Official Title:	A Phase 3 Global, Multicenter, Double-Blind Randomized Study of Carboplatin-Paclitaxel With INCMGA00012 or Placebo in Participants With Inoperable Locally Recurrent or Metastatic Squamous Cell Carcinoma of the Anal Canal Not Previously Treated With Systemic Chemotherapy (POD1UM-303/InterAACT 2)
NCT Number:	NCT04472429
Document Date:	INCMGA 0012-303 Statistical Analysis Plan Am 1: 08 NOV 2023

Statistical Analysis Plan

INCMGA 0012-303

A Phase 3 Global, Multicenter, Double-Blind Randomized Study of Carboplatin-Paclitaxel With INCMGA00012 or Placebo in Participants With Inoperable Locally Recurrent or Metastatic Squamous Cell Carcinoma of the Anal Canal Not Previously Treated With Systemic Chemotherapy (POD1UM-303/InterAACT 2)

IND Number:	**139,333**
EudraCT Number:	**2020-000826-24**
Sponsor:	**Incyte Corporation** **1801 Augustine Cut-Off** **Wilmington, DE 19803** **United States**
Protocol Version:	**Protocol Amendment 2 dated 25 MAY 2022**
CRF Approval Date:	**17 JAN 2023**
SAP Version:	**Amendment 1**
SAP Author:	[redacted] **Principal Biostatistician**
Date of Plan:	**08 NOV 2023**

This study is being conducted in compliance with Good Clinical Practice, including the archiving of essential documents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Term
ADA	antidrug antibody
ADSL	subject-level analysis dataset
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine transaminase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC_{0-t}	area under the plasma or serum concentration-time curve from time = 0 to the last measurable concentration at time = t
AUS	Australia
BICR	blinded independent central radiographic review
BMI	body mass index
BSA	body surface area
CAS	crossover analysis set
CI	confidence interval
C_{max}	maximum observed plasma or serum concentration
CMH	Cochran–Mantel–Haenszel
C_{min}	minimum observed plasma or serum concentration over the dose interval
COVID-19	coronavirus disease 2019
CR	complete response
CrCl	creatinine clearance
CRF	case report form
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DCR	disease control rate
DMC	Data Monitoring Committee
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
ELISA	enzyme-linked immunosorbent assay
EORTC-QLQ-C30	European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire
EQ-5D	EuroQol-5D

Abbreviation	Term
EQ-VAS	EuroQol Visual Analogue Scale
EU	Europe
FAS	full analysis set
FDA	Food and Drug Administration
GFR	glomerular filtration rate
H_0	null hypothesis
H_1	alternative hypothesis
HIV	human immunodeficiency virus
HPV	human papillomavirus
HR	hazard ratio
HR-PRO	health-related patient reported outcome
ICF	informed consent form
irAE	immune-related adverse event
iRECIST	modified RECIST for immune-based therapeutics
IRR	infusion-related reaction
IV	intravenous(ly)
MedDRA	Medical Dictionary for Regulatory Activities
NA	North America
NCI	National Cancer Institute
NE	not evaluable
ODWG	Organ Dysfunction Working Group
ORR	objective response rate
ORR-CO	objective response rate – crossover period
OS	overall survival
PD	progressive disease
PD-L1	programmed death-ligand 1
PFS	progression-free survival
PFS2	progression-free survival 2
PK	pharmacokinetic(s)
PR	partial response
PRO	patient reported outcome
PT	preferred term
Q4W	every 4 weeks
QLQ-ANL27	Quality of Life Questionnaire for Anal Cancer
QOL	quality of life
QTc	corrected QT interval

Abbreviation	Term
QTcF	QT interval corrected by Fridericia
QTcB	QT interval corrected by Bazett
RECIST	Response Evaluation Criteria in Solid Tumors
ROW	rest of world
SAP	Statistical Analysis Plan
SCAC	squamous cell carcinoma of the anal canal
SD	stable disease
SOC	system organ class
TEAE	treatment-emergent adverse event
t_{max}	time to maximum concentration
T/N/M	tumor/node/metastasis
UK	United Kingdom
ULN	upper limit of normal
WHO	World Health Organization

1. INTRODUCTION

This study is a Phase 3 global, multicenter, placebo-controlled double-blind randomized study that will enroll participants with inoperable locally recurrent or metastatic SCAC not previously treated with systemic chemotherapy. Section 2 of the Protocol provides a detailed description of the investigational product, target patient population, rationale for doses to be examined, and potential risks and benefits of treatment with retifanlimab. The study planned to enroll approximately [redacted] participants in a 1:1 randomization between retifanlimab and placebo. The primary endpoint is PFS, and the key secondary endpoint OS will follow gatekeeping procedure on type I error spending. Participants may receive study drug for up to [redacted] administrations of study drug), and those participants who received placebo in combination with chemotherapy and who experience documented disease progression verified by BICR will have the opportunity to receive retifanlimab in the Crossover Period.

The purpose of this SAP is to provide details of the statistical analyses that have been outlined in the INCMGA 0012-303 Protocol. The scope of this plan includes the interim and final analyses that will be executed by the Department of Biostatistics or designee. The PK, ADA, and pharmacodynamics analyses will be executed by the Department of Clinical Pharmacokinetics and the Department of Translational Sciences.

2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS

2.1. Protocol and Case Report Form Version

This SAP is based on INCMGA 0012-303 Protocol Amendment 2 dated 25 MAY 2022 and CRFs approved on 17 JAN 2023. Unless superseded by an amendment, this SAP will be effective for all subsequent Protocol amendments and eCRF versions.

2.2. Study Objectives and Endpoints

Table 1 presents the objectives and endpoints.

Table 1: Objectives and Endpoints

Objectives	Endpoints
Primary	
To compare the efficacy of carboplatin-paclitaxel with retifanlimab versus carboplatin-paclitaxel with placebo in participants with inoperable locally advanced or metastatic SCAC not previously treated with systemic chemotherapy.	PFS, defined as the time from the date of randomization until disease progression according to RECIST v1.1 by BICR or death due to any cause.
Key Secondary	
To compare the efficacy of carboplatin-paclitaxel with retifanlimab versus carboplatin-paclitaxel with placebo in participants with inoperable locally advanced or metastatic SCAC not previously treated with systemic chemotherapy.	OS, defined as the time from the date of randomization until death due to any cause.
Secondary	
To determine additional measures of clinical benefit.	ORR, defined as the percentage of participants having a CR or PR, according to RECIST v1.1 as determined by BICR.
	DOR, defined as the time from the first documented response (CR or PR) according to RECIST v1.1 until disease progression as determined by BICR or death due to any cause.
	DCR, defined as the number of participants maintaining either an ORR or stable disease according to RECIST v1.1 as determined by BICR.
To evaluate the safety of carboplatin-paclitaxel with retifanlimab or placebo in participants with SCAC not previously treated with systemic chemotherapy.	Number of participants experiencing AEs and number of participants discontinuing study drug due to AEs.
To determine the PK of retifanlimab when administered with carboplatin-paclitaxel to participants with SCAC not previously treated with systemic chemotherapy.	Population PK, including C_{max}, t_{max}, C_{min}, and AUC_{0-t}, will be summarized.

Table 1: Objectives and Endpoints (Continued)

Objectives	Endpoints
Exploratory	
[redacted]	

3. STUDY DESIGN

3.1. Overall Study Design

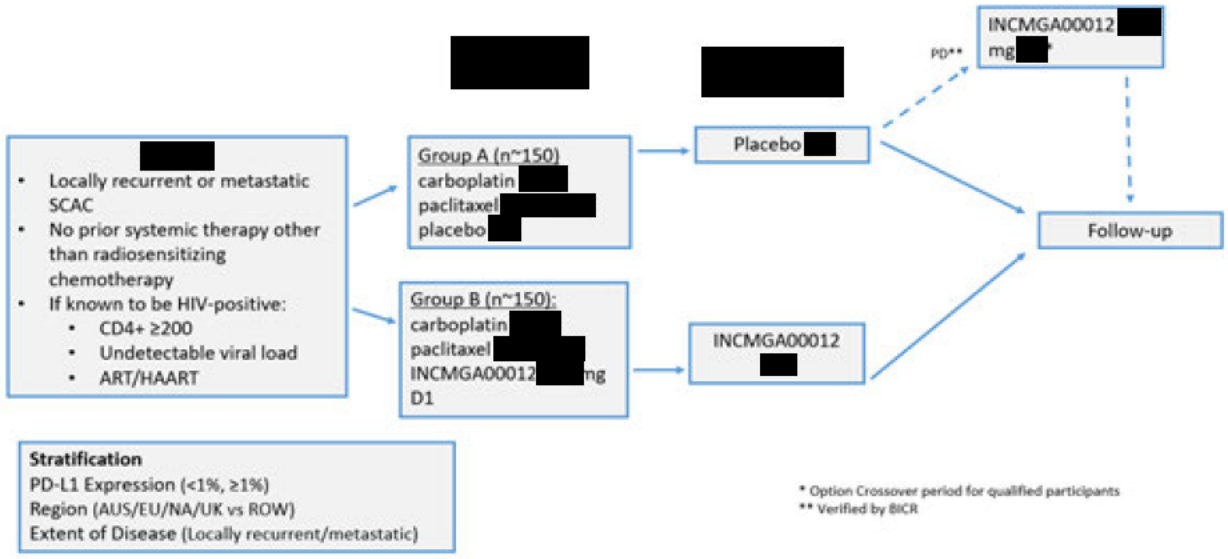
This study is a Phase 3 global, multicenter, placebo-controlled, double-blind, randomized study that will enroll participants with inoperable locally recurrent or metastatic SCAC not previously treated with systemic chemotherapy. Participants with well-controlled HIV infection will be eligible. Participants will receive up to [redacted] induction cycles ([redacted] weeks) of carboplatin ([redacted] on Day 1) and paclitaxel ([redacted] mg/m^2 on Days [redacted] concurrently with retifanlimab or placebo. Retifanlimab [redacted] mg Q4W or placebo will be administered for up to [redacted] cycles ([redacted]) in the absence of disease progression, intolerable toxicity, death, withdrawal of consent, lost to follow-up, or premature discontinuation of all assigned study drug administration for any other reason. Approximately [redacted] participants will be randomized. Randomization to retifanlimab or placebo will be 1:1 and blinded. Stratification factors are presented in Table 2. For the stratification category of PD-L1 status, participants with tumor results of PD-L1 negative and nonevaluable are grouped in the PD-L1 < 1% stratum. Participants with nonevaluable tumors for PD-L1 status will be capped at 10% of the total population.

