Official Title: **A PHASE III, OPEN-LABEL, RANDOMIZED STUDY OF ATEZOLIZUMAB (MPDL3280A, ANTI−PD-L1 ANTIBODY) IN COMBINATION WITH CARBOPLATIN + PACLITAXEL WITH OR WITHOUT BEVACIZUMAB COMPARED WITH CARBOPLATIN + PACLITAXEL + BEVACIZUMAB IN CHEMOTHERAPY-NAIVE PATIENTS WITH STAGE IV NON-SQUAMOUS NON−SMALL CELL LUNG CANCER**

NCT Number: NCT02366143

Document Date: SAP Version 1: 20-Apr-17

STATISTICAL ANALYSIS PLAN

TITLE: **A PHASE III, OPEN-LABEL, RANDOMIZED STUDY OF ATEZOLIZUMAB (MPDL3280A, ANTI–PD-L1 ANTIBODY) IN COMBINATION WITH CARBOPLATIN + PACLITAXEL WITH OR WITHOUT BEVACIZUMAB COMPARED WITH CARBOPLATIN + PACLITAXEL + BEVACIZUMAB IN CHEMOTHERAPY-NAIVE PATIENTS WITH STAGE IV NON-SQUAMOUS NON–SMALL CELL LUNG CANCER**

PROTOCOL NUMBER: GO29436

STUDY DRUG: Atezolizumab (RO5541267)

VERSION NUMBER: 1

IND NUMBER: 117296

EUDRACT NUMBER: 2014-003207-30

SPONSOR: F. Hoffmann-La Roche Ltd

PLAN PREPARED BY: [redacted]

DATE FINAL: See electronic date stamp below.

Name	Reason for Signing	Date and Time (UTC)
[redacted]	Company Signatory (Clinical)	20-Apr-2017 21:37:13

STATISTICAL ANALYSIS PLAN APPROVAL

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 GO29436 (IMpower150), "A Phase III, Open-Label, Randomized Study of Atezolizumab (MPDL3280A, Anti–PD-L1 Antibody) in Combination with Carboplatin + Paclitaxel with or without Bevacizumab Compared with Carboplatin + Paclitaxel + Bevacizumab in Chemotherapy-Naive Patients with Stage IV Non-Squamous Non–Small Cell Lung Cancer." The background for the study can be found in the study protocol.

2. STUDY DESIGN

This randomized, Phase III, multicenter, open-label study is designed to evaluate the safety and efficacy of TECENTRIQ® (atezolizumab) in combination with carboplatin + paclitaxel with or without bevacizumab compared with treatment with carboplatin + paclitaxel + bevacizumab in approximately 1200 chemotherapy-naive patients with Stage IV non-squamous non–small cell lung cancer (NSCLC).

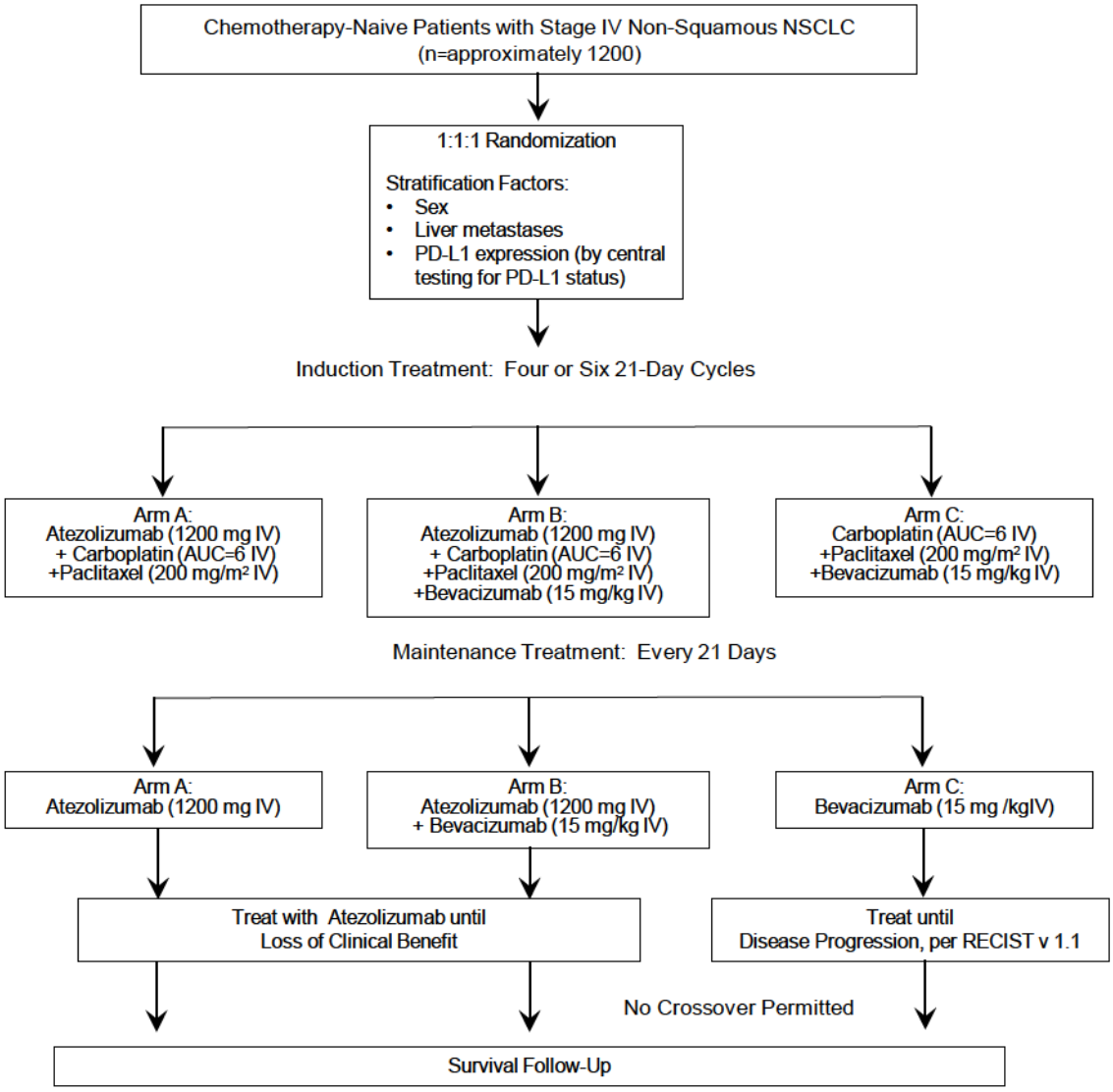
Eligible patients were stratified by sex (male vs. female), presence of liver metastases at baseline (yes vs. no), and programmed death–ligand 1 (PD-L1) tumor expression by immunohistochemistry (IHC) (tumor cell [TC] 3 and any tumor-filtrating immune cell [IC] vs. TC0/1/2 and IC2/3 vs. TC0/1/2 and IC0/1) tested by a central laboratory. TC0/1/2/3 are defined as PD-L1 staining in $<1\%$, $\geq 1\%$ to $<5\%$, $\geq 5\%$ to 50%, and $\geq 50\%$ of the TCs, respectively. IC0/1/2/3 are defined as PD-L1–stained ICs covering $<1\%$, $\geq 1\%$ to $<5\%$, $\geq 5\%$ to 10%, and $\geq 10\%$ of the tumor area, respectively.

Eligible patients were randomized in a 1:1:1 ratio to one of the following treatment regimens:

- Treatment Arm A: atezolizumab + carboplatin + paclitaxel (induction: four or six 21-day cycles); atezolizumab (maintenance: 21-day cycles)
- Treatment Arm B: atezolizumab + carboplatin + paclitaxel + bevacizumab (induction: four or six 21-day cycles); atezolizumab + bevacizumab (maintenance: 21-day cycles)
- Treatment Arm C: carboplatin + paclitaxel + bevacizumab (induction: four or six 21-day cycles); bevacizumab (maintenance: 21-day cycles)

Patients who were randomized to Arm A or Arm B may continue treatment with atezolizumab beyond radiographic progression by Response Evaluation Criteria in Solid Tumors Version 1.1 (RECIST v1.1), provided they were experiencing 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). Figure 1 illustrates the study design. The details for the study design can be found in the study protocol.

Figure 1 Study Schema



AUC = area under the concentration–time curve; IV = intravenously; NSCLC = non–small cell lung cancer; PD-L1 = programmed death–ligand 1; RECIST v.1.1 = Response Evaluation Criteria in Solid Tumors Version 1.1.

Patients undergo tumor assessments at baseline and every 6 weeks (±7 days) for the first 48 weeks after Cycle 1, Day 1, regardless of dose delays. After 48 weeks, tumor assessment is required every 9 weeks (±7 days). Patients undergo tumor assessments until radiographic disease progression per RECIST v1.1 or loss of clinical benefit (for atezolizumab-treated patients who continue treatment after radiographic disease progression in accordance with RECIST v1.1), withdrawal of consent, study termination by Sponsor, or death, whichever occurs first. Patients who discontinue treatment for reasons other than radiographic disease progression per RECIST v1.1 (e.g., toxicity) continue scheduled tumor assessments until radiographic disease progression per

RECIST v1.1 or loss of clinical benefit (for atezolizumab-treated patients who continue treatment after radiographic disease progression in accordance with RECIST v1.1), withdrawal of consent, study termination by Sponsor, or death, whichever occurs first, regardless of whether patients start a new anti-cancer therapy.

The co-primary efficacy endpoints are progression-free survival (PFS) as assessed by the investigator according to RECIST v1.1, and overall survival (OS). See Section 2.2 for details on co-primary efficacy endpoints, secondary endpoints, and other endpoints such as safety, pharmacokinetic (PK), and exploratory outcome measures.

The primary analyses of PFS will be performed on the biomarker-selected population defined using expression of a T-effector (Teff) and PD-L1 gene signature in tumor tissue (thereafter referred to as tGE) and excluding patients with a sensitizing *EGFR* mutation or *ALK* translocation (wild type, tGE–WT) and on all randomized patients (the intent-to-treat [ITT] population) excluding patients with a sensitizing *EGFR* mutation or *ALK* translocation (ITT–WT). Sensitizing *EGFR* mutations include all *EGFR* activating mutations in exons 18 through 21. The primary analyses of OS will be performed on the ITT–WT population. See Section 4.1 for details on analysis populations.

There are no interim analyses planned for PFS and there is one interim analysis planned for OS in this study. See Section 2.4 for detailed analysis timing.

An external independent Data Monitoring Committee (iDMC) will be set up to evaluate safety data on an ongoing basis.

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is in Appendix 1. For additional details, see the Schedules of Assessments in Appendix 2 and Appendix 3.

2.2 OUTCOME MEASURES

2.2.1 Primary Efficacy Outcome Measures

The co-primary efficacy outcome measures for this study are the following:

- PFS, defined as the time from randomization to the first occurrence of disease progression as determined by the investigator according to RECIST v1.1 or death, whichever occurs first, in the tGE–WT population and the ITT–WT population
- OS, defined as the time from randomization to death from any cause in the ITT–WT population

2.2.2 Secondary Efficacy Outcome Measures

The secondary efficacy outcome measures for this study are the following:

- OS in the tGE–WT population
- PFS, as determined by the investigator according to RECIST v1.1, and OS in the TC2/3 or IC2/3 WT population and the TC1/2/3 or IC1/2/3 WT population

- PFS, as determined by the investigator according to RECIST v1.1, and OS in the tGE population and the ITT population
- Objective response, defined as either complete response (CR) or partial response (PR) as determined by the investigator according to RECIST v1.1 in the tGE–WT population and the ITT–WT population
- Duration of response (DOR), defined for patients with an objective response as the time from the first occurrence of a documented objective response to the time of documented disease progression as determined by the investigator according to RECIST v1.1 or death from any cause, whichever occurs first, in the tGE–WT population and the ITT–WT population
- PFS as assessed by the independent review facility (IRF), defined as the time from randomization to the first occurrence of disease progression as determined by the IRF according to RECIST v1.1 or death from any cause, whichever occurs first, in the tGE–WT population and the ITT–WT population
- OS rates at 1 and 2 years, defined as the proportion of patients alive at 1 and 2 years after randomization estimated using Kaplan-Meier (KM) methodology for the tGE–WT population and the ITT–WT population
- PFS as assessed by the investigator according to RECIST v1.1 and OS in the two atezolizumab-containing arms in the tGE–WT population and the ITT–WT population
- Time to deterioration (TTD) in patient-reported lung cancer symptoms, defined as time from randomization to deterioration (10-point change) on each of the European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire–Core (EORTC QLQ-C30) and the supplemental lung cancer module (EORTC QLQ-LC13) symptom subscales in the tGE–WT population and the ITT–WT population
- Change from baseline in patient-reported lung cancer symptoms (cough, dyspnea, and chest pain) on the symptom severity score of the Symptoms in Lung Cancer (SILC) scale for the tGE–WT population and the ITT–WT population

2.2.3 Exploratory Efficacy Outcome Measures

The exploratory efficacy outcome measures for this study are the following:

- Time to response (TTR), defined as the time from randomization to the first occurrence of a CR or PR as determined by the investigator according to RECIST v1.1
- Time in response (TIR), which shares the definition of DOR for patients with an objective response as assessed by investigator and which is defined as 1 day for patients without an objective response
- Objective response rate (ORR) and DOR as assessed by the IRF according to RECIST v1.1
- PFS, ORR, and DOR as assessed by the investigator in accordance with modified RECIST (Arms A and B only)

- PFS rates at 6 months and 1 year, defined as the proportion of patients alive and without progression as assessed by the investigator according to RECIST v1.1 at 6 months and 1 year after randomization, estimated using KM methodology
- OS rate at 3 years, defined as the proportion of patients alive at 3 years after randomization, estimated using KM methodology
- Change from baseline in patient-reported outcomes (PROs) of health-related quality of life (HRQoL), lung cancer–related symptoms, and health status as assessed by the EORTC QLQ-C30 and EORTC QLQ-LC13
- Utility scores of the Euro QoL5 Dimensions 3-Level Version (EQ-5D-3L)
- 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 in combination with chemotherapy
- Status of ICs and other exploratory biomarkers in mandatory biopsy specimens and blood collected at progression

2.2.4 Pharmacokinetic Outcome Measures

The PK outcome measures for this study are the following:

- Maximum observed serum atezolizumab concentration (C_{max}) after infusion (Arms A and B)
- Minimum observed serum atezolizumab concentration (C_{min}) prior to infusion at selected cycles, at treatment discontinuation, and at 120 days ($\pm$30 days) after the last dose of atezolizumab (Arms A and B)
- Plasma concentrations for carboplatin (Arms A, B, and C)
- Plasma concentrations for paclitaxel (Arms A, B, and C)
- Bevacizumab C_{max} and C_{min} (Arms B and C)

2.2.5 Safety Outcome Measures

The safety outcome measures for this study are the following:

