Official Title: **A PHASE III, OPEN-LABEL, MULTICENTER, RANDOMIZED STUDY EVALUATING THE EFFICACY AND SAFETY OF ATEZOLIZUMAB (MPDL3280A, ANTI-PD-L1 ANTIBODY) IN COMBINATION WITH CARBOPLATIN +PACLITAXEL OR ATEZOLIZUMAB IN COMBINATION WITH CARBOPLATIN+NAB-PACLITAXEL VERSUS CARBOPLATIN+NAB-PACLITAXEL IN CHEMOTHERAPY-NAIVE PATIENTS WITH STAGE IV SQUAMOUS NON-SMALL CELL LUNG CANCER**

NCT Number: NCT02367794

Document Date: SAP Version 3: 11-January-2018

STATISTICAL ANALYSIS PLAN

TITLE: **A PHASE III, OPEN-LABEL, MULTICENTER, RANDOMIZED STUDY EVALUATING THE EFFICACY AND SAFETY OF ATEZOLIZUMAB (MPDL3280A, ANTI–PD-L1 ANTIBODY) IN COMBINATION WITH CARBOPLATIN + PACLITAXEL OR ATEZOLIZUMAB IN COMBINATION WITH CARBOPLATIN + NAB-PACLITAXEL VERSUS CARBOPLATIN + NAB-PACLITAXEL IN CHEMOTHERAPY–NAÏVE PATIENTS WITH STAGE IV SQUAMOUS NON–SMALL CELL LUNG CANCER**

PROTOCOL NUMBER:	GO29437
STUDY DRUG:	Atezolizumab (RO5541267)
VERSION NUMBER:	3
IND NUMBER:	117296
EUDRACT NUMBER:	2014-003208-59
SPONSOR:	F. Hoffmann-La Roche Ltd
PLAN PREPARED BY:	[redacted]
DATE FINAL:	19 May 2017
DATES AMENDED	Version 2: 26 September 2017 Version 3: See electronic date stamp below.

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

Name	Reason for Signing	Date and Time (UTC)
[redacted]	Company Signatory	11-Jan-2018 07:27:08

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

The following changes have been made in Statistical Analysis Plan (SAP) GO29437, Version 3:

1. Per modification plan in Appendix 4 of SAP GO29437, Version 2 for primary analysis, Statistical Analysis Plan (SAP) GO29437, Version 2 has been amended to reflect the final statistical testing strategy prior to study database lock and unblinding. In SAP GO29437, Version 2, based on recently available data from Study GO29436 that are external to Study GO29437.
   - The primary analysis of progression-free survival (PFS) was planned to be conducted hierarchically: first, in patients with high expression of the T-effector (Teff) gene signature (Teff-high, previously referred to as tGE) and, subsequently, in the intent-to-treat (ITT) population (Section 2.3). Based on recently available data from Study GO29436 suggest that the treatment effect on PFS is observed in the ITT population, thus in accord with the modification scenarios, the primary analysis of PFS in SAP GO29437, Version 3 will be conducted in the ITT population directly, i.e., testing of PFS in the Teff-high population has been removed from the primary testing hierarchy.
   - Recent external data from Study GO29436 also suggest that the treatment effect on overall survival (OS) may be delayed such that an accurate assessment of the magnitude of the effect will take longer than initially expected. Thus, a second OS interim analysis with α spending has been added for when approximately 410 OS events have occurred in Arm B and Arm C combined. The previously planned initial interim analysis at the time of primary PFS analysis will still be conducted but with a nominal α spent (Section 2.4.2).
2. Based on recent data from Study GO29436, the patient-reported outcomes (PRO) time to deterioration (TTD) analysis was modified to form a composite symptom endpoint rather than assessing each individual symptom. PRO responder analyses were also added to help interpret the results from the change from baseline analyses.

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

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. BACKGROUND

This Statistical Analysis Plan (SAP) provides details of the planned analyses and statistical methods for Study GO29437 (IMpower131), "A Phase III, Open-label, Multicenter, Randomized Study Evaluating the Efficacy and Safety of Atezolizumab (MPDL3280A, Anti–PD-L1 Antibody) in Combination with Carboplatin + Paclitaxel or Atezolizumab in Combination with Carboplatin + Nab-Paclitaxel versus Carboplatin + Nab-Paclitaxel in Chemotherapy–Naïve Patients with Stage IV Squamous Non–Small Cell Lung Cancer." The background for the study can be found in the study protocol.

2. STUDY DESIGN

This is a randomized, Phase III, multicenter, open-label study designed to evaluate the safety and efficacy of atezolizumab in combination with carboplatin + paclitaxel or with carboplatin + nab-paclitaxel as compared with treatment with carboplatin + nab-paclitaxel in approximately 1025 chemotherapy–naïve patients with Stage IV 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-infiltrating immune cell [IC] vs. TC0/1/2 and IC2/3 vs. TC0/1/2 and IC0/1) tested by a central laboratory. TC0, TC1, TC2, and TC3 are defined as PD-L1 staining in $< 1\%, \geq 1\%$ to $< 5\%, \geq 5\%$ to $< 50\%$, and $\geq 50\%$ of the TCs, respectively. IC0, IC1, IC2, and IC3 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.

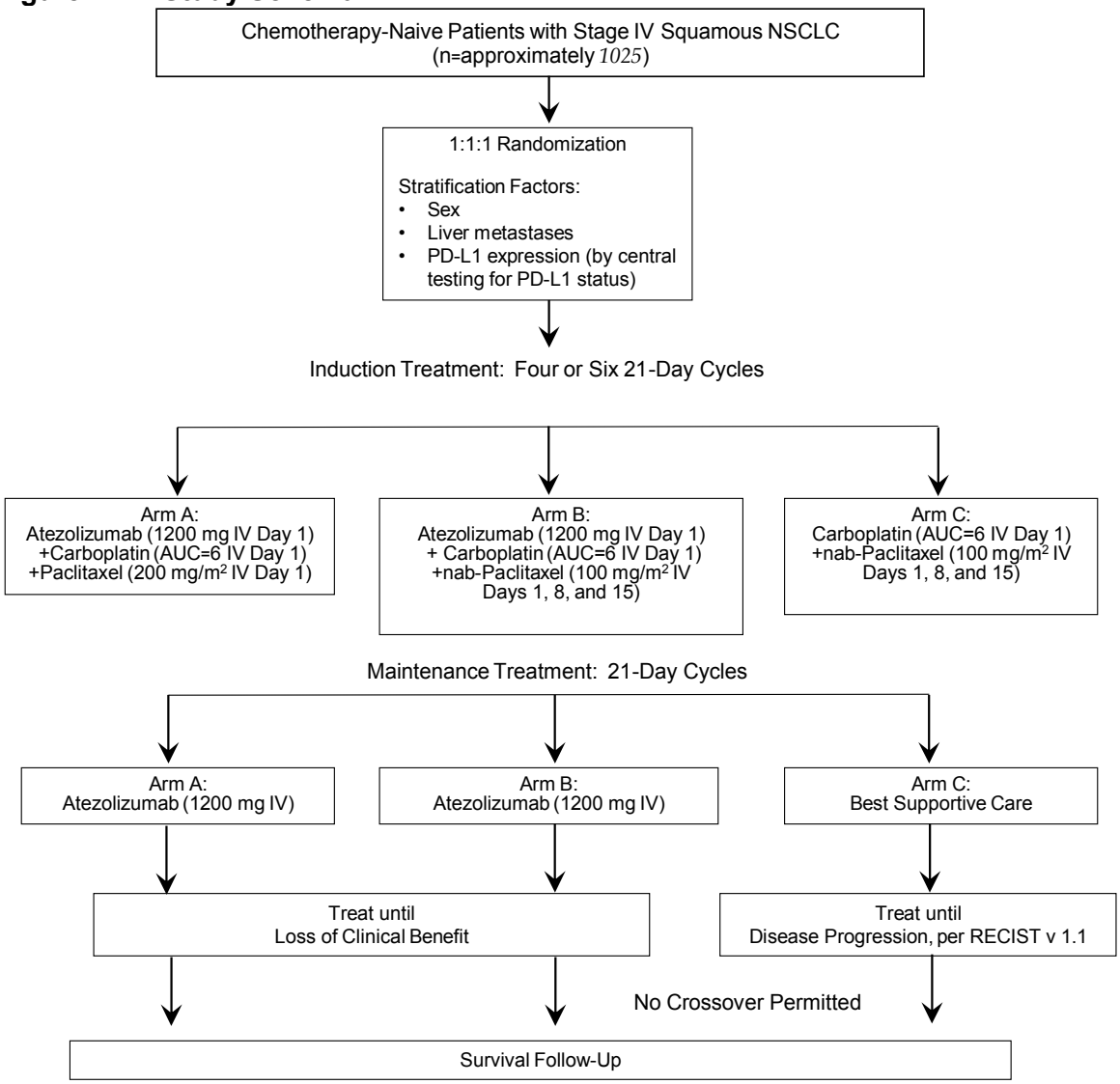
Eligible patients were randomized in a 1:1:1 ratio to one of the following treatment regimens:

- Treatment Arm A: Atezolizumab + carboplatin + paclitaxel (induction: four or six 21-day cycles); atezolizumab (maintenance: 21-day cycles)
- Treatment Arm B: Atezolizumab + carboplatin + nab-paclitaxel (induction: four or six 21-day cycles); atezolizumab (maintenance: 21-day cycles)
- Treatment Arm C: Carboplatin + nab-paclitaxel (induction: four or six 21-day cycles); best supportive care

Patients who were randomized to Arms A or Arm B may continue treatment with atezolizumab beyond radiographic disease 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 of which details 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 v1.1 = Response Evaluation Criteria in Solid Tumors, Version 1.1.

Note: In Arm A, patients with Asian race/ethnicity receive paclitaxel at a dose of 175 mg/m^2 (Protocol Amendment v6 dated 1 March 2017).

Patients undergo tumor assessments at baseline and every 6 weeks (±7 days) for the first 48 weeks following 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 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 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 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.

The co-primary efficacy endpoints are progression-free survival (PFS), as assessed by the investigator using 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 and all OS analyses will be performed on all randomized patients (the intent-to-treat [ITT] population). See Section 4.1 for details on analysis populations.

There are no interim analyses planned for PFS. Two interim analyses are 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 Measures

The co-primary efficacy endpoints are PFS, as assessed by the investigator using RECIST v1.1, and OS.

- PFS is defined as the time between the date of randomization and the date of first documented disease progression or death, whichever occurs first, in the ITT population.
- OS is defined as the time between the date of randomization and date of death from any cause in the ITT population.

2.2.2 Secondary Efficacy Measures

The secondary efficacy outcome measures for this study are:

- PFS, as determined by the investigator according to RECIST v1.1, and OS in the PD-L1 IHC populations defined as TC2/3 or IC2/3 and TC1/2/3 or IC1/2/3 using the SP142 assay

- PFS, as determined by the investigator according to RECIST v1.1, and OS in the biomarker populations defined by expression of a T-effector (Teff) gene signature in tumor tissue, as described in Section 4.4.5 (previously referred to as tGE)
- Objective response, defined as either complete response (CR) or partial response (PR) as determined by the investigator according to RECIST v1.1 in the ITT population
- Duration of response (DOR), defined for patients with objective response as the time from the first documented objective response to documented disease progression as determined by the investigator using RECIST v1.1 or death from any cause, whichever occurs first, in the ITT 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 ITT population
- PFS, as determined by the investigator according to RECIST v1.1, and OS in the two atezolizumab-containing arms in the ITT 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 Research and Treatment of Cancer (EORTC) Quality-of-Life Questionnaire Core 30 (QLQ-30) and supplemental lung cancer module (EORTC QLQ-LC13) symptom subscales (cough, dyspnea [single item], dyspnea [multi-item subscale], chest pain, and arm/shoulder pain) in the ITT 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 in the ITT population

2.2.3 Exploratory Efficacy Measures

The exploratory efficacy measures for this study are:

- 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
- 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
- Utility scores of the Euro QoL5 Dimensions 3-Level Version (EQ-5D-3L)

- Change from baseline and the number and proportion of patients with a clinically meaningful change 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

2.2.4 Pharmacokinetic Outcome Measures

The PK outcome measures for this study are:

- 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 (Arm A)
- Plasma concentrations for nab-paclitaxel (Arms B and C)

2.2.5 Safety Outcome Measures

The safety outcome measures for this study are:

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

2.3 DETERMINATION OF SAMPLE SIZE

This study will randomize approximately 1025 patients.

To control the overall type I error rate for the one-sided test at 0.025 for the primary comparison of atezolizumab + carboplatin + nab-paclitaxel (Arm B) against the control arm of carboplatin + nab-paclitaxel (Arm C), a one-sided α of 0.003 and a one-sided α of 0.022 are allocated to PFS and OS, respectively. The primary comparison of PFS will be tested at a one-sided α level of 0.003 in the ITT population. The primary comparison of OS will be tested in the ITT 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 population is statistically significant between the comparison of Arm B versus Arm C, PFS in the ITT population and OS in the ITT 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 3:22 between the

PFS and OS tests. Depending on the outcome of the PFS testing of Arm A versus Arm C in the ITT population, the α from the PFS comparison will be recycled back to the OS comparison in the ITT population for Arm A versus Arm C.

The PFS and OS analysis hierarchy and α allocation (Burman et al. 2009), including possible α recycling, are shown in Figure 2.