Table 2: Stratification Factors

Stratification Factor Category	Stratum
PD-L1 expression	< 1%, ≥ 1%
Extent of disease	Locally recurrent, metastatic
Region	AUS/EU/NA/UK, ROW

The study consists of 4 periods: screening, study drug administration, crossover, and follow-up. The date of Cycle 1 Day 1 is the date of administration of the first dose of retifanlimab or placebo. Figure 1 presents the study design schema.

Figure 1: Study Design Schema



Crossover will be allowed for participants who received placebo in combination with chemotherapy (following BICR verification of disease progression) and meet eligibility criteria for the crossover period.

3.2. Randomization

Participants will be randomized centrally in a 1:1 ratio to receive retifanlimab or placebo along with carboplatin and paclitaxel concurrently. These participants will be stratified by PD-L1 expression level (< 1%, ≥ 1%), extent of disease (locally recurrent, metastatic), and region (AUS/EU/NA/UK, ROW). As this is a randomized, double-blind, placebo-controlled study, neither the investigators nor the sponsor will be aware of the treatment to which a participant is randomized until the crossover period.

3.3. [redacted]

[redacted]

3.4. [redacted]

[redacted]

3.5. Schedule of Assessments

Refer to Protocol Amendment 2 dated 25 MAY 2022 for a full description of all study procedures and assessment schedules for this study.

4. DATA HANDLING DEFINITIONS AND CONVENTIONS

4.1. Scheduled Study Evaluations and Study Periods

4.1.1. Study Drug, Study Treatment and Treatment Groups

Retifanlimab or placebo is the study drug in this study.

Retifanlimab, placebo, and chemotherapy (ie, carboplatin-paclitaxel) are all considered as study treatments in this study.

There are 2 treatment groups during the randomized period:

- Group A: Placebo + Chemotherapy
- Group B: Retifanlimab + Chemotherapy

and 1 treatment group during the crossover period:

- Group C is open-label retifanlimab monotherapy.

4.1.2. Study Period

There are 3 periods in this study. The randomized period starts with randomization and continues until the end of the treatment period for Group B or until the first dose of crossover treatment for Group A. The crossover period starts on the first date of crossover treatment and is only applicable to Group A participants who crossed over to retifanlimab treatment, if eligible as per Protocol Section 8.7. The overall period is the same as the randomized period for participants without crossover treatment (ie, Group B), but for participants with crossover treatment (ie, Group A), this period covers from the first date of the randomized period until the last date of crossover period.

4.1.3. Day 1

Day 1 of the safety endpoints for the randomized period is the date the first dose of any component of the study treatment is administered to a participant. Day 1 of other endpoints for the randomized period is the date of randomization. Day 1 for the crossover period is the date of the first dose of crossover treatment. Day 1 for the overall period is the same as the randomized period.

4.1.4. Study Day

Study day will be separately defined for the randomized, crossover, and overall periods following the same definition; however, it will refer to a different Day 1 associated with each period/endpoint (ie, Day 1 of the safety endpoints during the randomized period/overall period, Day 1 of the nonsafety endpoints during the randomized period/overall period, and Day 1 of the crossover period).

If a visit/reporting date is on or after Day 1, then the study day at the visit/reporting date will be calculated as

Day # = (visit/reporting date – Day 1 date + 1).

If the visit/reporting date is before Day 1, then the study day at the visit/reporting date will be calculated as

Day # = (visit/reporting date – Day 1 date).

A study day of –1 indicates 1 day before Day 1.

4.1.5. Baseline Value

Baseline is the last nonmissing assessment obtained prior to or on Day 1 for each period and endpoint. Baseline of the safety endpoints for the randomized/overall period is the last nonmissing assessment obtained before the first administration of study treatment (ie, Day 1 of the safety endpoint). Baseline of the other endpoints for the randomized/overall period is the last nonmissing assessment obtained prior to or on the randomization date (ie, Day 1 of the nonsafety endpoints). Baseline for the crossover period is the last nonmissing assessment obtained before the first administration of crossover treatment (ie, Day 1 of the crossover period).

When scheduled assessments and unscheduled assessments occur on the same day and time of the assessment or time of first dose is not available, use the following conventions to determine baseline:

- If both a scheduled and an unscheduled visit are available on the day of the first dose and the time is missing, use the scheduled assessment as baseline.
- If all scheduled assessments are missing on the day of the first dose and an unscheduled assessment is available, use the unscheduled assessment as baseline.

4.1.6. [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

4.1.7. Cycle Length and Duration

One cycle is defined as 28 days (4 weeks) for participants in this study. Cycle 1 Day 1 is the day of the first infusion of either retifanlimab or placebo. Day 1 of subsequent cycles will correspond with the infusion date of retifanlimab or placebo as reported in the dataset. Scheduled visits after Cycle 1 Day 1 will have a ± 3-day window. Tumor assessments will be performed every 8 weeks (56 days) with a ± 7-day window during study drug administration, crossover, and/or disease follow-up regardless of any treatment delays, until second disease progression.

4.1.8. [redacted]

[redacted]

[redacted]

4.2. Variable Definitions

The following variables will only be calculated if not reported on the eCRF.

4.2.1. Body Mass Index

Body mass index will be calculated as follows:

BMI (kg/m^2) = [weight (kg)] / [height (m)]^2.

4.2.2. Body Surface Area

Body surface area will be calculated based on the Mosteller (1987) formula as follows:

BSA (m^2) = {[weight (kg) × height (cm)] / 3600}^{½}.

Sites will also record the BSA calculated per institutional standards.

4.2.3. Prior and Concomitant Medication

Prior medication is defined as any nonstudy medication started before the randomization date.

Concomitant medication is defined as any nonstudy medication that is started accordingly:

- Before the randomization date and is ongoing throughout the study or ends on/after the date of first study treatment administration.
- On/after the randomization date and is ongoing or ends during the course of the study.

A prior medication could also be classified as "both prior and concomitant medication" if the end date is on or after the randomization date. In the listing, it will be indicated whether a medication is only prior, only concomitant, or both prior and concomitant.

For the purposes of analysis, all medications will be considered concomitant medications unless the medications can unequivocally be defined as not concomitant, that is, if the start and end dates are all missing, then the medication will be considered concomitant medication.

5. STATISTICAL METHODOLOGY

5.1. General Methodology

Unless otherwise noted, SAS® software (SAS Institute Inc, Cary, NC; v9.4 or later) will be used for the generation of all tables, figures, and statistical analyses. Descriptive summaries for continuous variables will include, but not be limited to, the number of observations, mean, standard deviation, median, minimum, and maximum. Descriptive summaries for categorical variables will include the number and percentage of participants in each category.

Interim analyses for OS endpoints are planned for this study as defined in Section 10.

5.2. Treatment Groups

This is a placebo-controlled, double-blind, randomized study with an optional crossover period for randomized placebo-treated participants. Participants will be summarized by treatment group. The 2 treatment groups during the randomized period are defined as follows:

- Group A: placebo + carboplatin + paclitaxel.
- Group B: retifanlimab + carboplatin + paclitaxel.

The treatment group during the crossover period is defined as Group C (retifanlimab).

5.3. Analysis Populations

5.3.1. Full Analysis Set

The FAS will include all randomized participants. According to the intention-to-treat principle, participants will be analyzed according to the treatment and strata they have been assigned during randomization. The FAS will be used for the summary of demographics, baseline characteristics, participant disposition, and for all efficacy analyses.

5.3.2. Per-Protocol Population

Not applicable.

5.3.3. Safety Population

The safety population will include all randomized participants who received at least 1 dose of study treatment. Treatment groups for this population will be determined according to the actual treatment the participant received regardless of treatment assignment at randomization. All safety analyses will be conducted using the safety population.

5.3.4. Pharmacokinetic Evaluable Population

The PK evaluable population will include all participants who received at least 1 dose of study treatment and provided at least 1 postdose plasma sample.

5.3.5. Translational Evaluable Population

The translational evaluable population will include all participants who received at least 1 dose of study drug and provided at least 1 postdose sample.

5.3.6. Crossover Analysis Set

The CAS will include all participants randomized to Group A, who receive at least 1 dose of placebo and then crossover and receive at least 1 dose of retifanlimab. The CAS will be used for all analyses during crossover period.

6. BASELINE, EXPOSURE, AND DISPOSITION

Appendix A provides a list of data displays. Sample data displays are included in a separate document.

6.1. Demographics, Baseline Characteristics, and Disease History

6.1.1. Demographics and Baseline Characteristics

Demographic and baseline characteristics will be listed in detail. The following demographics and baseline characteristics will be summarized for the FAS: age, sex, race, ethnicity, body weight, height, BMI, BSA, region (AUS/EU/NA/UK, ROW). Qualitative data will be summarized by contingency tables; quantitative data will be summarized by descriptive summary statistics.

6.1.2. Baseline Disease Characteristics and Disease History

The time since initial diagnosis; time since current diagnosis; time since inoperable locally recurrent/metastatic diagnosis; tumor stage at initial diagnosis; PD-L1 expression ($< 1\%$, $\geq 1\%$); microsatellite instability status; HPV status; p16/INK4A tumor status; HIV infection status; ECOG performance status; current disease staging (overall, T/N/M staging); current site of disease; and extent of disease (eg, locally recurrent, metastatic) will be summarized for all participants in the FAS.

Time since diagnosis will be calculated as follows:

Time since diagnosis (years) = (date of randomization – date of diagnosis + 1) / 365.25.

6.1.3. Prior Therapy

The number and percentage of participants recording prior radiation or prior therapy for SCAC will be summarized in the FAS. Detailed information regarding prior radiation and prior surgery will be in 2 separate listings.

6.1.4. Medical History

Medical history will be summarized for all participants in the FAS; this summary will be presented by primary SOC and PT.

6.2. Disposition of Participants

The number and percentage of participants who were randomized, who were treated, who were ongoing with treatment, who discontinued study treatment with a primary reason for discontinuation, who were still in the study, and who withdrew from the study with a primary reason for withdrawal will be summarized for all participants in the FAS for the randomized and crossover periods. The number of participants randomized by country and site will also be provided by treatment group.

6.3. Protocol Deviations

Protocol deviations will be presented in the participant data listings and categorized by clinical review. Protocol deviations will be summarized by treatment group descriptively. Deviations related to COVID-19 will be listed by treatment group.

6.4. Exposure

For participants in the safety population, exposure will be summarized by treatment group, descriptively, for the randomized/crossover period as follows:

- **Total number of infusions of retifanlimab/placebo:** total number of infusions of retifanlimab/placebo per participant with an infusion.
- **Total number of infusions of carboplatin or paclitaxel:** total number of infusions of carboplatin or paclitaxel per participant with an infusion.
- **Total dose administered of retifanlimab/placebo (mg):** Accumulated actual administered dose of retifanlimab/placebo at each infusion.