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

2.3 DETERMINATION OF SAMPLE SIZE

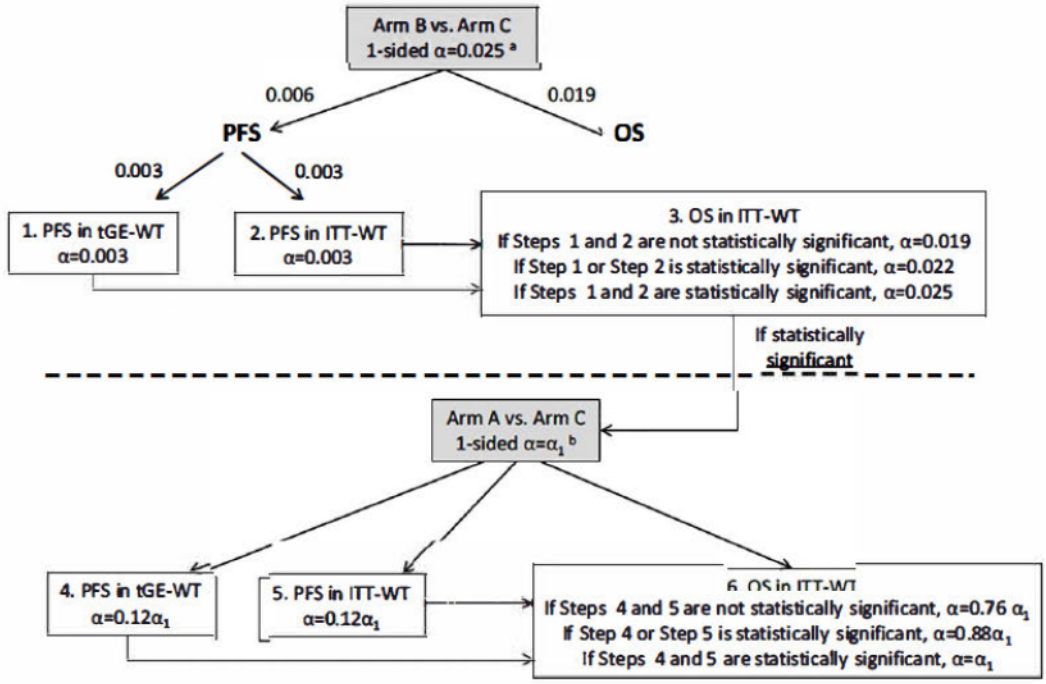
This study will randomize approximately 1200 patients, including approximately 600 patients in the tGE biomarker–selected population. Patients with sensitizing *EGFR* mutations or *ALK* translocations will be excluded from the primary analysis populations. There are expected to be approximately 1080 patients in the ITT–WT population given the approximately 10% prevalence for these genetic variations. The biomarker-selected tGE–WT population will therefore include approximately 540 patients assuming 50% prevalence with the chosen cutoff value for expression.

To control the overall type I error rate for treatment comparisons in the primary endpoints at a one-sided test at 0.025, a one-sided α of 0.006 and a one-sided α of 0.019 will be first allocated to PFS and OS, respectively, for the comparison of atezolizumab + carboplatin + paclitaxel + bevacizumab (Arm B) against the control arm of carboplatin + paclitaxel + bevacizumab (Arm C). The primary comparison of PFS will be tested at a one-sided α level of 0.003 in the tGE–WT and ITT–WT populations in parallel. The primary comparison of OS will be tested in the ITT–WT population at the allocated α together with the α recycled from the PFS analysis if the PFS comparison is statistically significant.

If OS in the ITT–WT population is statistically significant between the comparison of Arm B versus Arm C, PFS in the tGE–WT and ITT–WT populations and OS in the ITT–WT population for the primary comparison of atezolizumab + carboplatin + paclitaxel (Arm A) against the control arm (Arm C) will be tested with the same level of α split with the ratio of 3:3:19 between these three tests. Depending on the outcomes of the PFS testing of Arm A versus Arm C in the tGE–WT and ITT–WT populations, the α from these two PFS comparisons will be recycled back to the OS comparison in the ITT–WT population for Arm A versus Arm C.

The PFS and OS analysis and α allocation (Burman et al. 2009), including possible α recycling, are shown in Figure 2. If OS in the ITT–WT population is statistically significant between the comparison of Arm A versus Arm C, the remaining α will be further allocated to the comparison of secondary efficacy endpoints where PFS and OS are tested in the tGE and/or ITT populations. See Section 4.4.2.3 for details.

Figure 2 Progression-Free Survival and Overall Survival Analysis Hierarchy, α Allocation, and α Recycling



α_1 = type I error passed down to the comparison of Arm A vs. Arm C (i.e., 0.019, 0.022, or 0.025); ITT = intent to treat; OS = overall survival; PFS = progression-free survival; tGE = tumor gene expression; WT = wild type.

[a] To control the overall type I error rate for the one-sided test at 0.025, a one-sided type I error (α) will be allocated to PFS in the tGE-WT population, PFS in the ITT-WT population, and OS in the ITT-WT population in a 3:3:19 ratio for comparison of Arm B versus Arm C.

[b] If the difference in OS between Arm B and Arm C in the ITT-WT population is statistically significant at an α of 0.019, 0.022, or 0.025 (Step 3), that same α will become the overall one-sided type I error rate for the comparison of Arm A versus Arm C, with α allocated to PFS in the tGE-WT population, PFS in the ITT-WT population, and OS in the ITT-WT population at the same 3:3:19 ratio.

OS will be tested using the group sequential method at the interim and final OS analyses based on the α allocated to the primary comparisons of OS, as described above. Statistical significance at the interim analyses of OS will be tested as described in Section 2.4.2.

The sample size of this study is determined on the basis of the number of events required to demonstrate efficacy with regard to both PFS and OS (co-primary endpoints) for the comparison of Arm B versus Arm C.

The estimates of the number of events required to demonstrate efficacy with regard to PFS in the comparison for Arm B versus Arm C are based on the following assumptions:

- One-sided significance level of 0.003 for the primary comparison in the tGE–WT population
- One-sided significance level of 0.003 for the primary comparison in the ITT–WT population
- 98% power to detect a hazard ratio (HR) of 0.55, corresponding to an improvement in median PFS from 6 months to 10.9 months in the tGE–WT population
- 98% power to detect an HR of 0.65, corresponding to an improvement in median PFS from 6 months to 9.2 months in the ITT–WT population
- No interim analysis for PFS
- Dropout rate of 5% per 12 months

The estimates of the number of events required with regard to OS in the primary comparison for Arm B versus Arm C are based on the following assumptions:

- One-sided significance level of 0.019 for the primary comparison in the ITT–WT population
- 87% power to detect an HR of 0.75, corresponding to an improvement in median OS from 12 months to 16 months in the ITT–WT population
- One OS interim analysis performed at the time of the PFS final analysis; it is estimated that approximately 73% of the total OS events required for the final analysis have occurred using the Lan-DeMets approximation to the O'Brien-Fleming boundary
- Dropout rate of 5% per 24 months

The estimate of the number of events required to demonstrate efficacy with regard to PFS and OS in the comparison of Arm A versus Arm C is based on assumptions similar to those outlined above for Arm B versus Arm C.

With these assumptions, approximately 1200 patients in total will be randomized into this study. The ITT–WT population will include approximately 1080 patients, with approximately 720 patients in each primary comparison (i.e., Arm A vs. Arm C, Arm B vs. Arm C). The tGE–WT population will include approximately 540 patients, with approximately 360 patients in each primary comparison (i.e., Arm A vs. Arm C, Arm B vs. Arm C).

2.4 ANALYSIS TIMING

2.4.1 Primary Analysis Timing for PFS

No interim analyses are planned for the co-primary endpoint of PFS in this study. The PFS final analysis for the primary comparison of Arm B versus Arm C will be conducted when there are approximately 516 PFS events in the ITT–WT population in the combined Arm B and Arm C or when the last patient is randomized, whichever occurs

last. At the time of the final PFS analysis, it is expected that approximately 249 events would have occurred in the tGE–WT population. These numbers of events correspond to a minimum detectable difference (MDD) in HR of approximately 0.70 in the tGE–WT population and of approximately 0.78 in the ITT–WT population. It is projected that an observed HR of 0.70 or lower will result in a statistically significant difference between treatment arms in the tGE–WT population at this analysis. The PFS final analysis is expected to occur approximately 29 months after the first patient is randomized, based on the study assumptions and the projected accrual rate (see Table 1).

Table 1 Final Analysis for Progression-Free Survival for Arm B versus Arm C Comparison

Population	Analysis Timing (Months from FPI)	No. of Events (Event Ratio, %)	MDD in Hazard Ratio	Power, %
tGE–WT	29	249 (69)	0.70	98
ITT–WT	29	516 (72)	0.78	98

FPI = first patient in; ITT = intent-to-treat; MDD = minimum detectable difference; PFS = progression-free survival; tGE = tumor gene expression; WT = wild type.

The PFS analysis between Arm A and Arm C will be performed at the same time as the Arm B and Arm C comparison. However, the statistical significance will be claimed in accordance with the testing strategy in Figure 2 depending on the α recycled to the comparison of Arm A versus Arm C.

2.4.2 Interim and Final Analysis Timing for OS

There is one interim analysis planned for the co-primary endpoint of OS.

The OS interim analysis for the primary comparison of Arm B versus Arm C will be conducted by the Sponsor at the time of the final PFS analysis for the primary comparisons. It is expected that there will be approximately 370 OS events in the ITT-WT population in the combined Arm B and Arm C at this timepoint, in which case an interim OS analysis will be conducted with the stopping boundaries for the OS interim and final analyses computed using the Lan-DeMets approximation to the O'Brien-Fleming boundary. If there are significantly fewer than 370 OS events at the PFS final analysis, a nominal α of 0.01% (negligible impact on overall type I error rate) could be spent on the OS analysis at the time of the PFS final analysis. The next interim will be conducted after approximately 370 OS events are observed, with the stopping boundaries for the OS interim and final analyses calculated the same way as above.

The final OS analysis for each of the primary comparison of Arm B versus Arm C will be conducted when there are approximately 507 OS events in the ITT–WT population in the combined Arm B and Arm C. This number of events corresponds to an MDD in HR of

approximately 0.83 in the ITT–WT population. The OS final analysis is expected to occur approximately 40 months after the first patient is randomized.

The expected analysis timing for OS interim and final analyses for the comparison of Arm B versus Arm C is shown in Table 2.

Table 2 Analysis for Overall Survival for the Comparison of Arm B versus Arm C

	Analysis Timing		ITT–WT Population	
Type of Analysis	Months from FPI	Percentage Information, %	No. of Events (Event Ratio, %)	Power, % [a]
OS interim analysis	29	73	370 (51)	60
OS final analysis	40	100	507 (70)	87

FPI=first patient in; ITT=intent-to-treat; OS=overall survival; WT=wild type.

[a] Power is calculated using one-sided α of 0.019.

The stopping boundaries for the interim and final OS analyses will be calculated using the Lan-DeMets approximation to the O'Brien-Fleming boundary. The stopping boundaries for the comparison of Arm B versus Arm C in the ITT–WT population, assuming the specified observed number of events (370 and 507, respectively), are shown in Table 3. The p-value will be used to claim crossing of the boundaries. For example, if 370 events have occurred at the time of the OS interim analysis in the ITT–WT population for Arm B and Arm C combined, statistical significance of the OS endpoint in the ITT–WT population will be declared for Arm B if $p \leq 0.0073$ when PFS has claimed significance in either the tGE–WT or the ITT–WT population.

Table 3 Analysis Timing and Stopping Boundary for Overall Survival in the Intent-to-Treat Wild-Type Population for the Comparison of Arm B versus Arm C

	Stopping Boundary in HR (p-value)		
Analysis	If $\alpha=0.019$	If $\alpha=0.022$	If $\alpha=0.025$
OS interim analysis	HR≤0.770 (p-value≤0.0064)	HR≤0.776 (p-value≤0.0073)	HR≤0.781 p-value≤0.0087)
OS final analysis	HR≤0.829 (p-value≤0.0172)	HR≤0.833 (p-value≤0.0198)	HR≤0.837 (p-value≤0.0224)

HR=hazard ratio; OS=overall survival.

Note: α and p-value are one-sided.

The OS analysis between Arm A and Arm C will be performed at the same time as the Arm B versus Arm C comparison. However, the statistical significance will be claimed in accordance with the testing strategy in Figure 2 depending on the α recycled to the

comparison of Arm A versus Arm C. If the difference in OS between Arm B and Arm C in the ITT–WT population is statistically significant at an α of 0.019, 0.022, or 0.025, that same α will become the overall one-sided type I error rate for the comparison of Arm A versus Arm C, as described in Section 2.3. The stopping boundaries for the interim and final OS analyses for the comparison of Arm A versus Arm C in the ITT–WT population, assuming the specified observed number of events (370 and 507, respectively), are provided for the three scenarios ($\alpha =$ 0.019, 0.022, or 0.025) in Table 4.

Table 4 Analysis Timing and Stopping Boundary for Overall Survival in the Intent-to-Treat Wild-Type Population for the Comparison of Arm A versus Arm C

A. OS Is Tested Statistically Significant at $\alpha = 0.019$ in the ITT–WT Population for the Arm B versus Arm C Comparison

	Stopping Boundary in HR (p-value)		
Analysis Timing	If $\alpha = 0.01444$	If $\alpha = 0.01672$	If $\alpha = 0.019$
OS interim analysis	HR $\leq$ 0.760 (p-value $\leq$ 0.0042)	HR $\leq$ 0.766 (p-value $\leq$ 0.0051)	HR $\leq$ 0.770 (p-value $\leq$ 0.0064)
OS final analysis	HR $\leq$ 0.821 (p-value $\leq$ 0.0132)	HR $\leq$ 0.825 (p-value $\leq$ 0.0152)	HR $\leq$ 0.829 (p-value $\leq$ 0.0172)

B. OS Is Tested Statistically Significant at $\alpha = 0.022$ in the ITT–WT Population for the Arm B versus Arm C Comparison

	Stopping Boundary in HR (p-value)		
Analysis Timing	If $\alpha = 0.01672$	If $\alpha = 0.01936$	If $\alpha = 0.022$
OS interim analysis	HR $\leq$ 0.766 (p-value $\leq$ 0.0051)	HR $\leq$ 0.771 (p-value $\leq$ 0.0062)	HR $\leq$ 0.776 (p-value $\leq$ 0.0073)
OS final analysis	HR $\leq$ 0.825 (p-value $\leq$ 0.0152)	HR $\leq$ 0.829 (p-value $\leq$ 0.0175)	HR $\leq$ 0.833 (p-value $\leq$ 0.0198)

C. OS Is Tested Statistically Significant at $\alpha = 0.025$ in the ITT–WT Population for the Arm B versus Arm C Comparison

	Stopping Boundary in HR (p-value)		
Analysis Timing	If $\alpha = 0.019$	If $\alpha = 0.022$	If $\alpha = 0.025$
OS interim analysis	HR $\leq$ 0.770 (p-value $\leq$ 0.0064)	HR $\leq$ 0.776 (p-value $\leq$ 0.0073)	HR $\leq$ 0.781 (p-value $\leq$ 0.0087)
OS final analysis	HR $\leq$ 0.829 (p-value $\leq$ 0.0172)	HR $\leq$ 0.833 (p-value $\leq$ 0.0198)	HR $\leq$ 0.837 (p-value $\leq$ 0.0224)

ITT = intent to treat; HR = hazard ratio; OS = overall survival; WT = wild type.
Note: α and p-value are one-sided.

