician will review the patients in the provided spreadsheet to determine whether each one will be considered a treatment failure. The criteria will be based on their medical judgement.

Step 3: The study Medical Monitor and Safety Physician will provide a list back to programming indicating which events should be flagged as a treatment failure.

Step 4: The data will be integrated into the ADAM data sets so that treatment failure definitions are defined.

4.1.5.2. Disease Flare/DILI

The definition of Disease Flare can be seen in Section 3.1.6. A patient is only eligible to be deemed as having a disease flare after they have achieved CR or PR with a successful

glucocorticoid taper. Disease Flare assessment will occur in a blinded manner by the Medical Monitor and Safety Physician by the following steps:

Step 1a-Disease Flare: The study programming team will programmatically identify all patients that had an elevation of ALT after achieving [CR] or [PR with a successful glucocorticoid taper] that meets all the following criteria:

- is ≥25% above the [CR] or [PR with a successful glucocorticoid taper] value
- is ≥1.25 ULN
- is sustained for ≥1 week as verified via repeat laboratory assessments
- patient to re-start or escalate glucocorticoid or immunosuppressant therapy

and produce an Excel spreadsheet with all necessary data.

Step 1b-DILI: The study programming team will programmatically identify all patients that meet following criteria for the definition of a DILI:

- >3 × ULN elevation in aspartate aminotransferase (AST) and/or alanine aminotransferase (ALT);
- total bilirubin >2 ×ULN in the absence of findings of cholestasis (ie, absence of alkaline phosphatase elevation to >2 × ULN);

Step 2: The study Medical Monitor and Safety Physician will review the patients in the provided spreadsheet to determine if the patient's flare was not due to other identifiable causes such as viral hepatitis, DILI, concomitant medications, alcoholic liver injury or other reasons not related to AIH based on their medical judgement.

Step 3: The study Medical Monitor and Safety Physician will provide a list back to programming indicating which events should be flagged as a disease flare.

Step 4: The data will be integrated into the ADAM data sets so that disease flares and DILIs are identified.

4.1.5.3. Protocol Deviations

Protocol deviations will be identified prior to Database Freeze for the DBTP analysis. Methods for identifying protocol deviations are described below:

Protocol deviations assessment will occur in a blinded manner by the Medical Monitor, Clinical Operations and Biostatistics prior to the Database Freeze by the following steps:

Step 1: Clinical Operations will provide an Excel spreadsheet listing all patients who had protocol deviations along with all the necessary information prior to the database freeze.

Step 2: Study team members including the Medical Monitor and Biostatistician will review the patients in the provided spreadsheet to determine whether the patient will be included in the Per Protocol Population.

Step 3: The study Biostatistician will provide a list back to programming indicating which patients should be flagged as being removed from the Per Protocol Population.

Step 4: The data will be integrated into the ADAM data sets so that Per Protocol Population can be defined.

4.1.5.4. Opportunistic Infection

Opportunistic infections will be identified prior to Database Freeze for the DBTP analysis. Method for identifying Opportunistic Infections is described in Section 15.2.

4.1.6. Blinding/Unblinding/Database Freeze

Blinding and Unblinding procedures are described separately in the Blinding Management Plan Ver 2.0 for this study.

The site personnel, monitors, study patients, and personnel involved in the project, except for the designated unblinded persons, will remain blinded to the DBTP IMP assignment until the DBTP completion and the DBTP data are "frozen". At this point in time, select team members may become unblinded using the "UNBLINDED DURING OLE PERIOD FORM" found in the Blinding Management Plan. This is a pre-defined unblinding procedure to allow for Kezar Life Sciences, Inc. (Kezar) business decisions while maintaining the blind in operations of the study.

The Analysis of the data for the DBTP (ie, Primary Analysis) will occur after all patients have or would have completed the Week 24 visit or EOS visit prior to the OLE, whichever occurs later, in the DBTP, and all the DBTP data is fully cleaned, signed off by the Principal Investigator, and "frozen" to limit changes.

The final data analysis will be conducted after all patients have completed the study (the DBTP and OLE Period) and the database has been locked.

4.1.7. Rescue Therapy

Rescue therapy is defined as modification to the background AIH treatment, including dose or frequency increases in concomitant medications or initiation of a new medication. Rescue medications may include, but are not limited to systemic glucocorticoids, AZA, MMF, calcineurin inhibitor, and other immunosuppressants. While rescue therapy is typically restricted or prohibited per protocol (Protocol Section 6.2.4), patients may receive rescue therapy during the study at the discretion of the Investigator when deemed medically necessary for the safety of the patient. Examples of situations that might require rescue therapy include:

- Patients who meet the definition for treatment failure or disease flare that requires higher than the maximum protocol-allowed prednisone/prednisone equivalent dose, increase in dose(s) of background immunosuppressant medications, or addition of other immunosuppressant medication(s)
- Patients requiring hospitalization due to signs of acute liver failure
- Patients with evidence of a coagulation abnormality (eg, INR ≥1.5)
- Patients with any degree of mental alteration (encephalopathy) without pre-existing cirrhosis and with an illness of <26 weeks duration.

If an Investigator determines that a patient needs rescue therapy, the Medical Monitor should be informed; where possible, a discussion should occur between the Investigator and Medical

Monitor prior to implementation of the rescue therapy. Any rescue medications administered should be recorded in the eCRF 'Prior and Concomitant Medication/Therapy'. Patients who receive rescue therapy will temporarily discontinue IMP, complete the Double-blind Period visits, and remain eligible for the OLE Period.

5. SAMPLE SIZE

The statistical analyses of the study data will be descriptive in nature, and no formal hypothesis testing will be performed. No formal statistical sample size or power estimation has been performed.

The sample size of approximately 24 patients was set to allow a preliminary assessment of the potential effect of zetomipzomib with a glucocorticoid, and glucocorticoid treatment with placebo, in patients diagnosed with AIH. This is a signal-seeking study designed to collect data and perform preliminary assessments on safety and efficacy endpoints between treatment groups.

6. ANALYSIS POPULATIONS

6.1. Full Analysis Set

The Full Analysis Set (FAS) is defined as all randomized patients who receive at least 1 dose of IMP in this study and have Baseline and at least one post-treatment value for all the following measures: ALT, AST, and IgG. The FAS will be analyzed as randomized.

6.2. Per Protocol Set

The Per Protocol (PP) Set will consist of FAS patients who complete the DBTP and have no protocol violations considered to affect the evaluability of efficacy, as determined through blinded review of deviations prior to unblinding (Section 4.1.5.3). The PP Set will be analyzed as treated.

6.3. Safety Set

The Safety Set will include all patients who received at least one dose of IMP. The Safety Set will be analyzed as treated.

6.4. OLE Safety Set

The OLE Safety Set will include all patients who receive at least one dose of OLE IMP.
The OLE Safety Set will be analyzed as treated.

6.5. Intent to Treat Population

The Intent to Treat (ITT) data set will include all randomized patients. The ITT data set will be analyzed as randomized.

7. GENERAL STATISTICAL CONSIDERATIONS

7.1. Overview

7.1.1. Timing of Analyses

The first planned unblinded analysis of the data (ie, Primary Analysis) will be conducted after all patients have or would have completed the Week 24 visit or EOS visit prior to the OLE, whichever occurs later, in the DBTP, data being fully cleaned and database being "frozen" to limit changes. The final data analysis will be conducted after all patients have completed the study (DBTP and OLE Period). The Table of Contents for the Tables Listings and Figures will be defined in a separate document which will require signatures prior to the first Dry Run Delivery.

7.1.2. Subgroups

The DBTP data will be summarized by treatment group (zetomipzomib, placebo), unless specified otherwise. The OLE Period data will be summarized for all patients and the following groups:

- Group 1: Patients receiving zetomipzomib in the DBTP and continuing with zetomipzomib in the OLE Period
- Group 2: Patients receiving placebo in the DBTP followed by zetomipzomib in the OLE Period

7.1.3. Assessing Normality for Continuous Measures

Normality will be assessed graphically and/or via the Shapiro-Wilk statistic. If substantial departure from normality is observed, transformations, eg, natural log, or non-parametric statistical tests will be explored.

7.2. Missing or Inconclusive Data Handling

7.2.1. Missing Efficacy Assessments

Refer to Section 14.3.1 and Section 14.3.3.

7.2.2. Missing Safety Assessments

There will be imputation of missing data for the following:

1. Partial or missing dates where complete dates are required to flag data as treatment-emergent or concomitant with treatment. Partial/missing start and end dates for AEs and concomitant medications will be imputed as follows:

   Partial/missing start date:

   - Dates with missing day only will be imputed as the 1st of the month unless the month and year are same as the month and year of first dose of IMP, in which case missing day will be imputed as the first dose day of study medication.

- Dates with both day and month missing will be imputed as 1 Jan unless the year is same as the year of first dose of IMP, in which case missing day and month will be imputed as the first dose day and month of IMP.
- Completely missing dates will be imputed as the first dose date of IMP unless the end date is on or before the first dose date of IMP, in which case missing date will be imputed as 1 Jan of the same year as the end date.

Partial/missing end date:

- Dates with missing day only will be imputed as the last day of the month unless the month and year are the same as the month and year of the last dose of IMP, in which case missing day will be imputed as the last dose day of IMP.
- Dates with both day and month missing will be imputed as 31 Dec unless the year is same as the year of the last dose of IMP, in which case missing day and month will be imputed as the last dose day and month of IMP.
- If the ongoing flag is missing or "Yes," then the date will not be imputed unless death date is available, in which case the missing date will be imputed as the death date. If ongoing is "No," then the missing end date will be imputed as the last dose date.
- If the imputed date is after the date of death, then the end date will be set equal to the date of death.

2. For Prior AIH Medications or Medical history CRFs, partial or missing dates where complete dates are required to flag data as prior to treatment or calculate duration of AIH. Partial/missing start and end dates for prior medications and medical history will be imputed as follows:

   Partial/missing start date:

   - Dates with missing day only will be imputed as the 1st of the month.
   - Dates with both day and month missing will be imputed as 1 Jan.
   - Completely missing dates will be imputed as the screening visit date unless the end date is on or before the screening visit date, in which case missing date will be imputed as 1 Jan of the same year as the end date.

   Partial/missing end date:

   - If the ongoing flag is "No", dates with missing day only will be imputed as the last day of the month unless the month and year are the same as the month and year of the screening visit date, in which case missing day will be imputed as the screening visit day.
   - If the ongoing flag is "No", dates with both day and month missing will be imputed as 31 Dec unless the year is same as the year of the screening visit date, in which case missing day and month will be imputed as the screening visit day and month of IMP.

- If the ongoing flag is missing or "Yes," then the date will not be imputed unless death date is available, in which case the missing end date will be imputed as the death date. If ongoing is "No," then the missing end date will be imputed as the screening visit date.
- If the imputed date is after the date of death, then the end date will be set equal to the date of death.

7.2.3. QOL Missing Data

7.2.3.1. EQ-5D-5L Index Value

The index value will be set to missing if any dimension value is missing.

7.2.3.2. CLDQ Domain Scores and Overall Score

If value for one of the items is missing, the domain score will not be calculated and set to missing. If any domain score is missing, the overall score will be set to missing.

7.3. Scheduled/Unscheduled Visits

Once analysis visit windows are assigned, all visits, including scheduled visits, unscheduled visits, ET visit, and EOS visits will be eligible for being flagged as the "analyzed record" within the analysis visit window as described in Table 2 for the DBTP and Table 3 for the OLE Period. A subject's individual analysis visit window could potentially contain more than 1 visit. In the event multiple visits fall within an analysis visit window or in case of a tie, the following rules will be used in sequence to determine the "analyzed record" for the analysis visit window:

- If there is a scheduled visit/week for the analysis visit window, then the scheduled visit/week data will be used.
- If there is no scheduled visit/week for the analysis visit window, the data closest to the scheduled day will be used.
- If there is no scheduled visit/week for the analysis visit window and there is a tie in the number of days before and after the scheduled day, the later data will be used.

The data not flagged as the "analyzed record" will also be listed in subject listings.

Table 2: Analysis Visit Windows for DBTP

Analysis Visit	Analysis Week	Target Study Day	Study Day Interval for Analysis (inclusive)
2	0	Day 1	Prior to first dose of IMP
4	2	Day 15	2, 22
6	4	Day 29	23, 36
8	6	Day 43	37, 50
10	8	Day 57	51, 64
12	10	Day 71	65, 78

Analysis Visit	Analysis Week	Target Study Day	Study Day Interval for Analysis (inclusive)
14	12	Day 85	79, 92
16	14	Day 99	93, 106
18	16	Day 113	107, 120
20	18	Day 127	121, 134
22	20	Day 141	135, 148
24	22	Day 155	149, 162
26	24	Day 169	163, 176

Note: Study Day = Visit date – Date of first dose of IMP in DBTP + 1. Scheduled and unscheduled visits will be presented in patient listings.

Table 3: Analysis Visit Windows for Open Label Extension Period

Analysis Visit	Analysis Week	Target Study Day	OLE Study Day Interval for Analysis (inclusive)
27	None	None	Prior to first dose of OLE IMP
28	OLE Week 1	OLE Day 1	First dose of OLE IMP
31	OLE Week 4	OLE Day 22	2, 36
35	OLE Week 8	OLE Day 50	37, 64
39	OLE Week 12	OLE Day 78	65, 92
43	OLE Week 16	OLE Day 106	93, 120
47	OLE Week 20	OLE Day 134	121, 148
52	OLE Week 25	OLE Day 169	149, 183

Note: Study Day = Visit date – Date of first dose of IMP in OLE + 1.
Scheduled and unscheduled visits will be presented in patient listings.

Because there is only one EOT form for both study periods, the last value evaluated at last visit up to and including week 24 will be displayed instead of EOT form, which will be used for the OLE study period.

7.4. Data Analysis Conventions

Data analyses will be performed by [redacted] for unblinded analysis after the DBTP is completed and the database has been frozen or final analysis after the study is completed and the database has been locked. Analysis datasets will be created using data obtained from electronic data capture and external data if applicable. Statistical programming and analyses will be performed using SAS® Version 9.4 or higher. Outputs will be provided in portable document format for tables, listings, and figures using landscape format.

Study data will be listed by patient, treatment, and visit (as applicable) based on all randomized patients, unless otherwise specified. Listings will be sorted by patient number, visit date, time, and parameter as applicable.

Summaries for continuous and ordinal variables will include the number of observations (n), arithmetic mean, SD, median, Q1 (first quartile), Q3 (third quartile), IQR (Interquartile range), minimum, and maximum. Minima and maxima will be reported with the same precision as the raw values; means and medians will be presented to one additional decimal place than reported in the raw values. Standard deviations will be presented to two additional decimal places than reported in the raw values. Summaries for discrete variables will include frequency counts and percentages. All percentages will be rounded to one decimal place (ie, xx.x%).

All p-values will be rounded to four decimal places; p-values less than 0.0001 will be presented as "<0.0001"; and p-values greater than 0.9999 will be presented as ">0.9999".

Continuous efficacy endpoints (ie, change or percentage change from Baseline in ALT, AST, IgG by visit) will be analyzed using the longitudinal data analysis method based on a mixed model for repeated measures (MMRM) approach.

For time-to-event variables (eg, time to CR, time to [CR] or [PR with a successful glucocorticoid taper], time to flare, duration of CR, duration of [CR] or [PR with a successful glucocorticoid taper], and time to patient experiencing their first SIR) the Kaplan Meier method will be used.

7.5. Adjustments for Multiplicity

Due to the explorative nature of this study, no adjustments will be applied to perform multiple comparisons.

8. DISPOSITION OF PATIENTS

A summary for patient disposition by treatment group and overall for all patients will include:

DBTP:

- Screen failures
- Randomized patients
- Patients in each of the study populations (FAS, PP Set, Safety Set, OLE Safety Set)
- Patients completed treatment in DBTP
- Patients discontinued treatment in DBTP and the primary reason for discontinuation of treatment
- Patients completed DBTP portion of the study
- Patients discontinued from DBTP portion of the study and the primary reason for discontinuation of study

OLE Period:

- Patients completed treatment in OLE Period
- Patients discontinued treatment in the OLE Period and the primary reason for the discontinuation of treatment
- Patients completed the study
- Patients discontinued from the study and the primary reason for the discontinuation of study

Percentages for the patient disposition summary are based on number of randomized patients for DBTP and number of patients in the OLE Safety Set for the OLE Period.

Patient listings will be provided that include disposition, informed consent date and reconsent date (if applicable), inclusion and exclusion criteria, and exclusions from the analysis sets and reason for exclusion from population. Details of the study randomization, including randomization date, randomized and actual treatment, randomization stratification, and reason for unblinding (if applicable) will also be included within a patient listing.

9. PROTOCOL DEVIATIONS

Protocol deviations will be determined prior to unblinding the data.

For the reporting of the protocol deviations in the DBTP, the number and percentage of patients with any deviation, important deviations, not-important deviations, deviations resulting in exclusion from PP Set, will be summarized by treatment group for the FAS. For the OLE Period, the number and percentage of patients with any deviation, important deviations, and not-important deviations will be summarized by treatment group for the OLE Safety set.

A patient listing will be provided for protocol deviations that includes the visit at which the deviation occurred, the deviation date, the deviation code, the deviation description, the classification of the deviation as important or not-important, and deviations resulting excluding from PP Set.

10. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Demographics and Baseline characteristics will include sex, age in years at signing of Informed Consent Form (ICF), age group (<65 years, ≥65 years), race, ethnicity, Baseline height, Baseline weight, and Baseline body mass index, Baseline ALT, Baseline AST, Baseline IgG, Prednisone-Equivalent Dose of Glucocorticoid (GC) at Screening visit (only for Patients Who Took Oral GC at screening). Patients who record more than one race are grouped into the single category denoted as multi-racial. Demographic data will be summarized by treatment group and overall for the FAS, the Safety Set, and the OLE Safety Set.

Patient listings that include all demographic data, history of substance use, and tuberculosis test will be provided.

11. MEDICAL HISTORY

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 27. Medical history will be summarized separately by system organ class (SOC), preferred term (PT), and treatment group for the Safety Set and the OLE Safety Set. If a patient has more than one SOC or PT within a SOC, they will only be summarized once for that SOC or PT.

Medical history data will be presented in a patient listing.

12. PRIOR AND CONCOMITANT MEDICATIONS

Prior and concomitant medications will be coded using World Health Organization Drug Dictionary (WHODrug; Global B3, March 2023). Concomitant medications used in the DBTP will be defined as medications used between the first dose date of IMP in the DBTP inclusive and the first dose of IMP in OLE exclusive. Concomitant medications used in the OLE period will be defined as medications used on/after the first dose of IMP in the OLE Period. Prior medications will be defined as medications used and stopped prior to the first dose date of IMP (ie, medications started before the first dose, **not** including the medications that continue during the treatment). Prior medications for the Safety Set and the OLE Safety Set, concomitant medications used in the DBTP for the Safety Set, and concomitant medications used in OLE for the OLE Safety Set will be summarized separately by the therapeutic drug class (Anatomical Therapeutic Chemical [ATC] 4 classification), preferred name, and treatment group for the Safety Set. If the ATC 4 classification is not provided, the next lowest classification that is provided in the coding dictionary will be used. The preferred name will be defined as the active ingredient; if the active ingredient is not provided or includes more than two ingredients (eg, multivitamins), then the drug name will be presented as the preferred name. In addition, all medications identified as AIH medications will be summarized by their assigned AIH category (Antimalarial, Biologics, Glucocorticoids, Immunosuppressant) and preferred term for the Safety Set, Full analysis Set, and the OLE Safety Set.