  For each infusion:

  If the entire infusion is administered, then the actual administered dose is [redacted] mg.

  If the infusion was not completely administered, then the actual administered dose is the planned dose × (the estimated volume delivered / prepared volume).
- **Total dose administered of carboplatin or paclitaxel (mg):** Accumulated actual administered dose of chemotherapy at each infusion.

  For each infusion:

  If the entire infusion was administered, then the actual administered dose is the administered dose.

  If the infusion was not completely administered, then the actual administered dose is the planned dose × (the estimated volume delivered / prepared volume).
- **Total dose prescribed of retifanlimab/placebo (mg):** accumulated prescribed dose of retifanlimab/placebo at each infusion.

- **Total dose prescribed of carboplatin or paclitaxel (mg):** accumulated prescribed dose of carboplatin or paclitaxel at each infusion.
- **Average dose of retifanlimab/placebo (mg):** total dose administered of retifanlimab/placebo (mg) / total number of infusions retifanlimab/placebo.
- **Average dose of carboplatin or paclitaxel (mg):** total dose administered of carboplatin or paclitaxel (mg) / total number of infusions of carboplatin or paclitaxel.
- **Duration of treatment of retifanlimab/placebo (months):** (date of last dose of retifanlimab/placebo – date of first dose of retifanlimab/placebo + 1) / 30.4375.
- **Duration of treatment of carboplatin or paclitaxel (months):** (date of last dose of carboplatin or paclitaxel – date of first dose of carboplatin or paclitaxel + 1) / 30.4375.
- **Compliance of retifanlimab/placebo (%):** 100 × total dose administered of retifanlimab/placebo / total dose prescribed of retifanlimab/placebo.
- **Compliance of carboplatin or paclitaxel (%):** 100 × total dose administered of carboplatin or paclitaxel / total dose prescribed of carboplatin or paclitaxel.

Infusion information collected in the eCRF will be listed. Dose delay and temporary infusion interruption of retifanlimab/placebo and carboplatin or paclitaxel, may be summarized/listed by treatment group, as needed.

6.5. Prior and Concomitant Medication and Post-Treatment Therapy

Prior medications and concomitant medications will be coded using the WHO Drug Dictionary. The number and percentage of participants in the FAS for each prior and concomitant medication will be summarized by WHO drug class and WHO drug term. For the summary of concomitant medication, only on–treatment medications will be included. Medications with missing start/end dates will be considered as concomitant medication in the summary. Post-treatment anticancer therapy will be summarized. Other medications will be provided in the listing. Drugs intended to manage irAEs, as well as prophylaxis/premedication used to prevent infusion reactions, may be summarized separately. Procedures and nondrug therapies will also be summarized/listed per CRF.

7. EFFICACY

Appendix A provides a list of data displays. Sample data displays are included in a separate document.

7.1. General Considerations

Efficacy endpoints of this study include PFS, ORR, DOR, and DCR assessed by BICR according to RECIST v1.1 and OS. Listings of response assessment at each visit will be provided. Summary statistics will be summarized by treatment group, while statistical testing on PFS and OS will be a stratified log-rank test on treatment group level, taking PD-L1 expression, extent of disease, and region as stratification factors.

7.2. Analysis for the Primary Efficacy Parameters

7.2.1. Response Criteria

Objective response will be categorized using RECIST v1.1 by BICR or the investigators. Participants will have their objective response evaluated as CR, PR, SD, PD, or NE at each postbaseline radiological assessment based on changes in target lesions, nontarget lesions, and appearance of new lesions.

7.2.2. Primary Efficacy Analysis

Progression-free survival is defined as the time from the date of randomization to the date of the first documented progression or death due to any cause. Progression-free survival will be assessed via BICR according to RECIST v1.1.

The primary efficacy analysis will compare the median PFS between the 2 treatment groups.

The following statistical hypotheses will be tested to address the primary efficacy objective:

$$H_0: PFS_B \leq PFS_A \text{ versus } H_1: PFS_B > PFS_A,$$

where PFS_A and PFS_B are the median PFSs for Group A and Group B, respectively.

The stratified log-rank test will be used to compare PFS between treatment groups in the FAS at a 1-sided 2.5% level of significance, stratified for region, PD-L1 expression, and extent of disease. The strata identified in the randomization process will be used for the analysis.

Progression-free survival will be analyzed using a stratified log-rank test at an overall 1-sided 2.5% level of significance in the FAS population; this will be done according to the treatment group participants were randomized and the strata they were assigned during randomization. The HR and its 95% CI will be estimated based on the stratified Cox regression model using the same stratification factors as for the log-rank test with Efron's method (1977) accounting for ties in event times.

Kaplan-Meier curves for PFS will be presented by treatment group. The Kaplan-Meier estimate of median PFS will be presented with its 95% CI. The 95% CI will be calculated using the generalization of Brookmeyer and Crowley's method (1982) with log-log transformation (Klein and Moeschberger 1997). Censoring for PFS will follow the algorithm outlined in Table 3, which is based on FDA Guidance (2015). Date of death will be determined using the Death Report, Survival Follow-Up, and Participant Status eCRFs.

Since PFS is the only primary endpoint, and there is no IA planned for PFS analysis, the study will be unblinded at the time of PFS analysis when around 207 PFS events are reached.

Table 3: Evaluation and Censoring of Progression-Free Survival

Situation	Outcome	Date of Progression or Censoring
No baseline tumor assessments	Censored	Day 1 (date of randomization)
No valid postbaseline response assessments in the absence of death prior to first scheduled tumor assessment	Censored	Day 1 (date of randomization)
Progression documented between scheduled response assessments	Progressed	Date of first objective response of PD
No progression	Censored	Date of last valid radiologic assessment (not NE or missing)
Study withdrawal for undocumented progression	Censored	Date of last valid radiologic assessment (not NE or missing)
Study withdrawal for toxicity or other reason	Censored	Date of last valid radiologic assessment (not NE and not missing)
New anticancer treatment started	Censored	Date of last valid radiologic assessment (not NE or missing) on/before starting a new anticancer treatment
Death before first progressive response assessment	Progressed	Date of death
Death between adequate response assessments	Progressed	Date of death
Death or documented progression after missing 2 or more consecutive scheduled tumor assessments	Censored	Date of last valid radiologic assessment (not NE or missing) prior to missed assessments

7.2.3. Subgroup Analyses for the Primary Endpoint

Subgroup analyses will be performed on the following based on the participant's baseline status:

- Sex: male, female
- Baseline ECOG performance status: 0 versus ≥ 1
- Age: < 65 years versus ≥ 65 years and < 75 years versus ≥ 75 years
- Race: Caucasian, Asian, other
- Ethnicity: Hispanic or Latino, Not Hispanic or Latino, unknown
- HIV status: positive, negative, or unknown
- Liver metastatic: yes, no
- PD-L1 assessment: < 1%, ≥ 1%
- Extent of disease: locally recurrent, metastatic
- Region: AUS/EU/NA/UK, ROW
- HPV status: positive, negative, or unknown

Subgroup analyses will only be performed if at least 5 participants are present in each subgroup. Some grouping of classes will be considered if there are too few participants in some subgroups.

Efficacy analyses in subgroups will generally be exploratory and are intended to explore the intrinsic consistency of any treatment effects found overall.

Subgroup analyses of the primary endpoint (PFS) will be performed on the FAS by presenting the point estimates of HR with its 95% CI of Group B versus Group A under the stratified Cox model in the subgroup. Summary tables and forest plots may be presented.

7.2.4. Sensitivity and Supportive Analyses for the Primary Endpoint

Progression-free survival, defined as the time from the date of randomization to the date of the first documented progression assessed via investigator or death due to any cause, may be used for supportive analysis of the primary endpoint.

The following can be considered as sensitivity analysis:

The unstratified log-rank test for PFS may also be considered as a sensitivity analysis.

In addition, study withdrawal due to clinical progression and death or documented progression by BICR after more than 1 missed scheduled tumor assessment may be considered events for PFS in the sensitivity analysis as illustrated in Table 4.

Table 4: Sensitivity Analysis on Evaluation and Censoring of Progression-Free Survival

Situation	Outcome	Date of Progression or Censoring
No baseline tumor assessments	Censored	Day 1 (date of randomization)
No valid postbaseline response assessments in the absence of death prior to first scheduled tumor assessment	Censored	Day 1 (date of randomization)
Progression documented between scheduled response assessments	Progressed	Date of first objective response of PD
No progression	Censored	Date of last valid radiologic assessment (not NE or missing)
Study withdrawal for undocumented progression for clinical progression	Progressed	Date of clinical progression
Study withdrawal for undocumented progression for reasons other than clinical progression	Censored	Date of last valid radiologic assessment (not NE or missing)
New anticancer treatment started	Censored	Date of last valid radiologic assessment (not NE or missing) on/before starting a new anticancer treatment
Death before first progressive response assessment	Progressed	Date of death
Death between adequate response assessments	Progressed	Date of death
Death or documented progression after missing 2 or more consecutive scheduled tumor assessment	Progressed	Date of progression or date of death

7.3. Key Secondary Efficacy Parameters

7.3.1. Overall Survival

Overall survival is the key secondary endpoint and is defined as the time from the date of randomization to the date of death due to any cause. The key secondary efficacy analysis is the comparison of median OS between the 2 treatment groups.

The following statistical hypotheses will be tested to address the key secondary efficacy objective:

H_0: $OS_B \leq OS_A$ versus H_1: $OS_B > OS_A$,

where OS_A and OS_B are the median OS in Group A and Group B, respectively.

The efficacy analysis of OS comparing the 2 treatment groups in the FAS will be by a stratified log-rank test at an overall 1-sided 2.5% level of significance stratified in the same manner as PFS (see Section 7.2.2). Kaplan-Meier curves, medians, and 95% CIs of the median OS will be

presented for each treatment group. The HR for OS will be calculated, along with its 95% CI, from a stratified Cox model using the same stratification factors as the log-rank test (ie, with Efron's likelihood approximation to account for ties in event times). The analysis of OS will follow gatekeeping procedures, that is, hypothesis testing for OS will only occur if the hypothesis for PFS is rejected. An interim analysis of OS will be performed at the time of the PFS analysis (see Section 10), when unblinding will occur. The final analysis of OS will be performed when there are at least [redacted] death events observed. Summary statistics of the OS endpoint will still be provided.

7.4. Analysis of Additional Secondary Efficacy Parameters

7.4.1. Objective Response Rate

Objective response rate is defined as the percentage of participants with CR or PR at any postbaseline visit before the first PD or new anticancer therapy, according to RECIST v1.1 (Eisenhauer et al 2009) as determined by BICR. The analysis of ORR will be based on the FAS.

