Official Title: A PHASE III, OPEN-LABEL, RANDOMIZED STUDY OF ATEZOLIZUMAB (ANTI–PD–L1 ANTIBODY) COMPARED WITH A PLATINUM AGENT (CISPLATIN OR CARBOPLATIN) IN COMBINATION WITH EITHER PEMETREXED OR GEMCITABINE FOR PD-L1–SELECTED, CHEMOTHERAPY-NAIVE PATIENTS WITH STAGE IV NON-SQUAMOUS OR SQUAMOUS NON–SMALL CELL LUNG CANCER

NCT Number: NCT02409342

Document Date: SAP Version 3: 02-April-2019

STATISTICAL ANALYSIS PLAN

TITLE: **A PHASE III, OPEN-LABEL, RANDOMIZED STUDY OF ATEZOLIZUMAB (ANTI–PD–L1 ANTIBODY) COMPARED WITH A PLATINUM AGENT (CISPLATIN OR CARBOPLATIN) IN COMBINATION WITH EITHER PEMETREXED OR GEMCITABINE FOR PD-L1–SELECTED, CHEMOTHERAPY-NAIVE PATIENTS WITH STAGE IV NON-SQUAMOUS OR SQUAMOUS NON–SMALL CELL LUNG CANCER**

PROTOCOL NUMBER: GO29431

STUDY DRUG: Atezolizumab (MPDL3280A)

VERSION NUMBER: 3

IND NUMBER: 117296

EUDRACT NUMBER: 2014-003083-21

SPONSOR: F. Hoffmann-La Roche Ltd

PLAN PREPARED BY: [redacted], M.Sc.

DATE FINAL: Version 1: 18 April 2017
Version 2: 19 December 2018

DATE(S) AMENDED See electronic date stamp below.

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

Name	Reason for Signing	Date and Time (UTC)
[redacted]	Company Signatory	02-Apr-2019 19:04:42

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

The Statistical Analysis Plan (SAP) for Study GO29431 has been amended to reflect changes made in the protocol which consists of the following main changes in Statistical Analysis Plan:

- The tumor cell (TC)3 or tumor infiltrating immune cell (IC)3 population excluding patients with a sensitizing epidermal growth factor receptor (*EGFR*) mutation or anaplastic lymphoma kinase (*ALK*) translocation (i.e., TC3 or IC3-wild type [WT]) has been added as the first test in the hierarchy (see Sections 2.3 and 2.4)
- The timing of the efficacy analysis has been updated to be when the pre–specified criteria are met for the TC3 or IC3-WT population (see Sections 2.3 and 2.4)

In addition, the cutoffs have been clarified to include ≥1%, ≥25%, and ≥50% of tumor cells for the programmed death–ligand 1 (PD-L1) SP263 immunohistochemistry (IHC) assay and include ≥10, ≥16, and ≥20 mutations for the blood tumor mutation burden (bTMB) assay (see Section 4.4.2.6).

This amendment represents cumulative changes to the original analysis plan. Additional minor changes have been made to improve clarity and consistency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. BACKGROUND

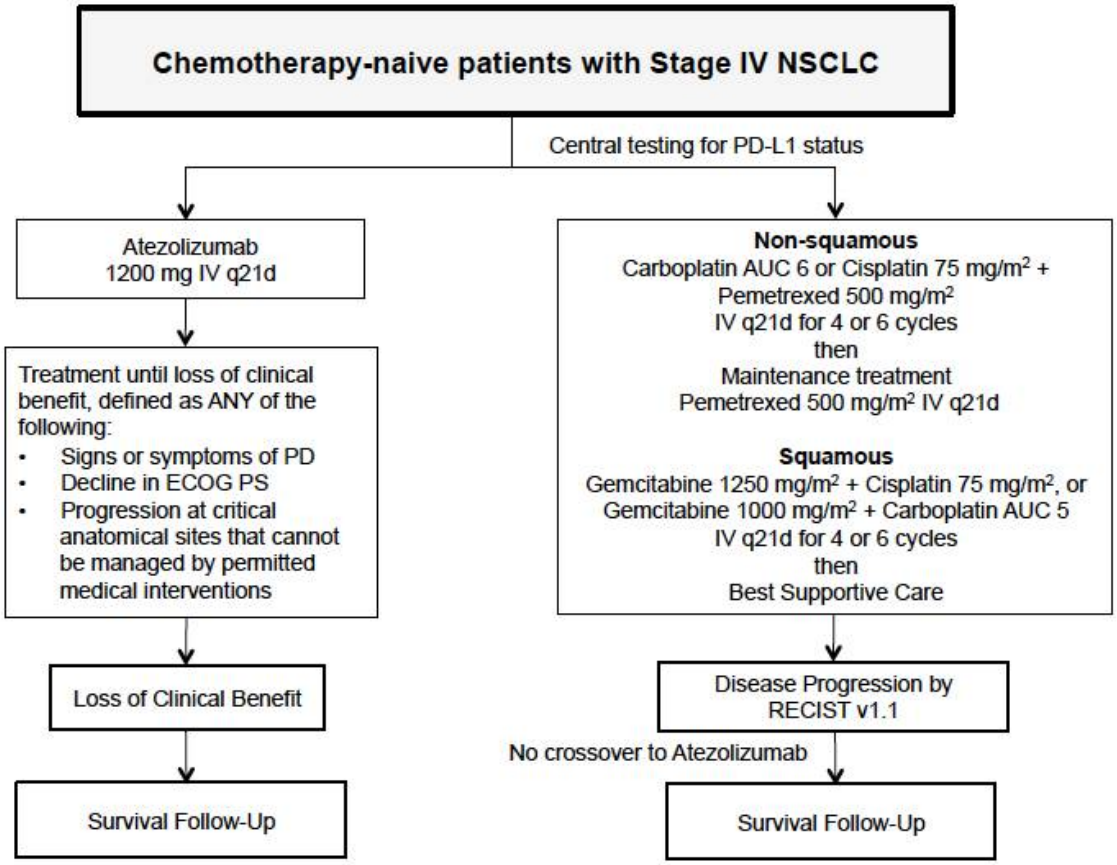
This Statistical Analysis Plan (SAP) provides details of the planned analyses and statistical methods for Study GO29431 (IMpower110), "A Phase III, Open-Label, Randomized Study of TECENTRIQ® (Atezolizumab, MPDL3280A, Anti–PD-L1 Antibody) Compared with a Platinum Agent (Cisplatin or Carboplatin) in Combination with Either Pemetrexed or Gemcitabine for PD-L1–Selected, Chemotherapy-Naive Patients with Stage IV Non-Squamous or Squamous Non-Small Cell Lung Cancer". The background for the study can be found in the study Protocol.

2. STUDY DESIGN

This is a randomized, Phase III, global, multicenter, open-label study designed to evaluate the safety and efficacy of atezolizumab compared with chemotherapy consisting of a platinum agent (cisplatin or carboplatin per investigator discretion) combined with either pemetrexed (non-squamous disease) or gemcitabine (squamous disease) in programmed death–ligand 1 (PD-L1)–selected patients, chemotherapy-naive patients with Stage IV non–small cell lung cancer (NSCLC).

Figure 1 illustrates the Study GO29431 study design. Eligible patients will be stratified by sex (male vs. female), Eastern Cooperative Oncology Group (ECOG) performance status (0 vs. 1), histology (non-squamous vs. squamous), and PD-L1 tumor expression status (tumor cell [TC]1/2/3 and any tumor-infiltrating immune cell [IC] vs. TC0 and IC1/2/3). Patients with non-squamous disease will be randomized in 1:1 ratio to receive either atezolizumab alone or pemetrexed in combination with cisplatin or carboplatin. Patients with squamous disease will be randomized in 1:1 ratio to receive either atezolizumab alone or gemcitabine in combination with cisplatin or carboplatin.

Figure 1 Study Schema for Study IMpower110



AUC = area under the concentration–time curve; ECOG PS = Eastern Cooperative Oncology Group performance status; IV = intravenous; NSCLC = non–small cell lung cancer; PD = progressive disease; PD-L1 = programmed death–ligand 1; q21d = every 21 days; RECIST = Response Evaluation Criteria in Solid Tumors.
Note: Gemcitabine is administered on Days 1 and 8.

Atezolizumab (fixed dose of 1200 mg) will be administered intravenously on Day 1 of each 21-day cycle. Atezolizumab treatment may continue as long as patients experience a clinical benefit as assessed by the investigator (i.e., in the absence of unacceptable toxicity or symptomatic deterioration attributed to disease progression as determined by the investigator after an integrated assessment of radiographic data, biopsy results [if available], and clinical status) or until unacceptable toxicity or death.

Patients randomized to receive pemetrexed in combination with either cisplatin or carboplatin (to treat non-squamous disease) will receive chemotherapy intravenously on Day 1 of each 21-day cycle for four or six cycles, followed by maintenance therapy with pemetrexed. Patients who are randomized to receive gemcitabine in combination with either cisplatin or carboplatin (squamous disease) will receive cisplatin or carboplatin intravenously on Day 1 and gemcitabine intravenously on Days 1 and 8 of each 21-day cycle for four or six cycles, followed by best supportive care. The intended number of

cycles planned for the platinum-based induction chemotherapy (i.e., four or six cycles) will be specified by the investigator prior to study randomization. Treatment will continue until disease progression, unacceptable toxicity, or death.

All patients will undergo tumor assessment at baseline and every 6 weeks (±7 days) for 48 weeks following Cycle 1, Day 1 regardless of treatment delays. After the completion of the Week 48 tumor assessment, tumor assessment will be required every 9 weeks (± 7 days) regardless of treatment delays, until radiographic disease progression per Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST v1.1; or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1), withdrawal of consent, death, or study termination by the Sponsor, whichever occurs first. Patients who discontinue treatment for reasons other than radiographic disease progression per RECIST v1.1 (e.g., toxicity, symptomatic deterioration) will continue scheduled tumor assessments until radiographic disease progression per RECIST v1.1 (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1), withdrawal of consent, death, or study termination by Sponsor, whichever occurs first. In the absence of radiographic disease progression per RECIST v1.1, tumor assessments should continue regardless of whether patients start a new anti-cancer therapy.

The primary efficacy endpoint is overall survival (OS). The secondary endpoints include investigator-assessed progression-free survival (PFS), objective response rate (ORR), duration of response (DOR) per RECIST v1.1, time to deterioration (TTD) and change from baseline per the Symptoms in Lung Cancer (SILC) scale, and TTD per European Organisation for Research and Treatment of Cancer (EORTC) scale. Other endpoints include safety, pharmacokinetics, and exploratory outcome measures (see Section 2.2).

An independent Data Monitoring Committee (iDMC) has been monitoring both safety and efficacy data (see Section 3.2).

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is in Appendix 1. For additional details, see the schedule of assessments in Appendix 2.

2.2 OUTCOME MEASURES

2.2.1 Primary Efficacy Outcome Measure

The primary efficacy outcome measure for this study is OS, defined as the time from randomization to death from any cause.