A summary of weekly average prednisone/prednisone equivalent dose (mg/day) will be summarized by treatment and week for both the DBTP and the OLE period in the Safety Set, Full Analysis Set, and the OLE Safety Set.

Prior and Concomitant medications, AIH medications, Prednisone-Equivalent Doses, and concomitant procedures will also be presented in patient listings.

13. DOSING COMPLIANCE AND TREATMENT EXPOSURE

13.1. Dosing Compliance

Dosing compliance (% compliance) for the DBTP and OLE period will be assessed by calculating the number of actual doses received and comparing that to the number of expected doses as follows:

$$\text{DBTP Compliance (\%)} = \frac{\text{Number of Actual Doses Received in DBTP} \times 100\%}{\text{Number of Expected Doses in DBTP}}$$

$$\text{OLE Compliance (\%)} = \frac{\text{Number of Actual Doses Received in OLE} \times 100\%}{\text{Number of Expected Doses in OLE}}$$

$$\text{Zetomipzomib Compliance during Study (\%)} = \frac{\text{Number of Actual Doses Received (DBTP+OLE)} \times 100\%}{\text{Number of Expected Doses in (DBTP + OLE)}}$$

The number of actual doses received will be calculated from the number of doses administered recorded in the Study Dosing Log eCRF. The number of expected doses that will be used for calculating compliances will be calculated as:

Number of Expected Doses in DBTP = Ceil [(Date of Last Dose in DBTP - Date of First Dose in DBTP +1) / 7]

Number of Expected Doses in OLE = Ceil [(Date of Last Dose in OLE - Date of First Dose in OLE +1) / 7]

Number of Expected Doses for Zetomipzomib during Study = Ceil [(Date of Last Dose of Zetomipzomib - Date of First Dose of Zetomipzomib +1) / 7]

for all patients, regardless of study completion status.

A categorical dosing compliance variable will also be derived as non-compliant (< 80%) and compliant (>80%).

Treatment compliance (%) for the DBTP using the Safety Set and for the OLE period using the OLE Safety Set will be summarized separately with continuous descriptive statistics for each treatment group. The compliance category defined above will be summarized with discrete summary statistics. Zetomipzomib compliance during the study will also be summarized for all patients in the Full-Analysis Set, Safety Set and the OLE-Safety Set.

13.2. Treatment Exposure

Extent of treatment exposure for completed or discontinued patients will be calculated in weeks using the following:

Extent of Exposure in DBTP (weeks) = (Date of Last Dose in DBTP - Date of First Dose in DBTP + 1) / 7

Extent of Exposure in OLE (weeks) = (Date of Last Dose in OLE - Date of First Dose in OLE + 1) / 7

Extent of Exposure for Zetomipzomib during Study (weeks) = (Date of Last Dose of Zetomipzomib - Date of First Dose of Zetomipzomib + 1) / 7

Extent of treatment exposure for patients who were lost to follow-up will be calculated in weeks using the following:

Extent of Exposure in DBTP (weeks) = (Date of Last Recorded Visit in DBTP - Date of First Dose in DBTP + 1) / 7

Extent of Exposure in OLE (weeks) = (Date of Last Recorded Visit in OLE - Date of First Dose in OLE + 1) / 7

Extent of Exposure for Zetomipzomib during Study (weeks) = (Date of Last Recorded Visit in the Study - Date of First Dose + 1) / 7

Extent of treatment exposure for each patient exposed to IMP for will be summarized with continuous descriptive statistics for each treatment group using the Full Analysis Set, Safety Set for the DBTP data and the OLE Safety Set for OLE data. In addition, total amount of IMP administered (mg) and dose Intensity (mg/week) for the DBTP and OLE will also be summarized.

Dose Intensity (mg/week) = Total Amount Administered / Extend of Exposure for the corresponding the DBTP, the OLE period, and zetomipzomib during study periods separately.

A patient listing of treatment exposure will also be produced which will include Reason Not Done if the treatment was not administered.

14. EFFICACY ANALYSES

14.1. General Considerations

All efficacy endpoints will be descriptively summarized by treatment group for each visit where applicable using observed data, unless specified otherwise. All efficacy data will be presented in patient listings.

No formal statistical sample size or power estimation has been performed. This study was not powered to detect a statistically significant difference between treatment arms for any endpoints. This study was designed as a signal seeking study. Normality will be assessed for all continuous measures using Shapiro Wilk Test and/or graphical assessments and appropriate transformation or non-parametric models may be used if data is not normally distributed.

14.2. Summary of Efficacy Endpoints

Summary of efficacy endpoints and analysis which includes important details and distinctions can be seen in Table 4. All efficacy endpoints will use the observed data for the analyses. If an endpoint is ambiguous because a patient achieved more than one response during a treatment period (DBTP or OLE) within the four responses listed below:

1. [CR]
2. [PR with a successful glucocorticoid taper]
3. Non-Response
4. TF (Treatment Failure)

then the Best Response will be chosen which is displayed above in order with best being number 1 and worst being number 4. The Sensitivity analyses to address missing data, other analyses methods, and different analysis populations are listed in Section 14.3.3.

Table 4: Efficacy Endpoints and Analysis Methods

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
Primary	DBTP	Proportion of patients who achieve CR (Section 3.1.2) by Week 24; analyses will also be performed on data collected at Weeks 12, 16, and 20.	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a CR during DBTP. Denominator: Number of subjects the analysis population.
	OLE	Proportion of patients experiencing a disease flare among the patients who achieved a CR during the DBTP	OLE Safety Set	Descriptive summary	Numerator: Number of subjects that experienced a CR with no flare during DBTP who roll into the OLE study and subsequently experience a disease flare in the OLE treatment period. Denominator: Number of subjects that experienced a CR with no flare up during DBTP who enroll into the OLE study
Secondary	DBTP	Change from Baseline in ALT (or log-transformed ALT if data is non-normal) at Weeks 12, 16, 20, and 24	FAS, ITT	MMRM	NA
	DBTP	Proportion of patients who achieve PR (Section 3.1.3) at Week 12, 16, 20, and 24	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a PR during DBTP (and did not reach criteria for CR at any point prior to Week 24) Denominator: Number of subjects the analysis population.
	DBTP	Time to CR	FAS, ITT	Kaplan-Meier	Duration from first dose of IMP to first CR during DBTP. Patients are censored at last non missing lab assessment up to and including Week 24.
	DBTP	Time to [CR] or [PR with a successful glucocorticoid taper]	FAS, ITT	Kaplan-Meier	Duration from first dose of IMP to first [CR] or [PR with a successful glucocorticoid taper] during DBTP. Patients are censored at last non missing lab assessment up to and including Week 24.

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
	DBTP	Proportion of patients experiencing a disease flare after CR	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who experience a disease flare during DBTP after their first CR. Denominator: Number of subjects in the analysis population that achieved CR during DBTP
	DBTP	Proportion of patients experiencing a disease flare after [CR] or [PR with a successful glucocorticoid taper]	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who experience a disease flare during DBTP after their first [CR] or [PR with a successful glucocorticoid taper] Denominator: Number of subjects in the analysis population that achieved [CR] or [PR with a successful glucocorticoid taper] during DBTP.
	DBTP	Proportion of patients who achieve PR with successful GC taper by Week 24; analyses will also be performed on data collected by Weeks 12, 16, and 20.	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a PR with a successful glucocorticoid taper during DBTP (and did not reach criteria for CR at any point prior to Week 24) Denominator: Number of subjects the analysis population.
	DBTP	Duration of CR during the DBTP	FAS, ITT	Kaplan-Meier	The start of the duration of CR begins when the patient achieves a CR and ends when the patient has a disease flare. Patients are censored at their last non missing lab assessment up to and including Week 24.
	DBTP	Duration of [CR] or [PR with a successful glucocorticoid taper] during the DBTP	FAS, ITT	Kaplan-Meier	The start of the duration of [CR] or [PR with a successful glucocorticoid taper] begins when the patient achieves a [CR] or [PR with a successful glucocorticoid taper]. The end of the duration of [CR] or [PR with a successful glucocorticoid taper] is either disease flare or Week 24 if the patient does not experience a disease flare before Week 24. Patients are

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
					censored at their last non missing lab assessment up to and including Week 24.
	DBTP	Proportion of patients who are treatment failures	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a treatment failure during DBTP Numerator: Number of subjects the analysis population.
	DBTP	Proportion of patients who achieve CR with successful glucocorticoid taper by Week 24; analyses will also be performed on data collected by Weeks 12, 16, and 20.	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a CR and successful glucocorticoid taper by Week 24 Denominator: Number of subjects the analysis population.
Secondary	DBTP	Proportion of patients who achieve CR with glucocorticoid taper to ≤ 5mg by Week 24	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a CR with glucocorticoid taper to ≤ 5mg by Week 24 Denominator: Number of subjects the analysis population.
	DBTP	Proportion of patients who achieve CR with glucocorticoid taper to 0mg by Week 24	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a CR with glucocorticoid taper to 0 mg by Week 24 Denominator: Number of subjects the analysis population.
	DBTP	Proportion of patients who achieve PR with glucocorticoid taper to ≤ 5mg by Week 24	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a PR (and did not reach criteria for CR with glucocorticoid taper to ≤ 5mg at any point prior to Week 24) with glucocorticoid taper to ≤ 5mg by Week 24 Denominator: Number of subjects the analysis population.

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
	DBTP	Proportion of patients who achieve PR with glucocorticoid taper to 0mg by Week 24	FAS, ITT	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population who have a PR (and did not reach criteria for CR with glucocorticoid taper to 0mg at any point prior to Week 24) with glucocorticoid taper to 0 mg by Week 24 Denominator: Number of subjects the analysis population.
Exploratory	DBTP	Proportion of patients who achieve ALT normalization, based on Prati criteria (males, 30 U/L; females, 19 U/L) (), with successful glucocorticoid taper, by Week 24	FAS	CMH or Fisher's exact	Numerator: Number of subjects in the analysis population with ALT normalization base on Prati criteria during DBTP Denominator: Number of subjects the analysis population.
Exploratory	DBTP	Change from Baseline in AST (or log-transformed AST if data is non-normal) at Weeks 12, 16, 20, and 24	FAS	MMRM	NA
	DBTP	Change from Baseline in IgG at Weeks 12, 16, 20, and 24	FAS	MMRM	NA
	DBTP	Change from Baseline in glucocorticoid dose at Weeks 12, 16, 20, and 24	FAS	ANCOVA	NA
	DBTP	Proportion of patients who achieve CR that had a baseline ALT >2x ULN	FAS	CMH or Fisher's exact	NA
	DBTP	GTI-CWS and GTI-AIS at Week 24	FAS	ANCOVA	NA
	DBTP	Change from Baseline in liver stiffness Measurement (LSM) and CAP score at Weeks 16 and 24	FAS	MMRM	NA
	DBTP	Descriptive Summary Liver Fibrosis stages shown in Table 6 at timepoints that LSM data was collected during DBTP	FAS	Descriptive summary	NA

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
	DBTP	Change from Baseline in and EQ-VAS and EQ-5D-5L index value at Weeks 16 and 24	FAS	ANCOVA	NA
	DBTP	Change from Baseline in CLDQ domain and overall scores at Weeks 16 and 24	FAS	ANCOVA	NA
	DBTP	Change from Baseline in liver histopathology at Week 24, based on Ishak score	FAS	ANCOVA	NA
	DBTP	Descriptive summary of liver histopathology characteristics at Baseline ~~a~~nd Week 24	FAS	Descriptive summary	NA
	OLE	Proportion of patients who have a CR during the OLE period of the study among the patients who did not achieve a CR during the DBTP	FAS	Descriptive summary	Numerator: Number of subjects who have a CR during OLE Denominator: Number of subjects the analysis population who did not achieve CR during DBTP.
	OLE	Proportion of patients who have a PR with a successful GC taper during the OLE period of the study among the patients who did not achieve a PR with a successful GC taper during the DBTP	FAS	Descriptive summary	Numerator: Number of subjects who have a PR with a successful GC taper (and did not reach criteria for CR at any point prior to OLE Week 25) during OLE Denominator: Number of subjects the analysis population who did not achieve PR with a successful GC taper during DBTP.
	OLE	Proportion of patients who have a CR during the OLE period of the study with a successful glucocorticoid taper who did not achieve a CR during the DBTP	FAS	Descriptive summary	Numerator: Number of subjects in the analysis population who have a CR with a successful glucocorticoid taper during the OLE period of the study who did not achieve a CR with a successful glucocorticoid taper during the DBTP Denominator: Number of subjects the analysis population.

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
	OLE	Proportion of patients who achieve CR with glucocorticoid taper to ≤5mg by OLE Week 25	FAS	Descriptive summary	Numerator: Number of subjects in the analysis population who achieve CR with glucocorticoid taper to ≤5mg by OLE Week 25 Denominator: Number of subjects the analysis population.
	OLE	Proportion of patients who achieve CR with glucocorticoid-taper to 0mg by OLE Week 25	FAS	Descriptive summary	Numerator: Number of subjects in the analysis population who achieve CR with glucocorticoid-taper to 0mg by OLE Week 25 Denominator: Number of subjects the analysis population.
	OLE	Proportion of patients who achieve PR with glucocorticoid taper to ≤5mg by OLE Week 25	FAS	Descriptive summary	Numerator: Number of subjects in the analysis population who achieve PR with glucocorticoid-taper to ≤5mg (and did not reach criteria for CR with glucocorticoid taper to ≤ 5mg at any point prior to OLE Week 25) by OLE Week 25 Denominator: Number of subjects the analysis population.
	OLE	Proportion of patients who achieve PR with glucocorticoid taper to 0mg by OLE Week 25	FAS	Descriptive summary	Numerator: Number of subjects in the analysis population who achieve PR with glucocorticoid-taper to 0mg (and did not reach criteria for CR with glucocorticoid taper to = 0mg at any point prior to OLE Week 25) by OLE Week 25 Denominator: Number of subjects the analysis population.
	OLE	Change from Baseline and Change from OLE-Baseline in ALT at OLE Week 25	FAS	Descriptive summary	NA

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
	OLE	Duration of CRs during the entire study (DBTP +OLE)	FAS	Kaplan-Meier	The start of the duration of CR begins when the patient achieves a CR. The end of the duration of CR is disease flare. Patients are censored at their last nonmissing lab assessment up to and including OLE Week 25.
	OLE	Duration of [CR] or [PR with a successful glucocorticoid taper] during the entire study (DBTP +OLE)	FAS	Kaplan-Meier	The start of the [CR] or [PR with a successful glucocorticoid taper] begins when the patient achieves a [CR] or [PR with a successful glucocorticoid taper]. The end of the duration of [CR] or [PR with a successful glucocorticoid taper] is disease flare. Patients are censored at their last nonmissing lab assessment up to and including OLE Week 25.
	OLE	Change from Baseline and Change from OLE-Baseline in AST at OLE Week 25	FAS	Descriptive summary	NA
	OLE	Change from Baseline and Change from OLE-Baseline in IgG at OLE Week 25	FAS	Descriptive summary	NA
	OLE	Change from Baseline and Change for OLE-Baseline in the EQ-5D-5L VAS and Index value at OLE Week 25	FAS	Descriptive summary	NA
	OLE	Change from Baseline and Change from OLE-Baseline in CLDC domain score and overall score at OLE Week 25	FAS	Descriptive summary	NA
	OLE	GTI-CWS and the GTI-AIS at OLE-Week25	FAS	Descriptive summary	NA
	OLE	Change from Baseline and change from OLE-Baseline in liver stiffness Measurement (LSM) and CAP score at OLE Week 25	FAS	MMRM	NA

Category	Period	Endpoint	Analysis Population(s)	Analysis Method	Definition (Additional Details)
	OLE	Descriptive Summary Liver Fibrosis stages shown in Table 6 at timepoints that LSM data was collected during the study	FAS	Descriptive Summary	NA
	OLE	Change from Baseline and change from OLE-Baseline in Ishak score	FAS	MMRM	NA

Notes: The terminology "during DBTP" used in any of the endpoints in Table 4 specifically means "the occurrence at any analysis visit up to and including Week 24." The terminology "OLE treatment period" used in any of the endpoints in Table 4 specifically means "from the first dose in OLE to OLE Week 25".
If a patient achieves more than one response during a treatment period within the four listed below:

1. [CR]
2. [PR with a successful glucocorticoid taper]
3. Non-Response
4. TF (Treatment Failure)

then the Best Response will be chosen which is displayed above in order.
Abbreviations: ALT = Alanine aminotransferase; AST = Aspartate aminotransferase; ANCOVA = Analysis of variance with fixed effect terms for randomized treatment and the randomization stratification, Baseline will be included as a covariate; CAP = controlled attenuation parameter; CLDQ = Chronic Liver Disease Questionnaire; CMH = Cochran-Mantel-Haenszel CMH test stratified by the randomization stratification factor (ie, glucocorticoid use at Screening, no glucocorticoid use at Screening); CR = complete biochemical response; DBTP = Double-Blind Treatment Period; EQ-5D-5L EuroQol 5-dimension 5-level; EQ-VAS = EuroQol Visual Analog Scale ;GTI= Glucocorticoid Toxicity Index; GTI-AIS Aggregate Improvement Score; GTI-CWS = GTI Cumulative Worsening Score; IgG = immunoglobulin G; LSM = Liver Stiffness Measurement; MMRM = Mixed model for repeated measures with fixed effect terms for randomized treatment, randomization stratification, timepoint (ie, analysis week) and randomized treatment by timepoint interaction, within patient error will be modelled by use of an unstructured variance-covariance matrix, Baseline value will be included as a covariate, and the Kenward-Roger method is used to adjust degrees of freedom and the test statistics; OLE = Open-Label Extension; PR = partial biochemical response

14.3. Analysis of Primary Endpoint

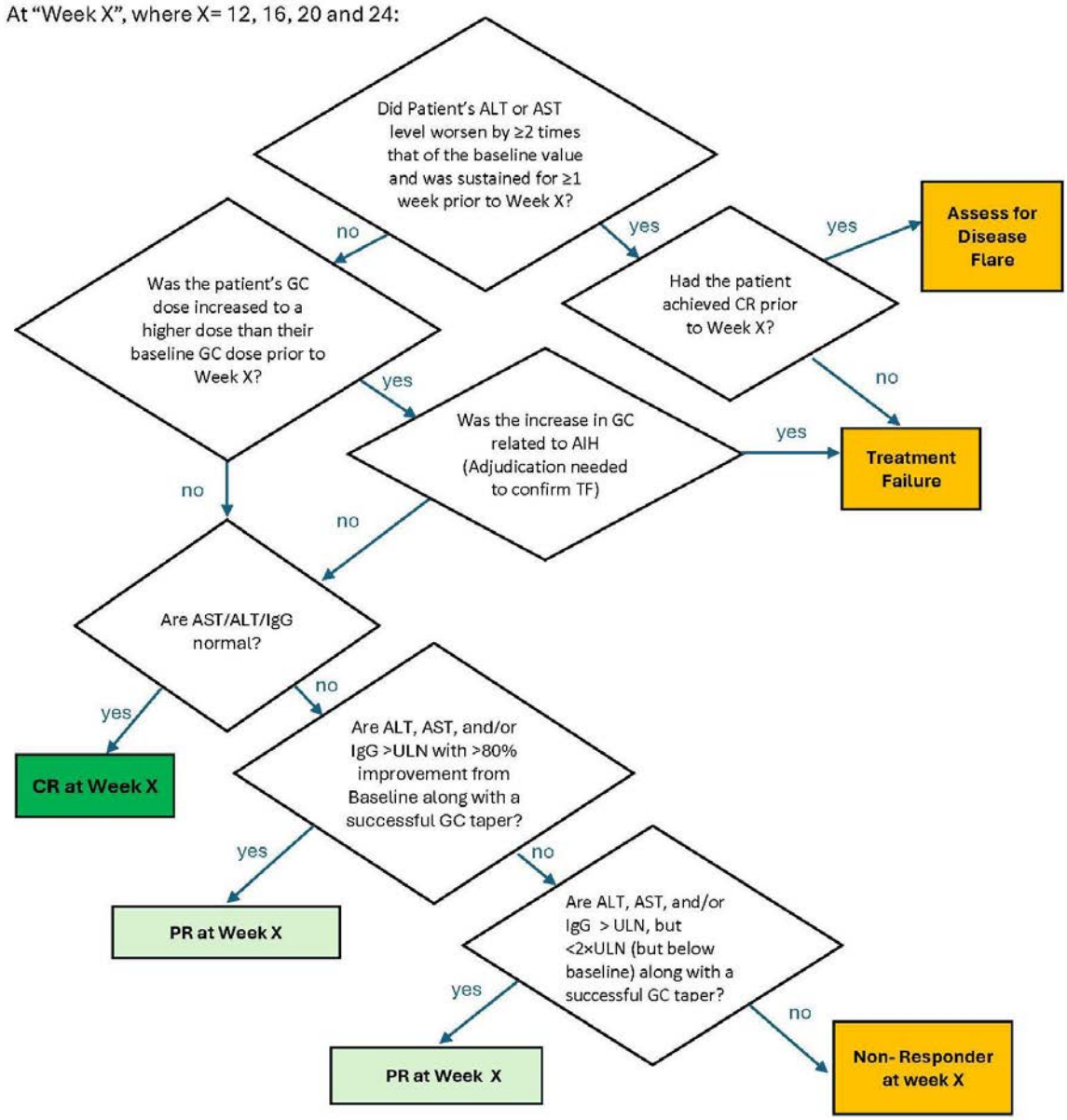
14.3.1. DBTP

The primary endpoint for the primary analysis is the proportion of patients who achieve CR (Section 3.1.2) by Week 24; analyses will also be performed on data collected by Weeks 12, 16, and 20.