The ORR and its exact 95% CI for each treatment group will be presented. Overall response rate will be compared between treatment arms using the stratified CMH test. The 95% CI for odds ratio between the 2 randomized treatment groups will be computed using the normal approximation to the binomial distribution. If both PFS and OS cross the boundary, ORR will be tested at 1-sided 2.5% level of significance, otherwise, the p-value from the CMH test will be presented as the nominal p-value without Type-1 error allocated for the testing. The ORR by investigator assessment may be provided. A waterfall plot of best percent change from baseline of sum of diameters in target lesions might be generated.

7.4.2. Duration of Response

Duration of response is defined as the time from first-documented response (CR or PR, determined by BICR) that is subsequently confirmed to the time of first-documented disease progression per RECIST v1.1 or death due to any cause. Analysis of DOR will be based on the FAS and will only be summarized for participants who responded to the treatment. If a participant does not have an event, DOR is censored at the date of the last adequate tumor assessment before data cutoff or new anticancer therapy, following the same algorithm as censoring of PFS (see Table 3). The Kaplan-Meier estimate of the distribution function will be constructed for DOR. The estimated median, along with 95% CIs, will be reported. Duration of response by investigator assessment may be provided. A swimmer plot for DOR may be generated.

7.4.3. Disease Control Rate

Disease control rate is defined as the proportion of participants with an objective response (CR or PR) or SD per RECIST v1.1, according to BICR. The DCR will be estimated, and the exact 95% CI will be reported.

7.5. [redacted]

7.5.1. [redacted]

[redacted]

7.5.2. [redacted]

[redacted]

8. SAFETY AND TOLERABILITY

All safety analyses during the randomized period and overall period will be based on the safety population. Some safety analyses will be done during the crossover treatment period and based on CAS. All listings and tables will be presented by actual treatment received. Safety analysis will be conducted by treatment group. Appendix A provides a list of data displays. Sample data displays are included in a separate document.

8.1. General Considerations

The clinical safety data (eg, vital signs, ECGs, routine laboratory tests, and AEs) will be summarized with descriptive statistics (eg, mean, frequency) using the safety population for the randomized and overall periods and some analyses will be repeated using the CAS for the crossover period. Some safety analyses, such as AEs and laboratory values, may be performed within a known HIV-positive subgroup if there are a sufficient number of participants.

Summary tables may be replaced with listings when appropriate, for instance, an AE frequency table may be replaced with a listing if it only contains a few unique PTs reported on relatively few participants.

The main comparative safety summaries will be presented by treatment group for the randomized period; separate safety summaries will be generated for the crossover and the overall periods.

8.2. Adverse Events

8.2.1. Adverse Event Definitions

Treatment-emergent AEs will be summarized by treatment group for the randomized, crossover, and overall periods. In the randomized and overall periods, a TEAE is any AE either reported for the first time or worsening of a pre-existing event after the first dose of study treatment until the end of the randomized/overall period. For the crossover period, a TEAE is any AE either reported for the first time or worsening of a pre-existing event after the first dose of crossover treatment until end of the crossover period or the start of new additional anticancer therapy. Any AE that occurs after the start of a new anticancer therapy (including the crossover treatment) will not be considered a TEAE for the randomized period. Analysis of AEs (as discussed below) will be limited to TEAEs; however, data listings will include all AEs, regardless of their timing in relation to study treatment administration.

Adverse events will be tabulated by MedDRA PT and SOC. Severity of AEs will be graded using the NCI CTCAE v5.0. The CTCAE reporting guidelines and grading details are available on the Cancer Therapy Evaluation Program website. Exposure-adjusted TEAEs may be summarized based on on–treatment duration of the randomized period.

The subset of AEs considered by the investigator to be related to study treatment will be considered to be treatment-related. If the investigator does not specify the relationship of the AE to study treatment, the AE will be considered to be treatment-related. The incidence of AEs and treatment-related AEs will be tabulated. In addition, serious TEAEs will also be tabulated.

Any missing onset date, causality, or severity must be queried for resolution. Unresolved missing causality and severity will be handled according to the following rules:

- An unresolved missing causality will be considered treatment-related.
- An unresolved missing severity will be identified as an unknown severity.

For the purpose of analysis, all AEs will be considered TEAEs unless the AE can unequivocally be defined as not treatment-emergent.

8.2.2. Adverse Events of Special Interest

Immune-related AEs (irAEs) and IRRs will be considered AESIs in this study. Adverse events of special interest identified in the AE dataset will be summarized. Predefined PTs will be grouped into AESI categories and used to identify irAEs or IRRs without consideration of the investigator's assessment of causality. Diagnosis of an IRR occurring anytime during the treatment period, or a symptom of an IRR that occurs within 1 day of infusion and that resolves within 2 days from AE onset, will be captured as an IRR.

Immune-related AEs as well as infusion reactions will be separately summarized to capture AESIs. An overall summary of AESIs will include number (%) of participants reporting any AESIs, any serious AESIs, any AESIs ≥ Grade 3, any treatment-related AESIs, any fatal AESIs, any AESIs that may lead to retifanlimab/placebo drug interruption/discontinuation, and other categories, as necessary.

8.2.3. Adverse Event Summaries

An overall summary of AEs during the randomized, crossover, and overall periods will include the following:

- Number (%) of participants reporting any TEAEs
- Number (%) of participants reporting any serious TEAEs
- Number (%) of participants reporting any Grade 3 or higher TEAEs
- Number (%) of participants reporting any TEAEs related to retifanlimab/placebo
- Number (%) of participants reporting any TEAEs related to chemotherapy (not in crossover treatment period)
- Number (%) of participants reporting any TEAEs related to COVID-19
- Number (%) of participants who had any retifanlimab/placebo treatment-related serious TEAEs
- Number (%) of participants who had any retifanlimab/placebo treatment-related Grade 3 or higher TEAEs
- Number (%) of participants who temporarily interrupted retifanlimab/placebo infusion because of TEAEs
- Number (%) of participants who delayed the next scheduled retifanlimab/placebo infusion because of TEAEs

- Number (%) of participants who permanently discontinued retifanlimab/placebo because of TEAEs
- Number (%) of participants who permanently discontinued chemotherapy because of TEAEs (not in crossover treatment period)
- Number (%) of participants who had a fatal TEAE

The following summaries will be produced by MedDRA terms (if 10 or fewer participants appear in a table, a listing may be appropriate):

- Summary of TEAEs by SOC and PT
- Summary of TEAEs by PT in decreasing order of frequency
- Summary of TEAEs by SOC, PT, and maximum severity
- Summary of Grade 3 or higher TEAEs by SOC and PT
- Summary of Grade 3 or higher TEAEs by PT in decreasing order of frequency
- Summary of retifanlimab/placebo treatment-related TEAEs by SOC and PT
- Summary of chemotherapy treatment-related TEAEs by SOC and PT
- Summary of TEAEs related to COVID-19 by SOC and PT
- Summary of retifanlimab/placebo treatment-related TEAEs by PT in decreasing order of frequency
- Summary of TEAEs with a fatal outcome by SOC and PT
- Summary of serious TEAEs by SOC and PT
- Summary of retifanlimab/placebo treatment-related serious TEAEs by SOC and PT
- Summary of serious TEAEs by PT in decreasing order of frequency
- Summary of TEAEs leading to retifanlimab/placebo dose delay by SOC and PT
- Summary of TEAEs leading to retifanlimab/placebo infusion interruption by SOC and PT
- Summary of TEAEs leading to discontinuation of retifanlimab/placebo by SOC and PT
- Summary of TEAEs leading to chemotherapy dose delay by SOC and PT
- Summary of TEAEs leading to chemotherapy infusion interruption by SOC and PT
- Summary of TEAEs leading to discontinuation of chemotherapy by SOC and PT

The following summaries for AESIs will be included:

- Overall summary of immune-related TEAEs
- Overall summary of treatment-emergent IRRs
- Summary of TEAEs of special interest (IRR and irAE separately) by group term and PT

- Summary of TEAEs of special interest (IRR and irAE separately) by PT in decreasing order of frequency
- Summary of TEAEs of special interest (IRR and irAE separately) by group term, PT, and maximum severity
- Summary of TEAEs of special interest (IRR and irAE separately) with a fatal outcome by group term and PT
- Summary of TEAEs of special interest (IRR and irAE separately) leading to retifanlimab/placebo dose delay by group term and PT
- Summary of TEAEs of special interest (IRR and irAE separately) leading to retifanlimab/placebo drug interruption by group term and PT
- Summary of TEAEs of special interest (IRR and irAE separately) leading to retifanlimab/placebo infusion interruption by group term and PT
- Summary of TEAEs of special interest (IRR and irAE separately) leading to retifanlimab/placebo discontinuation by group term and PT
- Summary of Grade 3 or higher TEAEs of special interest (IRR and irAE separately) by group term and PT

8.3. Clinical Laboratory Tests

8.3.1. Laboratory Value Definitions

Laboratory values and change from baseline values will be summarized descriptively by treatment group and by visit for the randomized period. Baseline will be determined according to Section 4.1.5. If there are multiple values that meet the criteria for baseline, Table 5 may be referred to as the tiebreaker in order to delineate which value will be defined as baseline.

Table 5: Baseline Laboratory Identification

Priority	Laboratory Visit	Proximity to Visit Window	Tiebreaker
1	Scheduled	In-window	Use smallest laboratory sequence number
2	Unscheduled	In-window	
3	Scheduled	Out-of-window	

Laboratory test values outside the normal range will be assessed for severity based on CTCAE grade (or similar criteria), when clinical intervention is required for CTCAE grading. The incidence of abnormal laboratory values and shift tables relative to baseline will be tabulated.

8.3.2. Laboratory Value Summaries

All test results and associated normal ranges from laboratories will be reported in SI units. Any laboratory test results and associated normal ranges from local laboratories will be converted to SI units. For test results that will be summarized with available normal ranges, the number and percentage of participants with the laboratory values being low (but never high), normal, high

(but never low), and both low and high will be calculated for each test. This shift summary may be produced for each test for the safety population during the randomized period and the CAS during the crossover period, respectively. The denominator for the percentage calculation will use the number of participants in the baseline category (ie, low, high, normal, missing) as the denominator for the percentage in each of the categories during the treatment period. Numeric laboratory values, the baseline value, postbaseline value, change from baseline, and percentage change from baseline will be summarized by visit.

Severity grades will be assigned to laboratory test values based on the numerical component of CTCAE. Shift tables will be presented showing change in CTCAE grade from baseline to worst postbaseline grade for the randomized and crossover periods. Separate summaries for abnormally high and abnormally low laboratory values will be provided when the laboratory parameter has both high and low grading criteria. The denominator for the percentage calculation will be the number of participants in the baseline category.