2.2.2 Secondary Efficacy Outcome Measures

The secondary efficacy outcome measures for this study are:

- PFS, defined as the time from randomization to the first occurrence of disease progression as determined by the investigator with use of RECIST v1.1 or death from any cause, whichever occurs first.
- Objective response (partial response [PR] plus complete response [CR]) as determined by the investigator according to RECIST v1.1
- DOR, defined as the time from the first occurrence of a documented objective response to the time of disease progression as determined by the investigator with use of RECIST v1.1 or death from any cause, whichever occurs first
- OS and Investigator-assessed PFS according to RECIST v1.1 in patients with PD-L1 expression defined by the SP263 immunohistochemistry (IHC) assay
- OS and investigator-assessed PFS according to RECIST v1.1 in patients with blood tumor mutational burden (bTMB)
- OS at 1-year and 2-year landmark timepoints
- TTD and change from baseline (i.e., improvement or deterioration based on presenting symptomatology) in each of the patient-reported lung cancer symptoms (cough, dyspnea, or chest pain) score as assessed by the SILC
- TTD in patient-reported lung cancer symptoms of cough, dyspnea (multi-item subscale), and chest pain, defined as time from randomization to deterioration (10-point change) as assessed respectively by the EORTC Quality-of-Life Questionnaire Core 30 (QLQ-C30) and supplementary EORTC Quality-of-Life Questionnaire Lung Cancer Module (QLQ-LC13) scale

2.2.3 Exploratory Efficacy Outcome Measures

The exploratory outcome measures for this study are:

- OS and investigator-assessed PFS according to RECIST v1.1 in patients with PD-L1 expression measured using the 22c3 PD-L1 IHC assay
- PFS at 6-month and at 1-year landmark timepoints
- OS at 3-year landmark timepoint
- OS and investigator-assessed PFS according to RECIST v1.1 in subgroups based on demographic and baseline characteristics
- Status of immune-cell infiltrate and other exploratory biomarkers in mandatory biopsy specimens collected at progression
- Status of PD-L1–, immune-, and NSCLC-related and other exploratory biomarkers in archival and/or freshly obtained tumor tissues and blood (or blood derivatives) collected before, during, or after treatment with atezolizumab or at progression and association with disease status and/or response to atezolizumab
- Utility scores of the EuroQoL 5 Dimension, 3 Level (EQ-5D-3L) questionnaire

- Change from baseline in patient-reported outcomes (PROs) of health-related quality of life (HRQoL), lung cancer-related symptoms, and functioning as assessed by the EORTC QLQ-C30 and QLQ-LC13

2.2.4 Pharmacokinetic Efficacy Outcome Measures

The pharmacokinetic (PK) outcome measures for this study are:

- Maximum serum atezolizumab concentration observed (C_{max}) after infusion on Day 1 of Cycle 1
- Minimum serum atezolizumab concentration observed (C_{min}) prior to infusion on Day 1 of Cycles 2, 3, 4, 8, 16, and every eighth cycle thereafter, at treatment discontinuation, and at 120 ($\pm$30) days after the last dose of atezolizumab

2.2.5 Safety Outcome Measures

The safety outcome measures for this study are:

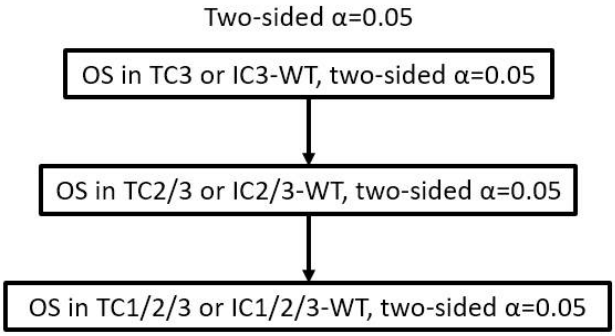
- Incidence, nature, and severity of adverse events graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, Version 4.0 (NCI CTCAE v4.0)
- Changes in vital signs, physical findings, and clinical laboratory test results during and following atezolizumab administration
- Incidence of anti-therapeutic antibody (ATA), also known as anti-drug antibody (ADA), response to atezolizumab and potential correlation with PK, pharmacodynamics (PD), safety, and efficacy parameters

2.3 DETERMINATION OF SAMPLE SIZE

2.3.1 Determination of Sample Size in the Global Enrollment Phase

Approximately 150 sites globally will participate in the study, and approximately 555 PD-L1–selected, chemotherapy-naive patients with Stage IV NSCLC are planned to be enrolled. For the primary comparison of OS, the Sponsor will control the overall type I error rate for the two-sided tests at 0.05 for the TC3 or IC3 subpopulation, the TC2/3 or IC2/3 subpopulation and the TC1/2/3 or IC1/2/3 population, with all populations excluding patients with a sensitizing *EGFR* mutation or *ALK* translocation. The hierarchy of α spending is specified in Figure 2.

Figure 2 Type I Error Control Plan



IC = tumor-infiltrating immune cell; OS = overall survival; TC = tumor cell; WT=wild type (i.e., excluding patients with a sensitizing epidermal growth factor receptor [*EGFR*] mutation or anaplastic lymphoma kinase [*ALK*] translocation).

Estimates of the number of events required to demonstrate efficacy in terms of OS are based on the following assumptions:

- 1:1 randomization ratio
- One interim analysis of OS in the TC3 or IC3-wild type (WT), the TC2/3 or IC2/3-WT, and the TC1/2/3 or IC1/2/3-WT population with stopping boundaries determined by the Lan-DeMets approximation to the Pocock boundary
- Two-sided significance level of 5% in the TC3 or IC3-WT, the TC2/3 or IC2/3-WT, and the TC1/2/3 or IC1/2/3-WT population
- 99% power to detect a hazard ratio (HR) of 0.45 for OS in the TC3 or IC3-WT subpopulation, 85% power to detect a HR of 0.65 for OS in the TC2/3 or IC2/3-WT subpopulation, and 77% power to detect a HR of 0.75 for OS in the TC1/2/3 or IC1/2/3-WT population
- Median survival of 14 months in the control arm (platinum-based chemotherapy)
- Event times exponentially distributed
- Dropout rate assumed for all treatment arms of 5% per 24 months
- Prevalence rate of 37% for TC3 or IC3 and 59% for TC2/3 or IC2/3 within the TC1/2/3 or IC1/2/3 population

With these assumptions, the final OS analysis will be conducted when approximately 135 deaths have occurred in the TC3 or IC3-WT subpopulation.

2.4 INTERIM ANALYSIS TIMING

The Sponsor plans to conduct one interim efficacy analysis for the primary endpoint of OS in the TC3 or IC3-WT, the TC2/3 or IC2/3-WT, and the TC1/2/3 or IC1/2/3-WT populations, respectively. With a lack of the final PD-L1 prevalence, to ensure the data maturity and have sufficient event-patient ratio for the evaluation of the OS benefit, an

interim analysis of OS in the TC3 or IC3-WT population will be conducted when both of the following criteria have been met:

- An approximately 45% event-patient ratio has been observed in the TC3 or IC3-WT subpopulation
- Approximately 96 deaths have occurred in the TC3 or IC3-WTsubpopulation

Based on the assumptions specified in Section 2.3.1, it is expected that approximately 154 OS events would have occurred in the TC2/3 or IC2/3-WT population at the time of the interim analysis of OS in the TC3 or IC3-WT subpopulation. If the OS interim analysis in the TC3 or IC3-WT population is claimed as statistically significant, the OS analysis in the TC2/3 or IC2/3-WT population will be tested under the overall type I error of 0.05. If there are significantly fewer than 154 OS events (i.e., <135 events) in the TC2/3 or IC2/3-WT population at the time of TC3 or IC3-WT OS interim analysis, a nominal two-sided alpha of 0.0001 (negligible impact on overall type I error rate) will be spent on the OS interim analysis in the TC2/3 or IC2/3-WT population. The next interim and final OS analysis in the TC2/3 or IC2/3-WT population will be conducted when approximately 154 and 216 events are observed, respectively. The interim and final analyses of OS in TC1/2/3 or IC1/2/3-WT would be conducted at the same time as those for the TC2/3 or IC2/3-WT population.

If the OS interim analysis in the TC3 or IC3-WT population is not claimed as statistically significant, the final analysis will be conducted when approximately 135 OS events have occurred in this population and if this final analysis is claimed statistically significant, the OS in the TC2/3 or TC2/3-WT and TC1/2/3 or IC1/2/3-WT populations will be tested at the planned interim and final analyses accordingly. Details are provided at the end of this section.

The interim and final analyses are expected to occur approximately 40 and 55 months, respectively, after the first patient is enrolled in the study, but the exact timing of these analyses will depend on the occurrence of OS events. The stopping boundaries for the interim and final analyses are shown in Table 1.

Table 1 Analysis Timing and Stopping Boundary for Interim and Final Analysis for Overall Survival

Interim and Final Analysis for OS			
Analysis Population		**Analysis Timing Approximate Time from First Patient in (months)**	
		40	**55**
TC3 or IC3-WT	Information Fraction (No. of events)	71% (96)	100% (135)
	Stopping Boundary [b] HR (p-value [a])	HR ≤ 0.657 (p ≤ 0.0399)	HR ≤ 0.678 (p ≤ 0.0242)
TC2/3 or IC2/3-WT	Information Fraction (No. of events)	71% (154)	100% (216)
	Stopping Boundary [b] HR (p-value [a])	HR ≤ 0.718 (p ≤ 0.0400)	HR ≤ 0.736 (p ≤ 0.0242)
TC1/2/3 or IC1/2/3-WT	Information Fraction (No. of events)	72% (281)	100% (390)
	Stopping Boundary [b] HR (p-value [a])	HR ≤ 0.783 (p ≤ 0.0403)	HR ≤ 0.796 (p ≤ 0.0241)

HR = hazard ratio; IC = tumor-infiltrating immune cell; OS = overall survival; TC = tumor cell; WT = wild type.

[a] Two-sided p-value.

[b] The Lan-DeMets approximation to the Pocock boundary

Boundaries will be adjusted based on observed number of OS events by using the Lan-DeMets approximation to the Pocock boundary. Hypothesis testing will be conducted in the order described below:

- Step 1: The interim analysis will test the null hypothesis of no difference in OS between the two arms in the TC3 or IC3-WT subpopulation. If the two-sided p-value that corresponds to the stratified log-rank test is ≤3.99%, the null hypothesis will be rejected and it will be concluded that atezolizumab prolongs the duration of OS relative to control treatment in the TC3 or IC3-WT subpopulation.
- [redacted]
- [redacted]

[redacted]

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

3. STUDY CONDUCT

3.1 RANDOMIZATION ISSUES

A permuted-block randomization was applied to ensure a balanced assignment to each treatment arm for the following stratification factors:

- Sex (male vs. female)
- ECOG performance status (0 vs. 1)
- Histology (non-squamous vs. squamous)
- PD-L1 tumor expression status (TC1/2/3 and any IC vs. TC0 and IC1/2/3)

3.2 DATA MONITORING

An iDMC will be used to evaluate safety data during the study on a regular basis and evaluate efficacy data at the prespecified OS interim analysis. Members of the iDMC will be external to the Sponsor and will follow a charter that outlines their roles and responsibilities. All summaries and analyses by treatment arm for the iDMC review will be prepared by an independent Data Coordinating Center (iDCC).

The safety data will include demographic data, adverse event data (including serious adverse events and adverse events of special interest), study conduct data, and relevant laboratory data. Following the safety data reviews, the iDMC will provide a recommendation to the Sponsor as to whether the study may continue, whether amendments to the protocol should be implemented, or whether the study should be stopped. The final decision will rest with the Sponsor.