The primary efficacy endpoint will be analyzed using the Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factor (ie, glucocorticoid use at Screening, no glucocorticoid use at Screening) or the Fisher exact test if there are too few subjects in either randomization stratum. Patients who have insufficient data for response determination for a time point will be considered as non-responders for that time point (primary analysis method).
A between-treatment difference value in an efficacy endpoint will be summarized using a point estimate and a two-sided 95% exact confidence interval (CI). Statistical uncertainty regarding the efficacy outcome estimate will be presented in a two-sided 95% Clopper-Pearson exact CI.
The primary efficacy analysis will be based on the FAS.

Figure 2 shows the flowchart for determining CR, PR, non-responder, or treatment failure.
It should be noted that if a patient's glucocorticoid dose is increased to a dose higher than their Baseline dose prior to the evaluable Week, then that patient may be considered a treatment failure. Treatment failures for cases like these will be determined in a blinded manner before Database Freeze occurs. Patients who have a treatment failure are not eligible for PR or CR during the treatment period (DBTP or OLE) that they experienced the treatment failure.

Figure 2: Flowchart Showing Paths that a Patient Can Take to Achieve CR, PR, Non-Responder, or Treatment Failure at Week 12, 16, 20, or 24



AIH = Autoimmune Hepatitis; ALT = Alanine aminotransferase; AST = Aspartate aminotransferase; CR = complete biochemical response; GC = Glucocorticoid; IgG = immunoglobulin G; PR = partial biochemical response; TF = Treatment Failure; ULN = Upper Limit of Normal

14.3.2. OLE Period

The primary efficacy endpoint is proportion of patients experiencing a disease flare among the patients who achieved a CR during the DBTP. Patients who have insufficient data for response determination will be considered non-responders. The endpoint will be summarized using frequency and percentages of patients based on the OLE Safety Set. Two-sided 95% Clopper-Pearson exact CIs for percentages will be provided.

14.3.3. Sensitivity Analyses

Sensitivity Analyses will be performed for the **DBTP data only.**

To evaluate the robustness of the primary analysis results, Table 5, describes the sensitivity analyses that will be performed.

Table 5: Sensitivity Analyses for Primary Analysis

Sensitivity Analysis #	Missing Values Imputation	Data Type and Analysis Populations(s)	Analysis Method
1	The same as the primary analysis (ie, primary analysis method)	Observed, PP	The same as the primary analysis
2	**Worst Observation Carried Forward**: see Section 14.3.3.1 for details	Imputed missing data, FAS	The same as the primary analysis
3	**Tipping Point Approach:** Multiple imputation, see Section 14.3.3.2 for details.	Imputed missing data, FAS	The same as the primary analysis for each imputation
4	**Worst Observation Carried Forward:** see Section 14.3.3.1 for details. See Section 14.4.6 for additional missing data details.	Imputed missing data, ITT	The same as the primary analysis

ALT= Alanine Aminotransferase; AST= Aspartate Aminotransferase: CR=Complete Biochemical Response, FAS=Full Analysis Set; IgG= Immunoglobulin G: PP=Per Protocol

14.3.3.1. Worst Observation Carried Forward

For determination of CR, missing post-Baseline value of ALT, AST, and IgG will be imputed using the worst value from previous post-Baseline timepoints. The worst value prior to the first dose of IMP will be used if all post-Baseline values from previous timepoints are missing.

14.3.3.2. Tipping Point Analysis

Tipping Point Analysis will only be done if the Primary Analysis is significant. Primary endpoint (CR) values will be derived based on the observed data (described in Table 5, Sensitivity Analysis #3). Missing CR values will be imputed using tipping point approach. First set all records in zetomipzomib with missing value as failures and all records in placebo with missing value as successes, that is the most stringent test case. The imputed data will be analyzed using the same analysis method as the primary analysis. If the p-value is less than or equal to 0.05 for the most stringent test case, then stop there. If the p-value is greater than 0.05 for the most stringent case, then set the missing records as failures or successes in all possible combinations from most stringent case to the least stringent case until the p-value turns from >0.05 to ≤0.05. If the p-value does not turn from >0.05 to ≤0.05, then there is no potential impact of missing data on the reliability of efficacy results. The inclusion of strata only occurs in most stringent case.

14.4. Analyses of Secondary Efficacy Endpoints

Normality will be assessed for continuous measures using Shapiro Wilk Test and/or graphical assessments and appropriate transformation or non-parametric models may be used if data is not normal.

14.4.1. Proportion of Patients Who Achieve PR, Disease Flare, or Treatment Failure

The primary analysis of the following secondary endpoints in the DBTP will be performed in the same manner that of primary efficacy endpoint.

- Proportion of patients who achieve PR with a successful GC taper at Weeks 12, 16, 20, and 24
- Proportion of patients experiencing a disease flare after CR among patients that achieved CR
- Proportion of patients experiencing a disease flare after PR with a successful GC taper among patients that achieved CR
- Proportion of patients who are treatment failures
- Proportion of patients who achieve CR with successful glucocorticoid taper by Week 24; analyses will also be performed on data collected by Weeks 12, 16, and 20
- Proportion of patients who achieve CR with glucocorticoid taper to ≤ 5mg by Week 24
- Proportion of patients who achieve CR with glucocorticoid taper to 0mg by Week 24
- Proportion of patients who achieve PR with glucocorticoid taper to ≤ 5mg by Week 24
- Proportion of patients who achieve PR with glucocorticoid taper to 0mg by Week 24

14.4.2. Change from Baseline in ALT, AST, IgG, Ishak, LSM, CAP scores

Mean changes from Baseline in ALT (or log-transformed ALT if data is non-normal) at Weeks 12, 16, 20, and 24 will be analyzed using the longitudinal data analysis method based on a MMRM approach. If data is not normally distributed, then the appropriate transformation of data may be applied. This model will include fixed effect terms for randomized treatment, randomization stratification (if strata occur in the data), timepoint (ie, analysis week), randomized treatment by timepoint interaction, and Baseline value as a covariate. Within patient error will be modelled by use of an unstructured variance-covariance matrix. Least-square (LS) means, LS mean differences between treatment and associate two-sided 95% CI, and p-values for treatment difference at each timepoint will be provided. Observed data in the FAS will be used for the primary analysis.

14.4.3. Time to [CR] or [PR with a glucocorticoid taper]

Time to CR (weeks) is duration from first dose of IMP to first CR calculated as: (Date of First CR – Date of First Dose of IMP + 1) / 7. Patients who have insufficient data for response determination for a time point will be considered non-responders for that time point. Time to CR

is censored at the date of last lab assessment before or on week 24. The median time to CR will be estimated using Kaplan-Meier methods and summarized. Kaplan-Meier curves by treatment will be provided. The analysis will be based on the FAS.

Time to [CR] or [PR with a successful glucocorticoid taper] (weeks) is duration from first dose of IMP to first [CR] or [PR with a successful glucocorticoid taper] calculated as: (Date of First [CR] or [PR with a successful glucocorticoid taper] – Date of First Dose of IMP + 1) / 7. Patients who have insufficient data for response determination for a time point will be considered non-responders for that time point. Time to [CR] or [PR with a successful glucocorticoid taper] is censored at the date of last lab assessment before or on week 24. The median time to PR with a glucocorticoid taper will be estimated using Kaplan-Meier methods and summarized. Kaplan-Meier curves by treatment will be provided. The analysis will be based on the FAS.

14.4.4. Duration of [CR] or [PR with successful glucocorticoid taper]

For the two measures:

1. Duration of [CR] during the DBTP
2. Duration of [CR] or [PR with successful glucocorticoid taper] during the DBTP

Duration of [CR] or [PR with successful glucocorticoid taper] is from the first time point that [CR] or [PR with successful glucocorticoid taper] is identified until the patient has a Disease Flare (Section 3.1.6) and is calculated as:

> Duration of CR during the DBTP = (Date of First Disease Flare on or before Week 24 – Date of First CR on or before Week 24 + 1)
>
> Duration of [CR] or [PR with successful glucocorticoid taper] = (Date of First Disease Flare on or before Week 24 – Date of the [CR] or [PR with successful glucocorticoid taper] on or before Week 24 + 1)

Duration during DBTP is censored at the date of last lab assessment before or on week 24.

For the two measures:

1. Duration of CR during the study (DBTP + OLE)
2. Duration of [CR] or [PR with a successful glucocorticoid taper] during the study (DBTP + OLE)

Duration of is from the first time point that [CR] or [PR with successful glucocorticoid taper] is identified in either period, DBTP or OLE until the patient has a Disease Flare (Section 3.1.6).

Duration for these measures is censored at the date of last lab assessment before or on OLE Week 25 (if the patient did not roll over to OLE, they are censored at the date of last lab assessment before or on week 24).

If a subject has [CR] or [PR with a successful glucocorticoid taper] at Week 24 visit but their Week 24 visit is same day as first OLE treatment, this [CR] or [PR with a successful glucocorticoid taper] would be counted in DBTP. The start of their duration of [CR] or [PR with a successful

glucocorticoid taper] would be that same date (Week 24 same day as first OLE) and would be measured in these measures:

1. Duration of CR during the entire study (DBTP +OLE)

2. Duration of [CR] or [PR with a successful glucocorticoid taper] during the entire study (DBTP +OLE)

14.4.5. Time to patient experiencing first Systemic Injection Reaction

Time to patient experiencing first SIR (days) is duration from first dose of IMP to first SIR calculated as: (Date of First SIR – Date of First Dose of IMP + 1). Time to SIR is censored at the date of last dose of IMP before or on week 24 + 2 days. Time to patient experiencing first SIR (days) in OLE is duration from first dose of OLE IMP to first SIR calculated as: (Date of First SIR – Date of First Dose of OLE IMP + 1). Time to SIR is censored at the date of last dose of OLE IMP + 2 days. The median time to SIR will be estimated using Kaplan-Meier methods and summarized. Kaplan-Meier curves by treatment will be provided. The analysis will be based on the FAS.

14.4.6. Sensitivity Analysis of Secondary Endpoints

To evaluate the robustness of the secondary analysis results, the sensitivity analyses will be performed on the population including all patients randomized in the study (ITT population), with missing data accounted for as follows:

1. If a patient is missing any dose of IMP, any combination of doses, or all doses of IMP, the patient will be analyzed in the arm that they were randomly assigned IMP throughout the Double-blind Treatment Period (DBTP) of the study.
2. If a patient is missing Baseline values for ALT, AST, and IgG data, imputation is addressed in Section 3.1.7 of the SAP.
3. If a patient is missing post-baseline values for ALT, AST, and IgG, then missing data will use the Worst Observation Carried Forward methodology as described in Section 14.3.3.1 of the SAP for imputation of the missing values.
4. If a patient was randomized, never dosed with IMP, and had no ALT, AST, and IgG measurements, the patient will be considered a non-responder for CR or PR and will be included in the analysis group for the additional ITT sensitivity analyses for the secondary endpoint.

14.5. Analyses of Exploratory Efficacy Endpoints

Normality will be assessed for continuous measures using Shapiro Wilk Test and/or graphical assessments and appropriate transformation or non-parametric models may be used if data is not normal.

14.5.1. DBTP Exploratory Efficacy Endpoints

The exploratory efficacy endpoints as described below will be analyzed using models including the following: MMRM model, ANCOVA model, t-test, and Wilcoxon rank sum test. The MMRM model will include fixed effect terms for randomized treatment, randomization

stratification, timepoint (ie, analysis week), randomized treatment by timepoint interaction, and Baseline value as a covariate. Within patient error will be modelled by use of an unstructured variance-covariance matrix. ANCOVA model will include fixed effect terms for randomized treatment and randomization stratification, Baseline value will be included as a covariate. Least-square (LS) means, LS mean differences between treatment and associate two-sided 95% CI, and p-values for treatment difference at each timepoint will be provided. The analyses will be performed using observed data based on the FAS.

- Change from Baseline in AST at Weeks 12, 16, 20, and 24
- Change from Baseline in IgG at Weeks 12, 16, 20, and 24
- Change from Baseline in liver stiffness (LSM) and controlled attenuation parameter (CAP) Score at Weeks 16 and 24, assessed by VCTE utilizing Fibroscan®

The exploratory efficacy endpoints, as below, will be analyzed using ANCOVA model, t-test, and Wilcoxon rank sum test. The ANCOVA model will include fixed effect terms for randomized treatment and randomization stratification, Baseline value will be included as a covariate. Least-square (LS) means, LS mean differences between treatment and associate two-sided 95% CI, and p-values for treatment difference at each timepoint will be provided. The analyses will be performed using observed data based on the FAS.

- Change from Baseline in glucocorticoid dose at Weeks 12, 16, 20, and 24
- GTI-CWS and GTI-AIS at Week 24
- Change from Baseline in the EQ-VAS and EQ-5D-5L index value at Weeks 16 and 24
- Change from Baseline in CLDQ domain and overall scores at Weeks 16 and 24
- Change from Baseline in Ishak score at Week 24 (modified Histological Activity Index Total, ranging from 0 to 18)

The exploratory efficacy endpoint below will be analyzed using CMH test stratified by the randomization stratification factor (ie, glucocorticoid use at Screening, no glucocorticoid use at Screening) or the Fisher exact test if there are too few subjects in either randomization stratum. Patients who have insufficient data for response determination for a time point will be considered non-responders for that time point. A between-treatment difference value in an efficacy endpoint (primary or secondary) will be summarized using a point estimate and a two-sided 95% exact CI. Statistical uncertainty regarding the efficacy outcome estimate will be presented in a two-sided 95% Clopper-Pearson exact CI.

- Proportion of patients who achieve ALT normalization based on Prati criteria (males, 30 U/L; females, 19 U/L), with successful glucocorticoid taper, by Week 24

Liver Fibrosis stages will be summarized descriptively at Weeks 0, 16, and 24 using number of subjects and percentages, by treatment arm based on the LSM scores shown in Table 6.

Table 6: Stages of Liver Fibrosis Based on FibroScan Liver Stiffness Measurement (LSM)

Stage of Liver Fibrosis[a]	LSM Range
Stage 1	> 5.0 to 6.3 kPa
Stage 2	> 6.3 to 8.2 kPa
Stage 3	> 8.2 to 12.7 kPa
Stage 4	> 12.7 kPa

[a] Source: (Guo et al., 2017)

Characteristic liver histopathology will be summarized descriptively for Weeks 0 and 24 identifying the number of subjects and percentages, by visit, characteristics and treatment arm. The histopathology characteristics include a yes/no status for: Interface hepatitis, Lymphoplasmacytic infiltrate, Rosette formation (acinar transformation of hepatocytes), Parenchymal collapse, Steatosis/steatohepatitis, Portal lymphoid aggregates, Germinal center formation in lymphoid aggregates, Bile duct injury, Fibrosis/cirrhosis, Centrilobular inflammation/necrosis, Other.

14.5.2. OLE Exploratory Efficacy Descriptive Endpoints

The exploratory efficacy endpoints listed below will be summarized descriptively using observed data based on the FAS set. Where appropriate, change from Baseline and change from OLE-Baseline measures will be summarized for Group 1, Group 2, and all patients.

- Proportion of patients who have CR during the OLE period of the study among the patients who did not achieve a CR during the DBTP
- Proportion of patients who have a CR during the OLE period of the study with successful glucocorticoid taper who did not achieve a CR during the DBTP
- Proportion of patients who achieve CR with glucocorticoid taper to ≤ 5mg by OLE Week 25
- Proportion of patients who achieve CR with glucocorticoid taper to 0mg by OLE Week 25
- Proportion of patients who achieve PR with glucocorticoid taper to ≤ 5mg by OLE Week 25
- Proportion of patients who achieve PR with glucocorticoid taper to 0mg by OLE Week 25
- Duration of CR during the entire study (DBTP + OLE)
- Duration of [CR] or [PR with a successful glucocorticoid taper] during the entire study (DBTP + OLE)
- Change from Baseline and change from OLE-Baseline in ALT
- Change from Baseline and change from OLE-Baseline in AST

- Change from Baseline and change from OLE-Baseline in IgG
- Change from Baseline and change from OLE-Baseline in the EQ-5D-5L Index Value and Visual Analog Scale (EQ-VAS)
- Change from Baseline and change from OLE-Baseline in CLDQ domain and overall scores
- GTI-CWS and the GTI AIS at OLE Week 25.
- Descriptive Summary Liver Fibrosis stages shown in Table 6 at timepoints that LSM data was collected during study

14.5.3. OLE Exploratory Efficacy Endpoints

The exploratory efficacy endpoints as described below will be analyzed using an MMRM model. The MMRM model will include fixed effect terms for randomized treatment, randomization stratification, timepoint (ie, analysis week), randomized treatment by timepoint interaction, and Baseline value as a covariate. The analyses will be performed using observed data based on the FAS.