The following summaries will be produced for laboratory data (by laboratory parameter) reported during the on–treatment period. All laboratory assessments will be listed, and those collected after the on–treatment period will be flagged in the listings.

- Number and percentage of participants with worst postbaseline CTCAE grade (regardless of the baseline status) will be summarized. Each participant will be counted only for the worst grade observed after baseline.
- Shift tables using CTCAE grades comparing baseline with the worst postbaseline value will be produced for hematology and biochemistry laboratory parameters with CTCAE grades.
- For laboratory parameters where CTCAE grades are not defined, shift tables to the worst postbaseline value will be produced using the low/normal/high classifications based on laboratory reference ranges.
- For laboratory parameters with any worsening postbaseline CTCAE grades than baseline, the worst postbaseline CTCAE grades will be summarized according to CTCAE severity grade levels.
- Number and percentage of participants meeting categorical liver function test criteria, including ALT, AST, and ALT or AST ($\geq 3 \times, 5 \times, 8 \times, 10 \times, 20 \times$ ULN), total bilirubin ($\geq 1 \times, 2 \times$ ULN), ALP ($\geq 1.5 \times, 2 \times, 3 \times, 5 \times, 8 \times, 10 \times$ ULN), combined categories of ALT/AST and total bilirubin (eg, ALT/AST $\geq 3 \times$ ULN and total bilirubin $\geq$ ULN), as well as potential Hy's Law criteria (ALT or AST $\geq 3 \times$ ULN and total bilirubin $\geq 2 \times$ ULN and ALP $< 2 \times$ ULN at the same visit). The worst values observed postbaseline for each participant will be used for each of the categories.
- Summary of TSH change during study treatment period and shift table based on TSH will be provided

8.4. Vital Signs

Values at each scheduled visit, change, and percentage change from baseline for vital signs, including systolic blood pressure, diastolic blood pressure, pulse, respiratory rate, body temperature, and weight will be summarized descriptively by treatment group for the randomized period.

Normal ranges for vital signs are defined in Table 6. For participants exhibiting vital sign abnormalities, the abnormal values will be listed. Alert vital signs are defined as an absolute value outside the defined range and percentage change from baseline greater than 25%. Note that the definition of alert vital signs does not apply for body temperature and weight. The values for participants exhibiting alert vital sign abnormalities will be listed.

Table 6: Normal Ranges for Vital Sign Values

Parameter	High Threshold	Low Threshold
Systolic blood pressure	≤ 155 mmHg	≥ 85 mmHg
Diastolic blood pressure	≤ 100 mmHg	≥ 40 mmHg
Pulse	≤ 100 bpm	≥ 45 bpm
Respiratory rate	≤ 24 breaths/min	≥ 8 breaths/min
Temperature	≤ 38°C	≥ 35.5°C

8.5. Electrocardiograms

Twelve-lead ECGs, including HR, PR, QRS, QT, and QTc (QTcF or QTcB) intervals, will be obtained for each participant during the study. Values at each scheduled visit, change, and percentage change from baseline will be summarized by treatment group for each ECG parameter for the randomized period. Baseline will be determined as the last nonmissing ECG measurements taken on or before Day 1 for safety endpoints.

Normal ranges for ECG values are defined in Table 7. Participants exhibiting ECG abnormalities will be listed with study visit. Abnormal values for participants with alert ECG values, defined as both the absolute value and the percentage change (30% for QRS, 25% for others) from baseline being outside normal ranges, will be identified and listed. Outliers of QT and QTc values, defined as absolute values ≥ 450 milliseconds, ≥ 480 milliseconds, and ≥ 500 milliseconds or change from baseline ≥ 30 milliseconds and ≥ 60 milliseconds, will be summarized.

Table 7: Normal Ranges for Electrocardiogram Intervals

Parameter	High Threshold	Low Threshold
QTc (QTcF/QTcB)	≤ 480 ms	≥ 295 ms
PR	≤ 220 ms	≥ 75 ms
QRS	≤ 120 ms	≥ 50 ms
QT	≤ 500 ms	≥ 300 ms
HR	≤ 100 bpm	≥ 50 bpm

9. [redacted]

9.1. [redacted]

[redacted]

9.2. [redacted]

[redacted]

10. INTERIM ANALYSES OF OVERALL SURVIVAL

10.1. Overview of Interim Analyses

If the primary analysis of PFS crosses the boundary, OS will be evaluated at 1-sided 2.5% with 2 looks. An interim analysis will be performed at the time of PFS analysis with the detailed alpha spending rule provided (see Table 8).

Table 8: Guidelines for Decisions in 2-Look Group Sequential Design – Overall Survival

	Interim Analysis at the Time of the PFS Analysis		**Final Analysis on OS**	
Projected timing[a]	45 months		55 months	
Projected enrollment	[redacted]		[redacted]	
Number of events	[redacted]		[redacted]	
Decision outcome	**Futility**	**Efficacy**	**Futility**	**Efficacy**
One-sided p-value	≥ 0.012	< 0.012	> 0.025	≤ 0.025
Estimated hazard rate reduction	≤ 32.5%	> 32.5%	< 27%	≥ 27%
Estimated median improvement (month)	≤ 9.6	> 9.6	< 7.4	≥ 7.4

[a] Randomized participants. Estimates based on exponential assumption where median OS is 20 months on control. The enrollment projection is 10 participants per month. Power is 73% under alternative hypothesis of a 0.67 HR with a dropout rate of 1%.

11. CHANGES AND MODIFICATIONS TO THE ANALYSIS PLAN

All versions of the SAP are listed in Table 9.

Table 9: Statistical Analysis Plan Versions

SAP Version	Date
Original	01 JUN 2022
Amendment 1	08 NOV 2023

11.1. Changes to Protocol-Defined Analyses

Not applicable.

11.2. Changes to the Statistical Analysis Plan

11.2.1. Amendment 1

This SAP version includes the following changes:

- Changed the randomized period from starting with the first dose of study treatment to starting with randomization, which also resulted in changes to start/stop dates.
- Removed the enrolled population.
- Updated the equations for calculation of total administered dose of study drug.
- Added Asian to the race subgroup analysis.
- Changed the sensitivity analysis.
- Changed overall response rate to objective response rate.
- Consolidated chemotherapy data tables and listings.
- Updated the projected timing, number of events, 1-sided p-value, estimated hazard rate reduction, and estimated median improvement values in Table 8.
- Updated the QLQ calculation in Appendix B.
- Updated Table 4 for sensitivity analysis on evaluation and censoring of PFS.
- Updated prior medication definition.

Other minor changes have been incorporated throughout the SAP and are noted in the redline version of the amendment.

12. REFERENCES

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics 1982;38:29-41.

Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228-247.

Efron B. The efficiency of Cox's likelihood function for censored data. J Am Stat Assoc 1977;72:557-565.

European Organisation for Research and Treatment of Cancer. EORTC QLQ-ANL27 Scoring Manual. Brussels: EORTC Quality Life Group. 2016.

Food and Drug Administration. Guidance for Industry: Clinical Trial Endpoints for the Approval of Non–Small Cell Lung Cancer Drugs and Biologics. 2015.

Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. New York: Springer-Verlag. 1997.

Mosteller RD. Simplified calculation of body surface area. N Engl J Med 1987;317:1098.

O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979;35:549-556.

Rao S, Sclafani F, Eng C, et al. InterAACT, a multicenter open label randomised phase II advanced anal cancer trial of cisplatin (CDDP) plus 5 fluorouracil (5FU) vs carboplatin (C) plus weekly paclitaxel (P) in patients (pts) with inoperable locally recurrent (ILR) or metastatic treatment naïve disease - an International Rare Cancers Initiative (IRCI) trial. Ann Oncol 2018;29:vii715-viii716 [abstrP669].

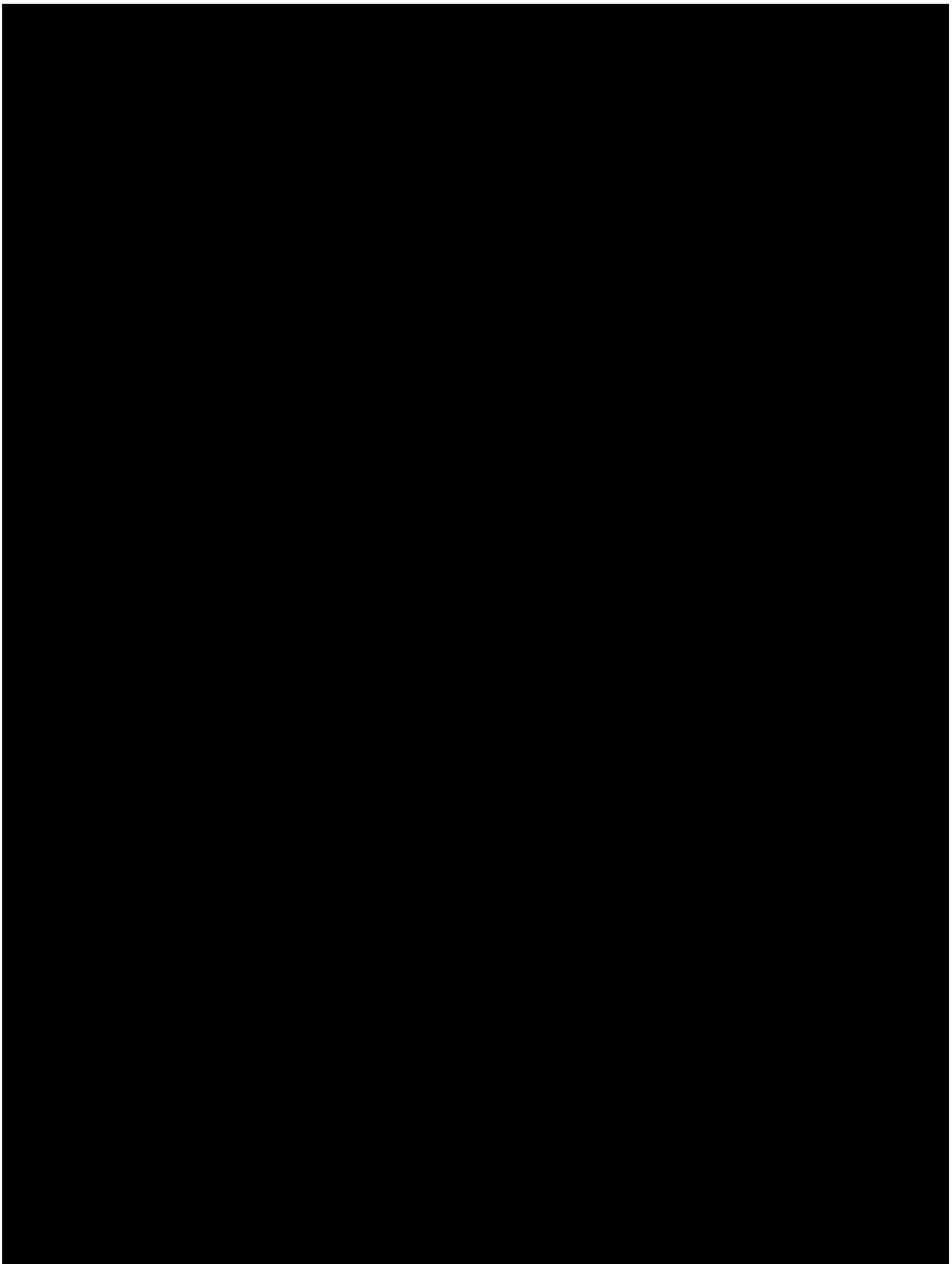
APPENDIX A. [redacted]