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

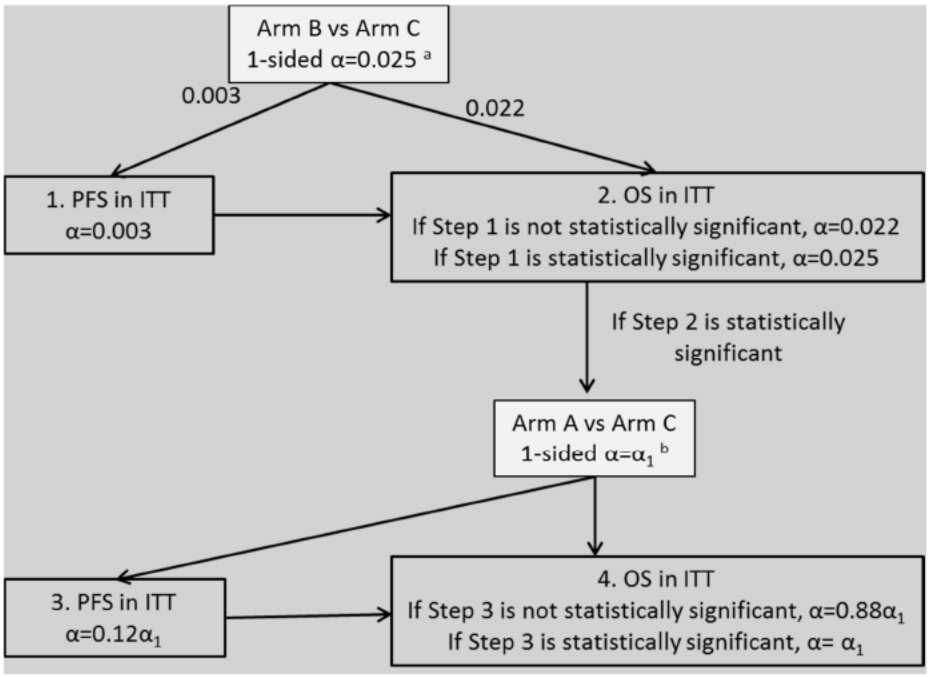
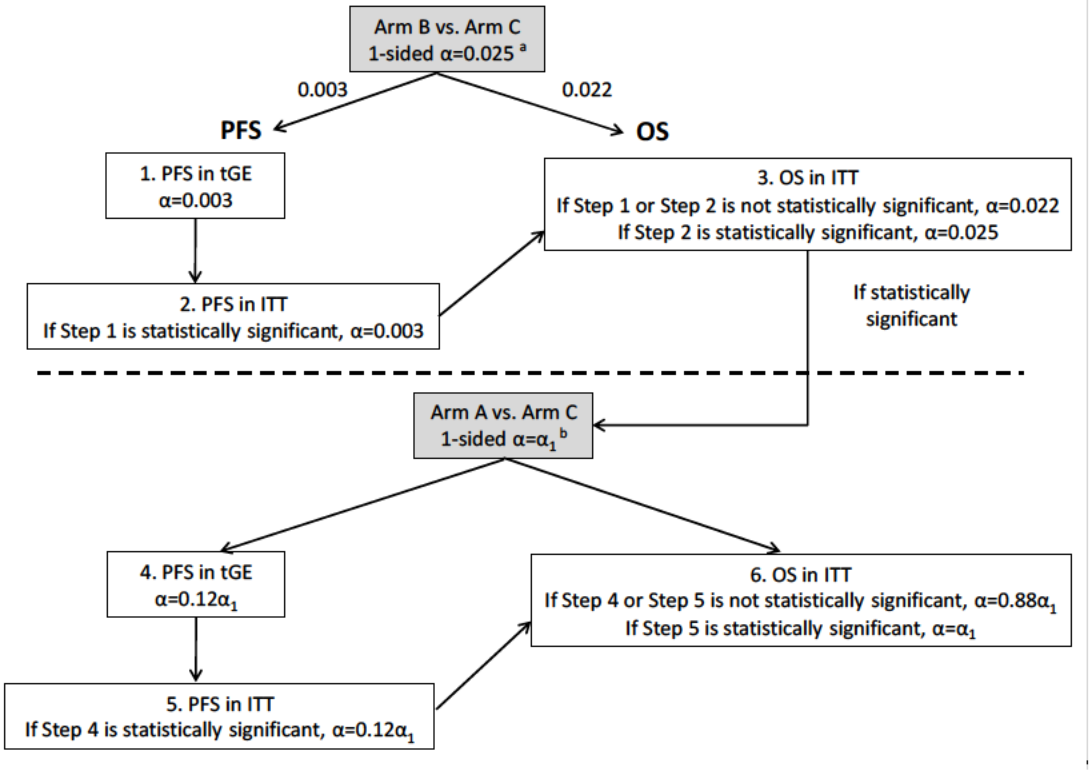


Figure 2 Progression-Free Survival and Overall Survival Analysis Hierarchy, α Allocation, and α Recycling (cont.)



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

α_1 = type I error passed down to the comparison of Arm A versus Arm C (i.e., 0.022 or 0.025)

[a] To control the overall type I error rate for the one-sided test at 0.025, a one-sided type I error (α) of 0.003 and 0.022 (ratio of 3:22) will be allocated to PFS and OS, respectively, for the comparison of Arm B versus Arm C.

[b] If the difference in OS between Arm B and Arm C in the ITT population is statistically significant at an α of 0.022 or 0.025 (Step 2), 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 and OS at the same 3:22 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 comparison of OS, as described above. Statistical significance at the interim and final analyses of OS will be tested as described in Section 2.4.2.

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) 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 of Am B versus Arm C are based on the following assumptions:

- One-sided significance level of 0.003 for the comparison of Arm B versus Arm C in the ITT 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 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 comparison of Am B versus Arm C are based on the following assumptions:

- One-sided significance level of 0.022 for the comparison of Arm B versus Arm C in the ITT population
- 86% power to detect an HR of 0.75, corresponding to an improvement in median OS from 12 months to 16 months in the ITT population
- First OS interim analysis performed at approximately 53% event-patient ratio (EPR) with nominal α spent (one sided $\alpha = 0.0001$), a second OS interim analysis performed at approximately 60% EPR or approximately 86% of the total OS events required for the final analysis using the Lan-DeMets approximation to the O'Brien-Fleming boundary for group sequential design
- Dropout rate of 5% per 24 months

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

With these assumptions, approximately 1025 patients in total will be randomized into this study, with approximately 684 patients in each 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 primary analysis for the primary comparison of Arm B versus Arm C will be conducted when approximately 365 OS events in the ITT population have occurred in Arm B and Arm C combined, which is when the first planned interim analysis of OS as described in Section 2.4.2 will be conducted. It is expected that there will be approximately 499 PFS events in the ITT population in Arm B and Arm C combined at this timepoint. These numbers of events correspond to a minimum detectable difference (MDD) in HR of approximately 0.78 in the ITT population. It is projected that an observed HR of 0.78 or lower will result in a statistically significant difference between treatment arms in the ITT population in this analysis. The PFS primary analysis is

expected to occur approximately 30 months after the first patient is randomized based on the study assumptions and the projected accrual rate.

Table 1 Primary 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, %
ITT	30	499 (73)	0.78	98

FPI = first patient in; ITT = intent-to-treat; MDD = minimum detectable difference.

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 are two interim analyses planned for the co-primary endpoint of OS.

The first OS interim analysis for the primary comparison of Arm B versus Arm C will be conducted when approximately 365 OS events in the ITT population have occurred in Arm B and Arm C combined, with nominal α spent (one sided $\alpha = 0.0001$) on the comparison. At the time of the first OS interim analysis, the final PFS analysis will be conducted by the Sponsor.

The second OS interim analysis for the primary comparison of Arm B versus Arm C will be conducted when approximately 410 OS events in the ITT population have occurred in Arm B and Arm C combined. This number of events corresponds to a MDD in HR of approximately 0.80 in the ITT population. The OS second interim analysis is expected to occur approximately 33 months after the first patient is randomized.

The final OS analysis for the primary comparison of Arm B versus Arm C will be conducted when there are approximately 477 OS events in the ITT population in Arm B and Arm C combined. This number of events corresponds to a MDD in HR of approximately 0.83 in the ITT population. The OS final analysis is expected to occur approximately 39 months after the first patient is randomized.

The expected analysis timing for the 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 Population	
Type of Analysis	Months from FPI	Percentage Information, %	No. of Events (Event Ratio, %)	Power, %[a]
OS first interim analysis	30	76	365 (53)	18
OS second interim analysis	33	86	410 (60)	76
OS final analysis	39	100	477 (70)	86

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

[a] Power is calculated using one-sided α of 0.022.

The stopping boundaries for the second 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 population, assuming the specified observed number of events at each analysis (365, 410, and 477, respectively), are shown in Table 3. The p-value will be used to claim crossing of the boundaries. For example, if approximately 410 events have occurred at the time of the OS second interim analysis in the ITT population for Arm B and Arm C combined, statistical significance of the OS endpoint in the ITT population will be declared for Arm B if one-sided $p \leq 0.0156$ when PFS has claimed significance in the ITT population.

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

	Stopping Boundary in HR (p-value)	
Analysis Timing	If $\alpha = 0.022$[a]	If $\alpha = 0.025$[b]
OS first interim analysis	HR ≤ 0.682 (p-value ≤ 0.0001)	HR ≤ 0.683 (p-value ≤ 0.0001)
OS second interim analysis	HR ≤ 0.804 (p-value ≤ 0.0134)	HR ≤ 0.808 (p-value ≤ 0.0156)
OS final analysis	HR ≤ 0.826 (p-value ≤ 0.0182)	HR ≤ 0.830 (p-value ≤ 0.0206)

HR = hazard ratio; OS = overall survival.

a. When the PFS result is not statistically significant in the ITT population for Arm B vs Arm C comparison.

b. When the PFS result is statistically significant in the ITT population for Arm B vs Arm C comparison

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 population is statistically significant at an one-sided α of 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 population, assuming the specified observed number of events at each analysis (365, 410, and 477, respectively), are provided for the two scenarios ($\alpha = 0.022$ or 0.025) in Table 4.