- Change from Baseline and change from OLE-Baseline in liver stiffness and controlled attenuation parameter score, assessed by VCTE utilizing Fibroscan® at OLE Week 25
- Change from Baseline and OLE-Baseline in Ishak score at OLE Week 25 (modified Histological Activity Index [HAI]) scores (Ishak et al., 1995)

15. SAFETY ANALYSIS

15.1. Adverse Events

AEs will be coded using the MedDRA Version 27 or later.

In the DBTP, for the subjects who rolled over into the OLE period, Treatment-Emergent Adverse Events (TEAEs) are defined as AEs that occurred from dates on or after the first DBTP dose of IMP administration and continuing for 30 days following the last dose received in DBTP or the day before administration of the first OLE dose of IMP (in the OLE period), whichever comes first.

In the DBTP, for the subjects who did not continue into the OLE period, TEAEs are defined as AEs that occurred from dates on or after the first DBTP dose of IMP administration and continuing for 30 days following the last dose received in the DBTP.

In the OLE period, TEAEs are defined as AEs that start on or after the first OLE dose of IMP administration but before 30 days following the last OLE dose received.

An overall summary by treatment group for TEAEs in the DBTP based on the Safety Set will be presented that includes the number and percentage of patients who experienced at least one TEAE, treatment-emergent serious adverse event (TESAE), treatment-emergent serious adverse reaction (TESAR), Adverse Events of Special Interest (AESI), TEAE leading to dose reduction,

TEAE leading to early treatment discontinuation, TEAEs leading to dose interruption, TEAE leading to death, TEAE related to IMP, and Grade 3 or 4 TEAEs.

Incidences and percentages of TEAEs in the DBTP will be summarized based on the Safety Set by SOC and PT for the categories below, unless specified otherwise. The number of events will also be summarized for Injection Site Reaction (ISR).

- TEAEs
- TEAEs related to IMP
- TESAEs
- TESAEs related to IMP
- TEAEs leading to early treatment discontinuation
- TEAE related to IMP leading to early treatment discontinuation
- TEAE leading to study withdrawal
- TEAE related to IMP leading to study withdrawal
- TEAEs leading to dose reduction
- TEAEs related to IMP leading to dose reduction
- TEAEs leading to death by category (Death On-Treatment, Death Post-Treatment, and All Death)
- TEAE related to IMP leading to death
- ISRs (also summarized based on the Safety Set by High-Level Term and Preferred Term)
- TEAEs by maximum severity
- Grade 3 or 4 TEAEs
- Grade 3 or 4 related to IMP TEAEs
- All TEAEs by PT in descending order of frequency
- Grade 3 or 4 TEAEs by PT in descending order of frequency
- AESIs
- AESI by the categories: systemic injection reaction (SIR) and thrombotic microangiopathy (preferred term) which also includes thrombotic thrombocytopenic purpura (preferred term) and hemolytic uremic syndrome (preferred term)
- AESIs related to IMP
- AESIs leading to dose reduction
- AESIs leading to early treatment discontinuation
- Grade 3 or 4 AESIs
- AESIs by maximum severity

- TEAEs by category
  - QTc prolongation and cardiac toxicities
  - liver toxicity
  - nephrotoxicity
  - abnormal immunoglobulin levels
  - abnormal vital signs
  - infection-related TEAEs
  - opportunistic infections
  - bone marrow toxicity
  - TEAEs having onset within 24 hours of the previous dose of study drug

Similar to summaries of TEAEs in the DBTP, summaries of TEAEs in the OLE period will also be performed based on the OLE Safety Set.

Similar to summaries of TEAEs in the DBTP, incidence and percentage of TEAEs with onset during zetomipzomib exposure in the entire study (the DBTP and the OLE Period) will be summarized for all patients treated with zetomipzomib.

AEs, serious adverse events (SAEs), ISRs, SIRs, AESIs, AEs leading to IMP withdrawal, AEs leading to dose reduction, and AEs leading to death will be presented separately in patient listings.

Summary measure(s) of TEAEs (eg, cumulative incidence rate, exposure-adjusted incidence rate) will be provided.

15.2. TEAE by category

Each TEAE category will be summarized in a stand-alone table for the Primary Analysis and the Final Data Analysis. The table will summarize all TEAEs identified as belonging to each category using the following methods:

Step 1: The study programming team will pull out all TEAEs and produce an Excel spreadsheet with all TEAEs recorded up until the data cut.

Step 2: The study Medical Monitor and Safety Physician will review the TEAEs in the provided spreadsheet to determine whether each one will be considered as belonging to a TEAE category. The criteria will be based on their medical judgement.

Step 3: The study Medical Monitor and Safety Physician will provide a list back to programming indicating which events should be flagged for each TEAE category.

Step 4: The data will be integrated into the ADAM data sets so that tables and listings can be created.

15.3. DILI

The CSR will include a summary of any identified DILIs as well as any testing/evaluations conducted for alternative etiologies of treatment emergent liver test elevations in narratives of possible cases of DILI.

Drug-induced liver injury (Hy's law) includes all of the following:

- >3 × ULN elevation in aspartate aminotransferase (AST) and/or alanine aminotransferase (ALT);
- total bilirubin >2 ×ULN in the absence of findings of cholestasis (ie, absence of alkaline phosphatase elevation to >2 × ULN);
- no other etiology to explain the combination of increased ALT/AST and total bilirubin which will be adjudicated by the study Medical Monitor and Safety Physician.

15.4. Systemic Injection Reactions

A Systemic Injection Reaction (SIR) associated with SC injections were reported as TEAEs, consisting of one or more of the following symptoms: hypotension, tachycardia, nausea, vomiting, dizziness, headache, pyrexia, rigors, and/or chills. In reporting TEAEs related to zetomipzomib tolerability, terms such as the NCI-CTCAE terms of 'infusion-related reaction,' 'cytokine release syndrome,' 'acute infusion reaction,' or 'allergic or hypersensitivity reaction' should not be used to identify SIRs.

Investigators were asked to report each individual SIR using the checkbox in the Adverse Event eCRF. SIRs will be analyzed using SAF as follows:

- Time to patient experiencing their first SIR
- Number of patients experiencing SIRs and number of SIRs by dose
- SIR by prednisone equivalent daily dose at time of SIR
- Listing of all SIRs including information on grade, actions taken for management, any rescue medicines or concomitant medications that were administered, and dose of steroids at the time of the SIR.

15.5. Clinical Laboratory Data

Clinical laboratory data including hematology, IgG, coagulation, serum chemistry, and urinalysis are collected at specified visits in the DBTP and the OLE period.

- Hematology: complete blood count with differential
- Coagulation: prothrombin time, INR, activated partial thromboplastin time, antiphospholipid antibodies (Screening only)
- Clinical chemistry: Non-fasting chemistry panel including electrolytes, AST, ALT, alkaline phosphatase (ALP) and total bilirubin; lactate dehydrogenase, blood urea nitrogen (BUN), albumin, total protein, creatinine, glucose, low-density lipoproteins (LDL), amylase, lipase, triglycerides, hemoglobin A1c (HgbA1c), and C-reactive protein. LDL, ALP, amylase, lipase, triglycerides, and HgbA1c are only required at

Day 1 (pre-dose) and Week 24 of the Double-blind Treatment Period, and OLE Week 25

- IgG: Immunoglobulin G
- Autoantibodies: Antinuclear antibodies (ANA), smooth muscle antibody (SMA), liver kidney microsome type 1 (anti-LKM-1) antibody, anti-liver cytosol type 1 (LC1) antibodies, anti-soluble liver antigen/liver pancreas (SLA/LP) antibodies
- Urinalysis: macroscopic and microscopic examination

The quantitative variables will be summarized by treatment group (ie, treatment group for DBTP data; Group 1, Group 2, and all patients for OLE Section data (Section 7.1.2)) with continuous descriptive statistics. The qualitative variables (counts and percentages) will be summarized by treatment group at each analysis week. Urinalysis parameters that are entered as negative or a negative value will be interpreted as 0. Laboratory values falling below the Lower Limit of Quantification (LLQ), for example a result <0.3, it will be treated as being equal to the LLQ value (in this case, 0.3) for reporting purposes. Change from Baseline will also be summarized by treatment group. Additionally for each applicable analyte of blood chemistry, hematology and coagulation, post-Baseline low and high values and shift from Baseline to post-Baseline low and high summary will be presented where the post-Baseline low or high is defined as the low or high value after the first dose of treatment. For OLE data, change from Baseline, change from OLE-Baseline, shift from Baseline, and shift from OLE-Baseline will also be summarized for Group 1, Group 2, and all patients. The summaries will be performed for the data in the DBTP based on the Safety Set and for the data in OLE based on the OLE Safety Set, separately.

A patient listing of the clinical laboratory data will also be produced.

Local laboratory data will be converted to the units reported by the central laboratory.

15.6. Vital Signs

Vital signs, including systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, temperature, weight, and Body Mass Index (BMI), will be summarized with continuous descriptive statistics at each analysis week by treatment group (ie, treatment group for DBTP data; Group 1, Group 2, and all patients for OLE data). Change from Baseline will also be summarized to each post-Baseline analysis week. For OLE data, change from OLE-Baseline and change from Baseline will also be summarized for Group 1, Group 2, and all patients.
The summaries will be performed for the data in the DBTP based on the Safety Set and for the data in the OLE period based on the OLE Safety Set separately.

A patient listing of the vital sign data will also be produced.

15.7. ECG 12-Lead

QT interval corrected for pulse rate by Fridericia's formula (QTcF) will be summarized with continuous descriptive statistics at each analysis week by treatment group (ie, treatment group for DBTP data; Group 1, Group 2, and all patients for OLE data). Change from Baseline will also be summarized to each post-Baseline analysis week. Overall assessment of ECG (normal, abnormal clinically significantly, abnormal not clinically significantly) will be summarized by treatment group at each analysis week using frequencies and percentages. Shift from baseline to

each post-Baseline analysis week in overall assessment of ECG will also be summarized. For OLE data, change from Baseline, change from OLE-Baseline, shift from Baseline, and shift from OLE-Baseline will also be summarized for Group 1, Group 2, and all patients. The summaries will be performed for the data in DBTP based on the Safety Set and the data in OLE based on the OLE Safety Set separately.

To evaluate QTc prolongation, the number and percentage of patients having notable ECG values according to the following categories will also be summarized by treatment group (ie, treatment group for DBTP data; Group 1, Group 2, and all patients for OLE data) for DBTP data and OLE data.

- QTcF: ≤450, >450 and ≤480, >480 and ≤500, >500 msec (for each analysis week)
- Change from Baseline in QTcF: ≤30, >30 and ≤60, >60 msec (for each post-Baseline analysis week)
- Maximum QTcF: ≤450, >450 and ≤480, >480 and ≤500, >500 msec (during post-Baseline and during post OLE-Baseline)
- Change from Baseline in QTcF: ≤30, >30 and ≤60, >60 msec (during post-Baseline and during post OLE-Baseline)

A patient listing of the ECG data will also be produced.

15.8. Physical Examination

Physical examination data will be presented in a patient listing only.

16. INTERIM ANALYSES (PRIMARY ANALYSES)

The first planned unblinded interim analysis (ie, Primary Analyses) of the data will be conducted after all patients have or would have completed the Week 24 visit or EOS visit, whichever occurs later, in the DBTP, data being fully cleaned and database being “frozen” to limit changes.
The Final Data Analysis will be conducted after all patients have completed the study (the DBTP and the OLE Period).

Periodic evaluations of safety data will be performed throughout this study by the IDMC.

17. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

Changes to the analysis originally planned in the protocol are found in Table 7.

Table 7: Changes to Analysis Planned in the KZR-616-208 Protocol

Item	SAP	Rationale
1	The cytokine analysis will be provided by the Biomarker Group at Kezar and results will be provided in a stand-alone separate report.	Description of these analyses is outside the scope of this SAP.
2	Section 3.1.8 Patients who go below the lower limit of normal (for ALT/AST/IgG) will still be considered complete biochemical response.	Low values of these labs are not relevant in this patient population.
3	The absolute change from baseline will be calculated for all measures and percent change where applicable.	To enable analyses of both the absolute change paired with percent change.
4	Descriptive liver histopathology characteristics were added to exploratory endpoints for DBTP.	Kezar is collecting liver histopathology characteristics separately from the Ishak Score and thus should be reported.
5	Changed wording of "Complete Remission" and "Partial Remission" to "Complete Response" and "Partial Response", respectively.	To avoid confusion and stay current with recommendations by the International Autoimmune Hepatitis Group's recommendations of terminology (Pape et al., 2022).
6	PR definition in Protocol: ALT, AST, and IgG values (if IgG level is elevated at Baseline) with glucocorticoid dose not higher than starting dose (at Baseline) meeting one of the following criteria: • > upper limit of normal (ULN) but <2 × ULN (values must be lower than Baseline) • >ULN with >80% improvement from Baseline. PR definition in SAP: ALT, AST, and IgG values (if IgG level is elevated at Baseline) with glucocorticoid dose not higher than starting dose (at Baseline) meeting one of the following criteria • <2 × ULN (values must be lower than Baseline) • >80% improvement from Baseline. Note: Each laboratory value (ALT, AST, IgG) must be considered separately for PR criteria above. For example, if AST is in the normal range, IgG is in the normal range, but ALT is <2 x ULN, then this qualifies as a PR. If all three laboratory values (ALT, AST, IgG) are	Modified definition of Partial Response to include more details for interpretability for the programmers. Added a note with an example to clarify to programmers the difference between CR and PR.

Item	SAP	Rationale
	less than the ULN, then this is considered a CR. A patient cannot be a PR and a CR at the same time A patient's response can improve from a PR to a CR. Once a patient has a treatment failure, they are not eligible for a PR during the treatment period in which they failed treatment (ie, the DBTP or the OLE Period).	
7	Protocol definition of Non-Responder: Patients who do not achieve a reduction from Baseline in ALT or AST of ≥25% with glucocorticoid dose not higher than starting dose (at Baseline). Patients who have insufficient data for response determination at a time point will be considered non-responders for that time point. SAP definition of Non-Responder: Patients who do not achieve a CR, PR with a successful glucocorticoid taper, or TF. Patients who have insufficient data for response determination at a time point will be considered non-responders for that time point.	The definition of "Non-Responder" was contradictory by design. For example, the study design allows for patients to enter the study with ALT = 1.25 x ULN. If this patient's ALT was reduced by 25%, they would meet the criteria of Complete Response. This doesn't make sense to have a patient be a "non-responder" and a "complete response". Kezar's intent is to simplify categories of response at each analysis time point for every patient enrolled to fall into one of the four categories listed below. 1. Complete Response 2. Partial Response 3. Non-Response 4. Treatment Failure
8	Protocol definition of Treatment failure: • Patient's ALT or AST level worsened ≥2 times that of the Baseline value that is sustained for ≥1 week as verified via repeat laboratory assessments, despite compliance with standard of care (ie, with regard to inclusion criteria) or protocol-defined therapy. • If a glucocorticoid dose is increased above the Baseline dose, it may be considered a treatment failure unless attributed to an adverse event not relating to AIH. SAP definition of Treatment Failure: Treatment Failure is defined by meeting one of the following criteria: • Patient's ALT, AST, or IgG (if IgG level is elevated at Baseline)	The SAP includes cases where "immunosuppressant dose is increased above the Baseline dose or if a new immunosuppressant is added during the study due to an elevation in liver enzymes" for Treatment Failure. This is in alignment with protocol rules stating: "The dose of these immunosuppressants must be stable for at least 4 weeks prior to Screening and must be held constant unless tolerability/AEs require dose adjustment and glucocorticoid taper is completed." (Page 8 of the protocol) The Note at the bottom of the definition was added to help distinguish the definition of Treatment Failure versus Disease Flare and emphasize the rules for defining the four categories of: 1. Complete Response 2. Partial Response

Item	SAP	Rationale
	level worsened ≥2 times that of the Baseline value that is sustained for ≥1 week as verified via repeat laboratory assessments, despite compliance with standard of care (ie, with regard to inclusion criteria) or protocol-defined therapy. OR • If the glucocorticoid or immunosuppressant dose is increased above the Baseline dose or if a new immunosuppressant is added during the study due to an elevation in liver enzymes, it may be considered a treatment failure unless it is attributed to an adverse event (AE) not relating to AIH. Patients falling under this criterium will be classified through a blinded evaluation prior to the DBTP database freeze (see Section 4.1.5.1). Note: Treatment failure can only occur if a patient has not already achieved a [CR] or [PR with a successful glucocorticoid taper].	3. Non-Response 4. Treatment Failure
9	<u>Protocol definition of Disease flare:</u> Elevation of ALT after achieving CR that meets all the following criteria: • is ≥25% above the CR value • is ≥1.25 ULN • is sustained for ≥1 week as verified via repeat laboratory assessments • requires the patient to re-start or escalate glucocorticoid therapy • is not due to other identifiable causes (eg, viral hepatitis, concomitant medications, alcoholic liver injury, etc.)	Added “achieving a PR with a successful GC taper” because we also added the endpoint “[CR] or [PR with successful glucocorticoid taper]” due to feedback from the KOLs. Disease flare marks the end of the duration of [CR] or [PR with successful glucocorticoid taper].

Item	SAP	Rationale
	SAP definition of Disease flare: Elevation of ALT after achieving [CR] or [PR with a successful glucocorticoid taper] that meets all the following criteria: • is ≥25% above the [CR] or [PR with a successful glucocorticoid taper] value • is ≥1.25 ULN • is sustained for ≥1 week as verified via repeat laboratory assessments • requires the patient to re-start or escalate glucocorticoid therapy or immunosuppressants • is not due to other identifiable causes [eg, viral hepatitis, drug-induced liver injury (DILI), concomitant medications, alcoholic liver injury, etc.] The last 2 bullets of the Disease Flare definition (above) will be adjudicated through a blinded evaluation prior to the DBTP database freeze (see Section 4.1.5.2).	
10	Added new endpoints for DBTP and OLE: • Duration of patients achieving CR • Duration of patients achieving [CR] or [PR with a successful glucocorticoid taper] • Time to [CR] or [PR with a successful glucocorticoid taper] • Proportion of patients who achieve CR with glucocorticoid taper to ≤5 mg by Week 24 • Proportion of patients who achieve CR with glucocorticoid taper to 0 mg by Week 24 • Proportion of patients who achieve PR with glucocorticoid taper to ≤5 mg by Week 24 • Proportion of patients who achieve PR with glucocorticoid taper to 0 mg by Week 24 • Proportion of patients who achieve CR that had a baseline ALT >2x ULN	Feedback from KOLs was provided that these endpoints are important and meaningful for the patient.

Item	SAP	Rationale
11	Added Section 15.3 to provide a summary (table and listing) of any identified DILI during the study	Information added to [redacted].
12	Added new subgroup analysis for the primary endpoint assessing "Patients who have a baseline ALT >2x ULN"	New endpoint will help distinguish treatment effect in different groups of patients with higher and lower ALT at baseline.
13	Added Section 15.4 to provide additional Safety Analyses for Systemic Injection Reactions SIR	[redacted]
14	Added Section 6.5 to define the Intent to Treat Population	Provide details of ITT population in response to [redacted]
15	Added ITT population to Table 4: Efficacy Endpoints and Analysis Methods for the Primary and Secondary endpoints for the DBTP and Table 5: Sensitivity Analyses for Primary Analysis	Provide Sensitivity Analyses based on all randomized participants (ITT population) in response to [redacted]
16	Added Section 14.4.6 Sensitivity Analysis of Secondary Endpoints	Provide Sensitivity Analyses based on all randomized participants (ITT population) in response to [redacted]

EOS = End of Study; DBTP = Double Blind Treatment Period; ITT = Intent To Treat; OLE = Open Label Extension; SAP = Statistical Analysis Plan.