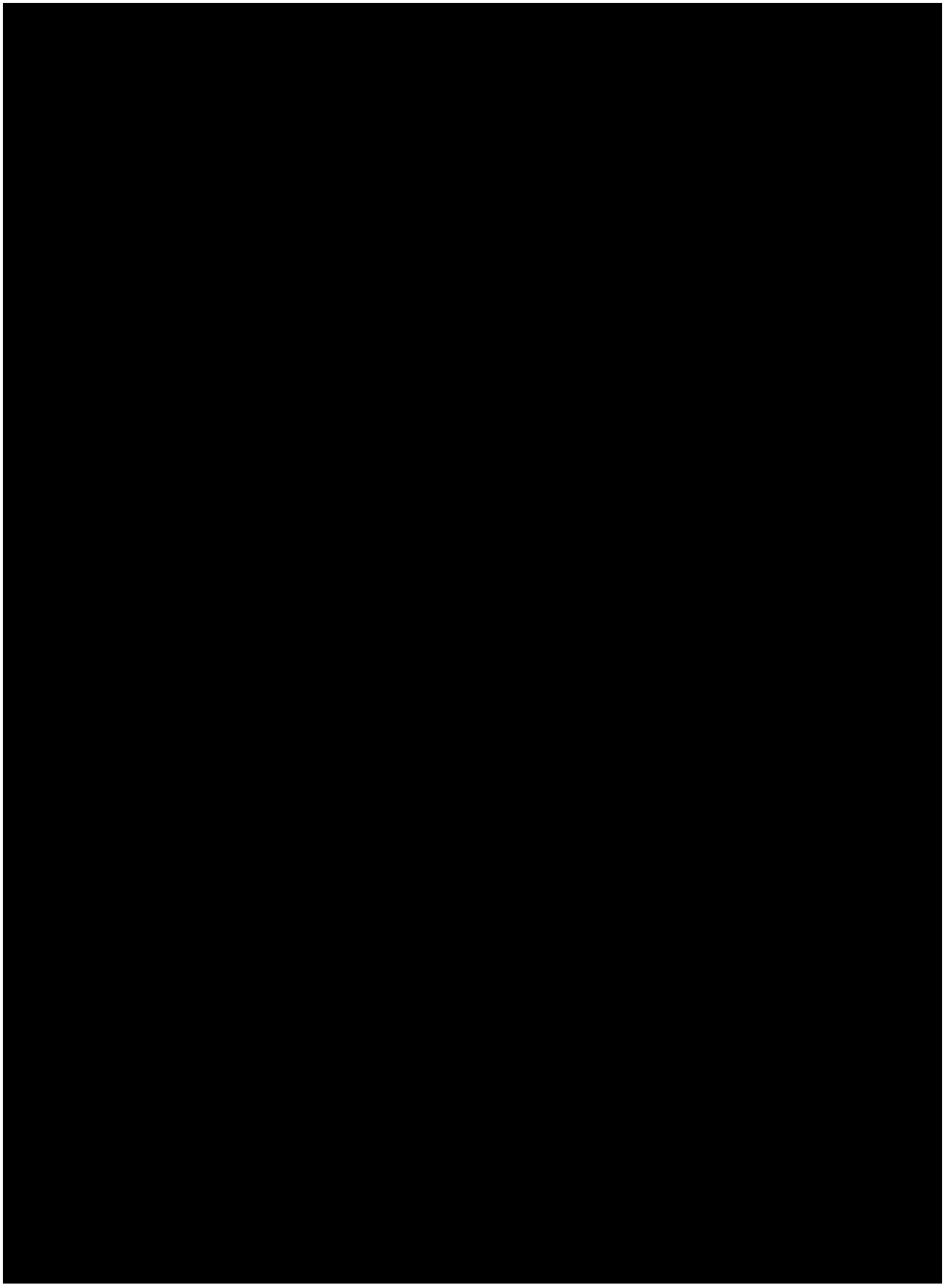
[redacted]

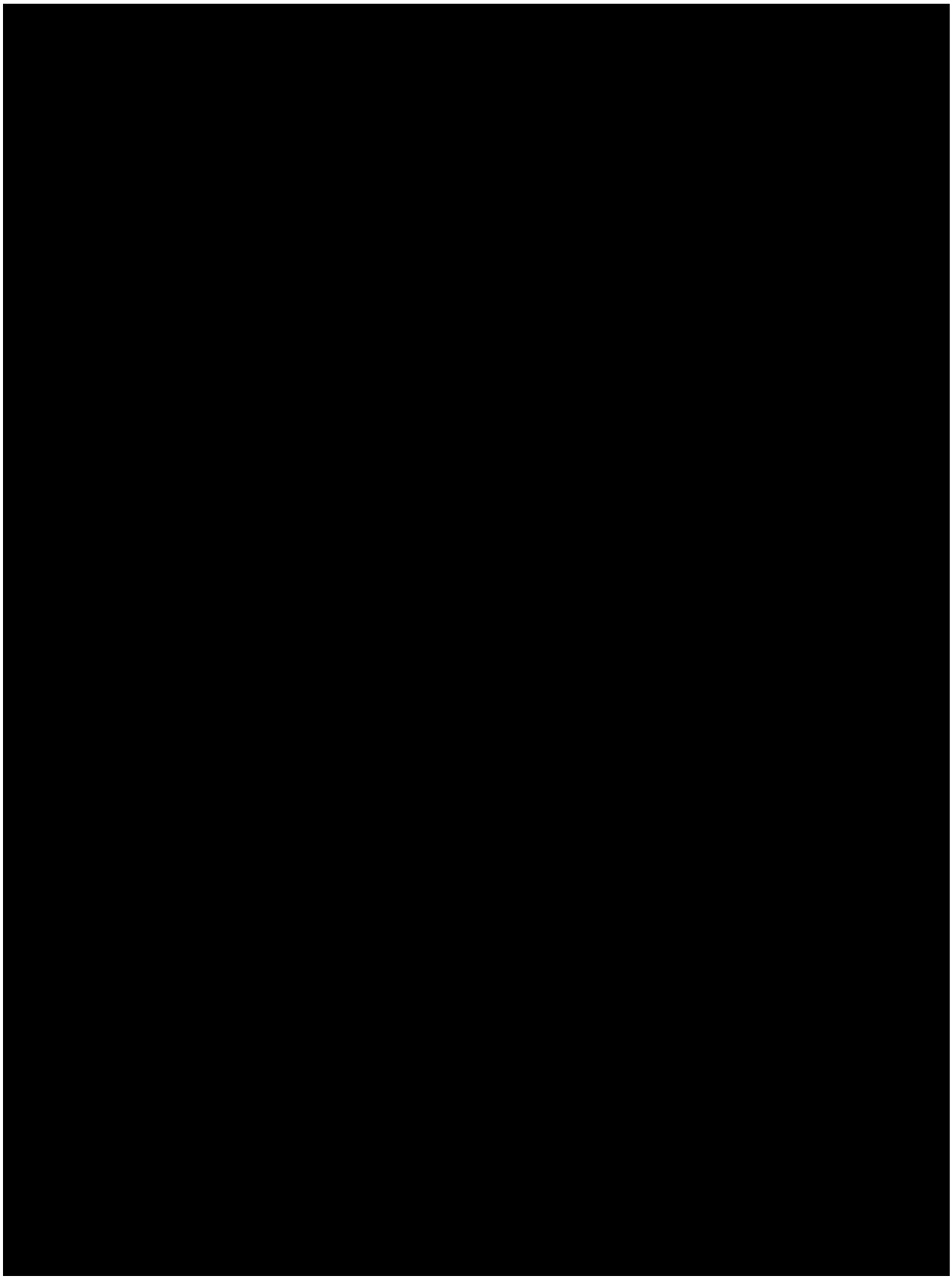
[redacted]