3. STUDY CONDUCT

3.1 RANDOMIZATION ISSUES

Randomization to the treatment and control arms will occur in a 1:1:1 ratio using a permuted-block randomization method. Randomization will be stratified by the following factors:

- Sex (male vs. female)
- Presence of liver metastases at baseline (yes vs. no)
- PD-L1 expression by IHC (TC3 and any IC vs. TC0/1/2 and IC2/3 vs. TC0/1/2 and IC0/1)

3.2 INDEPENDENT REVIEW FACILITY

An IRF will be used to conduct a blinded radiology review of the imaging data and will provide an independent assessment of tumor response and progression for all patients according to a separate IRF charter. IRF-assessed endpoints will be used for secondary and exploratory analyses.

3.3 DATA MONITORING

An external iDMC will be used to evaluate safety during the study on a periodic basis, approximately every 6 months from the point of first patient in until the time the database is locked and the study is unblinded for the primary PFS analysis and interim OS analysis. In addition, the iDMC will review safety data once 12 patients have been enrolled into each treatment arm and have received treatment for two cycles. All summaries and analyses by treatment arm for the iDMC review will be prepared by an independent data coordinating center. Following the data review, the iDMC will provide a recommendation 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.

Members of the iDMC will be external to the Sponsor and will follow a separate iDMC charter that outlines their roles and responsibilities, as well as a detailed monitoring plan.

4. STATISTICAL METHODS

The analyses described in this SAP will supersede those specified in Protocol GO29436 for the purposes of a regulatory filing.

4.1 ANALYSIS POPULATIONS

4.1.1 Efficacy Analysis Populations

The randomized population or ITT population is defined as all randomized patients, regardless of receipt of the assigned treatment.

The ITT-WT population is defined as the ITT population excluding patients with a sensitizing *EGFR* mutation or *ALK* translocation.

The tGE biomarker-selected population is defined as ITT patients with a signature score greater than or equal to –1.91 in baseline tumor tissues. The signature score captures the gene expression of *PD-L1(CD274)*, *CXCL9*, and *IFN-γ*, relative to a housekeeping gene, using RNA isolated from patients' formalin-fixed paraffin-embedded tumor tissue and is measured on a Roche Molecular System cobas 4800 platform.

The tGE–WT population is defined as the tGE population excluding patients with a sensitizing *EGFR* mutation or *ALK* translocation.

The PD-L1–selected TC2/3 or IC2/3 population is defined as ITT patients with PD-L1 TC2/3 or IC2/3 expression in baseline tumor tissue. Similarly the PD-L1–selected TC1/2/3 or IC1/2/3 population is defined as ITT patients with PD-L1 TC1/2/3 or IC1/2/3 expression in baseline tumor tissue.

The PD-L1–selected WT populations are defined as the PD-L1–selected populations (TC2/3 or IC2/3 population or TC1/2/3 or IC1/2/3 population) excluding patients with a sensitizing *EGFR* mutation or *ALK* translocation.

The primary analyses of PFS will be performed on the ITT–WT population and the tGE-WT populations, and the primary analyses of OS will be performed on the ITT–WT population. Patients are grouped according to the treatment assigned at randomization, regardless of whether they received any assigned study drug.

ORR will be analyzed for patients in the tGE–WT and ITT–WT populations who have measurable disease as assessed by investigator at baseline. DOR will be assessed in patients in the tGE–WT and ITT–WT populations who have an objective response and measurable disease as assessed by investigator at baseline. TTD analyses will be conducted on the tGE–WT and ITT-WT populations. Change from baseline analysis in PROs will be performed for patients in the tGE-WT and ITT–WT populations who have both a non-missing baseline assessment and at least one non-missing post-baseline assessment.

4.1.2 Pharmacokinetic-Evaluable Population

The pharmacokinetic-evaluable population is defined as all patients who received any dose of atezolizumab, bevacizumab, carboplatin, or paclitaxel and who have evaluable PK samples post-dose.

4.1.3 Safety Population

The safety population includes treated patients, defined as randomized patients who received any protocol 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 randomized to Arm C, if atezolizumab was received in error in addition to Arm C treatment, the patients will be grouped with Arm B for the safety analyses.

4.2 ANALYSIS OF STUDY CONDUCT

Study enrollment, major protocol deviations including major deviations of inclusion or exclusion criteria, and reasons for study termination will be summarized overall and by treatment arm for the biomarker-selected tGE, the tGE–WT, the ITT, and the ITT–WT populations. Study treatment administration and reasons for discontinuation from study treatment will be summarized for the safety population.

4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

Demographic characteristics, such as age, race/ethnicity, baseline disease characteristics (e.g., Eastern Cooperative Oncology Group [ECOG] performance status), and stratification factors (sex, presence of liver metastasis at baseline, PD-L1 tumor expression by IHC) will be summarized by treatment arms for the biomarker–selected tGE–WT and the ITT–WT populations. Descriptive statistics (mean, median, SD, and range) will be presented for continuous data, and frequencies and percentages will be presented for categorical data.

Unless otherwise stated, baseline values are the last available data obtained prior to the patient receiving the first dose of any component of study treatment on Cycle 1, Day 1. For ECOG performance status, the baseline is defined as the value at screening.

4.4 EFFICACY ANALYSIS

Patients will be grouped for efficacy analyses in accordance with the treatment assigned at randomization, whether or not the assigned treatment was received.

4.4.1 Primary Efficacy Endpoint

The co-primary efficacy endpoints are PFS and OS.

PFS is defined as the time from randomization to the first documented disease progression as determined by the investigator with the use of RECIST v1.1 or death from any cause, whichever occurs first. Data for patients who are alive and who have not experienced disease progression at the time of analysis will be censored at the date of the last tumor assessment. Data for patients with no post-baseline tumor assessment will be censored at the date of randomization plus 1 day.

OS is 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 analysis will be censored at the date last known to be alive. Data for patients who do not have post-baseline information will be censored at the date of randomization plus 1 day.

The null and alternative hypotheses regarding PFS and OS in each population (tGE–WT, ITT–WT) can be phrased in terms of the survival functions $S_A(t)$ or $S_B(t)$ for the atezolizumab arms (Arm A or Arm B, respectively) and $S_C(t)$ for the control arm (Arm C):

$$H_0: S_A(t) = S_C(t) \text{ versus } H_1: S_A(t) > S_C(t)$$

$$\text{or } H_0: S_B(t) = S_C(t) \text{ versus } H_1: S_B(t) > S_C(t)$$

The HRs, λ_A/λ_C and λ_B/λ_C (where λ_A, λ_B, and λ_C represent the hazard of having a PFS or death in Arm A, Arm B, and Arm C, respectively), comparing the treatment effect between the two treatment arms, will be estimated using a stratified Cox regression model with the same stratification variables used for the stratified log-rank test, and the 95% CI will be provided.

For PFS analysis in the tGE–WT population, the stratification factors will be sex (male vs. female) and presence of liver metastases at baseline (yes vs. no). For PFS and OS analyses in the ITT–WT population, the stratification factors will be sex (male vs. female), presence of liver metastases at baseline (yes vs. no), and PD-L1 tumor expression by IHC ([TC3 and any IC, TC0/1/2 and IC2/3 combined] vs. TC0/1/2 and IC 0/1). The stratification factors will be those used during randomization as recorded in the interactive voice and Web response system per the request of the U.S. Food and Drug Administration. The analyses based on stratification factors as recorded on the electronic Case Report Form may also be provided. Results from an unstratified analysis will also be provided.

Treatment comparisons will be conducted by first comparing Arm B versus Arm C and then comparing Arm A versus Arm C based on a stratified log-rank test in the biomarker-selected tGE–WT and ITT–WT populations for the PFS endpoint and in the ITT–WT population for the OS endpoint. For each comparison, analyses will be conducted according to an analysis hierarchy and an α-spending algorithm to control for the type I error rate (see Figure 2 in Section 2.3) and to account for an interim analysis (see Section 2.4.2).

To control the overall type I error rate for the one-sided test at 0.025, the hypothesis testing for the comparison of Arm B versus Arm C will be done at the specified significance levels in the order described below:

1. PFS in the tGE–WT population will be tested at $\alpha = 0.003$ (one sided). If the estimate of the HR is < 1 and the one-sided p-value corresponding to the stratified log-rank test is < 0.003, the null hypothesis will be rejected, and it will be concluded that atezolizumab + carboplatin + paclitaxel + bevacizumab prolongs the duration of PFS relative to the control arm in the tGE–WT population.
2. PFS in the ITT–WT population will be tested at $\alpha = 0.003$ (one sided).

3. α recycling from PFS to OS will be conducted as follows:

   If the PFS result is not statistically significant in either the tGE–WT population or the ITT–WT population, OS in the ITT–WT population will be tested at $\alpha = 0.019$ (one sided).

   If the PFS result is only statistically significant in the tGE–WT population or the ITT–WT population, OS in the ITT–WT population will be tested at $\alpha = 0.022$ (one sided).

   If the PFS result is statistically significant in both the tGE–WT population and the ITT–WT population, OS in the ITT–WT population will be tested at $\alpha = 0.025$ (one sided).

If the difference in OS between Arm B and Arm C in the ITT–WT population is statistically significant, the comparison of Arm A versus Arm C will be statistically tested at the same significance level, with α allocated to PFS in the tGE–WT population, to PFS in the ITT–WT population, and to OS in the ITT-WT population at the same 3:3:19 ratio (see Figure 2 in Section 2.3). Depending on the outcome of the PFS testing of Arm A versus Arm C in the tGE–WT and ITT–WT populations, the α from these two PFS comparisons will be recycled back to the OS comparison in the ITT–WT population for Arm A versus Arm C. The significance levels at which the PFS and OS will be tested for the comparison of Arm A versus Arm C are shown in Table 5 for all scenarios

Table 5 Significance Level for Comparison of Arm A versus Arm C

Arm B versus Arm C		Arm A versus Arm C
α level at which the OS is tested statistically significant in the ITT–WT population	α level at which the PFS will be tested in the tGE-WT and ITT-WT populations	α level at which the OS will be tested in the ITT-WT population
0.019	0.00228	0.01444, if the PFS result is not statistically significant in either PFS population 0.01672, if the PFS result is statistically significant in only one of the PFS populations 0.019, if the PFS results are statistically significant in both PFS populations
0.022	0.00264	0.01672, if the PFS result is not statistically significant in either PFS population 0.01936, if the PFS result is statistically significant in one of the PFS populations 0.022, if the PFS results are statistically significant in both PFS populations
0.025	0.003	0.019, if the PFS result is not statistically significant in either PFS population 0.022, if the PFS result is statistically significant in one of the PFS populations 0.025, if the PFS results are statistically significant in both PFS populations

ITT = intent to treat; OS = overall survival; PFS = progression-free survival; WT = wild type.

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

Follow-up for OS, defined as the time from randomization to death or last known date alive, will be summarized for all patients included in the analysis (whether patient is alive or has died). Follow-up will be summarized using the KM method with data for patients who died censored at the date of death.

4.4.2 Secondary Efficacy Endpoints

4.4.2.1 OS in the tGE–Wild-Type Population

OS in tGE–WT population will be analyzed using similar methods as described in Section 4.4.1.

4.4.2.2 PFS and OS in PD-L1–Selected Wild-Type Populations

PFS and OS in PD-L1-selected TC2/3 or IC2/3 WT and TC1/2/3 or IC1/2/3 WT populations will be analyzed using similar methods as described in Section 4.4.1.

4.4.2.3 PFS and OS in the tGE and ITT Populations

PFS and OS in the tGE and ITT populations will be analyzed using the same methods as described in Section 4.4.1.

If the difference in OS between Arm A and Arm C in the ITT–WT population is statistically significant, a comparison of Arm B versus Arm C and a comparison of Arm A versus Arm C will be conducted in the tGE population and the ITT population, with a comparison of Arm B versus Arm C prioritized similarly as shown in Figure 2. The α allocation will follow the same α-spending algorithm and allocation ratio (3:3:19) described for analysis of the co-primary efficacy endpoints as described in Section 2.3.

4.4.2.4 Objective Response Rate

ORR (confirmation not required) is defined as the proportion of patients with an objective response, either CR or PR, with the use of RECIST v1.1, as determined by the investigator. Patients not meeting this criteria, including those without any post-baseline assessment, will be considered non-responders. Confirmation of response according to RECIST v1.1 is not required, but for the exploratory purposes, ORR with confirmation may be reported. The analysis population for ORR will be the tGE–WT and ITT–WT populations with measurable disease at baseline, i.e., patients with at least one measureable lesion in accordance with RECIST v1.1 based on investigator assessment. ORR will be compared between treatment arms using the stratified Cochran-Mantel-Haenszel test. The 95% CI for the difference in ORRs between the two treatment arms will be computed using the normal approximation to the binomial distribution. An estimate of ORR and its 95% CI will be calculated for each treatment arm using the Clopper-Pearson method.

4.4.2.5 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 according to RECIST v1.1 or death from any cause, whichever occurs first. Data for patients who are alive and who have not experienced disease progression at the time of 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. DOR will be estimated using KM methodology. Comparisons between treatment arms will be made using the stratified and unstratified log-rank test for descriptive purposes only.

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

4.4.2.6 PFS as Assessed by Independent Review Facility

PFS as assessed by the IRF is defined as the time from randomization to the first documented disease progression as determined by the IRF using the RECIST v1.1 or death from any cause, whichever occurs first.

PFS as assessed by the IRF will be analyzed using the same methods that will be used for PFS as assessed by the investigator.

4.4.2.7 OS Rate at the 1- and 2-Year Landmark Time Points

The OS rate at the 1- and 2-year landmark time points after randomization within the tGE–WT and ITT–WT populations will be estimated for each treatment arm using KM methodology, along with 95% CI calculated with the standard error derived from the Greenwood formula. The 95% CI for the difference in OS rates between the two treatment arms will be estimated using the normal approximation method, with standard errors computed using the Greenwood method.

4.4.2.8 Investigator-Assessed PFS and OS Comparison between Atezolizumab-Containing Arms

The comparison between the atezolizumab-containing arms, i.e., Arm A and Arm B, will also be made using investigator-assessed PFS per RECIST v1.1 and OS endpoints.

4.4.2.9 Patient-Reported Outcomes

Through the use of the SILC and the EORTC QLQ-C30 and EORTC QLQ-LC13, lung cancer symptom data will be collected about symptoms commonly associated with cancer treatments, and disease and treatment impact on patients' functioning.

The EORTC QLQ-C30 and EORTC QLQ-LC13 will be scored according to the EORTC scoring manual (Fayers et al. 2001). The QLQ-C30 and QLQ-LC13 consist of both

multi-item scales and single-item measures such as functional scales, symptom scales, and a global health status/quality-of-life scale.

For multi-item subscales, if ≤50% of items within the multi-item subscale are missing at a given timepoint, the multi-item score will be calculated on the basis of the non-missing items. If >50% of items are missing or if a single-item measure is missing, the subscale is missing.

All of the EORTC scales and single-item measures will be linearly transformed so that each score has a range of 0–100. A high score for a functional scale represents a high or healthy level of functioning, and a high score for the global health status and HRQoL represents a high HRQoL; however, a high score for a symptom scale or item represents a high level of symptomatology or problems.