18. REFERENCES

Guo L, Zheng L, Hu L, et al. (2017) Transient elastography (FibroScan) performs better than non-invasive markers in assessing liver fibrosis and cirrhosis in autoimmune hepatitis patients. *Medical science monitor: international medical journal of experimental and clinical research* 23: 5106.

Ishak K, Baptista A, Bianchi L, et al. (1995) Histological grading and staging of chronic hepatitis. *J Hepatol* 22(6): 696-699.

Pape S, Snijders RJ, Gevers TJ, et al. (2022) Systematic review of response criteria and endpoints in autoimmune hepatitis by the International Autoimmune Hepatitis Group. *Journal of hepatology* 76(4): 841-849.

Prati D, Taioli E, Zanella A, et al. (2002) Updated definitions of healthy ranges for serum alanine aminotransferase levels. *Ann Intern Med* 137(1): 1-10.

19. APPENDICES

APPENDIX A. SAS® CODES FOR EQ-5D-5L INDEX VALUES

From https://euroqol.org/wp-content/uploads/2024/01/US_valueset_SAS.txt

```
/*Computing EQ-5D-5L index values with SAS using the United States (US)
Pickard value set
Version 1.2 (Updated 31/01/2022)

The variables for the 5 dimensions of the EQ-5D-5L descriptive system
should be named 'mobility', 'selfcare', 'activity', 'pain', and
'anxiety'.
If they are given different names the syntax code below will not work
properly. The 5 variables should contain the values for the different
dimensions
in the EQ-5D health profile (i.e. 1, 2, 3, 4 or 5). The variable
'EQindex' contains the values of the
EQ-5D-5L index values on the basis of the US set of weights.
*/

*****************************************************************
*SAS syntax code for the computation of index*
*values with the US TTO value set*
*****************************************************************

data WORK.CAT;
set WORK.CAT;


if mobility eq 1 then disut_mo=0;
else if mobility eq 2 then disut_mo=0.096;
else if mobility eq 3 then disut_mo=0.122;
else if mobility eq 4 then disut_mo=0.237;
else if mobility eq 5 then disut_mo=0.322;

if selfcare eq 1 then disut_sc=0;
else if selfcare eq 2 then disut_sc=0.089;
else if selfcare eq 3 then disut_sc=0.107;
else if selfcare eq 4 then disut_sc=0.220;
else if selfcare eq 5 then disut_sc=0.261;

if activity eq 1 then disut_ua=0;
else if activity eq 2 then disut_ua=0.068;
else if activity eq 3 then disut_ua=0.101;
else if activity eq 4 then disut_ua=0.255;
else if activity eq 5 then disut_ua=0.255;

if pain  eq 1 then disut_pd=0;
else if pain  eq 2 then disut_pd=0.060;
else if pain  eq 3 then disut_pd=0.098;
else if pain  eq 4 then disut_pd=0.318;
else if pain  eq 5 then disut_pd=0.414;

if anxiety eq 1 then disut_ad=0;
```

```
else if anxiety eq 2 then disut_ad=0.057;
else if anxiety eq 3 then disut_ad=0.123;
else if anxiety eq 4 then disut_ad=0.299;
else if anxiety eq 5 then disut_ad=0.321;

disut_total=disut_mo+disut_sc+disut_ua+disut_pd+disut_ad;
EQindex=1-disut_total;
run;
```

APPENDIX B. SCHEDULE OF VISITS AND ASSESSMENTS

Table 8: Schedule of Assessments (Visits with Study Assessments) – Double Blind Treatment Period

		Double-blind Treatment														
Visit Number	**Screen[a] 1**	**2**		**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**		**20**	**22**	**24**	**26[t]**
Start of Week	**-4**	**0**		**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**		**18**	**20**	**22**	**24**
		Pre	**Post**								**Pre**	**Post**				
Study Day ± Window (Days)	**-28 to -1**	**1**		**15±1**	**29±1**	**43±1**	**57±1**	**71±1**	**85±1**	**99±1**	**113±1**		**127±1**	**141±1**	**155±1**	**169±1**
Informed consent[a]	X															
Informed consent for genotyping[b]	X															
Inclusion/exclusion criteria	X															
Demographic data	X															
Medical history (including AIH, procedures, prior therapy, social history)	X															
Concomitant therapy	X	←										→				
Screening serology tests[c]	X															
TB test	X															
FSH[d]	X															
Pregnancy test[e]	X	X			X		X		X		X			X		X
Full physical examination[f]	X										X					X

		Double-blind Treatment														
Visit Number	**Screen[a] 1**	**2**		**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**		**20**	**22**	**24**	**26[t]**
Start of Week	**-4**	**0**		**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**		**18**	**20**	**22**	**24**
		Pre	**Post**								**Pre**	**Post**				
Study Day ± Window (Days)	**-28 to -1**	**1**		**15±1**	**29±1**	**43±1**	**57±1**	**71±1**	**85±1**	**99±1**	**113±1**		**127±1**	**141±1**	**155±1**	**169±1**
Abbreviated physical examination[f]		X			X		X		X					X		
Vital signs[g]	X	X		X	X	X	X	X	X	X	X		X	X	X	X
12-lead ECG[h]	X										X					X
EQ-5D-5L[i]	X										X					X
CLDQ[i]	X										X					X
GTI assessment		X														X
Hematology, IgG, autoantibodies[j], and coagulation[k]	X	X			X		X		X		X			X		X
Serum chemistry[k]	X	X		X	X	X	X	X	X	X	X		X	X	X	X
Plasma cytokines/proteomics[l]		X					X				X					X
Leukocyte subsets immunophenotyping[l]		X									X					X
Genotyping (DNA)[m]		X														
Gene expression (RNA)[m]		X					X				X					X
Blood sample for PK[n]		X	X								X	X				
Urinalysis[o]	X	X			X		X		X		X			X		X

		Double-blind Treatment												
Visit Number	**Screen[a] 1**	**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26[t]**
Start of Week	**-4**	**0**	**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**
		Pre **Post**								**Pre** **Post**				
Study Day ± Window (Days)	**-28 to -1**	**1**	**15±1**	**29±1**	**43±1**	**57±1**	**71±1**	**85±1**	**99±1**	**113±1**	**127±1**	**141±1**	**155±1**	**169±1**
SARS-CoV-2[p]	X													
FibroScan[q]	X									X				X
Liver biopsy	X[r]													X
Randomization		X												
IMP administration[s]		X	X	X	X	X	X	X	X	X	X	X	X	
Adverse events		←												→

Abbreviations: AIH=autoimmune hepatitis; CLDQ=Chronic Liver Disease Questionnaire; DNA=deoxyribonucleic acid; ECG=electrocardiogram; EQ-5D-5L=EuroQoL 5-dimension 5-level; FSH=follicle stimulating hormone; GTI=Glucocorticoid Toxicity Index; HBcAb=hepatitis B core antibody; HBsAb=hepatitis B surface antibody; HBsAg=hepatitis B surface antigen; hCG=beta-human chorionic gonadotropin; HCV=hepatitis C virus; HEENT=head, eyes, ears, nose, throat; HIV=human immunodeficiency virus; IgG=immunoglobulin G; IMP=investigational medicinal product; OLE=open-label extension; PK=pharmacokinetic; Post=post-dose; Pre=pre-dose; RT-PCR=real-time polymerase chain reaction; RNA=ribonucleic acid; SARS-CoV-2=serious acute respiratory syndrome coronavirus 2; TB=tuberculosis; VCTE= vibration-controlled transient elastography

a The screening period may last up to 4 weeks; however, laboratory assessments delayed due to logistic considerations may be obtained up to an additional 7 days later, upon approval by the Sponsor. Informed consent must be signed prior to any study-related activity in KZR-616-208.

b Patients who provide additional informed consent will undergo blood sampling for genetic analysis.

c Includes HIV, HBsAg, HBcAb, HBsAb, and HCV.

d FSH testing should be done for females who are not surgically sterile (see Protocol Section 7.2.7), with amenorrhea for >1 but ≤2 years, and without documented confirmatory FSH level in the postmenopausal range.

e Women of childbearing potential only (see Protocol Section 7.2.7). Positive urine pregnancy tests will be confirmed with a serum hCG test.

f A full physical examination includes, at a minimum, assessment of the following: general appearance, skin, HEENT, heart; chest/breast, abdomen, neurological system (briefly), lymph nodes, spine, and skeletal extremities. The examination will also include body weight and height (height at Visit 2 predose only). An abbreviated physical examination, symptom directed as determined by the Investigator, will include general appearance, cardiovascular, gastrointestinal, and pulmonary systems. The abbreviated physical examination may be performed at various unscheduled timepoints if clinically indicated.

g Vital signs (pulse rate, respiratory rate, blood pressure, and body temperature) will be collected prior to dosing. Pulse rate and blood pressure (using a calibrated sphygmomanometer) should be collected after the patient has been resting for at least 5 minutes in the seated position (pulse rate first, followed by blood pressure). If the blood pressure is elevated on the first measurement at Screening, it should be repeated after at least an additional 5 minutes of rest. It is recommended that blood pressure be measured using the same arm at each assessment. When the time of vital signs measurement coincides with a blood sample collection, the vital signs will be measured before blood sample collection.
h 12-lead ECG should be performed after the patient has been resting for at least 5 minutes in the supine position.
i It is recommended that the EQ-5D-5L and CLDQ be completed prior to any other procedures or assessments (other than signing of informed consent) at applicable visits.
j See Protocol Section 7.2.4.5 for the list of autoantibodies. In addition to the timepoints specified, samples for autoantibodies will be collected if a patient experiences a disease flare.
k Patients will be in a seated or supine position during blood collection. See Protocol Section 7.2.4.5 for assessments.
l Plasma samples for cytokines and proteomics and blood samples for immune cell profiling will be collected prior to dosing. Samples for proteomics and immune cell profiling will be stored for future analysis of chemokines and immune cell subsets.
m Samples will be collected prior to dosing in patients who have provided the proper informed consent for genetic analyses.
n Samples for PK analysis will be drawn from all patients prior to IMP administration, at 30 minutes (±10 minutes) after administration, and at 4 hours (±15 minutes) after administration. In addition, patients will be randomized 1:1:1 to 3 groups to obtain an additional PK sample at either 15 minutes (±5 minutes), 1 hour (±10 minutes), or 2 hours (±15 minutes) after IMP administration. The time of the PK blood draw must be recorded.
o A 30 mL aliquot should be obtained from the first morning void. All urinalyses will include both macroscopic and microscopic examination.
p Where local conditions and requirements mandate SARS-CoV-2 testing, these tests will be conducted prior to Visit 2. If testing is conducted, the test must be an RT-PCR assay for SARS-CoV-2, with results expected within 24 hours.
q FibroScan is a device using VCTE to measure liver stiffness. The scan will be performed at least 2 hours after food intake.
r Results of a liver biopsy performed within 6 months prior to Screening (or at Screening) will be used as the Baseline liver biopsy. Liver biopsies will be performed and analyzed by a local laboratory.
s See Protocol Table 2 for weekly schedule of IMP administration at visits where no study assessments are performed.
t For patients who roll over immediately to the OLE Period, Study Visit 26 (Protocol Table 1) may occur on the same day as Study Visits 27 and 28 (Protocol Table 3). Patients who are not eligible or choose not to roll over to the OLE Period will have an EOS visit (see Protocol Table 3) 4 weeks after completion of the Double-blind Treatment Period.

Table 9: Schedule of IMP Administration (Visits with no Study Assessments) – Double Blind Treatment Period

	Double-blind Treatment											
Visit Number	**3**	**5**	**7**	**9**	**11**	**13**	**15**	**17**	**19**	**21**	**23**	**25**
Start of Week	**1**	**3**	**5**	**7**	**9**	**11**	**13**	**15**	**17**	**19**	**21**	**23**
Study Day ± Window (Days)	**8±1**	**22±1**	**36±1**	**50±1**	**64±1**	**78±1**	**92±1**	**106±1**	**120±1**	**134±1**	**148±1**	**162±1**
IMP administration	X	X	X	X	X	X	X	X	X	X	X	X

Abbreviations: IMP=investigational medicinal product
NOTE: Home trial service will be arranged for patients who choose not to return to the site for these visits.

Table 10: Schedule of Assessments (Visits with Study Assessments) – Open Label Extension Period

	OLE Eligibility	Open-label Treatment								
Visit Number	**27[a]**	**28[a]**	**31**	**35**	**39**	**43**	**47**	**EOT**	**EOS[n]**	**ETV[o]**
Start of OLE Week		**OLE 1**	**OLE 4**	**OLE 8**	**OLE 12**	**OLE 16**	**OLE 20**	**OLE 25**	**N/A**	**N/A**
OLE Study Day ± Window (Days)	**0 (-28)**	**1**	**22±7**	**50±7**	**78±7**	**106±7**	**134±7**	**169±7**	**N/A**	**N/A**
OLE inclusion/exclusion criteria	X									
Concomitant therapy	←	—	—	—	—	—	—	→		X
Pregnancy test[b]	X		X	X	X	X	X	X	X	X
Full physical examination[c]	X					X		X		X
Abbreviated physical examination[c]			X	X	X		X		X	
Vital signs[d]	X		X	X	X	X	X	X	X	X
12-lead ECG[e]	X					X		X		X
EQ-5D-5L[f]	X					X		X		X
CLDQ[f]	X					X		X		X
GTI assessment	X							X		X
Hematology, IgG, autoantibodies[g], and coagulation[h]	X		X	X	X	X	X	X	X	X
Serum chemistry[h]	X		X	X	X	X	X	X	X	X
Urinalysis[i]	X		X	X	X	X	X	X		X
Plasma cytokines/proteomics[j]								X		X
Leukocyte subsets immunophenotyping[j]								X		X
Gene expression (RNA)[k]								X		X
FibroScan[l]	X							X		X

	OLE Eligibility	Open-label Treatment								
Visit Number	**27[a]**	**28[a]**	**31**	**35**	**39**	**43**	**47**	**EOT**	**EOS[n]**	**ETV[o]**
Start of OLE Week		**OLE 1**	**OLE 4**	**OLE 8**	**OLE 12**	**OLE 16**	**OLE 20**	**OLE 25**	**N/A**	**N/A**
OLE Study Day ± Window (Days)	**0 (-28)**	**1**	**22±7**	**50±7**	**78±7**	**106±7**	**134±7**	**169±7**	**N/A**	**N/A**
Optional liver biopsy								X		X
IMP administration[m]		X	X	X	X	X	X			
Adverse events	←								→	X

Abbreviations: AIH=autoimmune hepatitis; CLDQ=Chronic Liver Disease Questionnaire; ECG=electrocardiogram; EOS=End-of-Study; EOT=End-of-Treatment; EQ-5D-5L=EuroQoL 5-dimension 5-level; ETV=Early Termination Visit; GTI=Glucocorticoid Toxicity Index; hCG=beta-human chorionic gonadotropin; IgG=immunoglobulin G; IMP=investigational medicinal product; N/A=not applicable; OLE=open-label extension; RT-PCR=real time polymerase chain reaction; VCTE=vibration-controlled transient elastography.

a For patients who roll over immediately from the Double-blind Treatment Period to the OLE Period, the Week 24 Visit of the Double-blind Treatment Period (Study Visit 26) in Protocol Table 1 may occur on the same day as Study Visits 27 and 28 in Protocol Table 3. For patients who initiate the OLE Period within 4 weeks after the Week 24 Visit of the Double-blind Treatment Period, the most recently reported assessments from Week 20 or later in the Double-blind Treatment Period may serve as the eligibility assessments for the OLE Period (ie, OLE Visit 27 assessments do not need to be repeated). For patients who roll into the OLE Period >4 weeks and <12 weeks after the Week 24 Visit of the Double-blind Treatment Period, Visit 27 assessments must be performed within 28 days prior to first open-label IMP administration to confirm eligibility for the OLE. All new or changed concomitant medications and adverse events occurring after Week 24 of the Double-blind Treatment Period through Day 1 of the OLE Period should be recorded. If a patient does not meet eligibility criteria for the OLE Period, then the patient would proceed with the EOS visit. The patient would not be considered a screen fail for the OLE Period.

b Women of childbearing potential only. Positive urine pregnancy tests will be confirmed with a serum hCG test.

c A full physical examination includes, at a minimum, assessment of the following: general appearance, skin, HEENT, heart; chest/breast, abdomen, neurological system (briefly), lymph nodes, spine, and skeletal extremities. The examination will also include body weight. An abbreviated physical examination, symptom -directed as determined by the Investigator, will include general appearance, cardiovascular, gastrointestinal, and pulmonary systems. The abbreviated physical examination may be performed at various unscheduled timepoints if clinically indicated.

c Vital signs (pulse rate, respiratory rate, blood pressure, and body temperature) will be collected prior to dosing. Pulse rate and blood pressure (using a calibrated sphygmomanometer) should be collected after the patient has been resting for at least 5 minutes in the seated position (pulse rate first, followed by blood pressure). If the blood pressure is elevated on the first measurement at Screening, it should be repeated after at least an additional 5 minutes of rest. It is recommended that blood pressure be measured using the same arm at each assessment. When the time of vital signs measurement coincides with a blood sample collection, the vital signs will be measured before blood sample collection.

e 12-lead ECG should be performed after the patient has been resting for at least 5 minutes in the supine position.

f It is recommended that the EQ-5D-5L and CLDQ be completed prior to any other procedures or assessments at applicable visits.

g See Protocol Section 7.2.4.5 for the list of autoantibodies. In addition to the timepoints specified, samples for autoantibodies will be collected if a patient experiences a disease flare.

h Patients will be in a seated or supine position during blood collection. See Protocol Section 7.2.4.5 for assessments.

i A 30 mL aliquot should be obtained from the first morning void. All urinalyses will include both macroscopic and microscopic examination.
j Plasma samples for cytokines and proteomics and blood samples for immune cell profiling will be collected prior to dosing. Samples for proteomics and immune cell profiling will be stored for future analysis of chemokines and immune cell subsets.
k Samples will be collected prior to dosing in patients who have provided the proper informed consent for genetic analyses.
l FibroScan is a device using VCTE to measure liver stiffness. The scan will be performed at least 2 hours after food intake.
m See Protocol Table 4 for weekly schedule of IMP administration at visits where no study assessments are performed.
n The EOS (safety follow-up) visit has a ±7 day visit window.
o If a patient cannot continue with the planned assessments, the patient will complete the ETV followed by the End-of-Study (EOS) visit 4 weeks later.