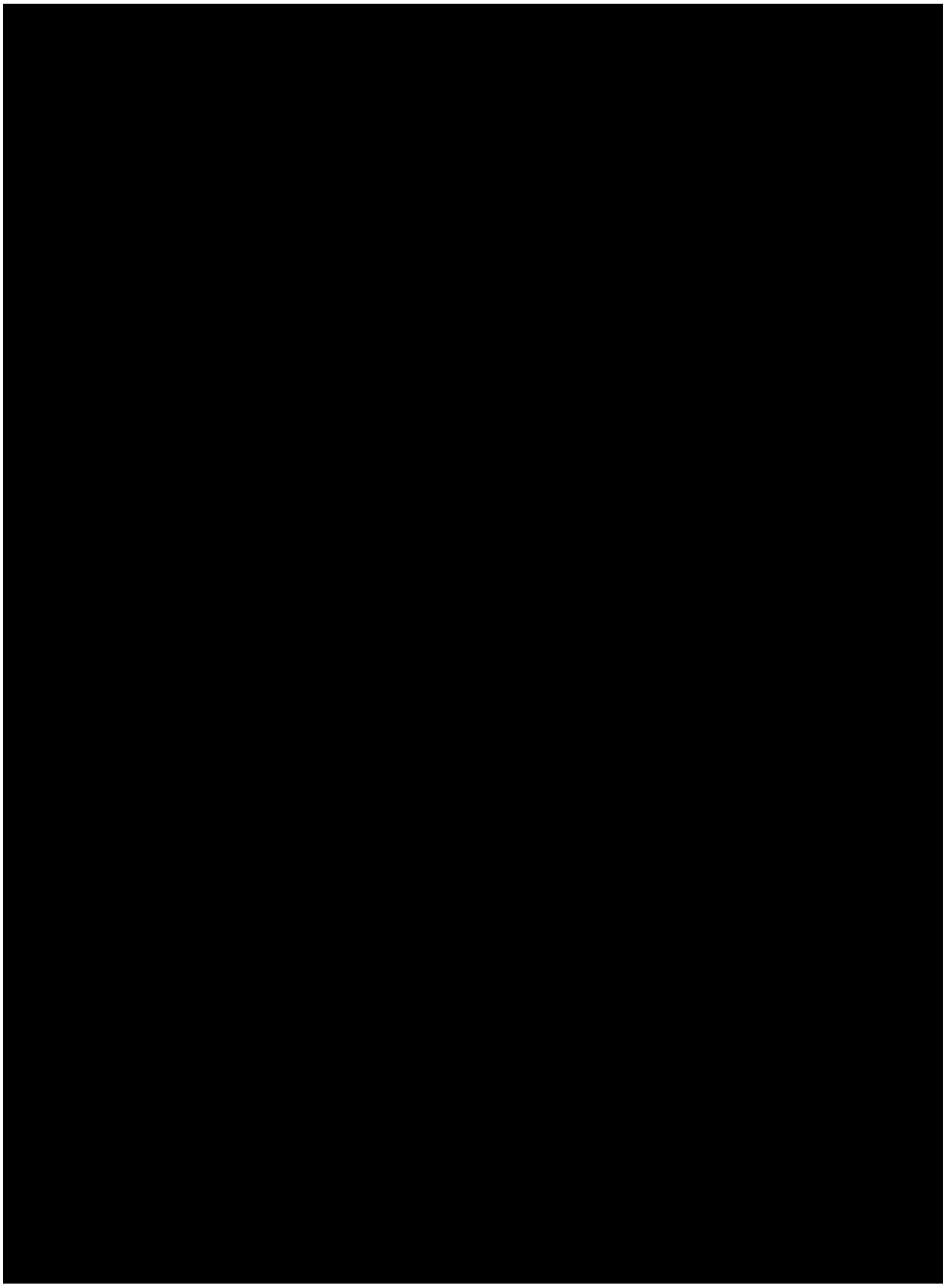
[redacted]

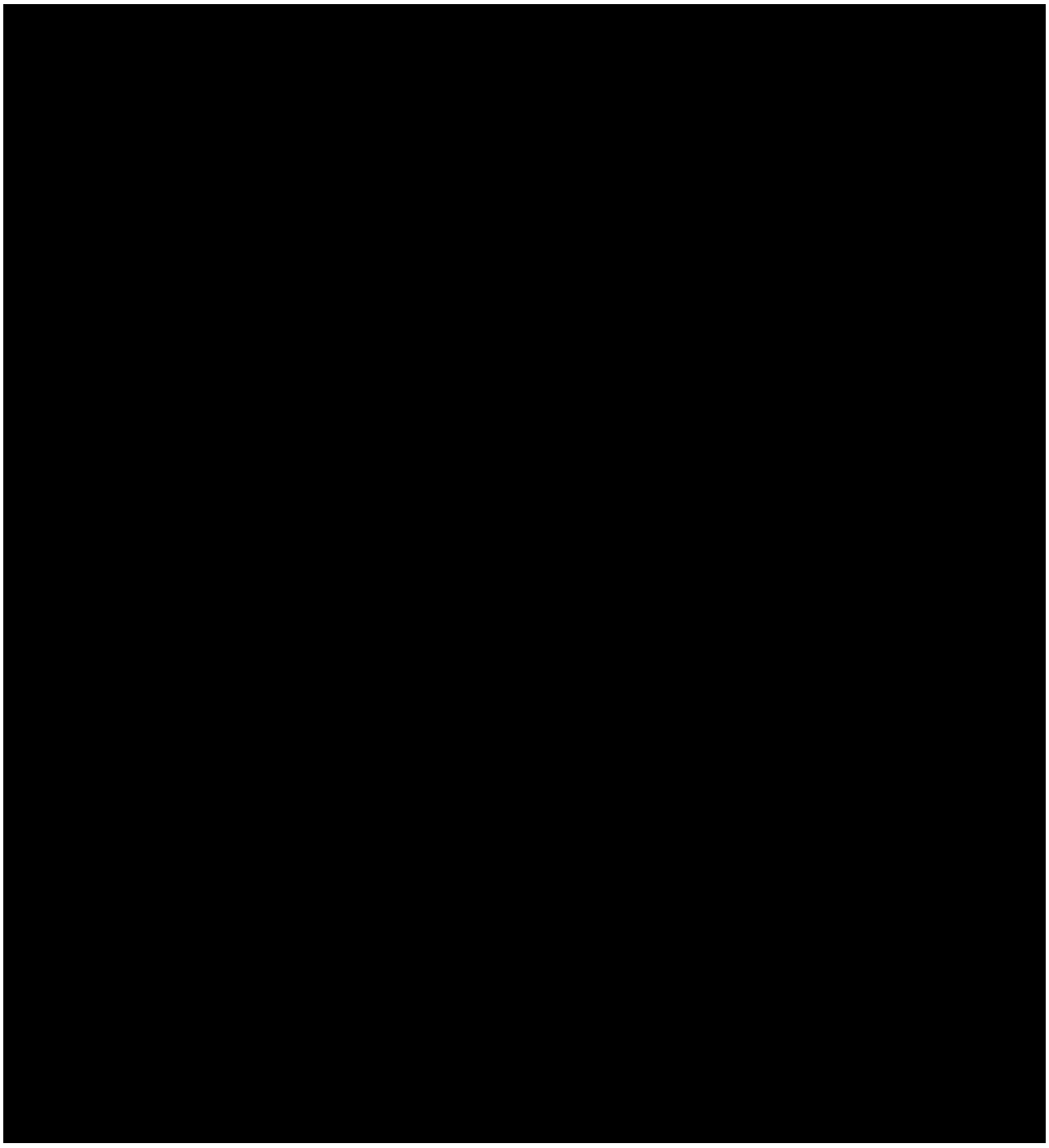
[redacted]