The interim analysis of efficacy data will be conducted in accordance with the methods that are specified in Section 2.4. The iDMC recommendations to unblind the study because of substantial evidence of efficacy of the study drug or to continue to the final analysis must be based on the specified interim analysis guidelines as specified in the iDMC Charter.

Any outcomes of these safety and efficacy reviews that affect study conduct will be communicated in a timely manner to the investigators for notification of the Institutional Review Boards and/or Ethics Committee.

4. STATISTICAL METHODS

The analyses described in this SAP will supersede those specified in the protocol for Study GO29431 for the purposes of a regulatory filing.

4.1 ANALYSIS POPULATIONS

4.1.1 Efficacy Analysis Populations

The randomized population or intent-to-treat (ITT) population is defined as all randomized patients, regardless of receipt of the assigned treatment.

The primary efficacy endpoint, OS, will be analyzed in three PD-L1–selected populations, the TC3 or IC3-WTsubpopulation, the TC2/3 or IC2/3-WT subpopulation and the TC1/2/3 or IC1/2/3-WT population.

Unless otherwise indicated, the secondary efficacy endpoints will be analyzed in TC1/2/3 or IC1/2/3-WT, TC2/3 or IC2/3-WT, and/or TC3 or IC3-WT (depending on the results of the primary endpoint analyses, referred to as the "secondary efficacy analysis populations" throughout Section 4.4.2).

Unless otherwise indicated, the exploratory efficacy endpoints will be analyzed in TC1/2/3 or IC1/2/3-WT, TC2/3 or IC2/3-WT, and/or TC3 or IC3-WT (depending on the results of the primary endpoint analyses, referred to as the "exploratory efficacy analysis populations" throughout Section 4.4.3).

For efficacy analyses, patients will be grouped per the treatment that was assigned at the time of randomization regardless whether they received any assigned study drug.

4.1.2 Pharmacokinetic–Evaluable Population

PK analyses will be based on PK observations from all randomized patients who received atezolizumab and provided at least one PK sample that was evaluable.

4.1.3 Safety Population

Safety analyses will include treated patients, defined as randomized patients who received any amount of study treatment.

For the safety analyses, patients will be grouped according to whether any amount of atezolizumab was received, including when atezolizumab was received in error. Specifically for patients who were randomized to the control arm, if atezolizumab was received in error, patients will be grouped to the atezolizumab arm for the safety analyses.

4.2 ANALYSIS OF STUDY CONDUCT

Enrollment, study drug administration, and discontinuation from the study will be summarized by treatment arm. The incidence of study drug discontinuation will similarly be tabulated. Protocol deviations, including major deviations of inclusion/exclusion criteria, will be summarized in a similar manner by treatment arm.

4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

Demographic and baseline characteristics, such as age, sex, race/ethnicity, and baseline disease characteristics, such as ECOG performance status and histology, will be summarized by treatment arm for the TC3 or IC3-WT subpopulation, the TC2/3 or IC2/3-WT subpopulation, and/or the TC1/2/3 or IC1/2/3-WT population (depending on the results of the primary endpoint analyses). Baseline measurements are the last available data obtained prior to Cycle 1, Day 1 visits unless otherwise stated. Descriptive statistics (mean, median, SD, and range) will be presented for continuous data, and frequencies and percentages will be presented for categorical data.

4.4 EFFICACY ANALYSIS

4.4.1 Primary Efficacy Endpoint

The primary efficacy endpoint is OS defined as the time from randomization to death from any cause. Data for patients who are not reported as having died at the time of the analysis will be censored at the date when they were last known to be alive. Patients

who do not have any post-baseline information will be censored at the date of randomization plus 1 day.

The primary efficacy analysis is the comparison of OS between the two treatment arms (atezolizumab arm and chemotherapy control arm). Treatment comparisons will be based on the stratified log-rank test. For the TC3 or IC3-WT subpopulation and the TC2/3 or IC2/3-WT subpopulation, the stratification factors will be those that were used during randomization (i.e., sex [male vs. female], ECOG performance status [0 vs. 1], histology [non-squamous vs. squamous]). For the TC1/2/3 or IC1/2/3-WT population, the stratification factors will be those that were used during randomization (i.e., sex [male vs. female], ECOG performance status [0 vs. 1], histology [non-squamous vs. squamous], and PD-L1 tumor expression status [TC1/2/3 and any IC vs. TC0 and IC1/2/3]) as recorded in the interactive Web/voice response system (IxRS). Due to the potential risk of over-stratification (Akazawa et al. 1997), if at least 1 stratum (i.e., for the TC1/2/3 or IC1/2/3-WT population, a combination of stratification factor levels across sex, ECOG performance status, histology, and PD-L1 tumor expression status per IxRS; for the TC3 or IC3-WT population and the TC2/3 or IC2/3-WT population, a combination of stratification factor levels across sex, ECOG performance status, and histology per IxRS) has less than 10 OS events, the stratification factor (1 of 4 stratification factors for the TC1/2/3 or IC1/2/3 population: sex, ECOG performance status, histology, and PD-L1 tumor expression by IHC per IxRS; 1 of 3 stratification factors for the TC3 or IC3-WT population and the TC2/3 or IC2/3-WT population: sex, ECOG performance status, and histology per IxRS) which contains the level with the smallest number of patients will be removed from the stratified analyses. The removal of the stratification factor will continue until there is no stratum with less than 10 OS events in the analysis population. The final set of stratification factors used in the stratified analyses of OS for a specific analysis population (e.g., TC2/3 or IC2/3) will be applied to all other efficacy endpoints where stratified analyses are planned for the same analysis population. The analyses that are based on stratification factors as recorded on the electronic Case Report Form (eCRF) may also be provided. Results from an unstratified analysis will also be provided.

The null and alternative hypotheses for the OS analysis can be phrased in terms of the survival functions $S_A(t)$ and $S_B(t)$ in the atezolizumab arm (Arm A) and the control arm (Arm B), respectively:

$$H0: \ S_A(t) = S_B(t) \text{ versus } H1: \ S_A(t) \neq S_B(t).$$

The Kaplan-Meier methodology will be used to estimate the median OS for each treatment arm and construct survival curves for the visual description of the difference between the treatment arms. The Brookmeyer-Crowley methodology will be used to construct the 95% CI for the median OS for each treatment arm (Brookmeyer and Crowley 1982).

The HR, λ_A/λ_B, where λ_A and λ_B represent the hazard of OS in the atezolizumab arm and the control arm, respectively, will be estimated with a stratified Cox regression model

with the same stratification variables that are used in the stratified log-rank test at the interim and final analyses. The unstratified HR will also be presented.

4.4.2 Secondary Efficacy Endpoints

Unless otherwise indicated, the secondary efficacy endpoints will be analyzed in the TC3 or IC3-WT subpopulation, the TC2/3 or IC2/3-WT subpopulation, and/or the TC1/2/3 or IC1/2/3-WT population (depending on the results of the primary endpoint analyses) (referred to as the "secondary efficacy analysis populations" throughout Section 4.4.2).

4.4.2.1 Progression-Free Survival

PFS is defined as the time (in months) from randomization to the first occurrence of disease progression, as determined by the investigator with use of RECIST v1.1, or death, whichever occurs first. Patients who have not experienced disease progression or death at the time of analysis will be censored at the time of last tumor assessment. Patients with no post-baseline tumor assessment will be censored at the date of randomization plus 1 day.

The PFS analyses will occur at the time of OS interim analyses (as defined in Section 2.4). [redacted] the PFS analyses will be performed based on the prespecified number of events in the TC3 or IC3-WT subpopulation, TC2/3 or IC2/3-WT subpopulation, and the TC1/2/3 or IC1/2/3-WT population. Specifically PFS for the TC3 or IC3-WT, TC2/3 or IC2/3-WT, and TC1/2/3 or IC1/2/3-WT populations will be tested in the prespecified order at a two-sided significance level of 0.05, respectively. [redacted],

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

PFS will be analyzed through use of the same methods described for the OS analysis (see Section 4.4.1).

4.4.2.2 Objective Response Rate

An objective response is defined as either CR or PR, as determined by the investigator with use of RECIST v1.1. Patients not meeting these criteria, including patients without any post-baseline tumor assessments, will be considered non-responders.

ORR is defined as the proportion of patients who had an objective response. The analysis of ORR will be performed for patients in the secondary efficacy analysis populations who have measurable disease at baseline. The confirmation of response in accordance with RECIST v.1.1 is not required, but for the exploratory purpose, rate of confirmed response may be reported as needed.

An estimate of ORR and its 95% CI will be calculated using the Clopper-Pearson method for each treatment arm. The 95% CIs for the difference in ORRs between the two arms will be computed using the normal approximation to the binomial distribution. The ORR will be compared between the two arms using the stratified Cochran-Mantel-Haenszel test, stratified by the same factors used in the primary OS analysis (see Section 4.4.1).

4.4.2.3 Overall Survival at 1- and 2-Year Landmark Timepoints

The 1- and 2-year OS rates are defined as the probabilities that patients are alive 1 and 2 years after randomization, respectively. The Kaplan–Meier method will be used to estimate the 1- and 2-year OS rates for each treatment arm. The 95% CI for the difference in OS rates between the two treatment arms will be estimated with the normal approximation method and standard errors computed using the Greenwood method.

4.4.2.4 Duration of Response

DOR is defined for patients with an objective response as the time from the first documented objective response to documented disease progression as determined by the investigator with use of RECIST v1.1 or death from any cause, whichever occurs first. Data for patients who are alive and have not experienced disease progression at the time of the analysis will be censored at the date of the last tumor assessment. If no tumor assessments were performed after the date of the first occurrence of the objective response (CR or PR), DOR will be censored at the date of the first occurrence of the objective response. Because the determination of DOR is based on a non-randomized subset of patients, formal hypothesis testing will not be performed. Comparisons between treatment arms will be made for descriptive purposes. The same methodologies detailed for the OS analysis will be used for the DOR analysis.

Confirmation of response in accordance with RECIST v1.1 will not be required but for exploratory purposes, DOR for patients with confirmed response may be reported.

4.4.2.5 Patient-Reported Outcomes

Patient-reported data regarding symptoms of lung cancer, symptoms commonly associated with cancer treatments, and disease and treatment's impact on patients' functioning and HRQoL (assessed using the SILC, EORTC QLQ-C30 and EORTC QLQ-LC13) will be collected. Sensitivity analyses (e.g., missing data, in the PRO-evaluable population) may be conducted as appropriate.

The EORTC QLQ-C30 and EORTC QLQ-LC13 will be scored according to the EORTC scoring manual 3rd edition (Fayers et al. 2001). The QLQ-C30 and QLQ-LC13 are composed of both multi-item scales and single-item measures including functional scales, symptom scales, and a global health status/quality-of-life (QoL) scale. For multi-item scales, if $\leq$ 50% of items within the scale are missing at a given timepoint, the scale score will be calculated on the basis of the non-missing items. If $>$ 50% of items in a multi-item scale are missing or if a single-item measure is missing, then the scale is considered missing. Each of the QLQ-C30 and QLQ-LC13 scale and single-item measure scores will be linearly transformed so that each score will range from 0−100. A high score for a functional scale represents a high/healthy level of functioning, a high score for the global health status/HRQoL scale represents high HRQoL; and a high score for a symptom scale/item represents a high level of symptomatology/problems. A $\geq$10-point change in the EORTC scale score is perceived by patients as clinically significant (Osoba et al. 1998).