Table 11: Schedule of IMP Administration (Visits with no Study Assessments) – Open Label Extension Period

	Open-label Treatment																	
Visit Number	**29**	**30**	**32**	**33**	**34**	**36**	**37**	**38**	**40**	**41**	**42**	**44**	**45**	**46**	**48**	**49**	**50**	**51**
Start of Week	**OLE 2**	**OLE 3**	**OLE 5**	**OLE 6**	**OLE 7**	**OLE 9**	**OLE 10**	**OLE 11**	**OLE 13**	**OLE 14**	**OLE 15**	**OLE 17**	**OLE 18**	**OLE 19**	**OLE 21**	**OLE 22**	**OLE 23**	**OLE 24**
Study Day ± Window (Days)	**8±3**	**15±3**	**29±3**	**36±3**	**43±3**	**57±3**	**64±3**	**71±3**	**85±3**	**92±3**	**99±3**	**113±3**	**120±3**	**127±3**	**141±3**	**148±3**	**155±3**	**162±3**
IMP administration	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

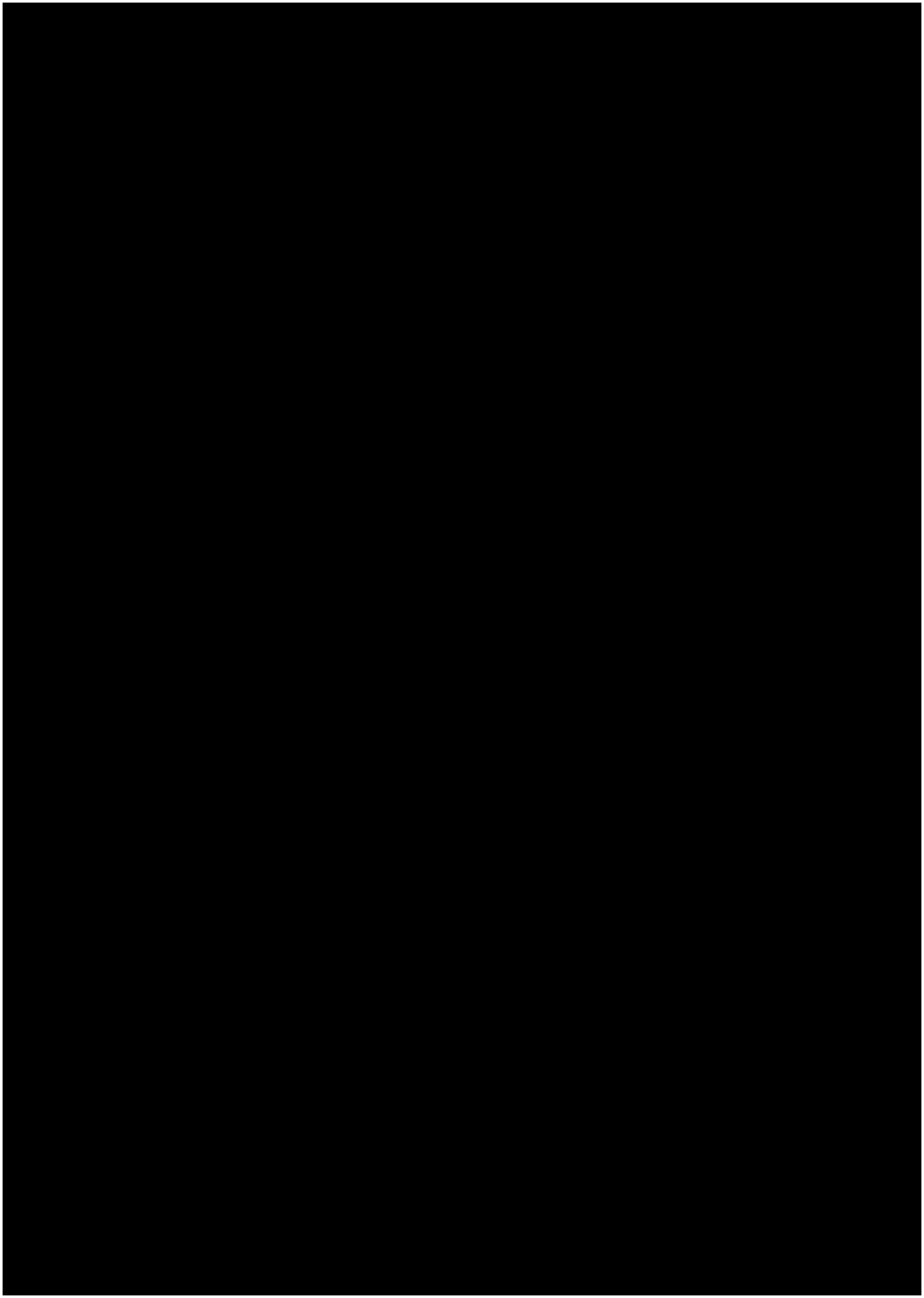
Abbreviations: IMP=investigational medicinal product.
NOTE: Home trial service will be arranged for patients who choose not to return to the site for these visits.

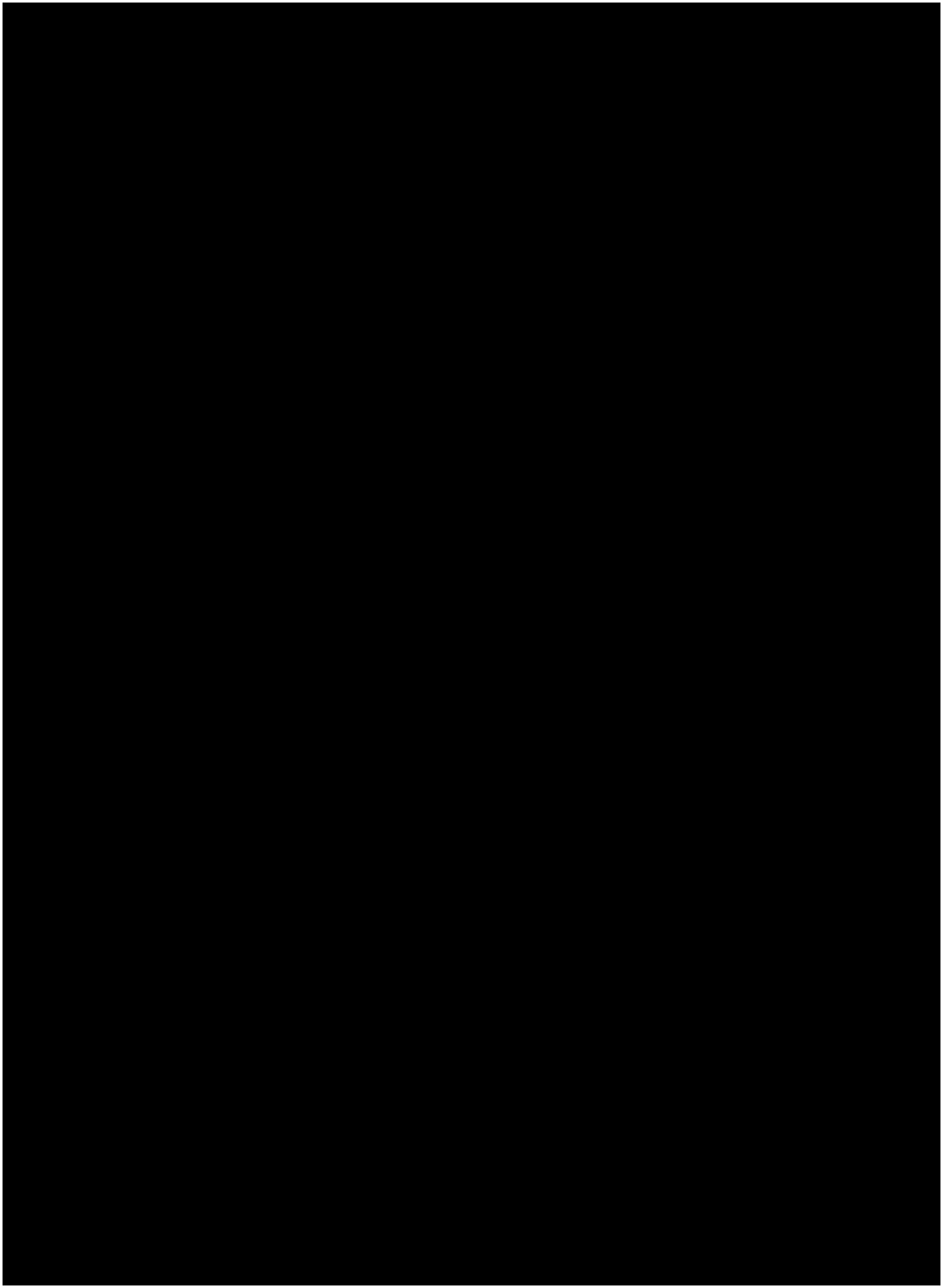
APPENDIX C. WEIGHTING, SCORING, AND ANALYSIS OF THE GTI

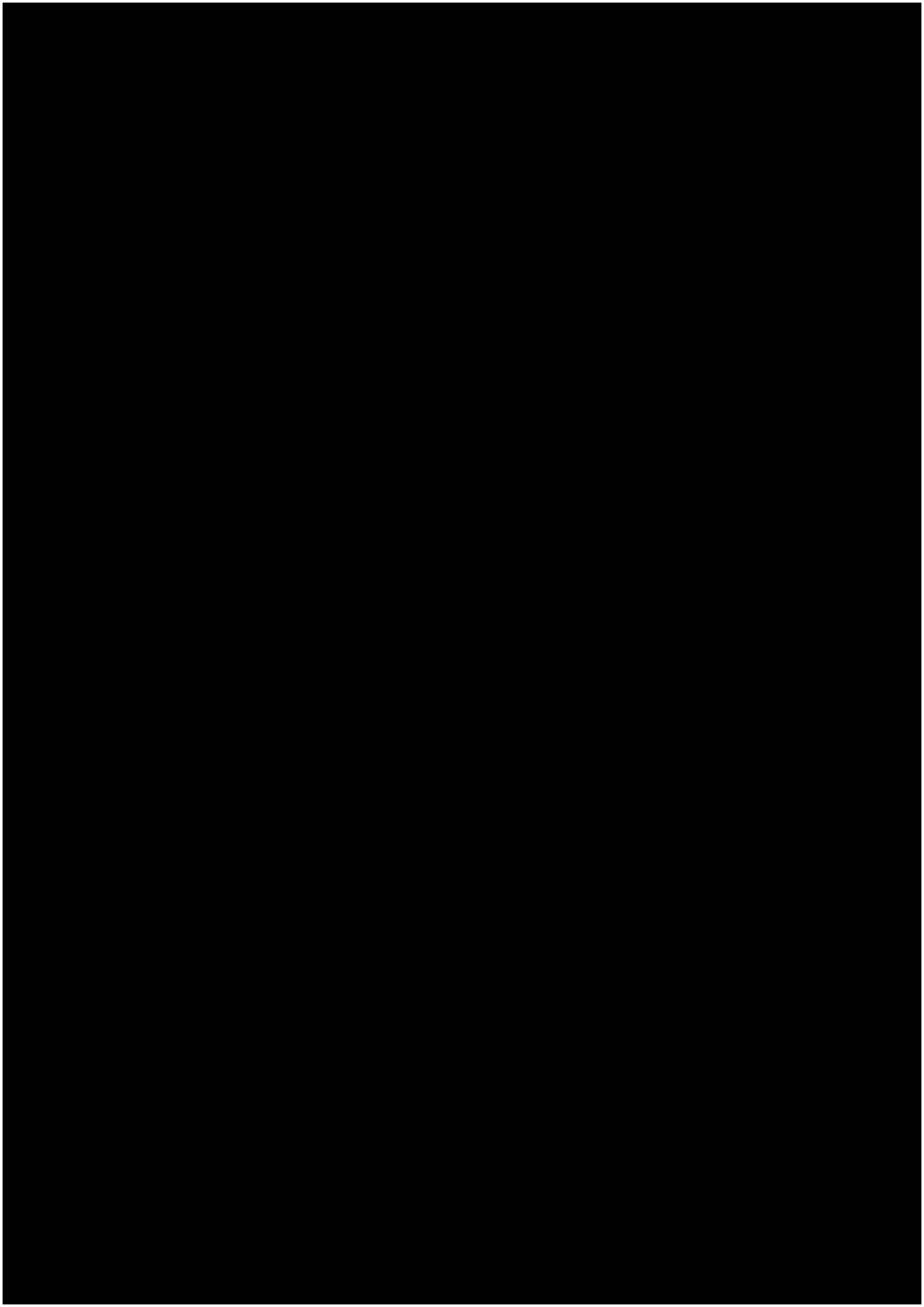
Weighting, Scoring, and Analysis of the GTI