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

A. OS tested for statistically significant at $\alpha = 0.022$ in the ITT population for Arm B vs. Arm C Comparison

	Stopping Boundary in HR (p-value)	
Analysis Timing	If $\alpha = 0.01936$[a]	If $\alpha = 0.022$[b]
OS first interim analysis	HR $\leq$ 0.682 (p-value $\leq$ 0.0001)	HR $\leq$ 0.682 (p-value $\leq$ 0.0001)
OS second interim analysis	HR $\leq$ 0.799 (p-value $\leq$ 0.0116)	HR $\leq$ 0.804 (p-value $\leq$ 0.0134)
OS final analysis	HR $\leq$ 0.822 (p-value $\leq$ 0.0160)	HR $\leq$ 0.826 (p-value $\leq$ 0.0182)

ITT = intent to treat; HR = hazard ratio; OS = overall survival.

a. When the PFS result is not statistically significant in the ITT population for Arm A vs Arm C comparison.

b. When the PFS result is statistically significant in the ITT population for Arm A vs Arm C comparison

Note: α and p-value are one-sided.

B. OS tested for statistically significant at $\alpha = 0.025$ in the ITT population for Arm B vs. Arm C Comparison

Analysis Timing	Stopping Boundary in HR (p-value)	
	If $\alpha = 0.022$[a]	If $\alpha = 0.025$[b]
OS first interim analysis	HR ≤ 0.682 (p-value ≤ 0.0001)	HR ≤ 0.683 (p-value ≤ 0.0001)
OS second interim analysis	HR ≤ 0.804 (p-value ≤ 0.0134)	HR ≤ 0.808 (p-value ≤ 0.0156)
OS final analysis	HR ≤ 0.826 (p-value ≤ 0.0182)	HR ≤ 0.830 (p-value ≤ 0.0206)

ITT = intent to treat; HR = hazard ratio; OS = overall survival.

a. When the PFS result is not statistically significant in the ITT population for Arm A vs Arm C comparison.

b. When the PFS result is statistically significant in the ITT population for Arm A vs Arm C comparison

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 Independent Review Facility (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 sensitivity 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 the first interim OS analysis. 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 amendment(s) 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 GO29437 for the purposes of a regulatory filing.

4.1 ANALYSIS POPULATIONS

4.1.1 Randomized Population/ITT Population

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

The Teff (-0.91) population is defined as ITT patients with Teff signature score $\geq$ -0.91 in baseline tumor tissues. The Teff (-1.91) population is defined as ITT patients with a Teff signature score ≥ -1.91. The Teff signature score captures the gene expression of PD-L1 (CD274), CXCL9, and interferon-gamma relative to a reference gene using RNA isolated from the patient's formalin–fixed paraffin-embedded tumor tissue, as measured on a Roche Molecular System Cobas 4800 platform.

The PD-L1 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 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 primary analyses of PFS and all OS analyses will be performed on the ITT population. Patients are grouped according to the treatment assigned at randomization, regardless of whether they received any assigned study drug.

Objective response rate (ORR) will be analyzed for patients in the ITT population who have measurable disease as assessed by investigator at baseline. DOR will be assessed in patients in the ITT population who have an objective response and measurable disease as assessed by investigator at baseline. TTD analyses will be conducted in the ITT population. Change from baseline analysis in PROs will be performed for patients in the ITT population 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, carboplatin, paclitaxel, or nab-paclitaxel and who have evaluable PK samples.

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, then the patients will be grouped to Arm B for the safety analyses.

4.2 ANALYSIS OF STUDY CONDUCT

Study enrollment, major protocol deviations including major deviations from inclusion/exclusion criteria, and reasons for discontinuation from the study will be summarized overall and by treatment arm for the ITT population. 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 ITT populations. Descriptive statistics (mean, median, standard deviation, 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.

4.4 EFFICACY ANALYSIS

Patients will be grouped for efficacy analyses according to 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 per RECIST v1.1 or death from any cause, whichever occurs first. Data for patients who are alive and have not experienced disease progression at the time of 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 ITT population can be phrased in terms of the survival functions $S_A(t)$ or $S_B(t)$ for the atezolizumab arms (Arms A or B, respectively) and $S_C(t)$ for the control arm (Arm C):

H_0: $S_A(t) = S_C(t)$ versus H_1: $S_A(t) > S_C(t)$

or H_0: $S_B(t) = S_C(t)$ 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 PD event 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 and OS analyses in the ITT 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 IC0/1). The stratification factors will be those used during randomization as recorded in the interactive web/voice response system (IxRS) 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 (CRF) may also be provided if the discrepancy between IxRS and CRF recorded strata is substantial. 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 ITT population for the PFS and OS endpoints. 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 interim analyses (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 ITT 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 + nab-paclitaxel prolongs the duration of PFS relative to the control arm in the ITT population.
2. α recycling from PFS to OS will be conducted as follows:

a) If PFS result is not statistically significant in the ITT population, then OS in the ITT population will be tested at $\alpha = 0.022$ (one-sided).

b) If PFS result is statistically significant in the ITT population, then OS in the ITT population will be tested at $\alpha = 0.025$ (one-sided).

If the difference in OS between Arm B and Arm C in the ITT 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 and OS endpoints at the same 3:22 ratio (see Figure 2 in Section 2.3). Depending on the outcome of the PFS testing of Arm A versus Arm C in the ITT population, the α from the PFS comparison will be recycled back to the OS comparison in the ITT 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 (one-sided) for Comparison of Arm A versus Arm C

Arm B vs. Arm C	Arm A vs. Arm C	
α level at which the OS is tested statistically significant in the ITT population	α level at which the PFS will be tested in the ITT population	α level at which the OS will be tested in the ITT population
0.022	0.00264	0.01936, if the PFS result is not statistically significant in the ITT population 0.022, if the PFS result is statistically significant in the ITT population
0.025	0.003	0.022, if the PFS result is not statistically significant in the ITT population 0.025, if the PFS result is statistically significant in the ITT population

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

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 PFS and OS in Teff Populations

PFS, as determined by the investigator according to RECIST v1.1, and OS in the Teff (formerly referred to as tGE) populations, will be analyzed using the same methods as described in Section 4.4.1.

4.4.2.2 PFS and OS in PD-L1 Populations

PFS, as determined by the investigator according to RECIST v1.1, and OS in the PD-L1 TC2/3 or IC2/3 and TC1/2/3 or IC1/2/3 populations will be analyzed using the same methods as described in Section 4.4.1.

4.4.2.3 Objective Response Rate

ORR (confirmation not required) is defined as the proportion of patients with an objective response, either CR or PR as determined by the investigator using RECIST v1.1. Patients not meeting these criteria, including those without any post-baseline assessment, will be considered non-responders. Confirmation of response according to RECIST is not required, but for exploratory purposes, ORR with confirmation may be reported. The analysis population for ORR will be the ITT population with measurable disease at baseline, i.e., patients with at least one measureable lesion as assessed by the investigator using RECIST v1.1. 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.

4.4.2.4 Duration of Response

DOR is defined for patients with an objective response as the time from the first documented objective response to documented disease progression as determined by the investigator using 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.

4.4.2.5 OS Rate at 1- and 2-Year Landmark Timepoints

The OS rate at 1- and 2-year landmark time points after randomization in the ITT population will be estimated for each treatment arm using KM methodology, along with 95% CI calculated with the standard error derived from Greenwood's 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 Greenwood's method.

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

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

4.4.2.7 Patient-Reported Outcomes

Through the SILC scale, the EORTC QLQ-C30, and the EORTC QLQ-LC13, lung cancer symptom data will be collected for symptoms commonly associated with cancer treatments and on disease and treatment impact on patient 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 $\leq$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 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.