Time to Deterioration in Patient-Reported Outcomes

TTD with use of the EORTC is defined as the time from baseline to the first confirmed clinically meaningful deterioration in the EORTC symptom score. Confirmed clinically meaningful deterioration in lung cancer symptoms is defined as a ≥10-point increase above baseline in a symptom score that must be held for at least two consecutive assessments or an initial ≥10-point increase above baseline followed by death within 3 weeks from the last assessment. A ≥10-point change in the EORTC scale score is perceived by patients as clinically significant (Osoba et al. 1998).

TTD will be documented for each of the following EORTC-based symptom scores:

- Cough (Question 31 on the EORTC QLQ-LC13)
- Dyspnea single item (Question 8 on the QLQ-C30)
- Dyspnea multi-item subscale (Questions 33-35 on the QLQ-LC13)
- Chest pain (Question 40 on the QLQ-LC13)
- Arm and/or shoulder pain (Question 41 on the QLQ-LC13)

TTD analyses will be performed for the ITT–WT population and will include all data collected through disease progression and survival follow-up. TTD analyses may be performed for the tGE–WT population when a PFS benefit has been demonstrated for this biomarker-selected population. The methodologies that are outlined for the analysis of OS will be used for the analyses of TTD for prespecified symptoms of the EORTC QLQ-C30 and EORTC QLQ-LC13 measures. TTD of the prespecified symptoms will be summarized using the KM method. Comparison of TTD using the EORTC QLQ-C30 and EORTC QLQ-LC13 measures between treatment arms will be performed using the stratified log-rank test; the stratified HRs and 95% CIs will also be reported. 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. There will be no imputation for missing baseline or post-baseline data for the TTD analysis.

Change from Baseline per SILC Scale

Summary statistics (mean, SD, median, 25th and 75th percentiles, and range) of the change from baseline per SILC scale will be provided. The analysis will be performed for patients in the ITT–WT population with a baseline and a post-baseline PRO assessment and may be performed for the tGE–WT population when a PFS benefit has been demonstrated for this biomarker-selected population. Graphs of the mean changes and standard errors over time from the baseline assessment for the total score and subscales will be provided for each treatment arm.

The analysis of SILC change from baseline 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 disease progression), at the last dose of treatment received before treatment discontinuation for any cause, and at the survival follow-up visits through 6 months.

4.4.3 Exploratory Efficacy Endpoints

4.4.3.1 Time in Response

TIR is defined as for DOR for patients with an objective response as determined by the investigator using RECIST v1.1 or as 1 day for patients without an objective response. For patients with an objective response, TIR will be analyzed as censored or an event according to whether their DOR was analyzed as censored or having a PFS event. Patients without an objective response will be analyzed as having an event on the day of randomization plus 1 day. TIR will be estimated using KM methodology. TIR will be compared between treatment arms for patients with measurable disease at baseline using the same analysis methods that will be used for OS.

4.4.3.2 Time to Response

TTR is defined for patients with an objective response as the time from randomization to the first occurrence of a CR or PR as determined by the investigator using RECIST v1.1. TTR will be summarized for descriptive purposes. The mean, standard error, median, and range of TTR will be provided. No formal treatment comparisons will be performed.

4.4.3.3 ORR and DOR as Assessed by the IRF per RECIST v1.1

The methodologies outlined for the secondary efficacy endpoint analyses will be used for the analyses of ORR and DOR as assessed by the IRF per RECIST v1.1.

4.4.3.4 PFS, ORR, and DOR as Assessed by the Investigator Using Modified RECIST

The endpoints assessed with modified RECIST will be limited to the atezolizumab-containing arms, i.e., Arm A and Arm B only, with no comparison with the control Arm C. Comparisons will be made between Arm A and Arm B.

PFS by modified RECIST is defined as the time from randomization to disease progression as determined by the investigator per modified RECIST or death from any cause, whichever occurs first. A patient is considered to have disease progression by modified RECIST if either of the following conditions are met:

a) Modified RECIST for progression are met at a tumor assessment and no subsequent tumor assessment is performed.

b) Modified RECIST for progression are met at a tumor assessment, and at the subsequent tumor assessment, the criteria for confirmed progression by modified RECIST are also met.

For patients who meet criterion a), the date of progression is the date of the tumor assessment that met the criteria for progression per modified RECIST. For patients who meet criterion b), the date of progression is the date of the tumor assessment at which the modified RECIST for progression are first met. Patients who do not meet either of the above conditions are not considered to have disease progression by modified RECIST.

ORR by modified RECIST is defined as the proportion of patients whose best overall response is either a PR or CR per modified RECIST.

DOR by modified RECIST is defined for patients who experienced an objective response (CR or PR per modified RECIST) as assessed by the investigator as the time from the first tumor assessment that supports the patient's objective response (CR or PR, whichever is recorded first) to diseases progression or death from any cause, whichever occurs first. Data for patients who are alive and who have not experienced disease progression at the time of analysis will be censored at the last tumor assessment date.

The methodologies outlined for the primary and secondary efficacy endpoint analyses will be used for the analyses of PFS, ORR, and DOR as assessed by the investigator using modified RECIST.

4.4.3.5 Progression-Free Survival Rates at 6 Months and 1 Year

PFS rates at 6 months and 1 year will be estimated and analyzed using the same method as described in Section 4.4.2.7.

4.4.3.6 Overall Survival Rate at 3 Years

OS rate at 3 years will be estimated and analyzed using the same method as described in Section 4.4.2.7.

4.4.3.7 EQ-5D-3L Health Status Data

The EQ-5D-3L questionnaire will also be collected to generate HRQoL and utility scores for use in economic models for reimbursement.

Health status is assessed by the EQ-5D-3L. 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 used in this study for economic modeling. This analysis will not be included in the Clinical Study Report (CSR) for this study.

4.4.3.1 Additional Patient-Reported Outcome Analyses

EORTC score changes from baseline will be descriptively analyzed using means, SDs, medians, and range by treatment arm for patients with a baseline assessment and at least one post-baseline assessment. The analyses will be performed at timepoints similar to those used for the analyses of change from baseline per SILC scale.

TTD may also be documented for a composite symptom endpoint (cough, dyspnea, chest pain) with use of the EORTC.

Compliance rates will be summarized by listing the number and proportion of patients in the PRO-evaluable subset who completed the PRO assessments at each timepoint by treatment arm. Reasons for non-completion will be summarized if available in the Case Report Form.

4.4.4 Sensitivity Analyses

4.4.4.1 Missing Tumor Assessment

The impact of missing scheduled tumor assessments on PFS will be assessed depending on the number of patients who missed assessments scheduled immediately prior to the date of disease progression per RECIST v1.1 or the data cutoff. If $>$ 5% of patients missed two or more assessments scheduled immediately prior to the date of disease progression per RECIST v1.1 or the data cutoff in any treatment arm, the following two sensitivity analyses will be performed:

- Patients who missed two or more scheduled assessments immediately prior to the date of disease progression per RECIST v1.1 or the data cutoff will be censored at the last tumor assessment prior to the missed visits.
- Patients who missed two or more scheduled assessments immediately prior to the date of disease progression per RECIST v1.1 or the data cutoff will be counted as having progressed on the date of the first of these missing assessments.

Statistical methodologies analogous to those used in the primary analysis of PFS as specified in Section 4.4.1 will be used for this sensitivity analysis.

4.4.4.2 Non-Proportional Hazard Non–Protocol-Specified Anti-Cancer Therapy

The impact of non–protocol-specified anti-cancer therapy (NPT) on PFS as determined by the investigator will be assessed, depending on the number of patients who received

NPT before a PFS event. If >5% of patients received NPT before a PFS event in the control arm, a sensitivity analysis will be performed for the comparisons between two treatment arms (i.e., Arm A vs. Arm C, Arm B vs. Arm C) in which data from patients who receive NPT before a PFS event will be censored at the last tumor assessment date before receipt of NPT.

The impact of subsequent NPT on OS will be assessed depending on the number of patients who received NPT. If >10% of patients received an NPT in the control arm, the following analyses may be performed to compare treatment arms:

- The discount method uses a "discounted" survival time after switching for patients who switch treatments based on a user-specified assumption for the effect on OS. OS will be discounted in accordance with a range of possible effects on OS of the subsequent NPT after treatment switching occurred (e.g., 10%, 20%, 30%).
- Rank-preserving structural failure time provides an estimate of the OS time for the control group had NPT not occurred (Robins and Tsiatis 1991). It estimates OS measured from the time of NPT by applying an estimate of the benefit of the NPT. The total overall survival time (sum of time to NPT and the estimated survival time after NPT started) will then be analyzed using the same methodology as for the primary analysis of OS.
- The inverse probability of censoring weighting method censors patients at start of NPT and uses the control arm patients to create weights that represent how NPT-treated–like a non–NPT-treated patient is (Robins and Finkelstein 2000). These time-varying weights are included in the OS analysis to correct the effect of NPT by giving increased weight to non-censored patients with similar characteristics to censored patients.

Delayed Clinical Effect

If a delayed separation of the KM curves is observed at the beginning of the curves and the delay is ≥3 months, the following analyses could be conducted to assess a potential delayed clinical effect for the treatment group.

Milestone OS Analysis

To assess the potential effect of long-term survival and delayed clinical effects, a milestone OS analysis will be conducted (Chen 2015). The milestone time points will be chosen such that the patients included in the analysis will achieve a certain patient-event ratio.

The milestone OS analysis will be conducted only when the milestone duration has elapsed from the time the last patient entered the study, using the same methods as those specified for the primary OS analysis.

Restricted Mean Survival Time

The restricted mean survival time (RMST) will be computed for OS using the area under the curve from baseline to several timepoints. RMST will be computed for each treatment arm and the difference with its 95% CI will be displayed.

Weighted Log-Rank Analysis

Where the delayed clinical effect is >10% of the median survival time of the control group, an analysis of OS may be performed using the weighted log-rank test (Fleming and Harrington 1991) that weights more heavily on late events to account for the delayed clinical effect (Fine 2007).

4.4.4.3 Loss to Follow-up

The impact of loss to follow-up on OS will be assessed depending on the number of patients who are lost to follow-up. If >5% of patients are lost to follow-up for OS in either treatment arm, a sensitivity analysis will be performed for the comparisons between two treatment arms (i.e., Arm A vs. Arm C, Arm B vs. Arm C) in which patients who are lost to follow-up will be considered as having died at the last date they were known to be alive.

4.4.4.4 Impact of Imbalance between Treatment Arms

Randomization is not stratified by factors defining the ITT–WT or tGE–WT analysis populations. This could cause a potential imbalance between treatment arms with respect to baseline characteristics, including prognostic factors, and the number of patients. Main baseline covariates (e.g., stratification factors) will be summarized by treatment arm in the ITT–WT population and/or the tGE–WT population (depending on the results of the primary endpoint analyses). When strong baseline imbalances are observed, the stratified Cox regression analysis for PFS and/or OS endpoints including such prognostic factors as covariates in the model will be performed to investigate the potential impact of such imbalances on the treatment effect.

4.4.5 Subgroup Analyses

The consistency of PFS and OS results in subgroups will be examined in the populations where PFS and/or OS benefit has been demonstrated. The subgroups are defined by the following:

- Demographics (age, sex, race/ethnicity)
- Baseline disease characteristics (e.g., ECOG performance status; presence of liver metastases at baseline; smoking status; metastatic sites such as brain, bone, etc.; *EGFR* mutation status; *KRAS* mutation status)
- PD-L1 IHC status (e.g., TC3 or IC3, TC2/3 or IC2/3, TC1/2/3 or IC1/2/3, and their corresponding complementary groups)

- Complementary biomarker population defined by a tGE cutoff value less than –1.91 and additional biomarker populations and corresponding complementary biomarker populations defined by the tGE cutoff values of –2.38 and –2.93

Summaries of PFS and OS, including the unstratified HR estimated from a Cox proportional hazards model and KM estimates of median PFS and OS, will be produced separately for each level of the subgroup for the comparisons between two treatment arms (i.e., Arm A vs. Arm C, Arm B vs. Arm C) and displayed in a forest plot (Lewis and Clarke 2001). KM plots of PFS and/or OS will also be produced for selected subgroups.

Summaries of ORR by subgroup will also be provided.

4.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

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

Plasma concentrations of carboplatin and paclitaxel will be collected in this study as outlined in Appendix 3. The concentrations of carboplatin and paclitaxel will be summarized using descriptive statistics as described above.

Additional PK analyses will be conducted, as appropriate, based on the availability of data. These additional analyses will not be included in the CSR for this study.

Exploratory biomarker analyses will be performed in an effort to understand the association of these markers with response to study drug, including efficacy and/or safety. The tumor biomarkers include but are not limited to PD-L1 and CD8, as defined by IHC, quantitative reverse transcriptase–polymerase chain reaction, or other methods. In addition, predictive, prognostic, and pharmacodynamic exploratory biomarkers in archival and/or fresh tumor tissue and/or blood will be examined for their association with disease status and/or clinical outcomes. These exploratory analyses will not be included in the CSR for this study.

4.6 SAFETY ANALYSES

The safety population includes treated patients, defined as randomized patients who received any protocol 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 randomized to Arm C, if atezolizumab was received in error in addition to Arm C treatment, the patients will be grouped with Arm B for the safety analyses.

4.6.1 Exposure of Study Medication

Study drug exposure, including 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 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 $\geq$10% of patients), serious adverse events, treatment-related serious adverse events, severe adverse events (Grade $\geq$3), adverse events of special interest, immune-mediated adverse events (defined as adverse events that require the use of systemic corticosteroids with no clear alternate etiology), and adverse events that lead to study drug discontinuation or interruption will be summarized.

"Treatment emergent" is defined for all events with onset on or after the first study drug treatment up to the data cutoff date.

In addition, adverse events with an onset on or after the first study drug treatment and up to 1 day before the date of the first dose of the maintenance therapy will be defined as adverse events that occurred during the induction therapy phase. Adverse events with an onset on or after the first dose of the maintenance therapy and up to the data cutoff date are defined as adverse events that occurred during the maintenance therapy phase. Key adverse events that occurred during the induction and maintenance therapy phases may also be summarized separately.

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 Data

Laboratory data will be summarized over time including change from baseline by treatment arm. Values outside the normal ranges will be summarized. In addition, selected laboratory data will be classified in accordance with NCI CTCAE and will be summarized by grade. Highest NCI CTCAE grade post-baseline will also be reported, and shift tables from baseline value 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 and by change over time including change from baseline.

ECOG performance status will also be summarized over time.

4.6.5 Anti-Drug Antibody

Serum levels and incidence of ADAs against atezolizumab and bevacizumab will be summarized. The analyses of pharmacokinetics, key efficacy, and safety by atezolizumab ADA status may be conducted to explore the potential impact of immunogenicity as appropriate.

4.7 MISSING DATA

See Section 4.4.1 and Section 4.4.2 for methods of handling missing data for the primary and secondary efficacy endpoints.

5. REFERENCES

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

Burman CF, Sonesson C, Guilbaud O, et al. A recycling framework for the construction of Bonferroni-based multiple tests. Stat Med 2009;28:739–61.

Chen TT. Milestone survival: a potential intermediate endpoint for immune checkpoint inhibitors. J Natl Cancer Inst 2015;107:djv156.

Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 scoring manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer, 2001.