Each SILC symptom scale (dyspnea, cough, chest pain) score will be calculated as the average of the component items (range 0 to 4). An increase in score suggests worsening in symptomatology. A symptom score change of 0.3 points for the dyspnea and cough scores is considered to be clinically significant; whereas a symptom score change of 0.5 points for the chest pain score is considered to be clinically significant.

Time to Deterioration in Patient-Reported Lung Cancer-Related Symptoms

TTD in each lung cancer-related symptom (cough, dyspnea, chest pain; assessed by the SILC) is defined as the time from randomization to confirmed clinically meaningful deterioration in each SILC symptom score. Confirmed clinically meaningful deterioration in lung cancer symptoms is defined as the first $\geq$0.3-point increase above baseline for the dyspnea and cough scores and the first $\geq$0.5-point increase above baseline for the chest pain score, which must be maintained for at least two consecutive assessments, or must be followed by death from any cause within one week from the last assessment.

TTD of lung cancer-related symptoms (cough, dyspnea, chest pain; assessed by the EORTC QLQ-LC13) as a 3-symptom composite endpoint is defined as the time from randomization to confirmed clinically meaningful deterioration in EORTC QLQ-LC13 cough, dyspnea, or chest pain score, whichever occurs first. Confirmed clinically meaningful deterioration in lung cancer symptoms is defined as the first $\geq$10-point increase above baseline in any of the three symptom scores (whichever occurs first), which must be held for at least two consecutive assessments; or the first $\geq$10-point increase above baseline must be followed by (1) death within 6 weeks from the last assessment through Week 48 or (2) death within 9 weeks from the last assessment from Week 48 thereafter. TTD in the individual QLQ-LC13 symptom scales will also be evaluated.

TTD analyses will be performed in the secondary efficacy analysis populations and will include all data collected through disease progression and survival follow-up. The

methodologies that are outlined for the analysis of OS will be used for these TTD analyses. If no baseline or post-baseline assessment is performed, patients will be censored at the randomization date plus 1 day. Patients without deterioration at the time of analysis will be censored at the last time they were known to have not deteriorated, i.e., the last assessment. There will be no imputation for missing baseline or post-baseline data for the TTD analysis. Sensitivity analyses around event definitions (e.g., time to permanent deterioration) may be conducted as appropriate.

Change from Baseline in Patient-Reported Lung Cancer-Related Symptoms

Summary statistics (mean, SD, median, interquartile range [IQR], range) of the change from baseline per SILC will be provided. The analysis will be performed for patients in the secondary efficacy analysis populations. Graphs of the mean changes and 95% CI over time from the baseline assessment for the total scale and subscales will be provided. The analyses will be performed at all on-treatment timepoints, as well as at the time of disease progression per RECIST v1.1 (PRO assessment completed within ±7 days of date of radiographic progressive disease), at the last dose of treatment received before treatment discontinuation of any cause, and the survival follow-up visits through 6 months. Longitudinal mixed models (e.g., repeated measures, linear) may be used to estimate least-square mean score changes from baseline by treatment arm and differences between treatment arms (including effect sizes). Within-patient change may be evaluated using cumulative distribution function (CDF) analyses by treatment arm.

4.4.2.6 Overall Survival and Investigator-Assessed Progression-Free Survival for SP263 PD-L1 and bTMB Subpopulation

OS and investigator-assessed PFS in patients with PD-L1 expression on tumor cells defined by the SP263 IHC assay (including ≥1%, ≥25%, and ≥50% of tumor cells) and in patients with bTMB (including bTMB ≥ 10, bTMB ≥ 16, and bTMB ≥ 20 mutations) will be analyzed and included in the Clinical Study Report (CSR; pending data availability). This will be analyzed through use of the same methods described for the OS analysis.

4.4.3 Exploratory Efficacy Endpoints

Unless otherwise indicated, the exploratory efficacy endpoints will be analyzed in the TC3 or IC3-WT subpopulation, the TC2/3 or IC2/3-WT subpopulation, and/or the TC1/2/3 or IC1/2/3-WT population (depending on the results of the primary endpoint analyses) (referred to as the “exploratory efficacy analysis populations” throughout Section 4.4.3).

4.4.3.1 Analyses of Overall Survival and Progression-Free Survival in the 22c3 PD-L1 IHC Assay Subpopulation

OS and investigator-assessed PFS, according to RECIST v1.1, will be analyzed in subpopulation of patients with PD-L1 expression using the 22c3 PD-L1 IHC assay, excluding patients with a sensitizing *EGFR* mutation or *ALK* translocation and will be analyzed with the same methods as described in Section 4.4.1. The details of the

analysis may be described in a separate Biomarker Analysis Plan (BAP). These exploratory analyses will not be included in the CSR.

4.4.3.2 PFS Analysis at 6-Month and 1-Year Landmark Timepoints

The 6-month and 1-year PFS rates are defined as the probability that patients are alive without disease progression at 6 months and 1 year after randomization, respectively. The Kaplan–Meier method will be used to estimate the 6-month and 1-year PFS rate for each treatment arm. The 95% CI for the difference in PFS rates between the two treatment arms will be estimated with the normal approximation method and the standard errors computed with the Greenwood method.

4.4.3.3 Overall Survival Analysis at 3-Year Landmark Timepoint

The 3-year OS rate is defined as defined as the probabilities that patients are alive 3 years after randomization. The Kaplan-Meier method will be used to estimate the 3-year OS rate for each treatment arm. The 95% CI for the difference in OS rates between the two treatment arms will be estimated with the normal approximation method and standard errors computed with the Greenwood method.

4.4.3.4 Impact of Demographic and Baseline Characteristics on Overall Survival and Progression-Free Survival

To assess the consistency of the study results in subgroups defined by demographic (e.g., age, sex, and race/ethnicity) and baseline prognostic characteristics (e.g., PD-L1 tumor expression status, histology, ECOG performance status, smoking history, number of metastatic sites, location of metastases, size of primary tumor, etc.), OS and investigator-assessed PFS according to RECIST v1.1 in these subgroups will be examined in the TC3 or IC3-WT, the TC2/3 or IC2/3-WT, and/or TC1/2/3 or IC1/2/3-WT populations, depending on the results of primary efficacy analyses. Summaries of OS and PFS, including unstratified HRs estimated from Cox proportional hazards models and Kaplan-Meier estimates of median OS and PFS, will be produced separately for each level of the categorical variables and displayed on a Forest plot (Lewis and Clarke 2001). Additionally, estimates of median OS, PFS, and survival curves will be constructed by histology and PD-L1 subgroups (TC3 or IC3-WT, TC2/3 or IC2/3-WT, and/or TC1/2/3 or IC1/2/3-WT, depending on the results of primary efficacy analyses), by the Kaplan-Meier methodology.

4.4.3.5 Exploratory Biomarker Analysis

Exploratory biomarker analyses in tumor tissue and/or blood will be performed in all randomized patients in an effort to understand the association of these biomarkers with study drug response, including efficacy and/or adverse events. The exploratory biomarkers may include but are not limited to PD-L1 and CD8, as defined by IHC, quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR), or other methods. Additional PD analyses may be conducted as appropriate. Results from these exploratory analyses will not be included in the CSR.

4.4.3.6 EQ-5D-3L Health Status Data

EQ-5D-3L data will be used to generate health status and utility scores for use in economic models. For the EQ-5D-3L health-state profiles, descriptive statistics that summarize the proportions of patients who reported having "no," "some," or "extreme" problems at each timepoint will be reported. Frequencies and percentages of missing data will also be reported at each timepoint. Patients without post-baseline assessments will be excluded from this analysis. A single summary index from the EQ-5D-3L health status will be utilized in this study for economic modeling. These results will not be reported in the CSR.

4.4.3.7 Additional Patient-Reported Outcome Analyses

EORTC QLQ-C30 and QLQ-LC13 scale score changes from baseline will be descriptively analyzed using summary statistics (mean, SD, median, IQR, range) by treatment arm for patients with a baseline assessment and at least one post-baseline assessment. Missing post-baseline items and subscales will not be imputed and will be summarized as observed. CDF analyses of global health status, physical function, and role function (as measured by the QLQ-C30) may be performed by treatment arm. The remaining QLQ-C30 scales (e.g., appetite loss, diarrhea) and QLQ-LC13 scales (e.g.,arm/shoulder pain, sore mouth) may be evaluated using longitudinal mixed models.

Completion rates will be calculated as the number of patients who completed a PRO assessment divided by the number of patients expected to complete a PRO assessment at each timepoint by treatment arm. Reasons for non-completion will be summarized if available in the CRF.

4.4.4 Sensitivity Analyses

Unless otherwise indicated, the sensitivity analysis will be conducted in the TC3 or IC3-WT subpopulation, the TC2/3 or IC2/3-WT subpopulation, and/or the TC1/2/3 or IC1/2/3-WT population (depending on the results of the primary endpoint analyses) (referred to as the "sensitivity analysis populations" throughout Section 4.4.4). In addition to the following sensitivity analyses, OS and investigator-assessed PFS according to RECIST v1.1 in the ITT population will be analyzed using the same methods described in Section 4.4.1.

4.4.4.1 Censoring for Non–Protocol-Specified Anti-Cancer Therapy

The impact of non–protocol-specified anti-cancer therapy on OS will be assessed depending on the number of patients who receive non–protocol-specified anti-cancer therapy. If $>10\%$ of patients received non–protocol-specified anti-cancer therapy before an OS event in any treatment arm, a "discounted" survival time will be generated based on a user-specified assumption for the OS effect for patients who started non–protocol-specified anti-cancer therapy. The duration from initiation of non–protocol-specified anti-cancer therapy to death or censoring date will be discounted in accordance with a range of possible effects on OS of subsequent non–protocol-specified anti-cancer therapy (e.g., 10%, 20%, 30%, etc.). Additional

sensitivity analyses such as an analysis with the rank that preserves a structural failure time model may be used to estimate patient OS without non–protocol-specified anti-cancer therapy, if deemed necessary. After adjustments are made for the effect of subsequent non–protocol-specified anti-cancer therapy on OS, the methods that outlined for the primary efficacy endpoint analyses will be used for these analyses (see Section 4.4.1).

4.4.4.2 Impact of Proportional Hazards Assumption

If the proportional hazards assumption is equivocal, restricted mean survival time may be assessed and the difference between treatment arms will be tested for several timepoints. The weighted log-rank test (Fleming and Harrington 1991) of OS may also be performed for treatment comparison with exploratory purpose. No multiplicity adjustment will be done for these sensitivity analyses.

4.4.4.3 Impact of Potential Imbalance between Treatment Arms

Since randomization is not stratified by the PD-L1 status of TC3 or IC3 and TC2/3 or IC2/3 at baseline, the treatment groups may be unbalanced with respect to some baseline disease characteristic and demographic factors within the subgroups. If substantial baseline factor imbalances are observed, the stratified Cox regression analysis including such prognostic factors may be performed to investigate the potential impact of such imbalances on the treatment effect.