TTD will be documented for a 3-symptom composite endpoint using the following EORTC QLQ-LC13 symptom scores: cough, chest pain, and dyspnea multi--item scale. In this instance, symptom deterioration will be determined as a $\geq$10-point increase above baseline in any of the listed symptom scores, whichever occurs first (cough, chest pain, and dyspnea multi-item scale). Confirmed clinically meaningful symptom deterioration will need to be held for the original symptom; a $\geq$10-point increase above baseline in a symptom score (i.e. cough) must be held for at least two consecutive assessments (i.e. cough for two assessments) or an initial $\geq$10-point increase above baseline followed by death within 3 weeks from the last assessment. A $\geq$10-point change in the EORTC scale score is perceived by patients as clinically significant (Osoba et al. 1998).

TTD analyses will be performed for the ITT population and will include all data collected through disease progression and survival follow-up. The methodologies that are outlined for the analysis of OS will be used for 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 may be provided. The analysis may be performed for patients in the ITT population with a baseline and a post-baseline PRO assessment. Graphs of the mean changes and standard errors over time from the baseline assessment for the total score and subscales may be provided for each treatment arm.

The analysis of SILC change from baseline may 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.

The analyses per SILC scale may not be included in the CSR for this study.

4.4.3 Exploratory Efficacy Endpoints

4.4.3.1 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 4.4.2.5.

4.4.3.2 Overall Survival Rate at 3 Years

OS rate at 3 years will be estimated and analyzed using the same method as described in Section 4.4.2.5.

4.4.3.3 EQ-5D-3L Health Status Data

Health status is assessed by the EQ-5D-3L. Data from the EQ-5D-3L questionnaire will be collected to generate HRQoL and utility scores for use in economic models for reimbursement.

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.

4.4.3.4 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 each of the following individual lung cancer symptoms (i.e., cough, dyspnea, chest pain) with use of the EORTC.

To help interpret the results from the change from baseline analyses, the number and proportion of patients with a clinically meaningful change may be summarized by treatment arm, for the EORTC QLQ-C30 global health status and physical function scale scores and for the SILC cough, chest pain, and dyspnea scale scores at each cycle post-baseline.

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 CRF.

These additional PRO analyses may not be included in the CSR for this study.

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 one 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 may 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 a 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 according to a range of possible effects on OS of the subsequent NPT after treatment switching occurred (e.g., 10%, 20%, and 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 OS 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 the start of NPT and uses patients in the control arm 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 into 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 effect of long-term survival and delayed clinical effects, a milestone OS analysis may be conducted (Chen 2015). The milestone timepoint(s) will be chosen such that the patients included in the analysis will achieve a certain percentage of patient-event ratio.

The milestone OS analysis may 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) may be computed for OS using the area under the curve from baseline to several timepoints. RMST may 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 arm, an analysis of OS may be performed using the weighted log-rank test (Fleming and Harrington 1991) that weight more on late events to account for the delayed clinical effect (Fine et al. 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 may be considered as having died at the last date they were known to be alive.

4.4.4.4 Progression-Free Survival 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 RECIST v1.1 or death from any cause, whichever occurs first.

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

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.; *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 as determined by the SP142 assay; subgroups with PD-L1 expression on at least 1%, 25%, or 50% of TC as determined by the SP263 assay may also be performed)
- Teff populations with the corresponding complementary biomarker populations defined by the Teff cutoff values of -0.91, and -1.91
- Intended cycles during induction phase (4 cycles vs. 6 cycles)

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

Atezolizumab PK samples will be collected in this study as outlined in Appendix 3. Atezolizumab serum concentration data (C_{min} and C_{max}) will be tabulated and summarized. Descriptive statistics will include means, medians, ranges, and SDs, as appropriate.

Plasma concentrations of carboplatin, paclitaxel, and nab-paclitaxel will be collected in this study as outlined in Appendix 3. The concentrations of carboplatin, paclitaxel, and nab-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 may 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 PD 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 full or partial dose 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 to 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 Medical Dictionary for Regulatory Activities thesaurus terms. Adverse events will be graded by the investigator according to 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, and adverse events that led 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.

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, deaths reported during the follow-up period after treatment completion/discontinuation, and causes of death will be summarized by treatment arm.

4.6.3 Laboratory Data

Laboratory data will be summarized over time by treatment arm. Change from baseline and values outside the normal ranges may be summarized. In addition, selected laboratory data will be classified according to NCI CTCAE and will be summarized by grade. Highest NCI CTCAE grade post-baseline will also be reported and shift tables from baseline to worst value during the study post-baseline will be presented.

4.6.4 Vital Signs

Changes in selected vital signs will be summarized by treatment arm and by change over time including change from baseline.

4.6.5 Anti-Drug Antibody

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

4.7 MISSING DATA

Please refer to 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, MULTICENTER, RANDOMIZED STUDY EVALUATING THE EFFICACY AND SAFETY OF ATEZOLIZUMAB (MPDL3280A, ANTI–PD-L1 ANTIBODY) IN COMBINATION WITH CARBOPLATIN + PACLITAXEL OR ATEZOLIZUMAB IN COMBINATION WITH CARBOPLATIN + NAB-PACLITAXEL VERSUS CARBOPLATIN + NAB-PACLITAXEL IN CHEMOTHERAPY–NAIVE PATIENTS WITH STAGE IV SQUAMOUS NON–SMALL CELL LUNG CANCER**
PROTOCOL NUMBER:	GO29437
VERSION NUMBER:	6
EUDRACT NUMBER:	2014-003208-59
IND NUMBER:	117296
TEST PRODUCT:	Atezolizumab (MPDL3280A, RO5541267)
PHASE:	III
INDICATION:	Squamous non–small cell lung cancer
SPONSOR:	F. Hoffmann-La Roche Ltd

Objectives

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

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

The term “tumor gene expression” (tGE) refers to randomized patients with a defined level of expression of a programmed death–ligand 1 (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 tumor-infiltrating immune cells (ICs), as analyzed through use of a centrally performed immunohistochemistry (IHC) test.

Efficacy Objectives

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 Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST v1.1) in the tGE population and the intent-to-treat (ITT) population

To evaluate the efficacy of atezolizumab as measured by overall survival (OS) in the ITT population

Appendix 1 Protocol Synopsis (cont.)

The secondary efficacy objectives for this study are the following:

To evaluate the efficacy of atezolizumab as measured by OS in the tGE 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 population and the TC1/2/3 or IC1/2/3 population

To evaluate the efficacy of atezolizumab as measured by investigator-assessed objective response rate (ORR) according to RECIST v1.1 in the tGE population and the ITT population

To evaluate the efficacy of atezolizumab as measured by investigator-assessed duration of response (DOR) according to RECIST v1.1 in the tGE population and the ITT population

To evaluate the OS rate at 1 and 2 years in each treatment arm for the tGE population and the ITT population n

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 population and the ITT population

To determine the impact of atezolizumab as measured by change from 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 in the tGE population and the ITT population

To evaluate the efficacy of the treatment regimen of atezolizumab + carboplatin + paclitaxel versus atezolizumab + carboplatin + nab-paclitaxel as measured by investigator-assessed PFS according to RECIST v1.1 and OS in the tGE population and the ITT population

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 + paclitaxel (Arm A) or with carboplatin + nab-paclitaxel (Arm B)
- To characterize the pharmacokinetics of carboplatin when given in combination with paclitaxel and atezolizumab (Arm A) or with nab-paclitaxel with and without atezolizumab (Arms B and C)
- To characterize the pharmacokinetics of paclitaxel when given in combination with atezolizumab and carboplatin (Arm A)
- To characterize the pharmacokinetics of nab-paclitaxel (reported as total paclitaxel) when given in combination with carboplatin with and without atezolizumab (Arms B and C)

Exploratory Objectives

The exploratory objectives for this study are the following:

- 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 (IMpower131) designed to evaluate the safety and efficacy of atezolizumab in combination with carboplatin + paclitaxel or with carboplatin + nab-paclitaxel compared with treatment with carboplatin + nab-paclitaxel in approximately 1025 chemotherapy–naive patients with Stage IV 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) and 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 cycle)

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

Treatment Arm C: Carboplatin + nab-paclitaxel (Induction: four or six 21-day cycles); best supportive care (Maintenance: 21-day cycle)

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 occurs (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 randomized to atezolizumab (Arms A and B) may continue treatment with atezolizumab beyond radiographic disease progression by 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. Patients randomized to carboplatin + nab-paclitaxel (Arm C) will be offered best supportive care provided they have non-progressive disease.