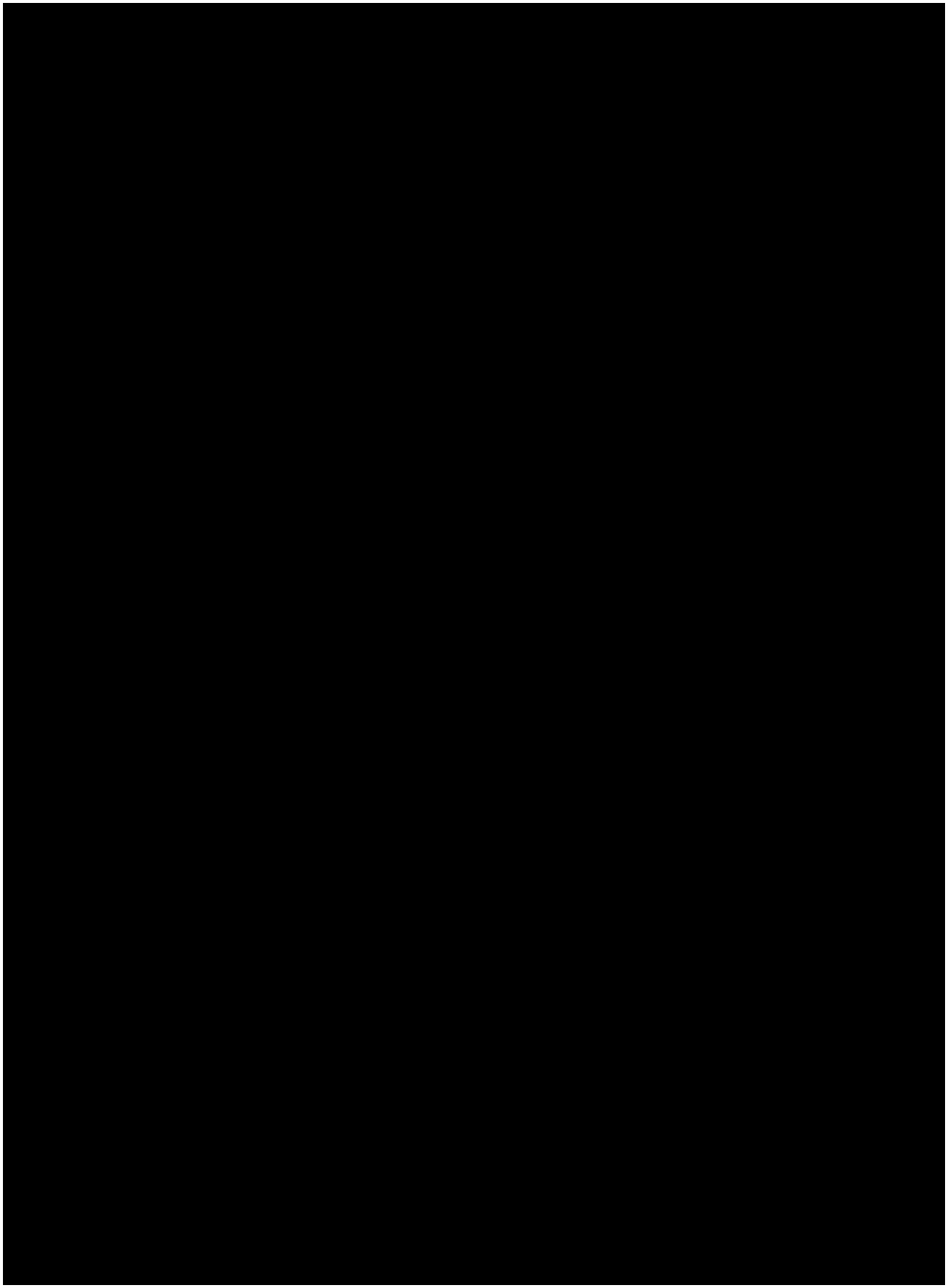




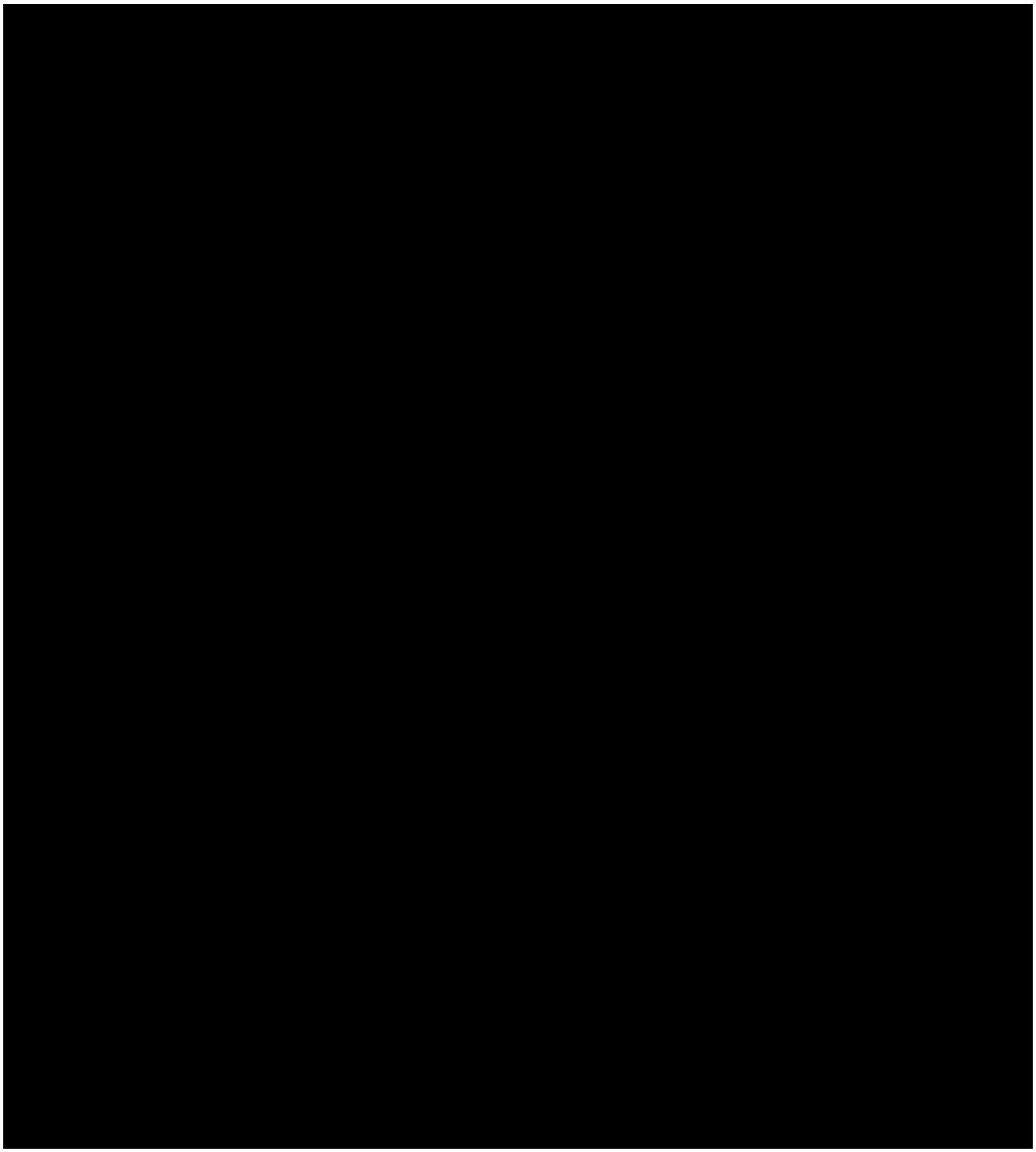


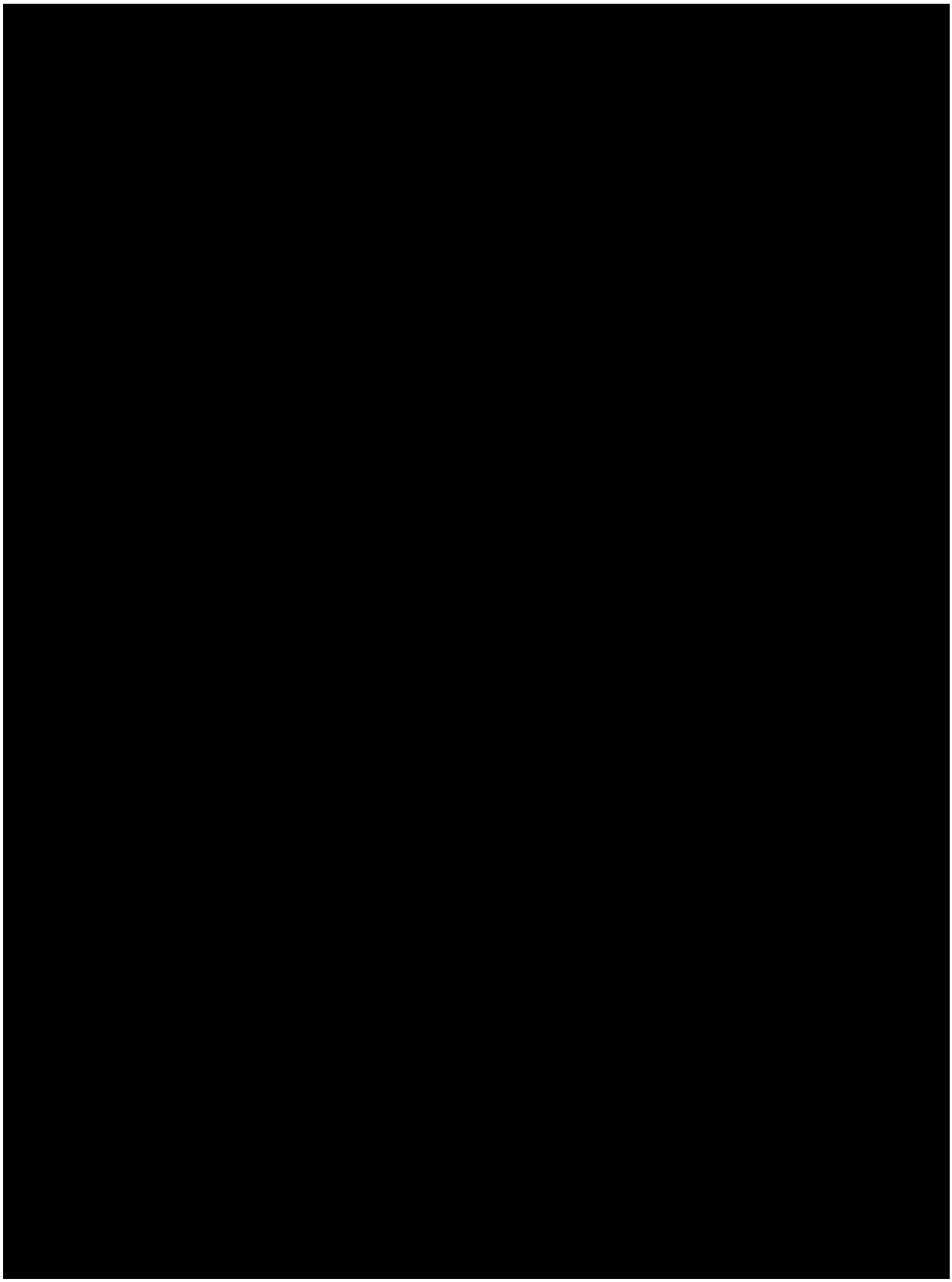


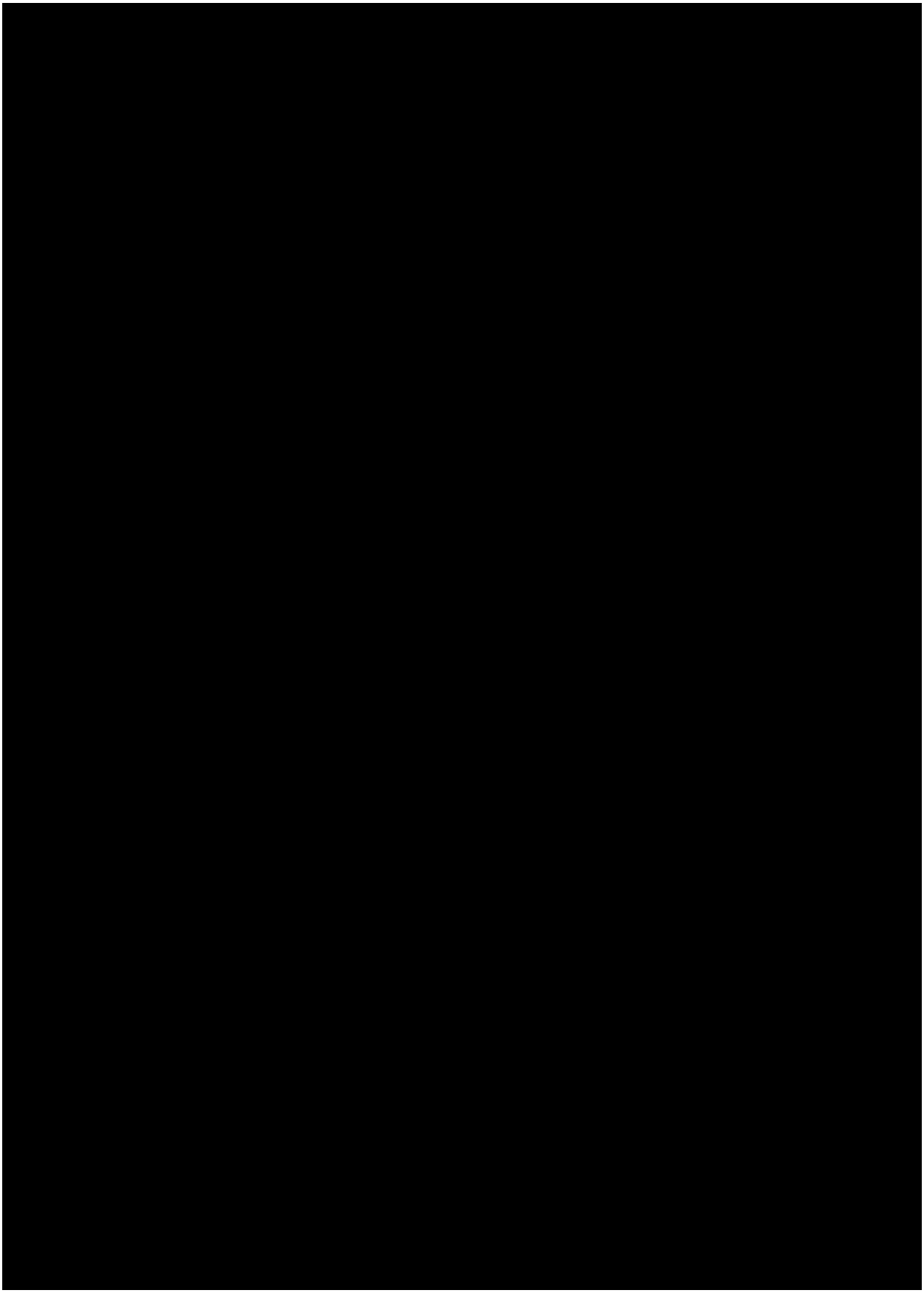




APPENDIX B. [redacted]







[redacted]