4.4.4.4 Impact of Expansion of PD-L1 Selection Criteria for Patient Eligibility

The demographics and baseline characteristics may be summarized for patients enrolled before and after expansion of PD-L1 selection criteria for patient eligibility. Descriptive statistics (mean, median, SD, and range) may be presented for continuous data, and frequencies and percentages may be presented for categorical data.

4.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

PK samples will be collected in this study as outlined in Appendix 2. Atezolizumab serum concentration data (C_{min} and C_{max}) will be tabulated and summarized. Descriptive statistics will include means, medians, ranges, and SDs, as appropriate.

Additional PK analyses will be conducted as appropriate on the basis of the availability of data.

4.6 SAFETY ANALYSES

Unless specified otherwise, the safety analyses that are described below will be conducted for the safety population and patients will be grouped in accordance whether any atezolizumab treatment was received (i.e., patients who received any dose of atezolizumab will be included in the atezolizumab arm). Patients who were randomized but did not receive any amount of any study drug will be not included in safety analyses.

4.6.1 Exposure of Study Medication

Study drug exposure with inclusion of treatment duration, number of cycles, and dose intensity will be summarized for each treatment arm with descriptive statistics.

4.6.2 Adverse Events

Verbatim description of adverse events will be mapped to the medical dictionary for regulatory activities (MedDRA) thesaurus terms. Adverse events will be graded by the investigator in accordance with the NCI CTCAE v4.0. Treatment-emergent adverse events will be summarized by mapped term, appropriate thesaurus level, NCI CTCAE grade, and treatment arm. Multiple occurrences of the same event will be counted once at the maximum grade. Adverse events, common adverse events (defined as adverse events that occur in at least 10% of patients), serious adverse events, treatment-related serious adverse events, severe adverse events (Grade ≥ 3), adverse events of special interest, and adverse events that lead to study drug discontinuation or interruption will be summarized.

Treatment-emergent is defined as all events with onset on or after the first study drug treatment up to the data cutoff date. Listings of adverse events will include all adverse events with onset on or after the first study drug treatment up to the data cutoff date.

Deaths during the study treatment period and those reported during the follow-up period after treatment completion or discontinuation and causes of death will be summarized by treatment arm.

4.6.3 Laboratory Test Result Data

Laboratory test result data will be summarized over time with inclusion of change from baseline by treatment arm. Values outside normal ranges will be summarized. Additionally, selected laboratory test result data will be classified according to NCI CTCAE and summarized by grade. The highest NCI CTCAE grade post-baseline will also be reported and shift tables from baseline to worst value during the study post-baseline will be presented.

4.6.4 Vital Signs

Changes in selected vital signs will be summarized by treatment arm.

4.6.5 Anti-Therapeutic Antibody

Serum levels and incidence of ATAs, also known as ADAs, against atezolizumab will be summarized. The analyses of pharmacokinetics, key efficacy, and safety by ATA status will be conducted to explore the potential impact of immunogenicity.

4.7 MISSING DATA

Refer to Sections 4.4.1 and 4.4.2 for methods on how to handle missing data for the primary and secondary efficacy endpoints.

5. REFERENCES

Akazawa K, Nakamura T, Palesch Y. Power of Logrank Test and Cox Regression Model in Clinical Trials with Heterogeneous Samples. Statistics in Medicine 1997;16:583-597.

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

Fayers PM, Aaronson NK, Bjordal K, et al. EORTC QLQ-C30 scoring manual (3rd edition). Brussels: EORTC, 2001.

Fleming TR, Harrington DP. Counting processes and survival analysis. New York: John Wiley & Sons, 1991.

Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. BMJ 2001;322:1479–80.

Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16:139–44.

Appendix 2
Schedule of Assessments

Procedure	Screening Days –28 to –1	All Treatment Cycles [a] Every 21 days (± 3 Days) [c]	Treatment Discontinuation Visit ≤30 Days after Last Dose	Survival Follow-Up [b] Every 3 Months after Disease Progression or Loss of Clinical Benefit
Informed consent Prescreening for PD-L1 testing [d] Main ICF for study participation	x			
Tumor tissue specimen for PD-L1 testing (15 FFPE slides required; blocks preferred) [e, f] Fresh or archival tissue can be used.	x			
EGFR and/or *ALK* assessment if status unknown (required for patients with non-squamous disease) [g]	x			
Demographic data	x			
Medical history and baseline conditions	x			
NSCLC cancer history	x			
Vital signs [h]	x	x [i]	x [i]	
Weight	x	x	x	
Height	x			
Complete physical examination	x			
Limited physical examination [j]		x	x	
ECOG Performance Status	x	x	x	
12-lead ECG	x	x [k]	x [k]	

Appendix 2
Schedule of Assessments (cont.)

Procedure	Screening Days –28 to –1	All Treatment Cycles [a] Every 21 days (± 3 Days) [c]	Treatment Discontinuation Visit ≤30 Days after Last Dose	Survival Follow-Up [b] Every 3 Months after Disease Progression or Loss of Clinical Benefit
Hematology [l]	x [m]	x	x	
Serum chemistry [n]	x [m]	x	x	
Coagulation test (INR or aPTT)	x [m]		x	
Pregnancy test (women of childbearing potential ONLY)	x [o]	x [p]	x [p]	
TSH, free T3, free T4 [q]	x	x [r]	x [r]	
HIV, HBV, HCV serology [s]	x			
Urinalysis [t]	x			
Study treatment administration		x [u]		
Tumor response assessment	x [v]	x [w]		x [x]
Serum sample for ATA assessment (atezolizumab-treated patients only) [y]		x	x	120 (±30) days after last dose
Serum sample for PK sampling (atezolizumab-treated patients only) [y]		x	x	120 (±30) days after last dose
Blood samples for PD biomarkers [y]	x	x	x	120 (±30) days after last dose
Informed consent to continue treatment beyond radiographic progression (atezolizumab-treated patients only)		At time of radiographic progression		

Appendix 2
Schedule of Assessments (cont.)

Procedure	Screening Days –28 to –1	All Treatment Cycles [a] Every 21 days (± 3 Days) [c]	Treatment Discontinuation Visit ≤30 Days after Last Dose	Survival Follow-Up [b] Every 3 Months after Disease Progression or Loss of Clinical Benefit
Tumor biopsy		At time of radiographic progression [z]		
Adverse events	x [aa]	x	x [bb]	
Concomitant medications	From 7 days before screening	x	x	
Patient-reported outcomes [cc]		x		x [dd]
Survival and anti-cancer therapy follow-up [b]				x

anti-HBc = antibody against hepatitis B core antigen; ATA = anti-therapeutic antibody; CT = computerized tomography; ECOG=Eastern Cooperative Oncology Group; EORTC = European Organisation for Research and Treatment of Cancer; ePRO = electronic patient-reported outcome; EQ-5D-3L = EuroQoL5 Dimension, 3 Level; FFPE = formalin-fixed paraffin-embedded; HBV = hepatitis B virus; HCV = hepatitis C virus; ICF = Informed Consent Form; IV = intravenous; QLQ-LC13 = Lung Cancer module; MRI = magnetic resonance imaging; NSCLC = non–small cell lung cancer; PD = pharmacodynamic; PD-L1 = programmed death–ligand 1; PK = pharmacokinetic; QLQ-C30 = Quality-of-Life Questionnaire Core 30; SILC = Symptoms in Lung Cancer; TSH = thyroid-stimulating hormone.

[a] Assessments should be performed before study drug infusion unless otherwise noted.

[b] Survival follow-up information will be collected via telephone calls, patient medical records, and/or clinic visits approximately every 3 months until death, loss to follow-up, or study termination by the Sponsor, whichever occurs first. All patients will be periodically contacted for survival and new anti-cancer therapy information unless the patient requests to be withdrawn from follow-up (this request must be documented in the source documents and signed by the investigator). If the patient withdraws from the study, study staff may use a public information source (e.g., county records) when permissible to obtain information about survival status only.

Appendix 2
Schedule of Assessments (cont.)

c Cycle 1 must be performed within 5 days after the patient is randomized. Screening assessments performed ≤ 96 hours before Cycle 1 Day 1 are not required to be repeated for Cycle 1 Day 1. In addition, ECOG Performance Status, limited physical examination, and local laboratory tests may be performed ≤ 96 hours before Day 1 of each cycle as specified in Section 4.5.13.2.

d Patients have the option to sign the Prescreening ICF to consent to PD-L1 tissue testing during prescreening, prior to signing the main ICF for study participation. Consent may be obtained more than 28 days before initiation of study treatment. For patients with non-squamous NSCLC, EGFR and/or ALK status if unknown may be assessed locally or at a central lab. Additional tissue will be required for central testing of EGFR and/or ALK.

e If a representative FFPE tumor specimen in paraffin block (preferred) or 15 (or more), freshly cut, unstained serial sections (on slides) from an FFPE tumor specimen are not available for PD-L1 testing, contact the Medical Monitor to discuss to determine if the patient may participate in the study. Fine-needle aspiration (defined as samples that do not preserve tissue architecture and yield cell suspension and/or cell smears), brushing, cell pellets from pleural effusion, and lavage samples are NOT acceptable. For core needle biopsy specimens, at least three cores should be submitted for evaluation. Retrieval of archival tumor sample can occur more than 28 days prior to randomization. See Section 4.5.7.1 and Section 4.1.1 for details.

f For patients whose initial archival tumor tissue sample is PD-L1 negative, a biopsy can be performed at screening to submit fresh tissue for the purposes of testing PD-L1 status. A positive result in any tumor tissue sample will satisfy this eligibility criterion.

g Patients with non-squamous NSCLC who have an unknown EGFR and/or ALK status will be required to be tested at prescreening/screening. Patients with squamous NSCLC who have an unknown status will not be required to be tested at prescreening/screening. EGFR and/or ALK may be assessed locally or at a central lab. Additional tissue will be required for central testing of EGFR and/or ALK. See Section 4.1.1 Inclusion criteria for details.

h Vital signs include pulse rate, respiratory rate, blood pressures, and temperature.

i For both study treatment arms, the patient's vital signs should be recorded within 60 minutes before infusion, and during and after the infusion if clinically indicated. For the atezolizumab arm, vital signs should also be collected during the first infusion every 15 (± 5) minutes and within 30 (± 10) minutes after the first infusion if clinically indicated. For subsequent infusions in both arms, vital signs will be collected within 60 minutes prior to the infusion and should be collected during the infusion and within 30 (± 10) minutes after the infusion, if clinically indicated or if symptoms occurred in the prior infusion. See Section 4.5.4 for details.

j Symptom-directed physical examinations; see Section 4.5.3 for details.

k ECG recordings will be obtained when clinically indicated.