Treatment with chemotherapy must be discontinued in all patients who exhibit evidence of progressive disease by RECIST v1.1.

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 assessments 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 continue 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 per RECIST v1.1 (e.g., toxicity, symptomatic deterioration) will continue scheduled tumor assessments until radiographic disease progression per RECIST v1.1 (or loss of clinical benefit for atezolizumab-treated patients who had continued treatment with atezolizumab after radiographic disease progression according to RECIST v1.1), withdrawal of consent, study termination by Sponsor, or death, whichever occurs first, regardless of whether patients start a new anti-cancer therapy.

All primary imaging data used for tumor assessment will be collected by the Sponsor to enable centralized, independent review of response endpoints. The independent reviews of the stored scans will be performed when requested.

For Treatment Arms A and B only:

During treatment (induction or maintenance), patients 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 study treatment atezolizumab at the time of initial progression

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 310 sites globally will participate in the study and approximately 1025 patients will be randomized.

Target Population

Inclusion Criteria

Patients may be eligible if they have chemotherapy–naive, Stage IV, 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

Appendix 1 Protocol Synopsis (cont.)

- Histologically or cytologically confirmed, Stage IV squamous NSCLC (per the Union Internationale contre le Cancer/American Joint Committee on Cancer staging system, 7th edition).

  Patients with tumors of mixed histology (squamous and non-squamous) are eligible if the major histological component appears to be squamous.

- No prior treatment for Stage IV squamous NSCLC

  Patients known to have 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* tyrosine kinase inhibitors (TKIs), such as erlotinib, gefitinib, or another *EGFR* TKI appropriate for the treatment of *EGFR*-mutant NSCLC. Patients with unknown *EGFR* mutation status do not require testing.

  Patients known to have 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 *ALK* mutation status do not require testing.

- 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 central nervous system (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.

Appendix 1 Protocol Synopsis (cont.)

- 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, for 30 days after the last dose of nab-paclitaxel, or for 6 months after the last dose of paclitaxel, whichever is later. 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 nab-paclitaxel, paclitaxel, and/or carboplatin. 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 or vasectomized partner (on the understanding that this is the only one partner during the whole study duration), and sexual abstinence. Men must refrain from donating sperm during the study and for 6 months after the last dose of nab-paclitaxel, paclitaxel, and/or carboplatin, whichever is latest.
- 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 trial if they have a partner of childbirth potential. Male patients must always use a condom.
- Women who are not postmenopausal ($\geq$ 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.

Appendix 1 Protocol Synopsis (cont.)

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 computed tomography (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.
- 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 anti–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 I 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 psoralen plus ultraviolet A radiation [PUVA], 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 the 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 render 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 21 days 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–cytotoxic T-lymphocyte-associated protein 4 (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, interferons and interleukin-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 Chemotherapy

- Known history of severe allergic reactions to platinum-containing compounds or mannitol

- Known sensitivity to any component of paclitaxel or nab-paclitaxel
- Grade ≥ 2 peripheral neuropathy as defined by NCI CTCAE v4.0 criteria (paclitaxel and nab-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 all of the following criteria have been met: approximately 240 PFS events have occurred in Arms B and C combined in the tGE population, approximately 499 PFS events have occurred in Arms B and C combined in the ITT population, and the last patient has been enrolled in the study. The final PFS analysis is expected to occur approximately 30 months after the first patient is enrolled.

With a sample size of 684 patients, approximately 477 OS events are expected to occur in Arms B and C combined in the ITT population for the final OS analysis. The final OS analysis is expected to occur approximately 39 months after the first patient is enrolled.

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

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

- OS in the tGE population
- PFS, as determined by the investigator according to RECIST v1.1, and OS in the TC2/3 or IC2/3 population and the TC1/2/3 or IC1/2/3 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 population and ITT population
- DOR, defined as the time interval from the 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 population and ITT population
- OS rates at 1 and 2 years for the tGE population and ITT 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 (cough, dyspnea [single item], dyspnea [multi-item subscale], chest pain, and arm/shoulder pain) in the tGE population and ITT population
- Change from baseline in patient-reported lung cancer symptoms (cough, dyspnea, and chest pain) on the symptom severity score of the SILC scale in the tGE population and ITT population

- PFS, as determined by the investigator according to RECIST v1.1, and OS in the two atezolizumab-containing arms in the tGE population and the ITT 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 study treatment 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 (Arm A)
- Plasma concentrations for nab-paclitaxel (Arms B and C)

Exploratory Outcome Measures

The exploratory outcome measures for this study are the following:

- 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)

The investigational medicinal products (IMPs) for this study are atezolizumab and erlotinib.

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

Switch maintenance treatment with erlotinib is no longer permitted. However, patients who had already started switch maintenance treatment with erlotinib under previous protocol versions may be allowed to continue treatment with erlotinib upon discussion of the risks, potential benefits and alternative treatment options with the investigator. For those patients who receive switch maintenance, institutions should follow the dosage and administration instructions in the prescribing information

Appendix 1 Protocol Synopsis (cont.)

.

Comparator

Nab-paclitaxel (100 mg/m^2 IV) will be administered on Days 1, 8, and 15 of each 21-day cycle for four or six cycles during the induction phase. Nab-paclitaxel will be considered an investigational medicinal product (IMP) for study purposes in countries where nab-paclitaxel is considered an IMP by local regulations.

Non-Investigational Medicinal Products

Comparator

- Carboplatin will be administered by IV infusion to achieve an initial target area under the concentration–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 randomized to Arm A. Carboplatin and nab-paclitaxel will be administered to patients in Arms B and C.

Statistical Methods

Primary Analysis

The co-primary efficacy endpoints are PFS as assessed by the investigator using RECIST v1.1, and OS. PFS will be analyzed in the tGE population and the ITT population. OS will be analyzed in the ITT population. Treatment comparisons will be conducted sequentially by first comparing Arm B versus Arm C and then comparing Arm A versus Arm C. 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 and to account for an interim analysis.

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 time 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 hazard ratio (HR) for PFS and OS for each comparison (i.e., Arm B vs. Arm C, Arm A 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) of 0.003 and 0.022 (ratio of 3:22) will be allocated to PFS and OS, respectively, for the comparison of Arm B versus Arm C.

1. PFS in the tGE 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 + nab-paclitaxel prolongs the duration of PFS relative to control treatment in the tGE population.
2. If the null hypothesis is rejected in the tGE population (Step 1), PFS in the ITT 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 null hypothesis is not rejected for PFS analysis in the ITT population (Step 2), OS in the ITT population will be tested at $\alpha = 0.022$ (one-sided).
   b. If the null hypothesis is rejected for PFS analysis in the ITT population (Step 2), OS in the ITT 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 population is statistically significant at an α of 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 and OS at the same 3:22 ratio (see Figure 3 in Section 6.1). If the difference in OS between Arm B and Arm C in the ITT 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.

1. 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 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 control treatment in the tGE population.
   b. If the null hypothesis is rejected in the tGE population (Step 1a), PFS in the ITT population will be tested at $\alpha = 0.00264$ (one sided).
   c. α recycling from PFS to OS will be conducted as follows:
      - If the null hypothesis is not rejected in the ITT population (Step 1b), OS in the ITT population will be tested at $\alpha = 0.01936$ (one sided).
      - If the null hypothesis is rejected in the ITT population (Step 1b), OS in the ITT population will be tested at $\alpha = 0.022$ (one sided).
2. 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 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 control treatment in the tGE population.
   b. If the null hypothesis is rejected in the tGE population (Step 2a), PFS in the ITT population will be tested at $\alpha = 0.003$ (one sided).
   c. α recycling from PFS to OS will be conducted as follows:
      - If the null hypothesis is not rejected in the ITT population (Step 2b), OS in the ITT population will be tested at $\alpha = 0.022$ (one sided).
      - If the null hypothesis is rejected in the ITT population (Step 2b), OS in the ITT 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/web response system.