Fine GD. Consequences of delayed treatment effects on analysis of time-to-event endpoints. Ther Innov Regul Sci 2007;41:535–9.

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.

Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Commun Stat Theory Methods 1991;20:2609–31.

Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics 2000;56:779–88.

Appendix 1
Protocol Synopsis

PROTOCOL SYNOPSIS

TITLE:	**A PHASE III, OPEN-LABEL, RANDOMIZED STUDY OF ATEZOLIZUMAB (MPDL3280A, ANTI–PD-L1 ANTIBODY) IN COMBINATION WITH CARBOPLATIN + PACLITAXEL WITH OR WITHOUT BEVACIZUMAB COMPARED WITH CARBOPLATIN + PACLITAXEL + BEVACIZUMAB IN CHEMOTHERAPY-NAIVE PATIENTS WITH STAGE IV NON-SQUAMOUS NON–SMALL CELL LUNG CANCER**
PROTOCOL NUMBER:	GO29436
VERSION NUMBER:	6
EUDRACT NUMBER:	2014-003207-30
IND NUMBER:	117296
TEST PRODUCT:	Atezolizumab (MPDL3280A, RO5541267)
PHASE:	III
INDICATION:	Non-squamous non–small cell lung cancer
SPONSOR:	F. Hoffmann-La Roche Ltd

Objectives

Efficacy Objectives

Unless otherwise specified, efficacy objectives will be analyzed for the following two treatment comparisons:

- Atezolizumab + carboplatin + paclitaxel + bevacizumab (Arm B) versus carboplatin + paclitaxel + bevacizumab (Arm C)
- Atezolizumab + carboplatin + paclitaxel (Arm A) versus carboplatin + paclitaxel + bevacizumab (Arm C)

The term "wild type" (WT) refers to randomized patients who do not have a sensitizing *EGFR* mutation or *ALK* translocation.

The term "tumor gene expression" (tGE) refers to randomized patients with a defined level of expression of a PD-L1 and T-effector gene signature in tumor tissue, as analyzed through use of a centrally performed RNA-based assay.

Some efficacy endpoints will be analyzed in a population of randomized patients with a defined level of PD-L1 expression on tumor cells (TCs) and immune cells (ICs), as analyzed through use of a centrally performed IHC test.

The co-primary objectives of this study are the following:

- To evaluate the efficacy of atezolizumab as measured by investigator-assessed progression-free survival (PFS) according to RECIST v1.1 in the tGE-WT population and the ITT-WT population
- To evaluate the efficacy of atezolizumab as measured by overall survival (OS) in the ITT-WT population

The secondary efficacy objectives for this study are the following:

- To evaluate the efficacy of atezolizumab as measured by OS in the tGE-WT population
- To evaluate the efficacy of atezolizumab as measured by investigator-assessed PFS according to RECIST v1.1 and OS in the TC2/3 or IC2/3 WT population and the TC1/2/3 or IC1/2/3 WT population
- To evaluate the efficacy of atezolizumab as measured by investigator-assessed PFS according to RECIST v1.1 and OS in the tGE population and the ITT population
- To evaluate the efficacy of atezolizumab as measured by investigator-assessed objective response rate (ORR) according to RECIST v1.1 in the tGE-WT population and the ITT-WT population
- To evaluate the efficacy of atezolizumab as measured by investigator-assessed duration of response (DOR) according to RECIST v1.1 in the tGE-WT population and the ITT-WT population
- To evaluate the efficacy of atezolizumab as measured by an Independent Review Facility (IRF)-assessed PFS according to RECIST v1.1 in the tGE-WT population and the ITT-WT population
- To evaluate the OS rate at 1 and 2 years in each treatment arm for the tGE-WT population and the ITT-WT population
- To compare the efficacy of the two atezolizumab-containing arms, Arm A versus Arm B, as measured by investigator-assessed PFS according to RECIST v1.1 and by OS in the tGE-WT population and the ITT-WT population
- To determine the impact of atezolizumab as measured by time to deterioration (TTD) in patient-reported lung cancer symptoms of cough, dyspnea (single-item and multi-item subscales), chest pain, or arm/shoulder pain, using the European Organisation for the Research and Treatment of Cancer (EORTC) Quality-of-Life Questionnaire–Core (QLQ-C30) and the supplemental lung cancer module (QLQ-LC13) in the tGE-WT population and the ITT-WT population
- To determine the impact of atezolizumab as measured by change in baseline (i.e., improvement or deterioration based upon presenting symptomatology) in patient-reported lung cancer symptom (chest pain, dyspnea, and cough) score using the Symptoms in Lung Cancer (SILC) scale symptom severity score for the tGE-WT population and the ITT-WT

Safety Objectives

The safety objectives for this study are the following:

- To evaluate the safety and tolerability of atezolizumab in each of the two treatment comparisons
- To evaluate the incidence and titers of anti-therapeutic antibodies (ATAs) against atezolizumab and to explore the potential relationship of the immunogenicity response with pharmacokinetics, safety, and efficacy

Pharmacokinetic Objectives

The pharmacokinetic (PK) objectives for this study are the following:

- To characterize the pharmacokinetics of atezolizumab when given in combination with carboplatin and paclitaxel with and without bevacizumab (Arms A and B)
- To characterize the pharmacokinetics of carboplatin when given in combination with paclitaxel with and without atezolizumab and/or bevacizumab (Arms A, B, and C)
- To characterize the pharmacokinetics of paclitaxel when given in combination with carboplatin with and without atezolizumab and/or bevacizumab (Arms A, B, and C)

- To characterize the pharmacokinetics of bevacizumab when given in combination with carboplatin and paclitaxel with and without atezolizumab (Arms B and C)

Exploratory Objectives

The exploratory objectives for this study are the following:

- To evaluate the efficacy of atezolizumab as measured by investigator-assessed time to response (TTR) and time-in-response (TIR) according to RECIST v1.1
- To evaluate ORR and DOR according to RECIST v1.1 as assessed by the IRF
- To evaluate investigator-assessed ORR, PFS, and DOR according to modified RECIST for the atezolizumab-containing treatment arms
- To evaluate PFS at 6 months and at 1 year in each treatment arm
- To evaluate the OS rate at 3 years in each treatment arm
- To assess predictive, prognostic, and pharmacodynamic exploratory biomarkers in archival and/or fresh tumor tissue and blood and their association with disease status, mechanisms of resistance, and/or response to study treatment
- To evaluate the utility of biopsy at the time of apparent disease progression to distinguish apparent increases in tumor volume related to the immunomodulatory activity of atezolizumab (i.e., pseudoprogression/tumor-immune infiltration) from true disease progression
- To evaluate and compare patient's health status as assessed by the EuroQoL 5 Dimensions 3-Level (EQ-5D-3L) questionnaire to generate utility scores for use in economic models for reimbursement
- To determine the impact of atezolizumab as measured by change from baseline in patient-reported outcomes of health-related quality of life, lung cancer–related symptoms, and functioning as assessed by the EORTC QLQ-C30 and LC13

Study Design

Description of Study

This is a randomized, Phase III, multicenter, open-label study (IMpower150) designed to evaluate the safety and efficacy of atezolizumab in combination with carboplatin + paclitaxel with or without bevacizumab compared with treatment with carboplatin + paclitaxel + bevacizumab in approximately 1200 chemotherapy-naive patients with Stage IV non-squamous non–small cell lung cancer (NSCLC).

Tumor specimens will be prospectively tested for PD-L1 expression by a central laboratory. Eligible patients will be stratified by sex (male vs. female), presence of liver metastases at baseline (yes vs. no), and by PD-L1 tumor expression by IHC (TC3 and any IC vs. TC0/1/2 and IC2/3 vs. TC0/1/2 and IC0/1). Patients will be randomized in a 1:1:1 ratio to receive one of the following treatment regimens:

Treatment Arm A: Atezolizumab + carboplatin + paclitaxel (Induction: four or six 21-day cycles); atezolizumab (Maintenance: 21-day cycles)

Treatment Arm B: Atezolizumab + carboplatin + paclitaxel + bevacizumab (Induction: four or six 21-day cycles); atezolizumab + bevacizumab (Maintenance: 21-day cycles)

Treatment Arm C: Carboplatin + paclitaxel + bevacizumab (Induction: four or six 21-day cycles); bevacizumab (Maintenance: 21-day cycles)

Appendix 1
Protocol Synopsis (cont.)

The number of cycles of induction treatment (four or six) will be at the discretion of the investigator and will be determined and documented prior to randomization. Induction treatment will be administered on a 21-day cycle until the following occur (whichever occurs first): 1) administration of four or six cycles or 2) disease progression (Arm C) or loss of clinical benefit (Arms A and B) is documented.

Following the induction phase, patients will continue treatment with maintenance therapy. Patients who are randomized to Arms B and C will continue treatment with bevacizumab until progressive disease, unacceptable toxicity, or death. Patients who are randomized to Arms A or B may continue treatment with atezolizumab beyond radiographic progression according to RECIST v1.1, provided they are experiencing 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).

Treatment with chemotherapy (Arms A, B, and C) and bevacizumab (Arms B and C) must be discontinued in all patients who exhibit evidence of progressive disease.

Patients will undergo tumor assessments at baseline and every 6 weeks for the first 48 weeks following Cycle 1, Day 1, regardless of dose delays. After 48 weeks, tumor assessment will be required every 9 weeks. Patients will undergo 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, study termination by Sponsor, or death, whichever occurs first. Patients who discontinue treatment for reasons other than radiographic disease progression (e.g., toxicity) 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 after radiographic disease progression according to RECIST v1.1), withdrawal of consent, study termination by Sponsor, or death, whichever occurs first, regardless of whether patients start a new anti-cancer therapy.

A secondary endpoint of this study is IRF-assessed PFS according to RECIST v1.1. An IRF will therefore conduct an independent review of the responses of all patients, including a blinded review of computed tomography (CT) scans. All primary imaging data used for tumor assessment will be collected by the Sponsor to enable centralized, independent review of response endpoints. These reviews will be performed prior to the final efficacy analyses.

For Treatment Arms A and B Only

At any point during treatment, patients receiving atezolizumab who show evidence of clinical benefit will be permitted to continue atezolizumab after RECIST v1.1 for progressive disease are met if they meet all of the following criteria:

- Evidence of clinical benefit as assessed by the investigator
- Absence of symptoms and signs (including worsening of laboratory values, e.g., new or worsening hypercalcemia) indicating unequivocal progression of disease
- No decline in Eastern Cooperative Oncology Group (ECOG) performance status that can be attributed to disease progression
- Absence of tumor progression at critical anatomical sites (e.g., leptomeningeal disease) that cannot be managed by protocol-allowed medical interventions
- Patients must provide written consent to acknowledge deferring other treatment options in favor of continuing atezolizumab at the time of initial progression

Appendix 1
Protocol Synopsis (cont.)

Patients in all treatment arms will undergo a mandatory tumor biopsy sample collection, unless not clinically feasible as assessed and documented by investigators, at the time of radiographic disease progression. These data will be used to explore whether radiographic findings are consistent with the presence of tumor. Additionally, these data will be analyzed to evaluate the association between changes in tumor tissue and clinical outcome and to understand further the potential mechanisms of progression and resistance to atezolizumab as compared with such mechanisms after treatment with chemotherapy alone. This exploratory biomarker evaluation will not be used for any treatment-related decisions. Patients in Arms A and B who are unable to undergo biopsy sample collection but who otherwise meet the criteria listed above may continue to receive atezolizumab.

Number of Patients

Approximately 270 sites globally will participate in the study, and approximately 1200 patients will be randomized.

Target Population

Inclusion Criteria

Patients may be eligible if they have chemotherapy-naive, Stage IV, non-squamous NSCLC.

Patients must meet all of the following criteria to be eligible for study entry:

- Signed Informed Consent Form
- Male or female, 18 years of age or older
- ECOG performance status of 0 or 1
- Histologically or cytologically confirmed, Stage IV non-squamous NSCLC (per the Union Internationale contre le Cancer/American Joint Committee on Cancer staging system, 7th edition).

  Patients with tumors of mixed histology (i.e., squamous and non-squamous) are eligible if the major histological component appears to be non-squamous.
- No prior treatment for Stage IV non-squamous NSCLC

  Patients with a sensitizing mutation in the epidermal growth factor receptor (*EGFR*) gene must have experienced disease progression (during or after treatment) or intolerance to treatment with one or more *EGFR* TKIs, such as erlotinib, gefitinib, or another *EGFR* tyrosine kinase inhibitor (TKI) appropriate for the treatment of *EGFR*-mutant NSCLC.

  Patients with an anaplastic lymphoma kinase (*ALK*) fusion oncogene must have experienced disease progression (during or after treatment) or intolerance to treatment with one or more *ALK* inhibitors (i.e., crizotinib) appropriate for the treatment of NSCLC in patients having an *ALK* fusion oncogene.

  Patients with unknown *EGFR* and/or *ALK* status require test results at screening. *ALK* and/or *EGFR* may be assessed locally or at a central laboratory.
- Patients who have received prior neo-adjuvant, adjuvant chemotherapy, radiotherapy, or chemoradiotherapy with curative intent for non-metastatic disease must have experienced a treatment-free interval of at least 6 months from randomization since the last chemotherapy, radiotherapy, or chemoradiotherapy.

- Patients with a history of treated asymptomatic CNS metastases are eligible, provided they meet all of the following criteria:

  Only supratentorial and cerebellar metastases allowed (i.e., no metastases to midbrain, pons, medulla or spinal cord)

  No ongoing requirement for corticosteroids as therapy for CNS disease

  No stereotactic radiation within 7 days or whole-brain radiation within 14 days prior to randomization

  No evidence of interim progression between the completion of CNS-directed therapy and the screening radiographic study

  Patients with new asymptomatic CNS metastases detected at the screening scan must receive radiation therapy and/or surgery for CNS metastases. Following treatment, these patients may then be eligible without the need for an additional brain scan prior to randomization, if all other criteria are met.
- Known PD-L1 tumor status as determined by an IHC assay performed by a central laboratory on previously obtained archival tumor tissue or tissue obtained from a biopsy at screening

  A representative formalin-fixed paraffin-embedded (FFPE) tumor specimen in paraffin block (preferred) or 15 or more unstained, freshly cut, serial sections on slides from an FFPE tumor specimen is required for participation in this study. If fewer than 15 slides are available at baseline (but no fewer than 10), the patient may still be eligible, upon discussion with the Medical Monitor. This specimen must be accompanied by the associated pathology report.

  Fine-needle aspiration (defined as samples that do not preserve tissue architecture and yield cell suspension and/or cell smears), brushing, cell pellet specimens (e.g., from pleural effusion, and lavage samples) are not acceptable.

  Tumor tissue from bone metastases that is subject to decalcification is not acceptable.

  For core needle biopsy specimens, preferably at least three cores embedded in a single paraffin block, should be submitted for evaluation.
- Measurable disease, as defined by RECIST v1.1

  Previously irradiated lesions can only be considered as measurable disease if disease progression has been unequivocally documented at that site since radiation and the previously irradiated lesion is not the only site of disease.
- Adequate hematologic and end organ function, defined by the following laboratory results obtained within 14 days prior to randomization:

  ANC $\geq$ 1500 cells/μL without granulocyte colony-stimulating factor support

  Lymphocyte count $\geq$ 500/μL

  Platelet count $\geq$ 100,000/μL without transfusion

  Hemoglobin $\geq$ 9.0 g/dL

  Patients may be transfused to meet this criterion.