Appendix 2
Schedule of Assessments (cont.)

l Hematology consists of CBC, including RBC count, hemoglobin, hematocrit, WBC count with differential (neutrophils, lymphocytes, eosinophils, monocytes, basophils, and other cells), and platelet count. Hematology tests must be performed prior to Day 1 infusions, and for gemcitabine administration, also prior to Day 8 infusions.

m At screening, the patient must have adequate hematologic and end-organ function defined by laboratory results obtained within 14 days prior to randomization, as described in Section 4.1.1.

n Serum chemistry includes BUN, creatinine, sodium, potassium, magnesium, chloride, bicarbonate or total CO2 (if considered standard of care for the region), calcium, phosphorus, glucose, total bilirubin, ALT, AST, alkaline phosphatase, LDH, total protein, and albumin.

o Serum pregnancy test within 14 days before Cycle 1, Day 1.

p Urine pregnancy tests; if a urine pregnancy test is positive, it must be confirmed by a serum pregnancy test.

q Total T3 will be tested only at sites where free T3 is not performed.

r Thyroid function testing (TSH, free T3, free T4) collected at Cycle 1, Day 1, and every fourth cycle thereafter.

s All patients will be tested for HIV prior to the inclusion into the study and HIV-positive patients will be excluded from the clinical study. Patients with active hepatitis B (chronic or acute; defined as having a positive HBsAg test result at screening) will be excluded from the study. Patients with past or resolved HBV infection (defined as the presence of HBcAb and absence of HBsAg) are eligible; HBV DNA must be performed prior to randomization in these patients. Patients with HCV will be excluded from the study; patients who test positive for HCV antibody are eligible only if polymerase chain reaction is negative for HCV RNA.

t Urinalysis by dipstick (specific gravity, pH, glucose, protein, ketones, and blood). Urinalysis is required at screening and will be obtained during study treatment when clinically indicated.

u For atezolizumab, the initial dose will be delivered over 60 (± 10) minutes. If the first infusion is well tolerated, all subsequent infusions may be delivered over 30 (± 10) minutes until disease progression per RECIST v1.1 or loss of clinical benefit. For chemotherapy, study drug will be administered according to the doses and suggested infusion times, including premedication, as described in Section 4.3.2.2-4.3.2.4.

v CT scans (with oral/IV contrast unless contraindicated) or MRI of the chest and abdomen. A CT or MRI scan of the pelvis is required at screening and as clinically indicated or as per local standard of care at subsequent response evaluations. A CT (with contrast) or MRI scan of the head must be done at screening to evaluate CNS metastasis in all patients. See Section 4.5.5 for details.

w Perform every 6 weeks (± 7 days; approximately every two cycles) for 48 weeks following Cycle 1, Day 1, and then every 9 weeks (± 7 days) thereafter, after completion of the Week 48 tumor assessment, regardless of treatment delays, until radiographic disease progression per RECIST v1.1 (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1), withdrawal of consent, death, or study termination by the Sponsor, whichever occurs first. CT scans may be repeated at any time if progressive disease is suspected. Tumor assessments will be performed according to RECIST v1.1 (see

Appendix 4) and modified RECIST (see Appendix 5) for patients in the atezolizumab arm, and only according to RECIST v1.1 for patients in the chemotherapy arm. See Section 4.5.5 f or details.

x If a patient discontinues study treatment for any reason other than radiographic disease progression per RECIST v1.1 (e.g., toxicity, symptomatic deterioration), tumor assessments will continue at the same frequency as would have been followed if the patient had remained on study treatment until radiographic disease progression per RECIST v1.1 (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1), withdrawal of consent, death, or study termination by the Sponsor, whichever occurs first, even if patient starts another anti-cancer therapy after study treatment discontinuation.

y See Appendix 2 for detailed schedule.

z Mandatory tumor biopsy at radiographic disease progression, if clinically feasible, within 40 days of radiographic progression or prior to the start of the next anti-cancer therapy, whichever is sooner (see Section 4.5.7.2).

[aa] Only serious adverse events caused by protocol-mandated intervention should be reported.

[bb] All serious adverse events and adverse events of special interest, regardless of relationship to study drug, will be reported until 90 days after the last dose of study drug or initiation of new systemic anti-cancer therapy after last dose of study treatment. All other adverse events, regardless of relationship to study drug, will be reported until 30 days after the last dose of study drug or initiation of new systemic anti-cancer therapy after last dose of study treatment. After this period, all deaths should continue to be reported. In addition, the Sponsor should be notified if the investigator becomes aware of any serious adverse event or adverse event of special interest that is believed to be related to prior exposure to study treatment (see Section 5.6).

[cc] EORTC QLQ-C30, EORTC QLQ-LC13, and EQ-5D-3L questionnaires will be completed by the patients on the ePRO tablet according to the tumor assessment schedule (i.e., every 6 weeks (± 7 days) for 48 weeks following Cycle 1, Day 1, and every 9 weeks (±7 days) thereafter after the completion of the Week 48 tumor assessment until radiographic disease progression per RECIST v1.1 (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1)prior to administration of study drug and prior to any other study assessments. The SILC scale will be completed using an electronic device at the patient's home on a weekly basis. Patients who discontinue study treatment for any reason other than disease progression per RECIST v1.1 (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1) will complete the EORTC QLQ-C30, EORTC QLQ-LC13, and EQ-5D-3L at each tumor assessment visit and will complete the SILC at home on a weekly basis, until radiographic disease progression per RECIST v1.1, unless the patient withdraws consent or the Sponsor terminates the study. Study personnel should review all questionnaires for completeness before the patient leaves the investigational site. Patients whose native language is not available in the ePRO device or who are deemed by the investigator incapable of inputting their ePRO assessment after undergoing appropriate training are exempt from all ePRO assessments.

Appendix 2
Schedule of Assessments (cont.)

[dd] During survival follow-up, the EORTC QLQ-C30, EORTC QLQ-LC13, and EQ-5D-3L will be completed at 3 and 6 months following disease progression (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1) *if the patient returns to the clinic*. The SILC scale will be completed monthly during survival follow-up for 6 months following disease progression (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1).

Appendix 2
Schedule of Assessments (cont.)

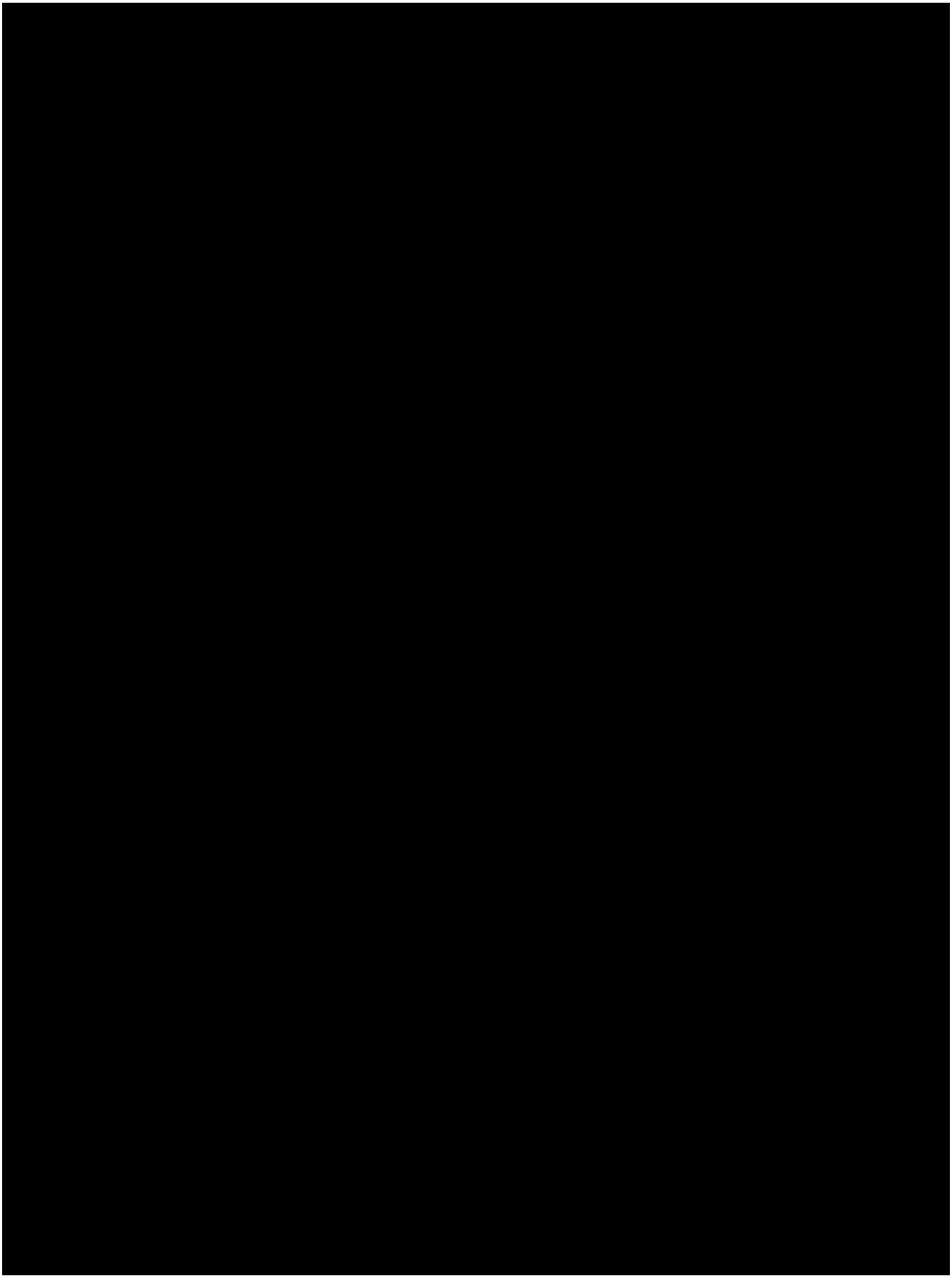
Schedule of Pharmacokinetic, Biomarker, and Anti-Therapeutic Antibody Assessments