Results from an unstratified analysis will also be presented. Kaplan-Meier methodology will be used to estimate median PFS and 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 randomize approximately 1025 patients, including approximately 513 patients in the tGE population, 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) for the comparison of Arm B versus Arm C.

The estimate of the number of events required to demonstrate efficacy with regard to PFS in the comparison of Arm B versus Arm C 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 population
- One-sided significance level of 0.003 for the comparison of Arm B versus Arm C in the ITT population
- 97% 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 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 population
- No interim analyses for PFS
- Dropout rate of 5% per 12 months

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

- One-sided significance level of 0.022 for the comparison of Arm B versus Arm C in the ITT population
- 86% power to detect an HR of 0.75, corresponding to an improvement in median OS from 12 months to 16 months in the ITT population
- One interim OS analysis performed at the time of the final PFS analysis, at which time approximately 76% 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 24 months

With these assumptions, approximately 1025 patients in total will be enrolled into this study, with approximately 684 patients in each comparison (i.e., Arm B vs. Arm C, Arm A vs. Arm C). The final PFS analysis will be conducted when all of the following criteria have been met: approximately 240 PFS events have occurred in Arms B and C combined in the tGE population, approximately 499 PFS events have occurred in Arms B and C combined in the ITT population, and the last patient has been enrolled in the study. The final PFS analysis is expected to occur approximately 30 months after the first patient is enrolled. These numbers of events would allow for a minimum detectable difference corresponding to an HR of approximately 0.70 in the tGE population and approximately 0.75 in the ITT population.

With a sample size of 684 patients, approximately 477 OS events are expected to occur in Arms B and C combined in the ITT population for the final OS analysis. The final OS analysis is expected to occur approximately 39 months after the first patient is enrolled.

Interim Analyses

There are 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. 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/Ethics Committees. A detailed plan will be included in the iDMC Charter.

If approximately 365 OS events have occurred in Arms B and C combined in the ITT population at the time of the final PFS analysis (see criteria for final PFS analysis in Section 6.1), an interim OS analysis will be conducted for Arm B versus Arm C in the ITT population. If there are significantly fewer than the expected 365 OS events at 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 365 OS events have occurred.

The final OS analysis for the comparison of Arm B versus Arm C will be conducted when approximately 447 OS events have occurred in Arms B and C combined in the ITT population. This is expected to occur approximately 39 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		
	Days –28 to –1	Every 21 days (±3 Days) [b]	Every 21 days (±3 Days)	≤30 Days after Last Dose of Study Treatment	Every 3 Months after Disease Progression or Loss of Clinical Benefit
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				
Demographic data	x				
Medical history and baseline conditions	x				
NSCLC cancer history	x				
Vital signs [d]	x	x	x	x	
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]	
Hematology [g]	x	x	x	x	

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		
	Days –28 to –1	Every 21 days (±3 Days) [b]	Every 21 days (±3 Days)	≤30 Days after Last Dose of Study Treatment	Every 3 Months after Disease Progression or Loss of Clinical Benefit
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 [y]
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 + nab-paclitaxel Arm C: Carboplatin + nab-paclitaxel)		x [n]			
Maintenance treatment administration Arm A: Atezolizumab Arm B: Atezolizumab Arm C: Best supportive care [o]			x [n]		
Tumor response assessment	x [p]	x [q]	x [q]		x [r]
Serum sample for ATA assessment		x	x	x	120 (±30) days after

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	≤30 Days after Last Dose of Study Treatment	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)		
(atezolizumab-treated patients only) [s]					last dose
Serum sample for PK sampling (atezolizumab-treated patients only) [s]		x	x	x	120 (±30) days after last dose
Carboplatin, paclitaxel, and nab-paclitaxel PK sampling as applicable (20 patients per arm) [s]		x			
Blood samples for PD biomarkers [s]	x	x	x	x	120 (±30) days after last dose
Optional blood for DNA extraction (RCR only) [s, t]		x			
Informed consent to continue treatment beyond radiographic progression (atezolizumab-treated patients only)		At time of radiographic progression			
Tumor biopsy		At time of radiographic progression [u]			
Optional tumor biopsy at other timepoints (RCR only)		Any time during study treatment or during survival follow-up			
Adverse events [v]	x	x	x	x [v]	
Concomitant medications	From 7 days before screening	x	x	x	
Patient-reported outcomes [w]		x [w]	x [w]		x [w]

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	≤30 Days after Last Dose of Study Treatment	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)		
Survival and anti-cancer therapy follow-up [x]					x

ATA = anti-therapeutic antibody; BSC = best supportive care; 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; ICF = Informed Consent Form; IV = intravenous; QLQ-LC13 = Lung Cancer module; NSCLC = non–small cell lung cancer; PD = pharmacodynamic; PD-L1 = programmed death–ligand 1; PGIS = Patient Global Impression of Severity; PK = pharmacokinetic; QLQ-C30 = Quality-of-Life Questionnaire Core 30; RECIST v1.1 = Response Evaluation Criteria in Solid Tumors Version 1.1 ; SILC = Symptoms in Lung Cancer; EQ-5D-3 Level version; TSH = thyroid-stimulating hormone.

Note: For patients who are randomized to Arm C who receive erlotinib switch maintenance allowed under previous protocol versions, patients will continue with cycle visits at the same frequency as those patients randomized to Arms A and B until disease progression, withdrawal of consent, death, or study termination by the Sponsor, whichever occurs first.

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

[b] Except for Cycle 1, Day 1, which must be performed within 5 days after the patient is randomized. In addition, ECOG performance status, limited physical examination, and local laboratory tests may be performed ≤96 hours before Day 1 of each cycle as specified in Section 4.5.12.2 of Protocol GO29437. Screening assessments performed ≤96 hours before Cycle 1, Day 1, do not need to be repeated for Cycle 1 Day 1.

[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 are not available for PD-L1 testing, contact the Medical Monitor to discuss to determine if the patient may participate in the study. Fine-needle aspiration (defined as samples that do not preserve tissue architecture and yield cell suspension and/or cell smears), brushing, cell pellets from pleural effusion, and lavage samples are NOT acceptable. For core needle biopsy specimens, at least three cores should be submitted for evaluation. Retrieval of archival tumor sample may occur outside the 28-day screening period prior to enrollment. See Section 4.5.7.1 and the inclusion criteria in Section 4.1.1 of Protocol GO29437.

[d] Vital signs include pulse rate, respiratory rate, blood pressures, and temperature. Vital signs should be recorded as described in Section 4.5.4 of Protocol GO29437 sites in Argentina, pulse oximetry will be performed at every visit; these data will not be recorded.

[e] Symptom-directed physical examinations; see Section 4.5.3 of Protocol GO29437 for details.

[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. Hematology tests must be performed prior to Day 1 infusions and for nab-paclitaxel administration, also prior to Day 8 and Day 15 infusions. See Section 5.1.10.1 of Protocol GO29437 for dose modifications due to hematological toxicities.

[h] Serum chemistry includes BUN, creatinine, sodium, potassium, magnesium, chloride, bicarbonate or total CO_2, calcium, phosphorus, glucose, total bilirubin, ALT, AST, alkaline phosphatase, LDH, total protein, and albumin. See Section 5.1.10.6 of Protocol GO29437 for dose modifications due to serum chemistry toxicities.