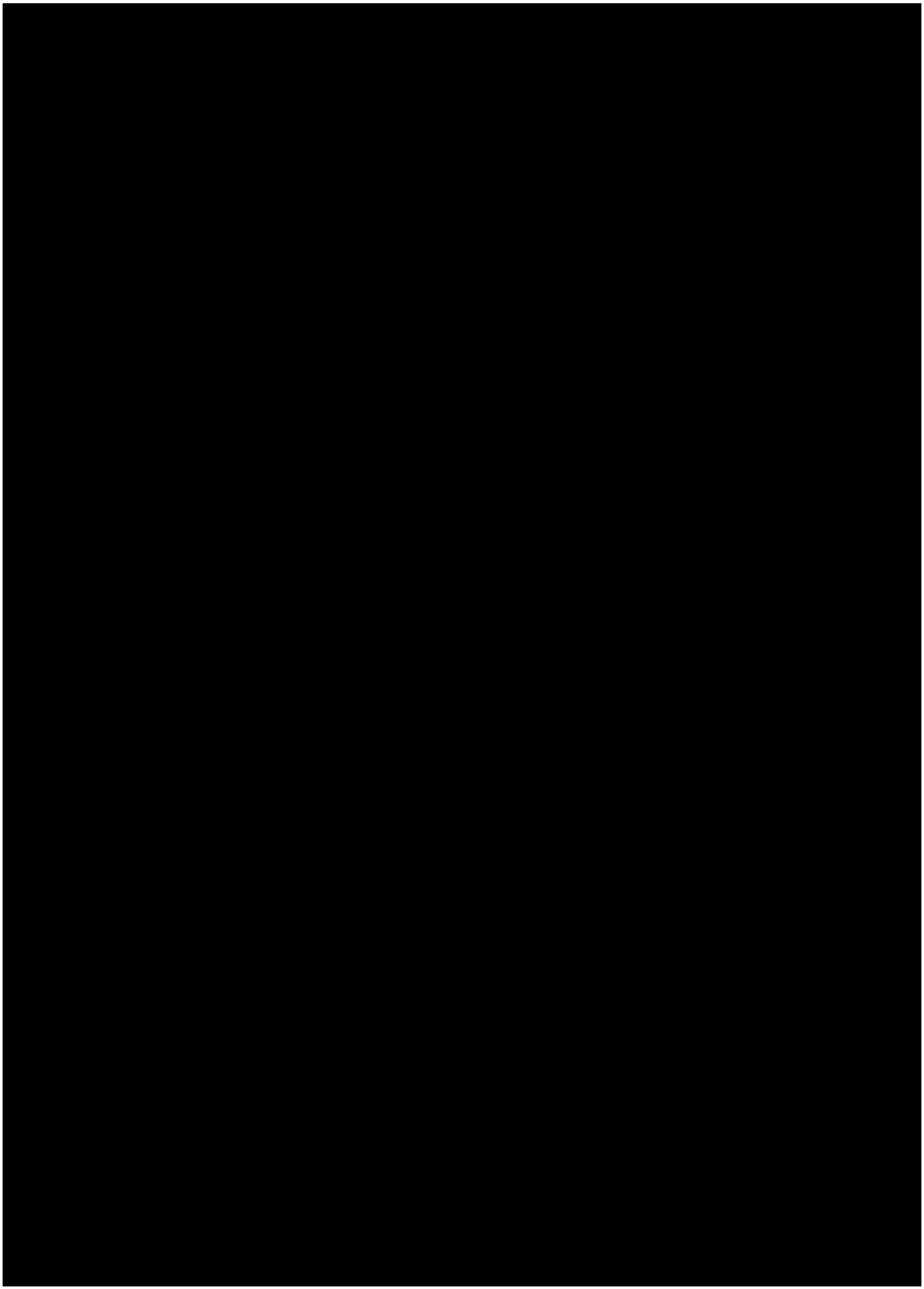
APPENDIX D. COMPOSITE GLUCOCORTICOID TOXICITY INDEX

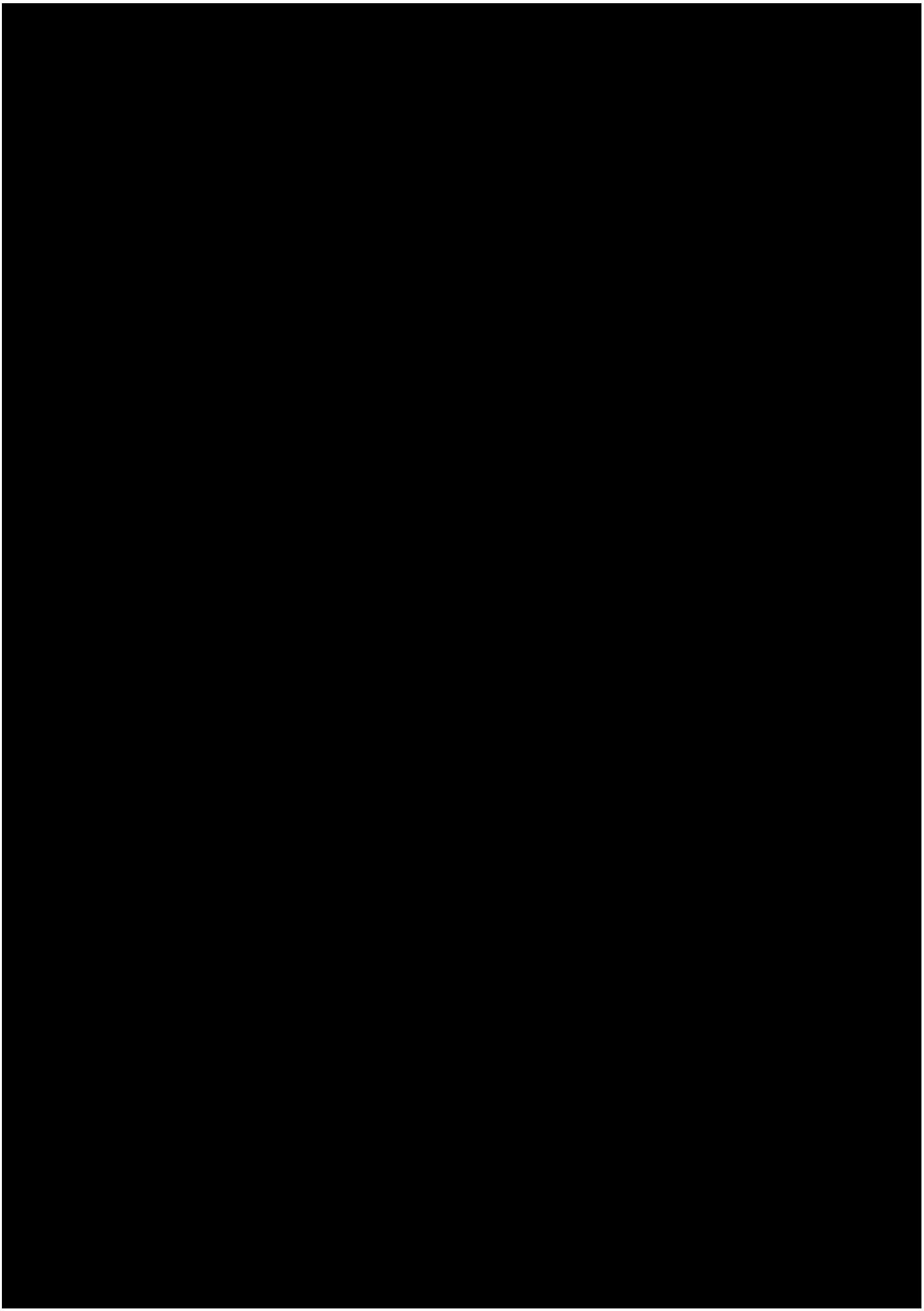
Composite Glucocorticoid Toxicity Index

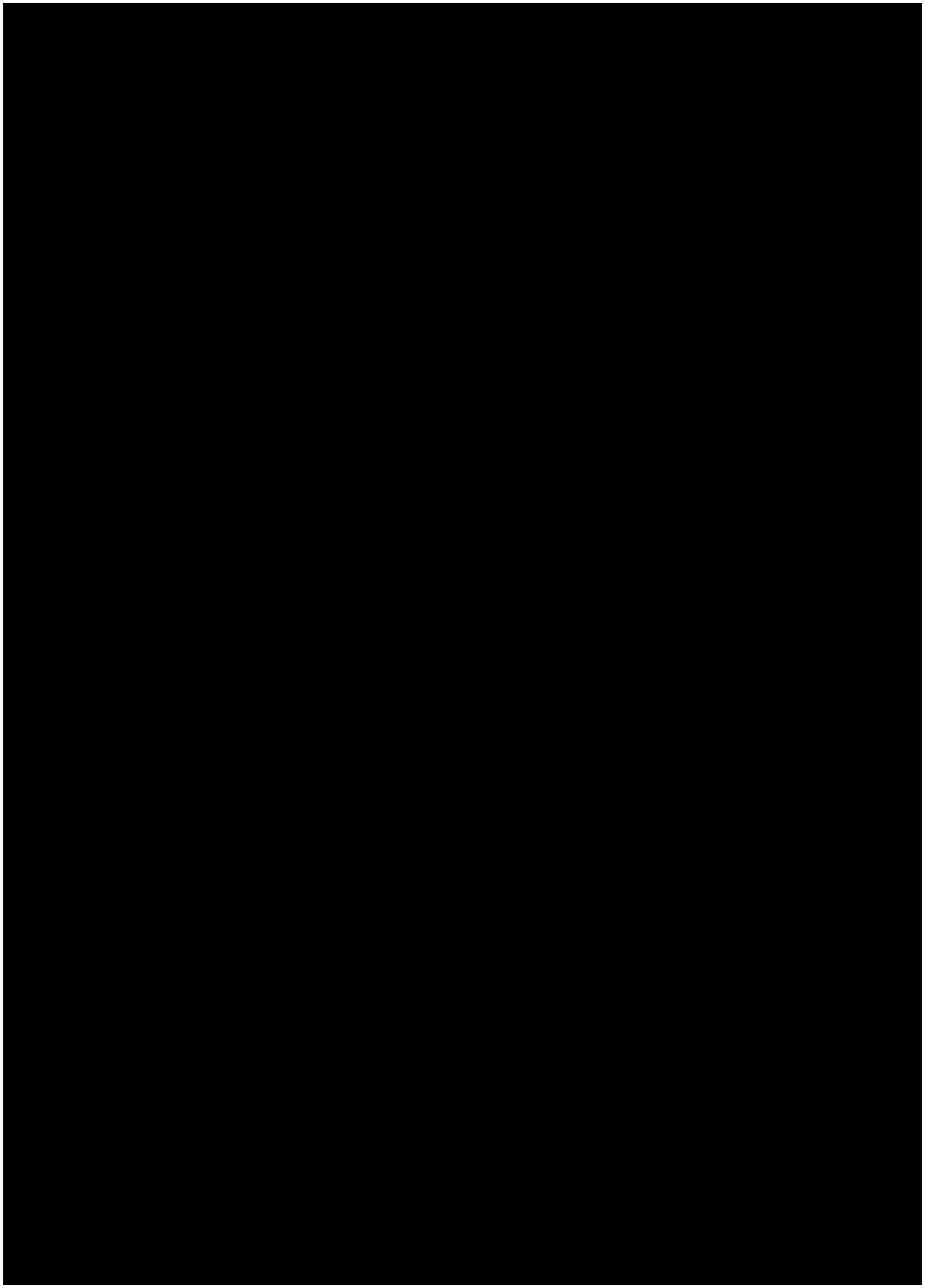


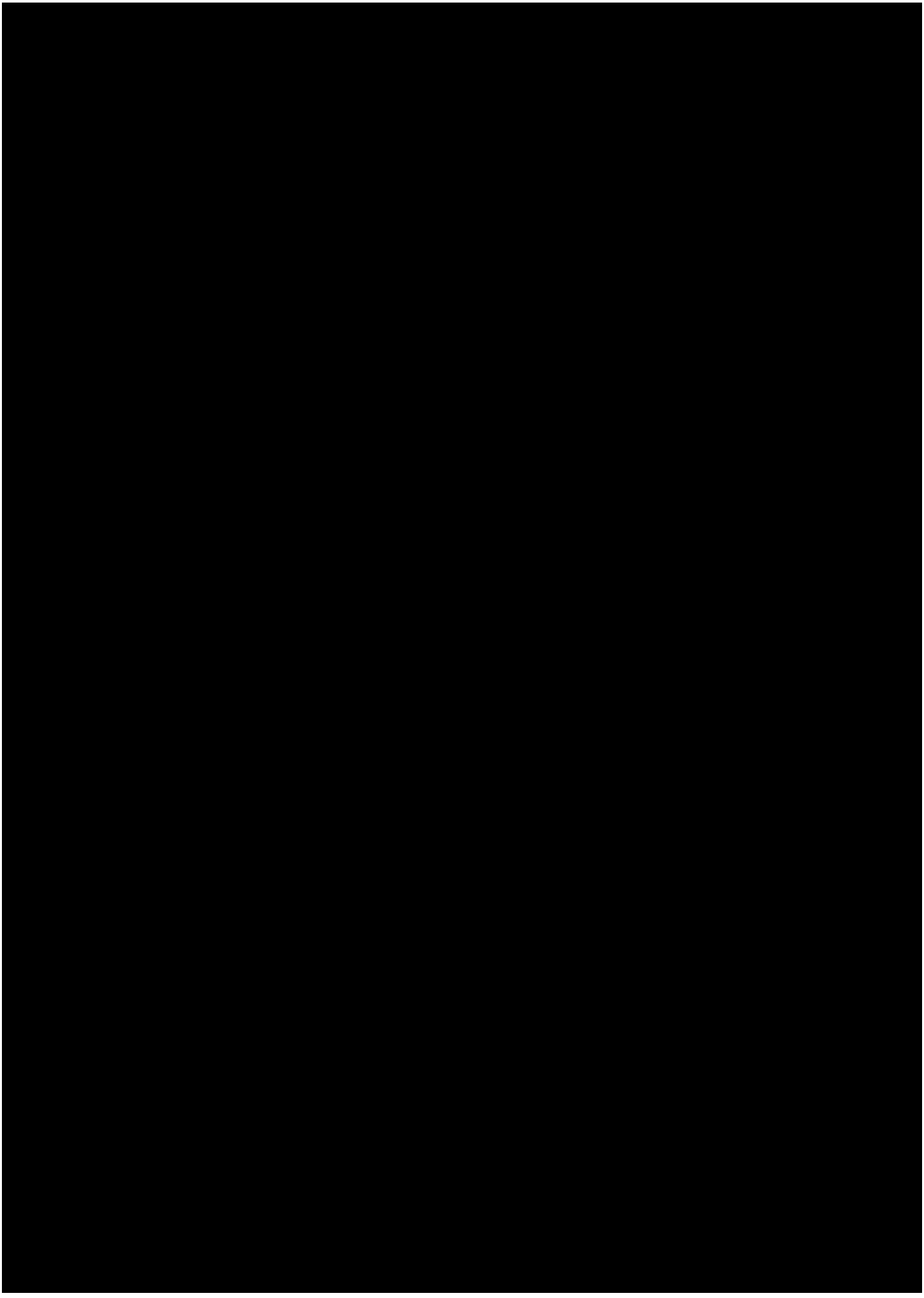


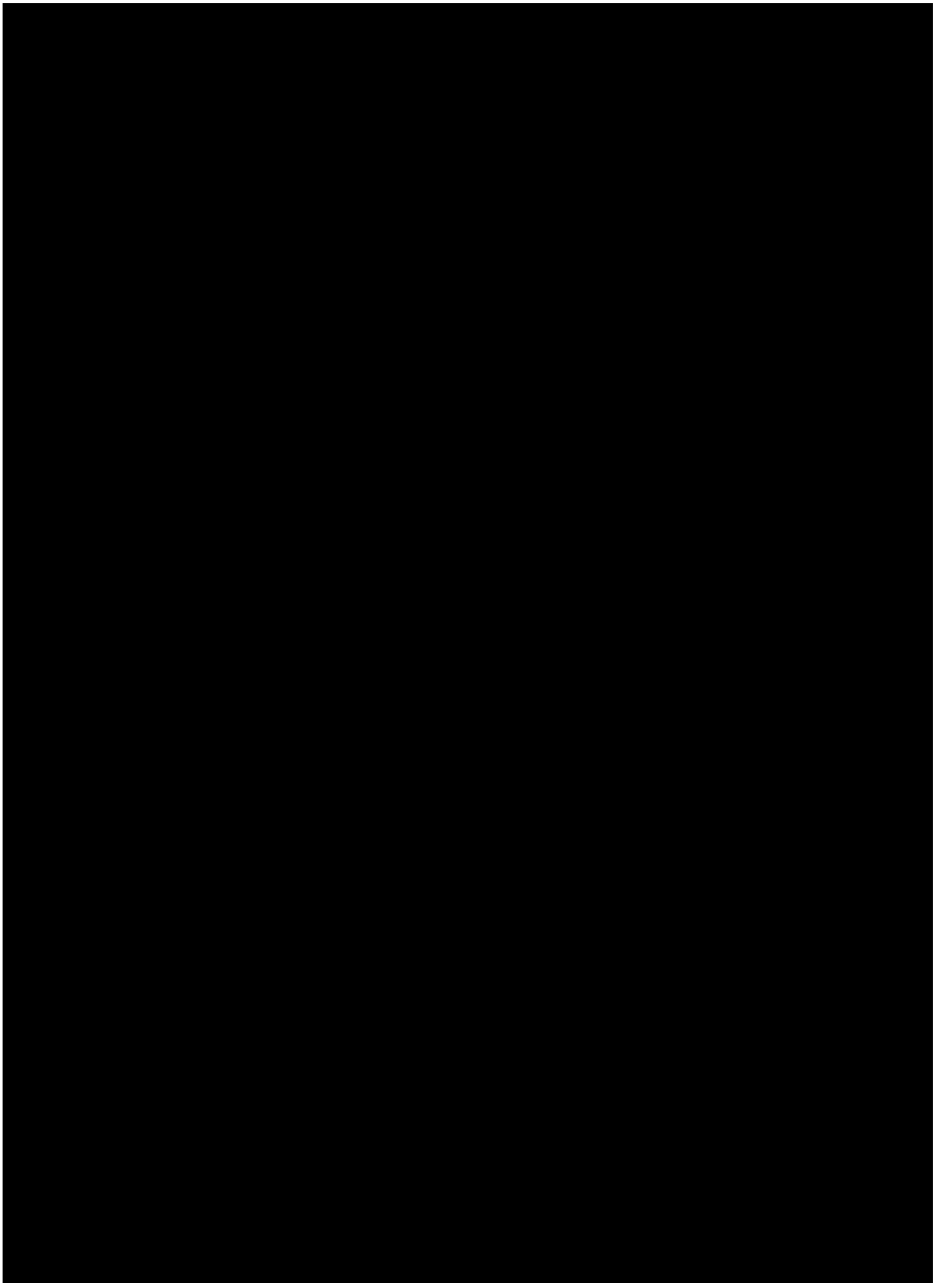


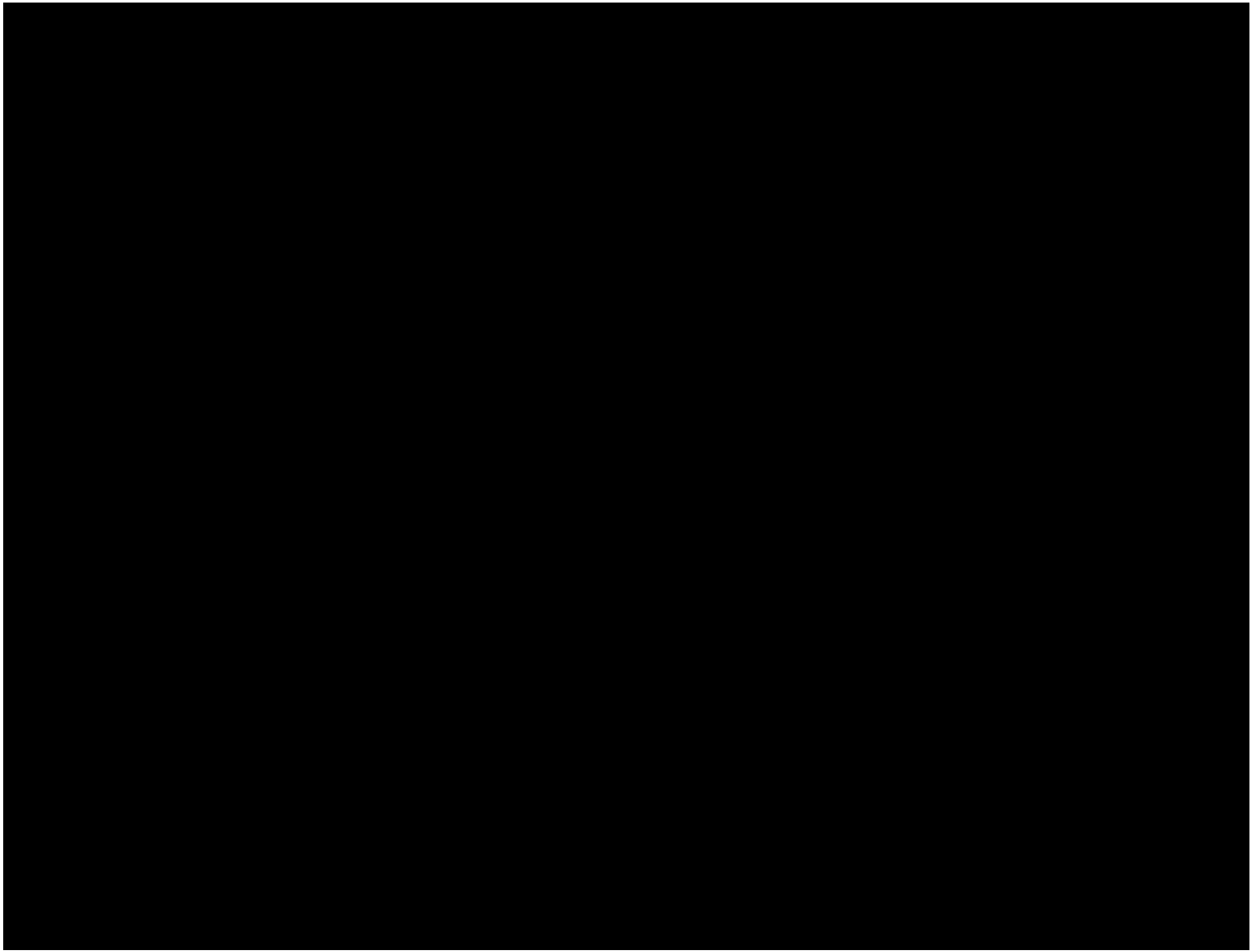


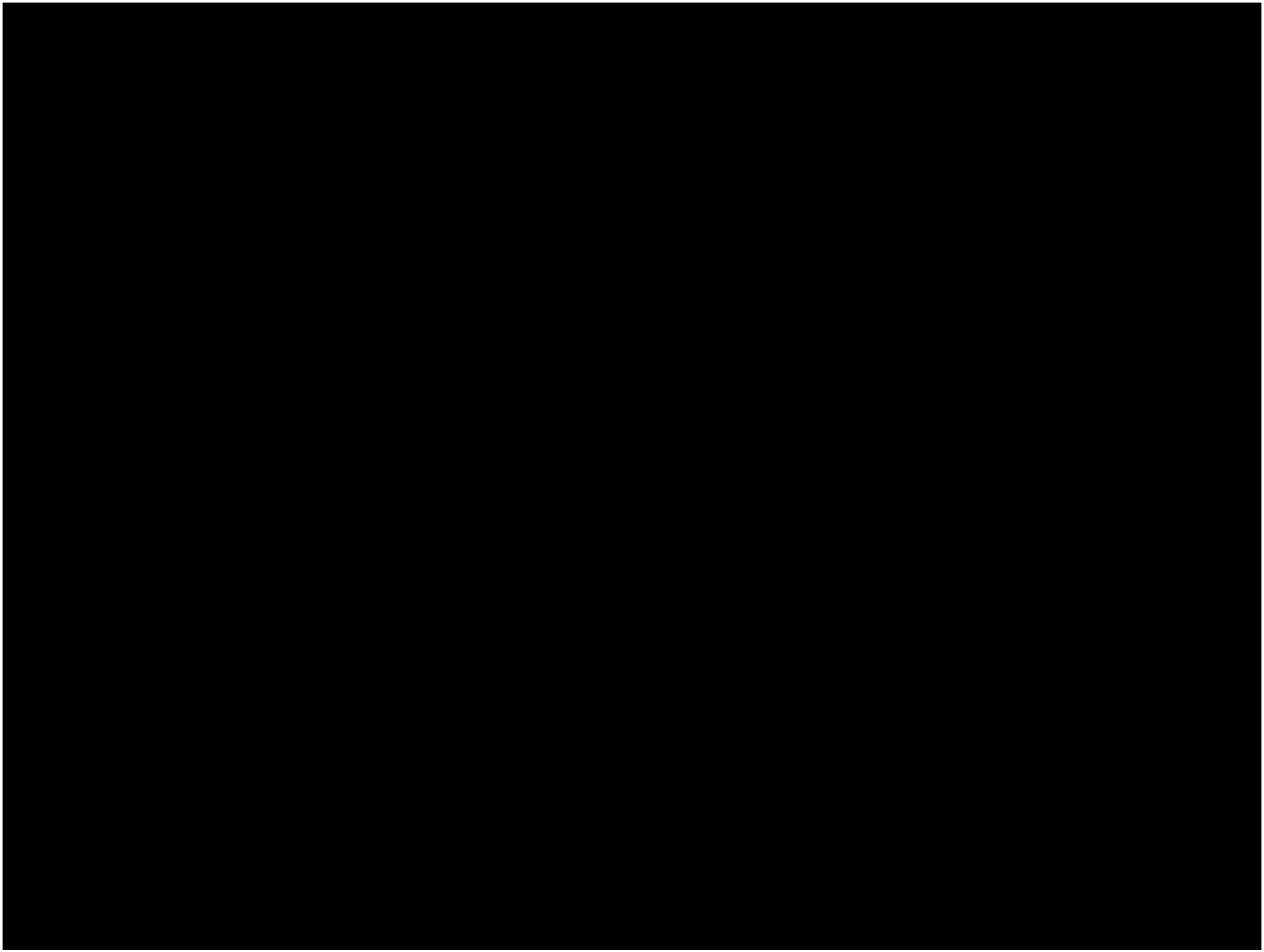


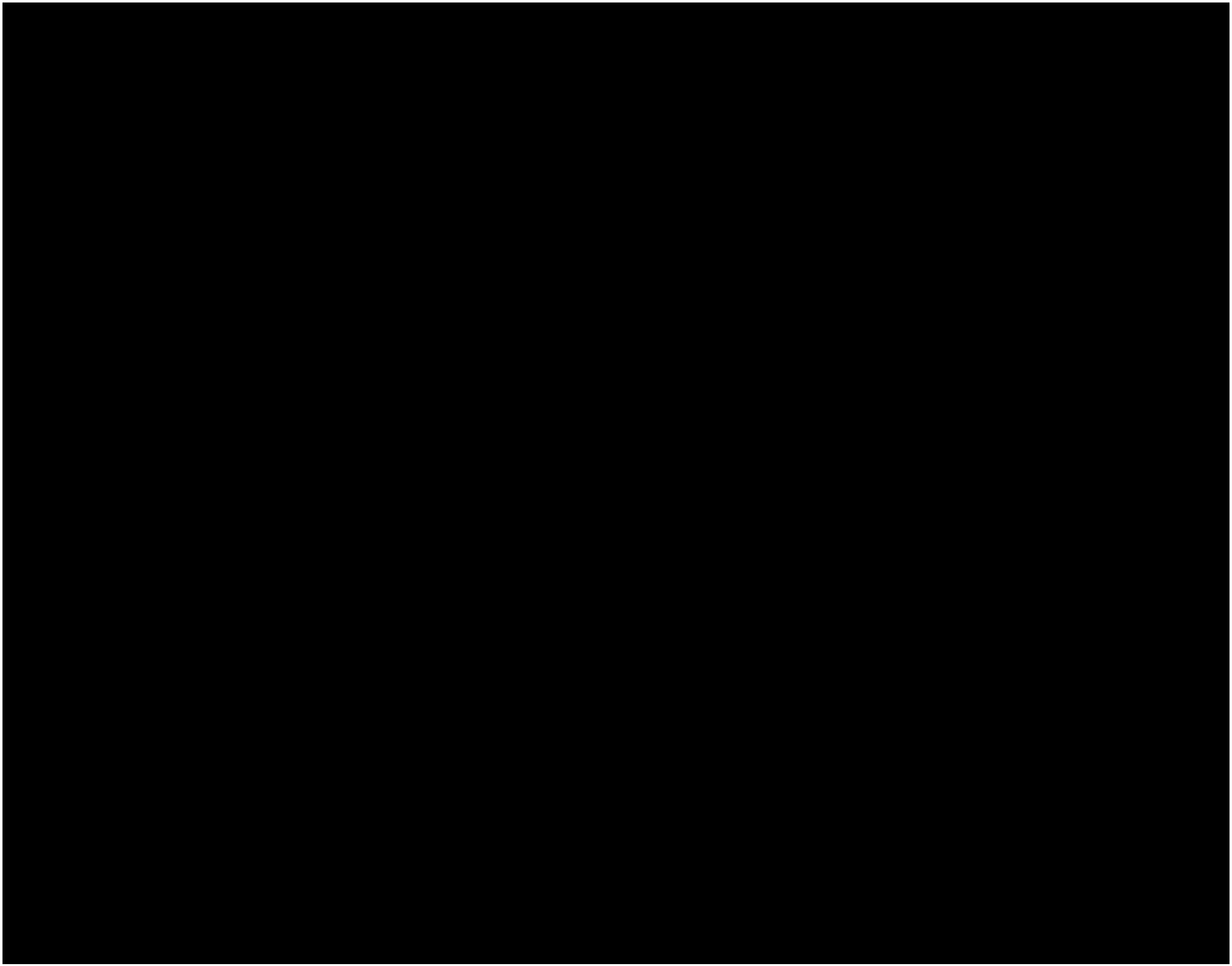