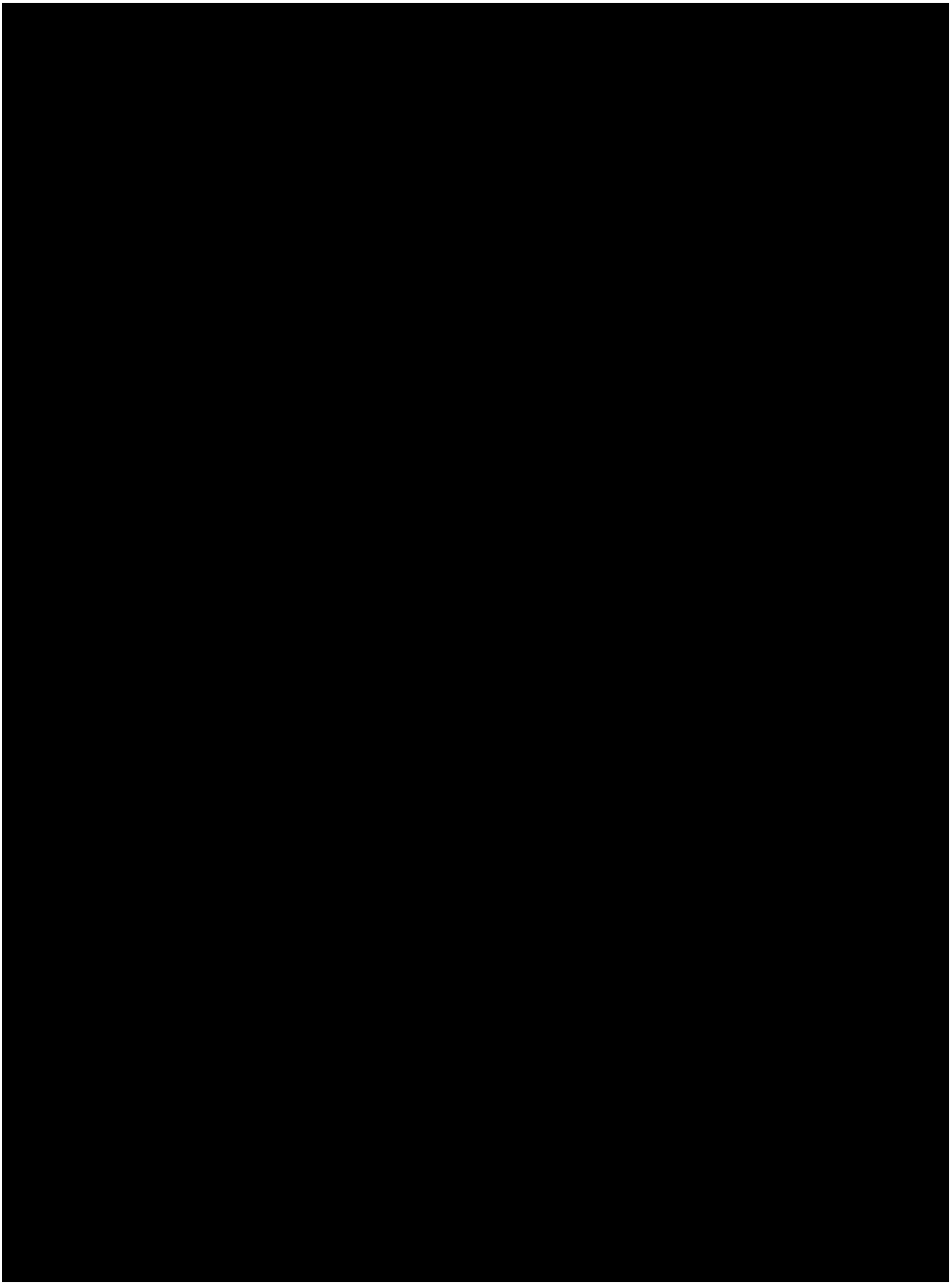
Study Visit	Time	Patients Randomized to Chemotherapy	Patients Randomized to Atezolizumab
Screening	—	Biomarkers [a]	Biomarkers [a]
Cycle 1, Day 1	Prior to dosing (same day as treatment administration)	Biomarkers [b]	ATA Atezolizumab pharmacokinetics Biomarkers b
	30 (± 10) minutes after end of atezolizumab infusion	—	Atezolizumab pharmacokinetics
Cycles 2, 3, 4, 8 and 16, Day 1	Prior to dosing (same day as treatment administration)	Biomarkers [b]	ATA Atezolizumab pharmacokinetics Biomarkers [b]
After Cycle 16, every eighth cycle, Day 1	Prior to dosing (same day as treatment administration)	Biomarkers [b]	ATA Atezolizumab pharmacokinetics Biomarkers [b]
At time of fresh biopsy (on-treatment, or at progression, including during follow-up)	At visit	Biomarkers [b]	Biomarkers [b]
Treatment discontinuation visit	At visit	Biomarkers [b]	ATA Atezolizumab pharmacokinetics Biomarkers [b]
120 ± 30 days after last dose of atezolizumab	At visit	—	ATA Atezolizumab pharmacokinetics Biomarkers [b]
Any time point during the study (RCR consent required)		Optional RCR blood (DNA extraction) [b]	Optional RCR blood (DNA extraction) [b]

ATA = anti-therapeutic antibody; RCR = Roche Clinical Repository.

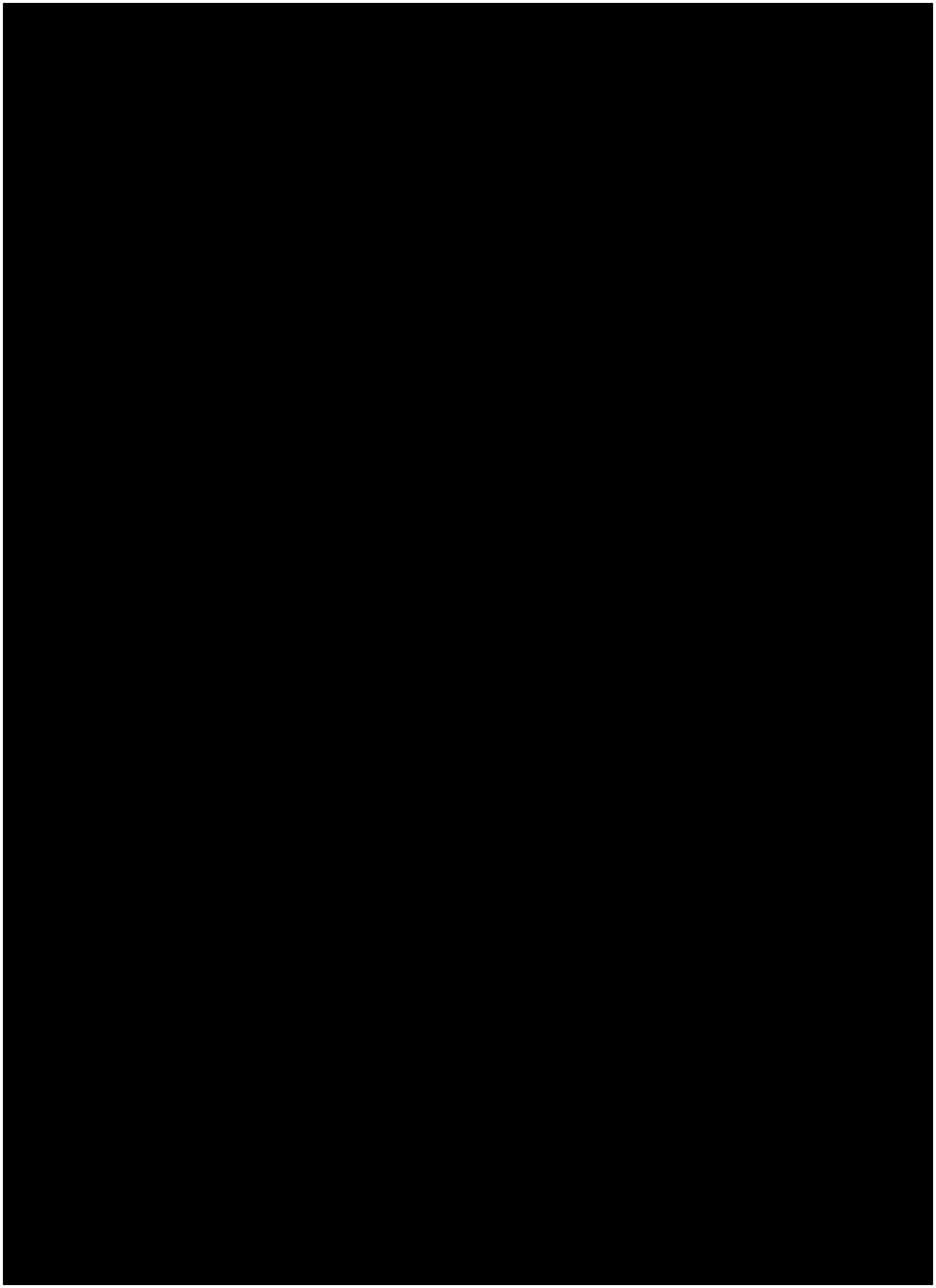
[a] Whole blood for biomarkers.

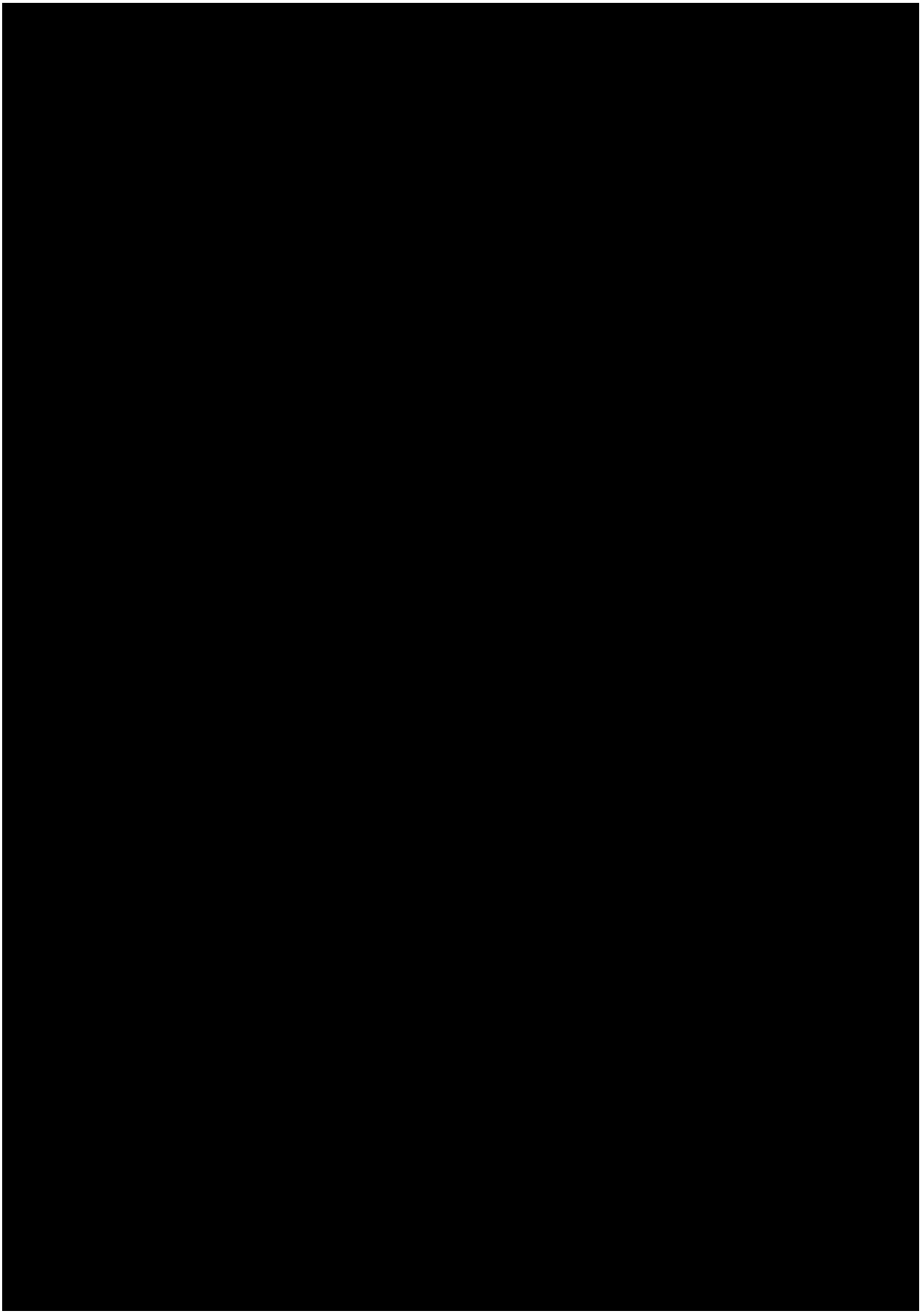
[b] Plasma and serum for biomarkers.