  INR or aPTT $\leq$ 1.5 $\times$ upper limit of normal (ULN)

  This applies only to patients who are not receiving therapeutic anticoagulation; patients receiving therapeutic anticoagulation should be on a stable dose.

  AST, ALT, and alkaline phosphatase $\leq$ 2.5 $\times$ ULN, with the following exceptions:

  Patients with documented liver metastases: AST and/or ALT $\leq$ 5 $\times$ ULN

Patients with documented liver or bone metastases: alkaline phosphatase $\leq 5 \times$ ULN.

Serum bilirubin $\leq 1.25 \times$ ULN

Patients with known Gilbert disease who have serum bilirubin level $\leq 3 \times$ ULN may be enrolled.

Serum creatinine $\leq 1.5 \times$ ULN

- For female patients of childbearing potential, agreement (by patient and/or partner) to use a highly effective form(s) of contraception that results in a low failure rate (< 1% per year) when used consistently and correctly, and to continue its use for 5 months after the last dose of atezolizumab and/or 6 months after the last dose of bevacizumab or paclitaxel, whichever is later). Such methods include: combined (estrogen and progestogen containing) hormonal contraception, progestogen-only hormonal contraception associated with inhibition of ovulation together with another additional barrier method always containing a spermicide, intrauterine device (IUD): intrauterine hormone-releasing system (IUS), bilateral tubal occlusion, vasectomized partner (on the understanding that this is the only one partner during the whole study duration), and sexual abstinence.
- For male patients with female partners of childbearing potential, agreement (by patient and/or partner) to use a highly effective form(s) of contraception that results in a low failure rate [< 1% per year] when used consistently and correctly, and to continue its use for 6 months after the last dose of bevacizumab, carboplatin, or paclitaxel. Male patients should not donate sperm during this study and for at least 6 months after the last dose of bevacizumab, carboplatin, or paclitaxel.
- Oral contraception should always be combined with an additional contraceptive method because of a potential interaction with the study drug. The same rules are valid for male patients involved in this clinical study if they have a partner of childbirth potential. Male patients must always use a condom.
- Women who are not postmenopausal (≥ 12 months of non–therapy-induced amenorrhea) or surgically sterile must have a negative serum pregnancy test result within 14 days prior to initiation of study drug

Exclusion Criteria

Patients who meet any of the criteria below will be excluded from study entry.

Cancer-Specific Exclusions

- Active or untreated CNS metastases as determined by CT or magnetic resonance imaging (MRI) evaluation during screening and prior radiographic assessments
- Spinal cord compression not definitively treated with surgery and/or radiation or previously diagnosed and treated spinal cord compression without evidence that disease has been clinically stable for > 2 weeks prior to randomization
- Leptomeningeal disease
- Uncontrolled tumor-related pain

  Patients requiring pain medication must be on a stable regimen at study entry.

  Symptomatic lesions amenable to palliative radiotherapy (e.g., bone metastases or metastases causing nerve impingement) should be treated prior to randomization. Patients should be recovered from the effects of radiation. There is no required minimum recovery period.

  Asymptomatic metastatic lesions whose further growth would likely cause functional deficits or intractable pain (e.g., epidural metastasis that is not currently associated with spinal cord compression) should be considered for locoregional therapy, if appropriate, prior to randomization.

Appendix 1
Protocol Synopsis (cont.)

- Uncontrolled pleural effusion, pericardial effusion, or ascites requiring recurrent drainage procedures (once monthly or more frequently)

  Patients with indwelling catheters (e.g., PleurX®) are allowed.

- Uncontrolled or symptomatic hypercalcemia (> 1.5 mmol/L ionized calcium or Ca > 12 mg/dL or corrected serum calcium > ULN)

  Patients who are receiving denosumab prior to randomization must be willing and eligible to receive a bisphosphonate instead while in the study.

- Malignancies other than NSCLC within 5 years prior to randomization, with the exception of those with a negligible risk of metastasis or death (e.g., expected 5-year OS > 90%) treated with expected curative outcome (such as adequately treated carcinoma in situ of the cervix, basal or squamous-cell skin cancer, localized prostate cancer treated surgically with curative intent, ductal carcinoma in situ treated surgically with curative intent)
- Known tumor PD-L1 expression status as determined by an IHC assay from other clinical studies (e.g., patients whose PD-L1 expression status was determined during screening for entry into a study with PD-1 or anti-PD-L1 antibodies but were not eligible are excluded)

General Medical Exclusions

- Women who are pregnant, lactating, or intending to become pregnant during the study
- History of severe allergic, anaphylactic, or other hypersensitivity reactions to chimeric or humanized antibodies or fusion proteins
- Known hypersensitivity or allergy to biopharmaceuticals produced in Chinese hamster ovary cells or any component of the atezolizumab formulation
- History of autoimmune disease, including but not limited to myasthenia gravis, myositis, autoimmune hepatitis, systemic lupus erythematosus, rheumatoid arthritis, inflammatory bowel disease, vascular thrombosis associated with antiphospholipid syndrome, Wegener's granulomatosis, Sjögren's syndrome, Guillain-Barré syndrome, multiple sclerosis, vasculitis, or glomerulonephritis

  Patients with a history of autoimmune-related hypothyroidism on a stable dose of thyroid replacement hormone are eligible for this study.

  Patients with controlled Type 1 diabetes mellitus on a stable dose of insulin regimen are eligible for this study

  Patients with eczema, psoriasis, lichen simplex chronicus, or vitiligo with dermatologic manifestations only (e.g., patients with psoriatic arthritis would be excluded) are permitted provided that they meet the following conditions:

  - Rash must cover less than 10% of body surface area (BSA).
  - Disease is well controlled at baseline and only requiring low-potency topical steroids.
  - No acute exacerbations of underlying condition within the previous 12 months (not requiring PUVA [psoralen plus ultraviolet A radiation], methotrexate, retinoids, biologic agents, oral calcineurin inhibitors, high-potency or oral steroids)

- History of idiopathic pulmonary fibrosis, organizing pneumonia (e.g., bronchiolitis obliterans), drug-induced pneumonitis, idiopathic pneumonitis, or evidence of active pneumonitis on screening chest CT scan

  History of radiation pneumonitis in the radiation field (fibrosis) is permitted.

- Positive test for HIV

All patients will be tested for HIV prior to inclusion into the study; patients who test positive for HIV will be excluded from the clinical study.

- Patients with active hepatitis B (chronic or acute; defined as having a positive hepatitis B surface antigen [HBsAg] test at screening) or hepatitis C

  Patients with past hepatitis B virus (HBV) infection or resolved HBV infection (defined as the presence of hepatitis B core antibody [HBcAb] and absence of HBsAg) are eligible only if they are negative for HBV DNA.

  Patients positive for hepatitis C virus (HCV) antibody are eligible only if PCR is negative for HCV RNA.
- Active tuberculosis
- Severe infections within 4 weeks prior to randomization, including but not limited to hospitalization for complications of infection, bacteremia, or severe pneumonia
- Received therapeutic oral or IV antibiotics within 2 weeks prior to randomization

  Patients receiving prophylactic antibiotics (e.g., for prevention of a urinary tract infection or to prevent chronic obstructive pulmonary disease exacerbation) are eligible.
- Significant cardiovascular disease, such as New York Heart Association cardiac disease (Class II or greater), myocardial infarction, or cerebrovascular accident within 3 months prior to randomization, unstable arrhythmias, or unstable angina

  Patients with known coronary artery disease, congestive heart failure not meeting the above criteria, or left ventricular ejection fraction $< 50\%$ must be on a stable medical regimen that is optimized in the opinion of the treating physician, in consultation with a cardiologist if appropriate.
- Major surgical procedure other than for diagnosis within 28 days prior to randomization or anticipation of need for a major surgical procedure during the course of the study
- Prior allogeneic bone marrow transplantation or solid organ transplant
- Administration of a live, attenuated vaccine within 4 weeks before randomization or anticipation that such a live attenuated vaccine will be required during the study
- Any other diseases, metabolic dysfunction, physical examination finding, or clinical laboratory finding giving reasonable suspicion of a disease or condition that contraindicates the use of an investigational drug or that may affect the interpretation of the results or renders the patient at high risk from treatment complications
- Patients with illnesses or conditions that interfere with their capacity to understand, follow and/or comply with study procedures

Exclusion Criteria Related to Medications

- Any approved anti-cancer therapy, including hormonal therapy, within 3 weeks prior to initiation of study treatment; the following exceptions are allowed:

  TKIs approved for treatment of NSCLC discontinued > 7 days prior to randomization; the baseline scan must be obtained after discontinuation of prior TKIs.
- Treatment with any other investigational agent with therapeutic intent within 28 days prior to randomization
- Prior treatment with CD137 agonists or immune checkpoint blockade therapies, anti–PD-1, and anti–PD-L1 therapeutic antibodies

Patients who have had prior anti–CTLA-4 treatment may be enrolled, provided the following requirements are met:

Last dose of anti–CTLA-4 at least 6 weeks prior to randomization

No history of severe immune-mediated adverse effects from anti–CTLA-4 (National Cancer Institute Common Terminology Criteria for Adverse Events [NCI CTCAE] Grade 3 and 4)

- Treatment with systemic immunostimulatory agents (including, but not limited to, IFNs, IL-2) within 4 weeks or five half-lives of the drug, whichever is longer, prior to randomization

  Prior treatment with cancer vaccines is allowed.

- Treatment with systemic immunosuppressive medications (including but not limited to corticosteroids, cyclophosphamide, azathioprine, methotrexate, thalidomide, and anti–tumor necrosis factor [anti-TNF] agents) within 2 weeks prior to randomization

  Patients who have received acute, low-dose (≤ 10 mg oral prednisone or equivalent), systemic immunosuppressant medications may be enrolled in the study.

  The use of corticosteroids (≤ 10 mg oral prednisone or equivalent) for chronic obstructive pulmonary disease, mineralocorticoids (e.g., fludrocortisone) for patients with orthostatic hypotension, and low-dose supplemental corticosteroids for adrenocortical insufficiency is allowed

Exclusions Related to Bevacizumab

- Inadequately controlled hypertension (defined as systolic blood pressure > 150 mmHg and/or diastolic blood pressure > 100 mmHg)

  Anti-hypertensive therapy to achieve these parameters is allowable.

- Prior history of hypertensive crisis or hypertensive encephalopathy
- Significant vascular disease (e.g., aortic aneurysm requiring surgical repair or recent peripheral arterial thrombosis) within 6 months prior to randomization
- History of hemoptysis (≥ one-half teaspoon of bright red blood per episode) within 1 month prior to randomization
- Evidence of bleeding diathesis or coagulopathy (in the absence of therapeutic anticoagulation)
- Current or recent (within 10 days of randomization) use of aspirin (> 325 mg/day) or treatment with dipyramidole, ticlopidine, clopidogrel, and clostazol
- Current use of full-dose oral or parenteral anticoagulants or thrombolytic agents for therapeutic purposes that has not been stable for > 2 weeks prior to randomization

  The use of full-dose oral or parenteral anticoagulants is permitted as long as the INR or aPTT is within therapeutic limits (according to the medical standard of the enrolling institution) and the patient has been on a stable dose of anticoagulants for at least 2 weeks prior to randomization.

  Prophylactic anticoagulation for the patency of venous access devices is allowed, provided the activity of the agent results in an INR < 1.5 × ULN and aPTT is within normal limits within 14 days prior to randomization.

  Prophylactic use of low-molecular-weight heparin (i.e., enoxaparin 40 mg/day) is permitted.

- Core biopsy or other minor surgical procedure, excluding placement of a vascular access device, within 7 days prior to the first dose of bevacizumab

Appendix 1
Protocol Synopsis (cont.)

- History of abdominal or tracheosphageal fistula or gastrointestinal perforation within 6 months prior to randomization
- Clinical signs of gastrointestinal obstruction or requirement for routine parenteral hydration, parenteral nutrition, or tube feeding
- Evidence of abdominal free air not explained by paracentesis or recent surgical procedure
- Serious, non-healing wound, active ulcer, or untreated bone fracture
- Proteinuria, as demonstrated by urine dipstick or >1.0 g of protein in a 24-hour urine collection

  All patients with $\geq 2+$ protein on dipstick urinalysis at baseline must undergo a 24-hour urine collection and must demonstrate ≤ 1 g of protein in 24 hours.
- Known sensitivity to any component of bevacizumab
- Clear tumor infiltration into the thoracic great vessels is seen on imaging
- Clear cavitation of pulmonary lesions is seen on imaging

Exclusions Related to Chemotherapy

- Known history of severe allergic reactions to platinum-containing compounds or mannitol
- Known sensitivity to any component of paclitaxel
- Grade ≥ 2 peripheral neuropathy as defined by NCI CTCAE v4.0 (paclitaxel)
- Known history of severe hypersensitivity reactions to products containing Cremophor® EL (e.g., cyclosporin for injection concentrate and teniposide for injection concentrate)

Length of Study

The final PFS analysis will be conducted when both of the following criteria have been met: approximately 516 PFS events have occurred in Arms B and C combined in the ITT-WT population and the last patient has been enrolled in the study. The final PFS analysis is expected to occur approximately 29 months after the first patient is enrolled. At the time of the final PFS analysis, it is expected that approximately 249 events will have occurred in the tGE-WT population.

With a sample size of 720 patients, approximately 507 OS events are expected to occur in Arms B and C combined in the ITT-WT population for the final OS analysis. The final OS analysis is expected to occur approximately 40 months after the first patient is enrolled. This number of events corresponds to a minimum detectable difference in HR of approximately 0.83 in the ITT-WT population

End of Study

The end of study is defined as when the required number of deaths for the final analysis of OS has been observed, which is expected to be approximately 44 months after the first patient is randomized. Additionally, the Sponsor may decide to terminate the study at any time. If the Sponsor decides to terminate the study, patients who are still receiving study treatment or undergoing survival follow-up may be enrolled into an extension study or a non-interventional study.