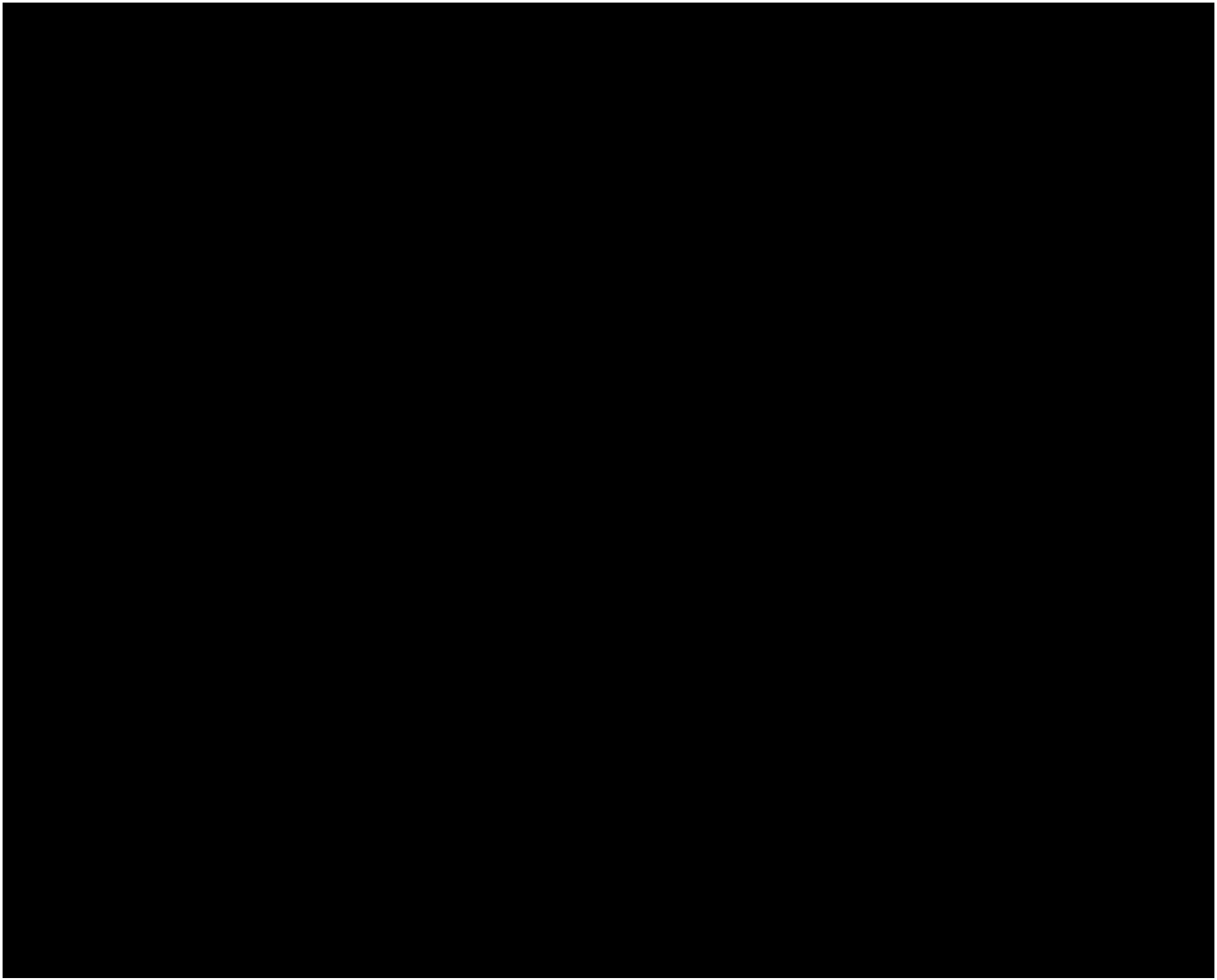


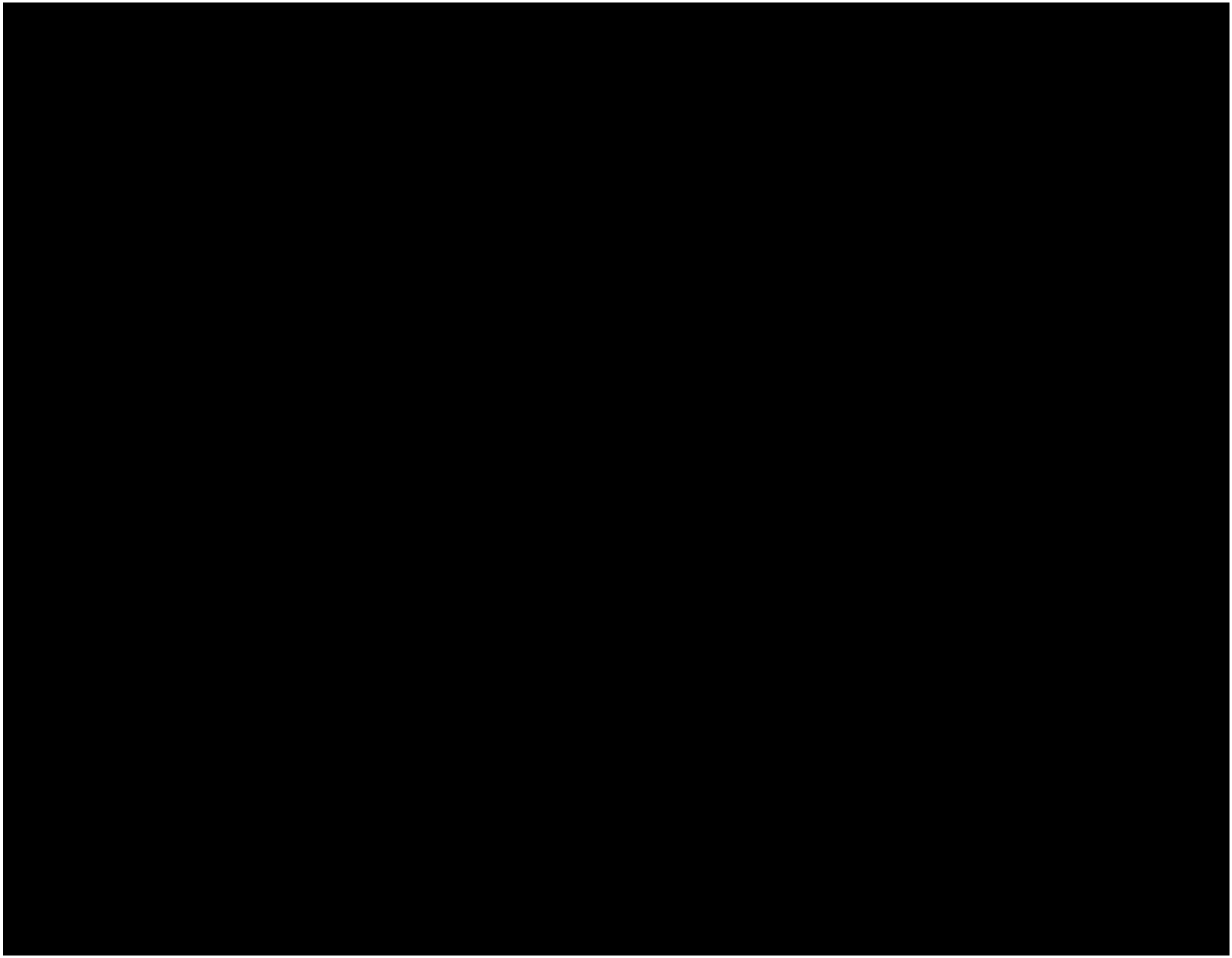


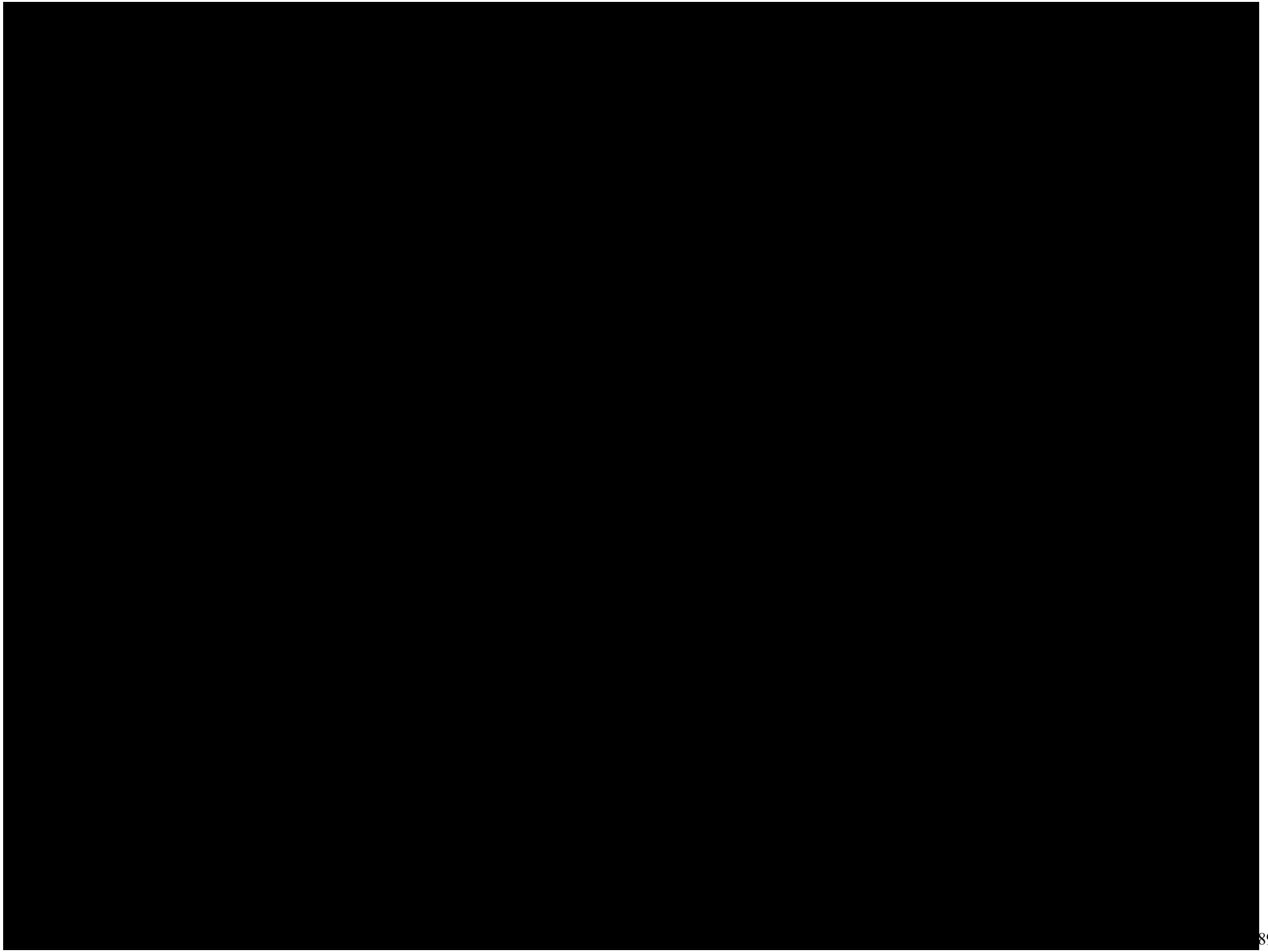


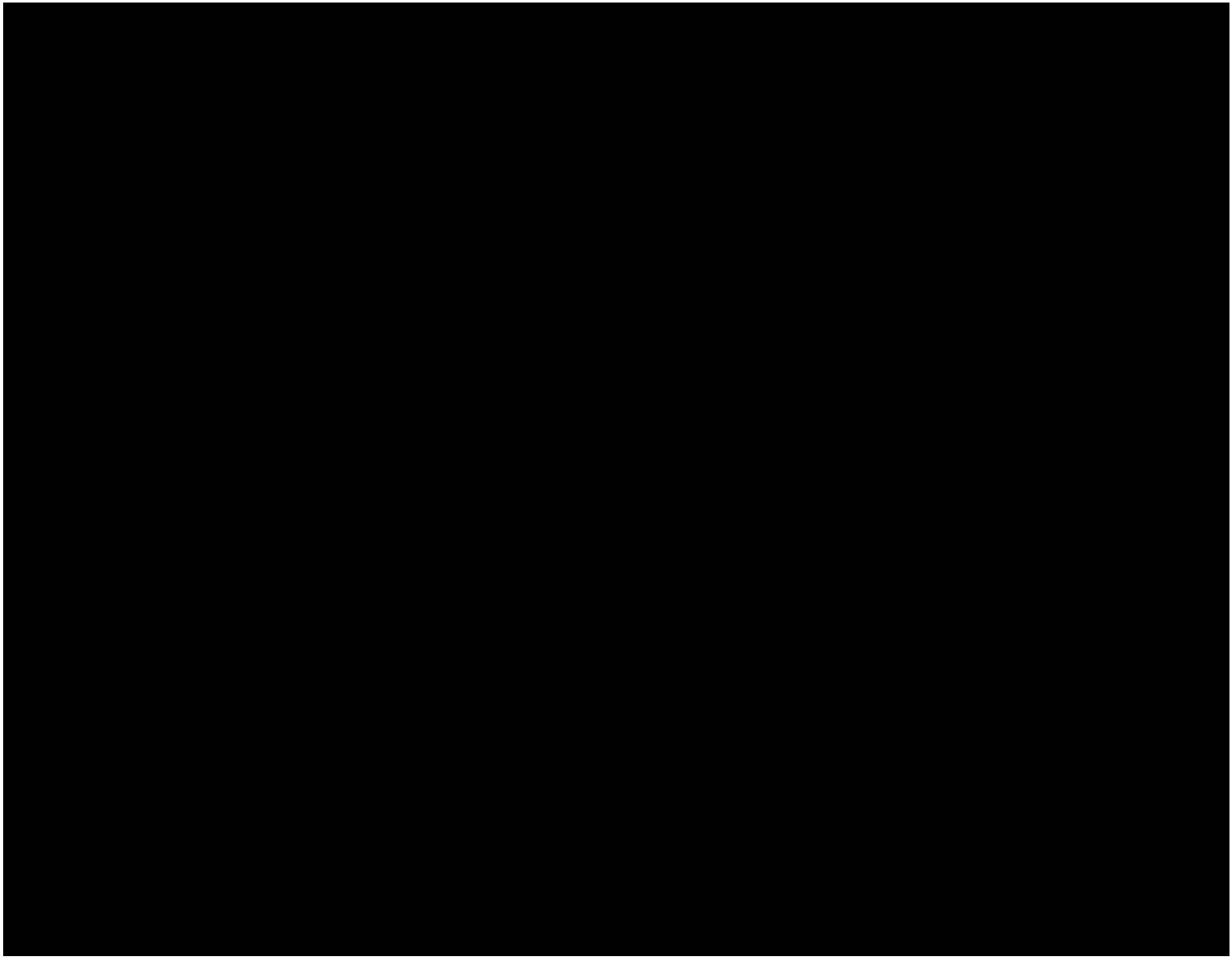












Signature Page for [redacted]

Approval Task	[redacted] 07-Feb-2025 01:34:24 GMT+0000

Approval Task	[redacted] 07-Feb-2025 01:48:08 GMT+0000

Approval Task	[redacted] 07-Feb-2025 01:50:17 GMT+0000

Signature Page for [redacted]