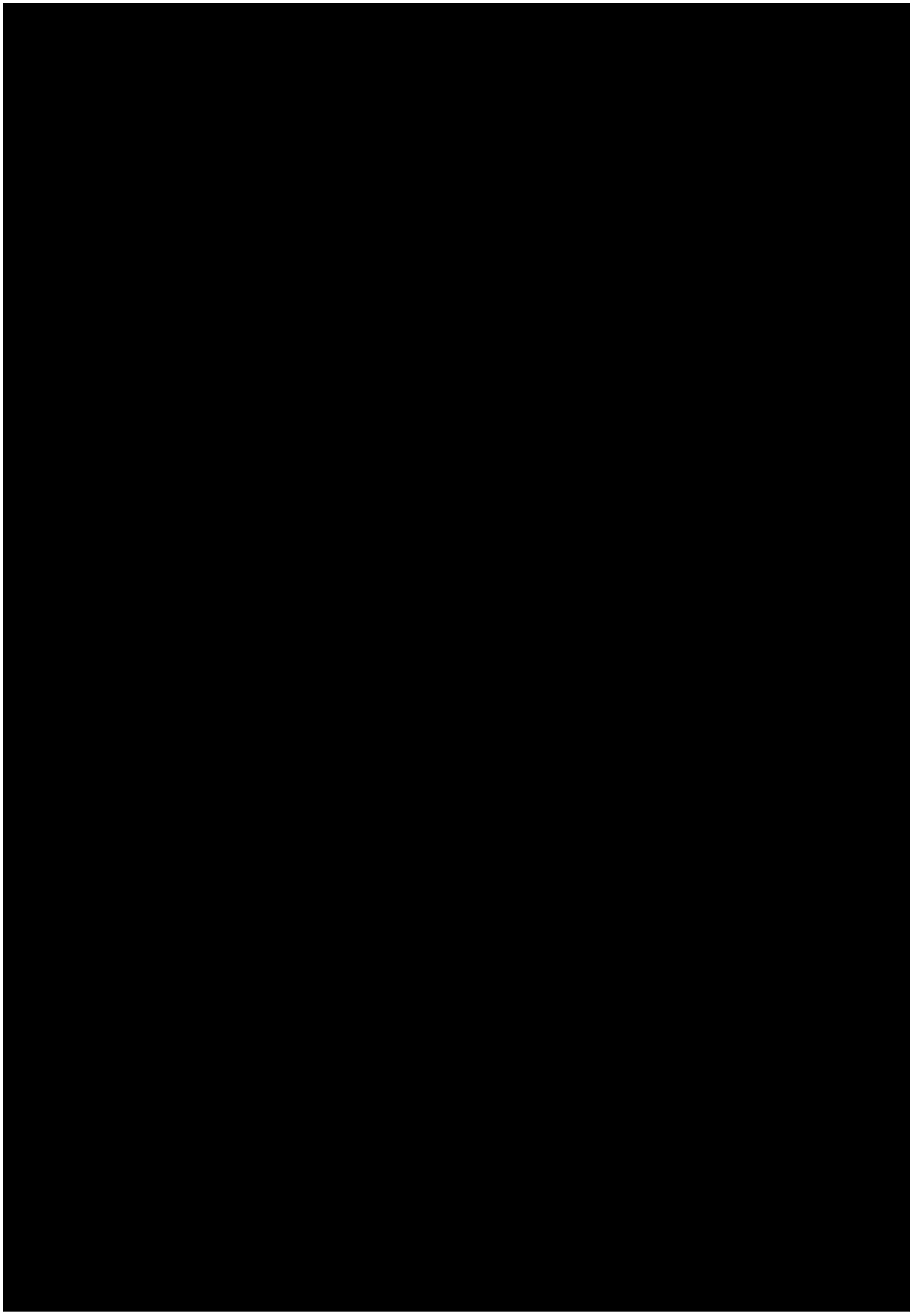


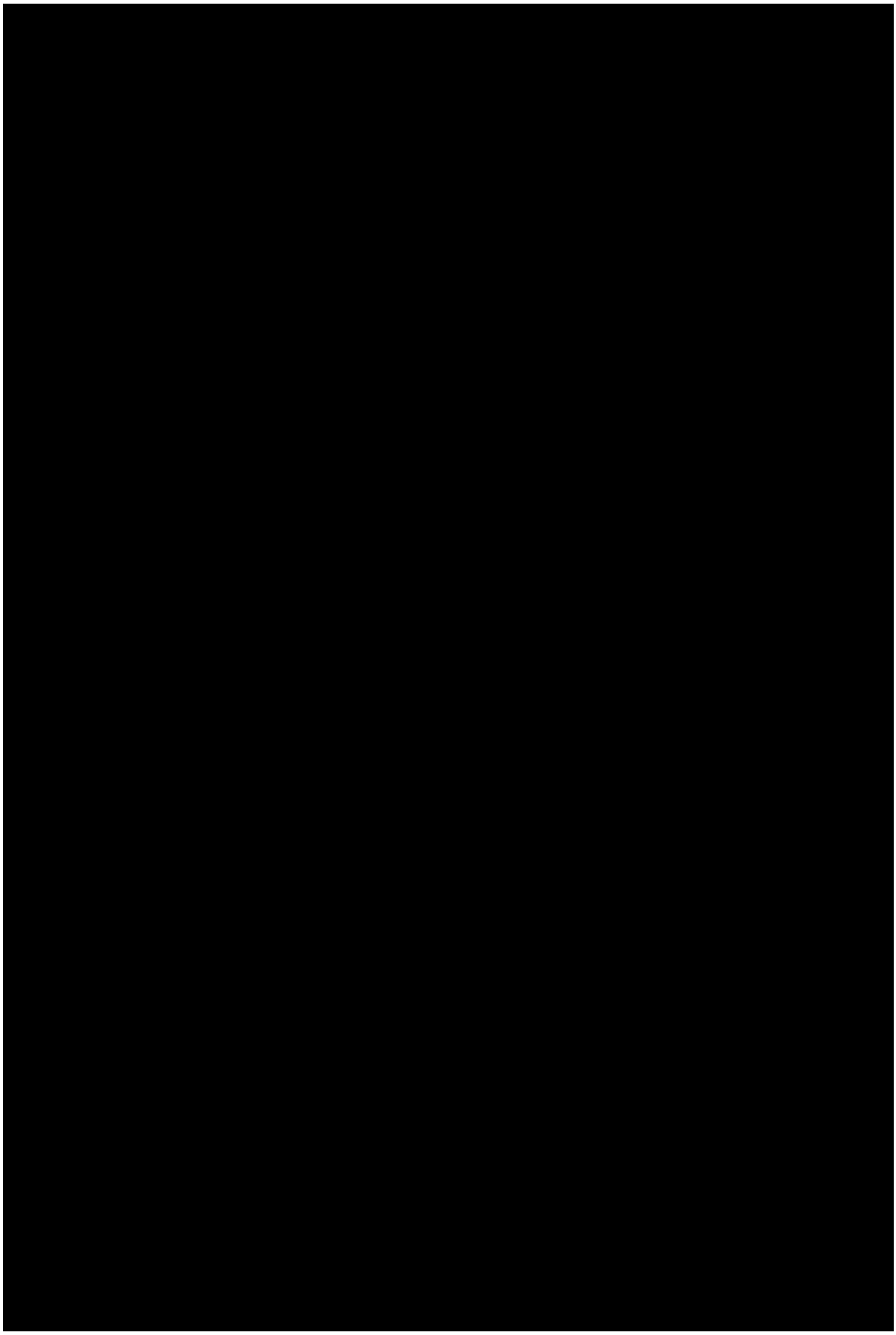
[redacted]

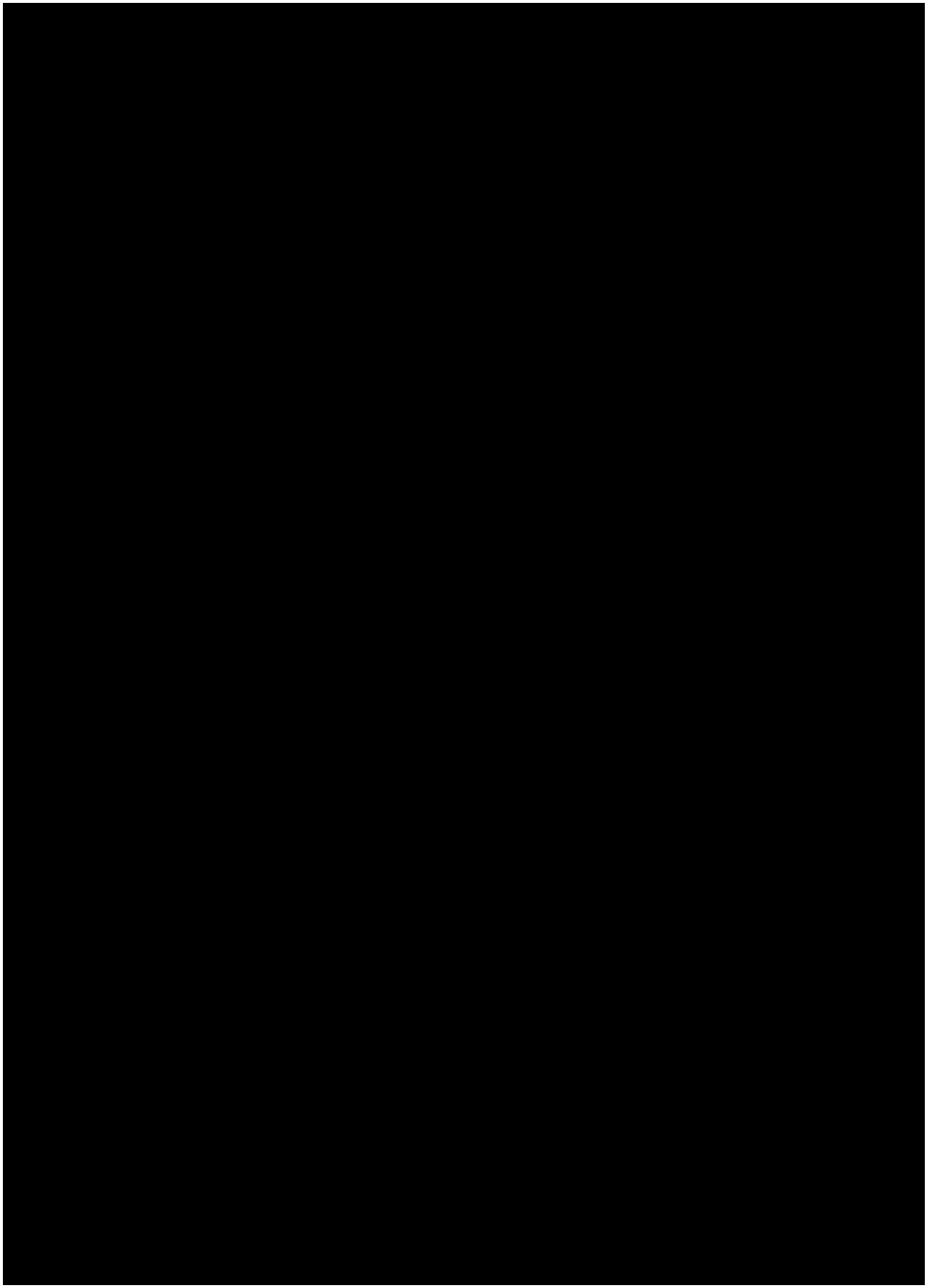




[redacted]







[redacted]

[redacted]

[redacted]