[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 to be tested only at sites where free T3 testing cannot be 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 ($\pm$15) minutes. If the first infusion is well tolerated, all subsequent infusions may be delivered over 30 ($\pm$10) minutes. For carboplatin, paclitaxel, and nab-paclitaxel, study drug will be administered according to the local prescribing information, including premedication with steroids (see Section 4.3.2 of Protocol GO29437).

[o] Only patients who are already receiving switch maintenance therapy with erlotinib, permitted under previous protocol versions, may be allowed to remain on treatment with erlotinib after discussion of the risks, potential benefits, and alternative treatment options with the investigator.

[p] CT scans (with oral/IV contrast unless contraindicated) or MRI of the chest and abdomen. A CT or MRI scan of the pelvis is required at screening

and as clinically indicated or as per local standard of care at subsequent response evaluations. A CT (with contrast) or MRI scan of the head must be done at screening to evaluate CNS metastasis in all patients. See Section 4.5.5 of Protocol GO29437 for details.

[q] Perform every 6 weeks (±7 days) (approximately every two cycles) for 48 weeks following Cycle 1, Day 1, and then, after completion of the Week 48 tumor assessment, every 9 weeks (±7 days) regardless of treatment delays, until radiographic disease progression (loss of clinical benefit for patients assigned to atezolizumab who continue treatment after disease progression according to RECIST v1.1), withdrawal of consent, death, or study termination by the Sponsor, whichever occurs first. See Section 4.5.5 of Protocol GO29437 for details.

[r] 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 (loss of clinical benefit for patients treated with atezolizumab who continue treatment after 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.

[s] See Appendix 2 of Protocol GO29437 for the detailed schedule.

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

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

[v] All serious adverse events and adverse events of special interest, regardless of relationship to study drug, will be reported until 90 days after the last dose of study treatment (inclusive of erlotinib switch maintenance for patients who are randomized to Arm C who receive switch maintenance allowed under previous protocol versions) or initiation of new systemic anti-cancer therapy after last dose of study treatment. All other adverse events, regardless of relationship to study drug, will be reported until 30 days after the last dose of study treatment (inclusive of erlotinib switch maintenance for patients who are randomized to Arm C who receive switch maintenance allowed under previous protocol versions) or initiation of new systemic anti-cancer therapy after last dose of study treatment. After this period, all deaths should continue to be reported. In addition, the Sponsor should be notified if the investigator becomes aware of any serious adverse event or adverse event of special interest that is believed to be related to prior exposure to study treatment (see Section 5.6 of protocol GO29437).

[w] 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 cycle visit during the induction period and then according to the tumor assessment schedule during the treatment maintenance phase prior to administration of study drug and prior to any other study assessment(s). SILC will be completed using an ePRO 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. The SILC would be completed monthly during survival follow-up for 6 months following disease progression or loss of clinical benefit for patients treated with atezolizumab who continue after 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 patient withdraws consent or the Sponsor terminates the study). Study personnel should review all questionnaires for completeness before the

patient leaves the investigational site. Patients whose native language is not available in the ePRO device or who are deemed by the investigator incapable of inputting their ePRO assessment after undergoing appropriate training are exempt from all ePRO assessments.

[x] 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.

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

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

Study Visit	Time	Arm A (Atezolizumab + Carboplatin + Paclitaxel)	Arm B (Atezolizumab + Carboplatin + Nab-Paclitaxel)	Arm C (Carboplatin + Nab-Paclitaxel)
Screening	N/A	Biomarkers [b]	Biomarkers [b]	Biomarkers [b]
Cycle 1, Day 1[e]	Predose (same day as treatment administration)	Atezolizumab ATA Atezolizumab pharmacokinetics Carboplatin pharmacokinetics [a] Paclitaxel pharmacokinetics [a] Biomarkers [d]	Atezolizumab ATA Atezolizumab pharmacokinetics Carboplatin pharmacokinetics [a] Nab-paclitaxel pharmacokinetics [a] Biomarkers [d]	Carboplatin pharmacokinetics [a] Nab-paclitaxel pharmacokinetics [a] Biomarkers [d]
	30 min (±10 min) after end of atezolizumab infusion	Atezolizumab pharmacokinetics	Atezolizumab pharmacokinetics	
	5 – 10 min before the end of paclitaxel infusion [a]	Paclitaxel pharmacokinetics [a]		
	1 hour after the end of paclitaxel infusion [a]	Paclitaxel pharmacokinetics [a]		
	5 – 10 min before the end of nab-paclitaxel infusion [a]		Nab-Paclitaxel pharmacokinetics [a]	Nab-Paclitaxel pharmacokinetics [a]
	1 hour after the end of nab-paclitaxel infusion [a]		Nab-Paclitaxel pharmacokinetics [a]	Nab-Paclitaxel pharmacokinetics [a]
	5 – 10 min before the end of carboplatin infusion [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]
	1 hour after the end of carboplatin infusion [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]
Cycle 2, Day 1 (±3 days)	Predose (same day as treatment administration)	Atezolizumab ATA Atezolizumab pharmacokinetics Biomarkers [d]	Atezolizumab ATA Atezolizumab pharmacokinetics Biomarkers [d]	Biomarkers [d]

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

Study Visit	Time	Arm A (Atezolizumab + Carboplatin + Paclitaxel)	Arm B (Atezolizumab + Carboplatin + Nab-Paclitaxel)	Arm C (Carboplatin + Nab-Paclitaxel)
Cycle 3, Day 1 (±3 days)	Predose (same day as treatment administration)	Atezolizumab ATA Atezolizumab pharmacokinetics Carboplatin pharmacokinetics [a] Paclitaxel pharmacokinetics [a] Biomarkers [d]	Atezolizumab ATA Atezolizumab pharmacokinetics Carboplatin pharmacokinetics [a] Nab-Paclitaxel pharmacokinetics [a] Biomarkers [d]	Biomarkers [d] Carboplatin pharmacokinetics [a] Nab-Paclitaxel pharmacokinetics [a] Biomarkers [d]
	30 min (±10 min) after end of atezolizumab infusion	Atezolizumab pharmacokinetics	Atezolizumab pharmacokinetics	
	5 – 10 min before the end of paclitaxel infusion [a]	Paclitaxel pharmacokinetics [a]		
	1 hour after the end of paclitaxel infusion [a]	Paclitaxel pharmacokinetics [a]		
	5 – 10 min before the end of nab-paclitaxel infusion [a]		Nab-Paclitaxel pharmacokinetics [a]	Nab-Paclitaxel pharmacokinetics [a]
	1 hour after the end of nab-paclitaxel infusion [a]		Nab-Paclitaxel pharmacokinetics [a]	Nab-Paclitaxel pharmacokinetics [a]
	5 – 10 min before the end of carboplatin infusion [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]
	1 hour after the end of carboplatin infusion [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]	Carboplatin pharmacokinetics [a]
Cycles 4, 8, and 16, Day 1 (±3 days)	Predose (same day as treatment administration)	Atezolizumab ATA Atezolizumab pharmacokinetics Biomarkers [d]	Atezolizumab ATA Atezolizumab pharmacokinetics Biomarkers [d]	Biomarkers [d]

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

Study Visit	Time	Arm A (Atezolizumab + Carboplatin + Paclitaxel)	Arm B (Atezolizumab + Carboplatin + Nab-Paclitaxel)	Arm C (Carboplatin + Nab-Paclitaxel)
After Cycle 16, every eighth cycle, Day 1 (±3 days)	Predose (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]
Treatment discontinuation visit	At visit	Atezolizumab ATA Atezolizumab pharmacokinetics Biomarkers [d]	Atezolizumab ATA Atezolizumab pharmacokinetics Biomarkers [d]	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.

Note: Serum PK samples for atezolizumab and plasma PK samples for carboplatin, paclitaxel, and nab-paclitaxel.

[a] Twenty patients in each treatment arm will undergo the additional PK assessments for carboplatin, paclitaxel, and nab-paclitaxel, where applicable.

[b] Whole blood for biomarkers.

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

[d] Plasma and serum for biomarkers.

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