Outcome Measures

Efficacy Outcome Measures

The co-primary efficacy outcome measures for this study are the following:

- PFS, defined as the time from randomization to the first occurrence of disease progression as determined by the investigator using RECIST v1.1 or death from any cause, whichever occurs first in the tGE-WT population and the ITT-WT population

- OS, defined as the time from randomization to death from any cause in the ITT-WT population

The secondary efficacy outcome measures for this study are the following:

- OS in the tGE-WT population
- PFS, as determined by the investigator according to RECIST v1.1, and OS in the TC2/3 or IC2/3 WT population and the TC1/2/3 or IC1/2/3 WT population
- PFS, as determined by the investigator according to RECIST v1.1, and OS in the tGE population and the ITT population
- Objective response, defined as partial response (PR) or complete response (CR) as determined by the investigator according to RECIST v1.1 in the tGE-WT population and the ITT-WT population
- DOR, defined as the time interval from first occurrence of a documented objective response to the time of disease progression as determined by the investigator using RECIST v1.1 or death from any cause, whichever occurs first in the tGE-WT population and the ITT-WT population
- PFS, defined as the time from randomization to the first occurrence of disease progression as determined by the IRF using RECIST v1.1 or death from any cause, whichever occurs first in the tGE-WT population and the ITT-WT population
- OS rates at 1 and 2 years in the tGE-WT population and the ITT-WT population
- PFS, as determined by the investigator according to RECIST v1.1, and OS in the two atezolizumab-containing arms in the tGE-WT population and the ITT-WT population
- TTD in patient-reported lung cancer symptoms, defined as time from randomization to deterioration (10-point change) on each of the EORTC QLQ-C30 and EORTC QLQ-LC13 symptom subscales in the tGE-WT population and the ITT-WT population
- Change from baseline in patient-reported lung cancer symptoms (chest pain, dyspnea, and cough) on the symptom severity score of the Symptoms in Lung Cancer scale in the tGE-WT population and the ITT-WT population

Safety Outcome Measures

The safety outcome measures for this study are the following:

- Incidence, nature, and severity of adverse events graded according to the NCI CTCAE v4.0
- Changes in vital signs, physical findings, and clinical laboratory results during and following atezolizumab administration
- Incidence of ATA response to atezolizumab and potential correlation with PK, pharmacodynamic, safety, and efficacy parameters

Pharmacokinetic Outcome Measures

The PK outcome measures for this study are the following:

- Maximum observed serum atezolizumab concentration (C_{max}) after infusion (Arms A and B)
- Minimum observed serum atezolizumab concentration (C_{min}) prior to infusion at selected cycles, at treatment discontinuation, and at 120 days ($\pm$30 days) after the last dose of atezolizumab (Arms A and B)
- Plasma concentrations for carboplatin (Arms A, B, and C)
- Plasma concentrations for paclitaxel (Arms A, B, and C)
- Bevacizumab C_{max} and C_{min} (Arms B and C)

Appendix 1
Protocol Synopsis (cont.)

Exploratory Outcome Measures

The exploratory outcome measures for this study are:

- TTR, defined as the time from randomization to first occurrence of a documented objective response as determined by the investigator according to RECIST v1.1
- Time in response (TIR), defined as 1 day for non-responders and defined the same as DOR for responders, as determined by the investigator according to RECIST v1.1
- Objective response and DOR, as determined by the IRF according to RECIST v1.1
- Objective response, PFS, and DOR, in the two atezolizumab-containing arms, as determined by the investigator according to modified RECIST (Arms A and B)
- PFS at 6 months and at 1 year
- OS rate at 3 years
- 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 in combination with chemotherapy
- Status of tumor-infiltrating immune cells and other exploratory biomarkers in mandatory biopsy specimens and blood collected at progression
- Utility scores of the EQ-5D-3L
- Change from baseline in patient-reported outcomes of health-related quality of life, lung cancer–related symptoms, and functioning as assessed by the EORTC QLQ-C30 and LC13

Investigational Medicinal Products

Test Product (Investigational Drug)

Atezolizumab (1200 mg IV) will be administered on Day 1 of each 21–day cycle. Atezolizumab will be administered to patients who are randomized to Arms A and B.

Comparator

Bevacizumab (15 mg/kg IV) will be administered on Day 1 of each 21-day cycle for four or six cycles during the induction phase and during the maintenance phase.

Bevacizumab will be administered to patients randomized to Arms B and C.

Non-Investigational Medicinal Products

- Carboplatin will be administered by IV infusion to achieve an initial target area under the concentration–time curve (AUC) of 6 mg/mL/min on Day 1 of each 21-day cycle for four or six cycles during the induction phase.
- Paclitaxel (200 mg/m^2 IV) will be administered on Day 1 of each 21-day cycle for four or six cycles during the induction phase.

Carboplatin and paclitaxel will be administered to patients in in all treatment arms.

Statistical Methods

Primary Analysis

The co-primary efficacy endpoints are PFS as assessed by the investigator using RECIST v1.1, and OS. The primary endpoint of PFS will be analyzed in the tGE-WT population and in the ITT-WT population, and the primary endpoint of OS will be analyzed in the ITT-WT population. PFS is defined as the time between the date of randomization and the date of first documented disease progression or death, whichever occurs first. Patients who have not experienced disease progression or died at the time of analysis will be censored at the time of the last tumor

assessment. Patients with no post-baseline tumor assessment will be censored at the date of randomization plus 1 day.

OS is defined as the time between the date of randomization and death from any cause. Data for patients who are not reported as having died at the date of analysis will be censored at the date when they were last known to be alive. Data for patients who do not have post-baseline information will be censored at the date of randomization plus 1 day.

The following analyses will be performed for both PFS endpoints described above and for OS. PFS and OS will be compared between treatment arms with the use of the stratified log-rank test. The HR for PFS and OS for each comparison (i.e., Arm A vs. Arm C, Arm B vs. Arm C) will be estimated using a stratified Cox regression model, respectively. The 95% CI for the HR will be provided.

The hypothesis testing will be done in the order described below:

Comparison of Arm B versus Arm C

To control the overall type I error rate for the one-sided test at 0.025, a one-sided type I error (α) will be allocated to PFS in the tGE-WT population, PFS in the ITT-WT population, and OS in the ITT-WT population in a 3:3:19 ratio for comparison of Arm B versus Arm C

1. PFS in the tGE-WT population will be tested at $\alpha = 0.003$ (one sided). If the estimate of the HR is < 1 and the one-sided p-value corresponding to the stratified log-rank test is < 0.003, the null hypothesis will be rejected, and it will be concluded that atezolizumab + carboplatin + paclitaxel + bevacizumab prolongs the duration of PFS relative to the control arm in the tGE-WT population.
2. PFS in the ITT-WT population will be tested at $\alpha = 0.003$ (one sided).
3. α recycling from PFS to OS will be conducted as follows:
   a. If the PFS result is not statistically significant in either the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.019$ (one sided).
   b. If the PFS result is only statistically significant in the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.022$ (one sided).
   c. If the PFS result is statistically significant in both the tGE-WT population and the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.025$ (one sided).

Comparison of Arm A versus Arm C

If the difference in OS between Arm B and Arm C in the ITT-WT population is statistically significant at an α of 0.019, 0.022, or 0.025 (Step 3 above), that same α will become the overall one-sided type I error rate for the comparison of Arm A versus Arm C, with α allocated to PFS in the tGE-WT population, PFS in the ITT-WT population, and OS in the ITT-WT population at the same 3:3:19 ratio. If the difference in OS between Arm B and Arm C in the ITT-WT population is not statistically significant, there will be no formal comparison of Arm A versus Arm C for the co-primary endpoints of PFS and OS.

Depending on the outcome of the PFS testing of Arm A vs. Arm C in the tGE-WT and ITT-WT populations, the α from these two PFS comparisons will be recycled back to the OS comparison in the ITT-WT population for Arm A vs. Arm C.

1. If the difference in OS between Arm B and Arm C is statistically significant at $\alpha = 0.019$ (one sided):
   a. PFS in the tGE-WT population will be tested at $\alpha = 0.00228$ (one sided). If the estimate of the HR is < 1 and the one-sided p-value corresponding to the stratified log-rank test is < 0.00228, the null hypothesis will be rejected, and it will be concluded that atezolizumab + carboplatin + paclitaxel prolongs the duration of PFS relative to the control arm in the tGE-WT population.
   b. PFS in the ITT-WT population will be tested at $\alpha = 0.00228$ (one sided)
   c. α recycling from PFS to OS will be conducted as follows:

- If the PFS result is not statistically significant in either the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.01444$ (one sided).
- If the PFS result is only statistically significant in the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.01672$ (one sided).
- If the PFS result is statistically significant in both the tGE-WT population and the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.019$ (one sided).

2. If the difference in OS between Arm B and Arm C is statistically significant at $\alpha = 0.022$ (one sided):
   a. PFS in the tGE-WT population will be tested at $\alpha = 0.00264$ (one sided). If the estimate of the HR is < 1 and the one-sided p-value corresponding to the stratified log-rank test is < 0.00264, the null hypothesis will be rejected, and it will be concluded that atezolizumab + carboplatin + paclitaxel prolongs the duration of PFS relative to the control arm in the tGE-WT population.
   b. PFS in the ITT-WT population will be tested at $\alpha = 0.00264$ (one sided)
   c. α recycling from PFS to OS will be conducted as follows:
      - If the PFS result is not statistically significant in either the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.01672$ (one sided).
      - If the PFS result is only statistically significant in the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.01936$ (one sided).
      - If the PFS result is statistically significant in both the tGE-WT population and the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.022$ (one sided).
3. If the difference in OS between Arm B and Arm C is statistically significant at $\alpha = 0.025$ (one sided):
   a. PFS in the tGE-WT population will be tested at $\alpha = 0.003$ (one sided). If the estimate of the HR is < 1 and the one-sided p-value corresponding to the stratified log-rank test is < 0.003, the null hypothesis will be rejected, and it will be concluded that atezolizumab + carboplatin + paclitaxel prolongs the duration of PFS relative to the control arm in the tGE-WT population.
   b. PFS in the ITT-WT population will be tested at $\alpha = 0.003$(one sided)
   c. α recycling from PFS to OS will be conducted as follows:
      - If the PFS result is not statistically significant in either the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.019$ (one sided).
      - If the PFS result is only statistically significant in the tGE-WT population or the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.022$ (one sided).
      - If the PFS result is statistically significant in both the tGE-WT population and the ITT-WT population, OS in the ITT-WT population will be tested at $\alpha = 0.025$ (one sided).

The stratification factors will be those used during randomization (i.e., sex [male vs. female], presence of liver metastases at baseline [yes vs. no], and PD-L1 tumor expression by IHC [TC3 and any IC vs. TC0/1/2 and IC2/3 vs. TC0/1/2 and IC0/1]), as recorded in the interactive voice and Web response system.

Appendix 1
Protocol Synopsis (cont.)

Results from an unstratified analysis will also be presented. Kaplan-Meier methodology will be used to estimate the median PFS and the median OS for each treatment arm, and a Kaplan-Meier curve will be constructed to provide a visual description of the difference between treatment arms. The Brookmeyer-Crowley methodology will be used to construct the 95% CI for the median PFS and the median OS for each treatment arm.

Determination of Sample Size

This study will enroll approximately 1200 patients. The ITT-WT population will include approximately 1080 patients, assuming a 10% prevalence for sensitizing *EGFR* mutations or *ALK* translocations. The tGE-WT population will include approximately 540 patients, assuming a 50% prevalence with the chosen tGE cutoff.

The sample size of this study is based on the number of events required to demonstrate efficacy with regard to both PFS and OS (co-primary endpoints).

The estimate of the number of events required to demonstrate efficacy with regard to PFS in the comparison of Arm A versus Arm B is based on the following assumptions:

- One-sided significance level of 0.003 for the comparison of Arm B versus Arm C in the tGE-WT population
- One-sided significance level of 0.003 for the comparison of Arm B versus Arm C in the ITT-WT population
- 98% power to detect an HR of 0.55, corresponding to an improvement in median PFS from 6 months to 10.9 months in the tGE-WT population
- 98% power to detect an HR of 0.65, corresponding to an improvement in median PFS from 6 months to 9.2 months in the ITT-WT population
- No interim analysis for PFS
- Dropout rate of 5% per 12 months

The estimate of the number of events required to demonstrate efficacy with regard to OS in the comparison of Arm A versus Arm B is based on the following assumptions:

- One-sided significance level of 0.019 for the comparison of Arm B versus Arm C in the ITT-WT population
- 87% power to detect an HR of 0.75, corresponding to an improvement in median OS from 12 months to 16 months in the ITT-WT population
- One interim OS analysis performed at the time of the final PFS analysis, at which time approximately 73% of the total number of OS events required for the final analysis are expected to have occurred as determined through use of the Lan-DeMets approximation to the O'Brien-Fleming boundary
- Dropout rate of 5% per 12 months

The estimate of the number of events required to demonstrate efficacy with regard to PFS and OS in the comparison of Arm A versus Arm C is based on assumptions similar to those outlined above for Arm B versus Arm C.

With these assumptions, approximately 1200 patients in total will be enrolled into this study, with approximately 720 patients in each comparison (i.e., Arm B vs. Arm C and Arm A vs. Arm C) in the ITT-WT population. The final PFS analysis will be conducted when both of the following criteria have been met: approximately 516 PFS events have occurred in Arms B and C combined in the ITT-WT population and the last patient has been enrolled in the study. The final PFS analysis is expected to occur approximately 29 months after the first patient is enrolled. At the time of the final PFS analysis, it is expected that approximately 249 events will have occurred in the tGE-WT population. These numbers of events would allow for a minimum detectable difference corresponding to an HR of approximately 0.70 in the tGE-WT population and 0.78 in the ITT-WT population.

Appendix 1
Protocol Synopsis (cont.)

With a sample size of 720 patients, approximately 507 OS events are expected to occur in Arms B and C combined in the ITT-WT population for the final OS analysis. The final OS analysis is expected to occur approximately 40 months after the first patient is enrolled. This number of events corresponds to a minimum detectable difference in HR of approximately 0.83 in the ITT-WT population

Interim Analyses

There will be no interim analyses planned for PFS in this study. An external independent Data Monitoring Committee (iDMC) will be set up to evaluate safety data on an ongoing basis. All summaries/analyses by treatment arm for the iDMC's review will be prepared by an independent Data Coordinating Center (iDCC). Members of the iDMC will be external to the Sponsor and will follow a charter that outlines their roles and responsibilities. Any outcomes of these safety reviews that affect study conduct will be communicated in a timely manner to the investigators for notification of the Institutional Review Boards (IRBs)/Ethics Committees (ECs). A detailed plan will be included in the iDMC Charter.

If approximately 370 OS events have occurred in Arms B and C combined in the ITT-WT population at the time of the final PFS analysis (see criteria for final PFS analysis), an interim OS analysis will be conducted for Arm B versus Arm C in the ITT-WT population. If there are significantly fewer than the expected 370 OS events at the time of the final PFS analysis, a nominal α of 0.01% (negligible impact on overall type I error rate) will be spent on the OS analysis at the time of the final PFS analysis and a second interim OS analysis will then be conducted after approximately 370 OS events have occurred.

The final OS analysis for the comparison of Arm B versus Arm C will be conducted when approximately 507 OS events have occurred in Arms B and C combined in the ITT-WT population. This is expected to occur approximately 40 months after the first patient is enrolled.

Appendix 2
Schedule of Assessments

Procedure	Screening	All Treatment Cycles [a]		Treatment Discontinuation Visit	Survival Follow-Up
		Induction Phase (Cycles 1 to 4 or 6)	Maintenance Phase		Every 3 Months after Disease Progression or Loss of Clinical Benefit
	Days –28 to –1	Every 21 days (± 3 Days) [b]	Every 21 days (± 3 Days)	≤ 30 Days after Last Dose	
Informed consent	x				
Tumor tissue specimen for PD-L1 testing (15 FFPE slides required; blocks preferred) [c] Fresh or archival tissue can be used.	x				
ALK and/or *EGFR* assessment if status is unknown (may be done locally or centrally)	x				
Demographic data	x				
Medical history and baseline conditions	x				
NSCLC cancer history	x				
Vital signs [d]	x	x [d]	x [d]	x [d]	
Weight	x	x	x	x	
Height	x				
Complete physical examination	x				
Limited physical examination [e]		x	x	x	
ECOG performance status	x	x	x	x	
12-Lead ECG	x	x [f]	x [f]	x [f]	

Appendix 2
Schedule of Assessments (cont.)

Procedure	Screening	All Treatment Cycles [a]		Treatment Discontinuation Visit	Survival Follow-Up
		Induction Phase (Cycles 1 to 4 or 6)	Maintenance Phase		Every 3 Months after Disease Progression or Loss of Clinical Benefit
	Days –28 to –1	Every 21 days (± 3 Days) [b]	Every 21 days (± 3 Days)	≤ 30 Days after Last Dose	
Hematology [g]	x	x	x	x	
Serum chemistry [h]	x	x	x	x	
Coagulation test (aPTT or INR)	x			x	
Pregnancy test (women of childbearing-potential ONLY)	x [i]	x [j]	x [j]	x [j]	x [x]
TSH, free T3, free T4 [k]	x	x [k]	x [k]	x	
HIV, HBV, HCV serology [l]	x				
Urinalysis [m]	x	x	x	x	
Determination of duration of induction treatment	x				
Induction treatment administration Arm A: Atezolizumab + carboplatin + paclitaxel Arm B: Atezolizumab + carboplatin + paclitaxel + bevacizumab Arm C: Carboplatin + paclitaxel + bevacizumab		x [n]			
Maintenance treatment administration Arm A: Atezolizumab Arm B: Atezolizumab + bevacizumab Arm C: Bevacizumab			x [n]		

Appendix 2
Schedule of Assessments (cont.)

Procedure	Screening	All Treatment Cycles [a]		Treatment Discontinuation Visit	Survival Follow-Up
		Induction Phase (Cycles 1 to 4 or 6)	Maintenance Phase		Every 3 Months after Disease Progression or Loss of Clinical Benefit
	Days –28 to –1	Every 21 days (± 3 Days) [b]	Every 21 days (± 3 Days)	≤ 30 Days after Last Dose	
Tumor response assessment	x [o]	x [p]	x [p]		x [q]
Serum sample for atezolizumab ATA assessment (atezolizumab patients only) [r]		x	x	x	120 (±30) days after last dose of atezolizumab
Serum sample for PK sampling (atezolizumab-treated patients only) [r]		x	x	x	120 (±30) days after last dose of atezolizumab
Carboplatin, paclitaxel, and bevacizumab (Arms B and C) PK sampling (20 patients per arm) [r]		x		x	
Bevacizumab ATA and PK sampling (Arms B and C patients) [r]		x		x	
Blood samples for PD biomarkers [r]	x	x	x	x	120 (±30) days after last dose of atezolizumab
Optional blood for DNA extraction (RCR only) [r, s]		x			
Informed consent to continue treatment beyond radiographic progression (atezolizumab-treated patients)		At time of radiographic progression			

Appendix 2
Schedule of Assessments (cont.)

Procedure	Screening	All Treatment Cycles [a]		Treatment Discontinuation Visit	Survival Follow-Up
		Induction Phase (Cycles 1 to 4 or 6)	Maintenance Phase		Every 3 Months after Disease Progression or Loss of Clinical Benefit
	Days –28 to –1	Every 21 days (± 3 Days) [b]	Every 21 days (± 3 Days)	≤ 30 Days after Last Dose	
Tumor biopsy		At time of radiographic progression [y]			
Optional tumor biopsy at other timepoints (RCR only)		Any time during study treatment or during survival follow-up			
Adverse events	x	x	x	x [t]	x [t]
Concomitant medications	x [u]	x	x	x	
Patient-reported outcomes (EORTC QLQ-C30, EORTC QLQ-LC13, SILC, PGIS, and EQ-5D-3L) [v]		x [v]	x [v]		x [v]
Survival and anti-cancer therapy follow-up					x [w]

ATA = anti-therapeutic antibody; CT = computed tomography; ECOG = Eastern Cooperative Oncology Group; EORTC = European Organization for Research and Treatment of Cancer; ePRO = electronic Patient-Reported Outcome; EQ-5D-3L = Euro QoL5 Dimensions 3-Level Version; FFPE = formalin fixed paraffin embedded; HBcAb = hepatitis B core antibody; HBV = hepatitis B virus; HCV = hepatitis C virus; IV = intravenous; LC13 = Lung Cancer module; MRI = magnetic resonance imaging; NSCLC = non–small cell lung cancer; PD = pharmacodynamic; PGIS = Patient Global Impression of Severity; PK = pharmacokinetic; QLQ-C30 = Quality-of-Life Questionnaire Core 30; RCR = Roche Clinical Repository; SILC = Symptoms in Lung Cancer; TSH = thyroid-stimulating hormone.

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

[b] Cycle 1, Day 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.

Appendix 2
Schedule of Assessments (cont.)

[c] If a representative FFPE tumor specimen in paraffin block (preferred) or 15 or more unstained, freshly cut, serial sections on slides from an FFPE tumor specimen is 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 a 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 outside the 28-day screening period prior to enrollment.

[d] Vital signs include pulse rate, respiratory rate, blood pressures, and temperature. Vital signs should be recorded as described. For all sites in Argentina, pulse oximetry will be performed at every visit and these data will not be recorded.

[e] Symptom-directed physical examinations.

[f] ECG recordings will be obtained when clinically indicated.

[g] Hematology consists of CBC, including RBC count, hemoglobin, hematocrit, WBC count with differential (neutrophils, lymphocytes, eosinophils, monocytes, basophils, and other cells), and platelet count.

[h] Serum chemistry includes BUN or urea, creatinine, sodium, potassium, magnesium, chloride, bicarbonate or total CO_2, calcium, phosphorus, glucose, total bilirubin, ALT, AST, alkaline phosphatase, LDH, total protein, and albumin.

[i] Serum pregnancy test within 14 days before Cycle 1, Day 1.

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

[k] Thyroid function testing (thyroid-stimulating hormone, free T3, free T4) collected at Cycle 1, Day 1 and every fourth cycle thereafter. Total T3 will be tested only at sites where free T3 is not performed.

[l] 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 only if their HBV DNA test is negative. 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.

[m] Urinalysis by dipstick (specific gravity, pH, glucose, protein, ketones, and blood).

[n] For atezolizumab, the initial dose will be delivered over 60 (±15) 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 bevacizumab, carboplatin, and paclitaxel, study drug will be administered according to the local prescribing information, including premedication with steroids.

[o] 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.

Appendix 2
Schedule of Assessments (cont.)

[p] 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 (or loss of clinical benefit for patients assigned to atezolizumab 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.

[q] 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 (or loss of clinical benefit for patients treated with atezolizumab 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.

[r] See Appendix 3 for the detailed schedule.

[s] The optional RCR whole blood sample requires an additional informed consent and the sample can be collected at any time during the course of the study.

[t] All serious adverse events and adverse events of special interest, regardless of relationship to study treatment, will be reported until 90 days after the last dose of study treatment or initiation of new non-protocol systemic anti-cancer therapy after the last dose of study treatment, whichever occurs first. All other adverse events, regardless of relationship to study treatment, will be reported until 30 days after the last dose of study treatment or initiation of new non-protocol systemic anti-cancer therapy after the last dose of study treatment, whichever occurs first. After this period, all deaths will 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

[u] From 7 days before screening.

[v] EORTC QLQ-C30, EORTC QLQ-LC13, PGIS, and EQ-5D-3L questionnaires will be completed by the patients on the ePRO tablet at each scheduled study visit prior to administration of study drug and prior to any other study assessment(s). SILC will be completed using an electronic device at the patient's home on a weekly basis. 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). The SILC 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). The PGIS is not required during survival follow-up. Patients who discontinue study treatment for any reason other than progressive disease or loss of clinical benefit will complete the EORTC QLQ-C30, EORTC QLQ-LC13, PGIS, 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 (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression) as determined by the investigator (unless the

Appendix 2
Schedule of Assessments (cont.)

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 on 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.

[w] Survival follow-up information will be collected via telephone calls, patient medical records, and/or clinic visits every 3 months or more frequently 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.

[x] For Argentina sites only: A urine pregnancy test is required monthly until 6 months after the last dose of study treatment.

[y] Mandatory biopsy, if clinically feasible, within 40 days of radiographic progression or prior to the start of the next anti-cancer therapy, whichever is sooner.

Appendix 3
Schedule of Pharmacokinetic, Biomarker, and Anti-Therapeutic Antibody Assessments

Study Visit	Time	Arm A (Atezolizumab + Carboplatin + Paclitaxel)	Arm B (Atezolizumab + Carboplatin + Paclitaxel + Bevacizumab)	Arm C (Carboplatin + Paclitaxel + Bevacizumab)
Screening	N/A	• Biomarkers [b]	• Biomarkers [b]	• Biomarkers [b]
Cycle 1, Day 1 [f]	Pre-dose (same day as treatment administration)	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Carboplatin pharmacokinetics [a] • Paclitaxel pharmacokinetics [a] • Biomarkers [d]	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Carboplatin pharmacokinetics [a] • Paclitaxel pharmacokinetics [a] • Bevacizumab pharmacokinetics • Bevacizumab ATA • Biomarkers [d]	• Carboplatin pharmacokinetics [a] • Paclitaxel pharmacokinetics [a] • Bevacizumab pharmacokinetics • Bevacizumab ATA • Biomarkers [d]
	30 min (±10 min) after end of atezolizumab infusion	• Atezolizumab pharmacokinetic	• Atezolizumab pharmacokinetics	
	30 min (±10 min) after end of bevacizumab infusion [a]		• Bevacizumab pharmacokinetics [a]	• Bevacizumab pharmacokinetics [a]
	5–10 min before the end of carboplatin infusion [a]	• Carboplatin pharmacokinetics [a]	• Carboplatin pharmacokinetics [a]	• Carboplatin pharmacokinetics [a]
	1 hr after end of carboplatin infusion [a]	• Carboplatin pharmacokinetics [a]	• Carboplatin pharmacokinetics [a]	• Carboplatin pharmacokinetics [a]
	5–10 min before the end of paclitaxel infusion [a]	• Paclitaxel pharmacokinetics [a]	• Paclitaxel pharmacokinetics [a]	• Paclitaxel pharmacokinetics [a]
	1 hr after end of paclitaxel infusion [a]	• Paclitaxel pharmacokinetics [a]	• Paclitaxel pharmacokinetics [a]	• Paclitaxel pharmacokinetics [a]

Appendix 3
Schedule of Pharmacokinetic, Biomarker, and Anti-Therapeutic Antibody Assessments (cont.)

Study Visit	Time	Arm A (Atezolizumab + Carboplatin + Paclitaxel)	Arm B (Atezolizumab + Carboplatin + Paclitaxel + Bevacizumab)	Arm C (Carboplatin + Paclitaxel + Bevacizumab)
Cycle 2, Day 1 (±3 days)	Pre-dose (same day as treatment administration)	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	• Biomarkers [d]
Cycle 3, Day 1 (±3 days)	Pre-dose (same day as treatment administration)	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Carboplatin pharmacokinetics [a] • Paclitaxel pharmacokinetics [a] • Biomarkers [d]	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Carboplatin pharmacokinetics [a] • Paclitaxel pharmacokinetics [a] • Bevacizumab pharmacokinetics • Bevacizumab ATA • Biomarkers [d]	• Carboplatin pharmacokinetics [a] • Paclitaxel pharmacokinetics [a] • Bevacizumab pharmacokinetics • Bevacizumab ATA • Biomarkers [d]
	30 min (±10 min) after end of atezolizumab infusion	• Atezolizumab pharmacokinetic	• Atezolizumab pharmacokinetics	
	30 min (±10 min) after end of bevacizumab infusion [a]		• Bevacizumab pharmacokinetics [a]	• Bevacizumab pharmacokinetics [a]
	5–10 min before the end of carboplatin infusion [a]	• Carboplatin pharmacokinetics [a]	• Carboplatin pharmacokinetics [a]	• Carboplatin pharmacokinetics [a]

Appendix 3
Schedule of Pharmacokinetic, Biomarker, and Anti-Therapeutic Antibody Assessments (cont.)

Study Visit	Time	Arm A (Atezolizumab + Carboplatin + Paclitaxel)	Arm B (Atezolizumab + Carboplatin + Paclitaxel + Bevacizumab)	Arm C (Carboplatin + Paclitaxel + Bevacizumab)
Cycle 3, Day 1 (±3 days) (cont.)	1 hr after end of carboplatin infusion [a]	• Carboplatin pharmacokinetics[a]	• Carboplatin pharmacokinetics[a]	• Carboplatin pharmacokinetics[a]
	5–10 min before the end of paclitaxel infusion [a]	• Paclitaxel pharmacokinetics[a]	• Paclitaxel pharmacokinetics[a]	• Paclitaxel pharmacokinetics[a]
	1 hr after end of paclitaxel infusion [a]	• Paclitaxel pharmacokinetic[a]	• Paclitaxel pharmacokinetics[a]	• Paclitaxel pharmacokinetics[a]
Cycles 4, 8, and 16, Day 1 (±3 days)	Pre-dose (same day as treatment administration)	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	• Atezolizumab ATA • Atezolizumab pharmacokinetic • Biomarkers [d]	• Biomarkers [d]
After Cycle 16 and every eighth cycle, Day 1 (±3 days)	Pre-dose (same day as treatment administration)	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	• Biomarkers [d]
At time of fresh biopsy (on-treatment or at progression, including during follow-up)	At visit	• Biomarkers [d]	• Biomarkers [d]	• Biomarkers [d]

Appendix 3
Schedule of Pharmacokinetic, Biomarker, and Anti-Therapeutic Antibody Assessments (cont.)

Study Visit	Time	Arm A (Atezolizumab + Carboplatin + Paclitaxel)	Arm B (Atezolizumab + Carboplatin + Paclitaxel + Bevacizumab)	Arm C (Carboplatin + Paclitaxel + Bevacizumab)
Treatment discontinuation visit	At visit	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Bevacizumab pharmacokinetics [e] • Bevacizumab ATA [e] • Biomarkers [d]	• Bevacizumab pharmacokinetics [e] • Bevacizumab ATA [e] • Biomarkers [d]
120 (±30 days) after last dose of atezolizumab	At visit	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	• Atezolizumab ATA • Atezolizumab pharmacokinetics • Biomarkers [d]	
Any time point during study (RCR consent required)		• Optional RCR blood (DNA extraction) [c]	• Optional RCR blood (DNA extraction) [c]	Optional RCR blood (DNA extraction) [c]

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

Notes: Serum PK samples for atezolizumab and bevacizumab; plasma PK samples for carboplatin, and paclitaxel.

All patients in Arm B and C will undergo the additional ATA and PK assessments for bevacizumab at Cycle 1 pre-dose, Cycle 3 pre-dose, and at the time of bevacizumab discontinuation.

[a] At selected centers, 20 patients in each treatment arm will undergo additional PK assessments for carboplatin, paclitaxel, and bevacizumab where applicable.

[b] Whole blood for biomarkers.

[c] The optional RCR blood sample (for DNA extraction) requires an additional informed consent and the sample can be collected at any time during the course of the study.

[d] Plasma and serum for biomarkers.

[e] Bevacizumab PK and ATA are required at time of bevacizumab discontinuation.

[f] Biomarker sampling before Cycle 1, Day 1 should be performed before patients are treated with the first dose of steroids.