

Official Title: A Phase III, Double-Blind, Multicenter, Randomized Study of Atezolizumab (Anti-PD-L1 Antibody) Versus Placebo as Adjuvant Therapy in Patients with High-Risk Muscle-Invasive Bladder Cancer who are ctDNA Positive Following Cystectomy

NCT Number: NCT04660344

Document Date: SAP Version 2: 11-Nov-2024

STATISTICAL ANALYSIS PLAN

STUDY TITLE: A PHASE III, DOUBLE-BLIND, MULTICENTER, RANDOMIZED STUDY OF ATEZOLIZUMAB (ANTI-PD-L1 ANTIBODY) VERSUS PLACEBO AS ADJUVANT THERAPY IN PATIENTS WITH HIGH-RISK MUSCLE-INVASIVE BLADDER CANCER WHO ARE ctDNA POSITIVE FOLLOWING CYSTECTOMY

STUDY NUMBER: BO42843

STUDY NAME: IMvigor011

VERSION NUMBER: 2

ROCHE COMPOUND: Atezolizumab (MPDL3280A; RO5541267)

EUDRACT NUMBER: 2020-004418-36

IND NUMBER: 120,827

NCT NUMBER: NCT04660344

PLAN PREPARED BY: [REDACTED] (Ph.D)

STATISTICAL ANALYSIS PLAN APPROVAL

SPONSOR: F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS: Grenzacherstrasse 124
4070 Basel, Switzerland

DATE FINAL: See electronic date stamp on the last page of this document

CONFIDENTIAL

This is an F. Hoffmann-La Roche Ltd document that contains confidential information. Nothing herein is to be disclosed without written consent from F. Hoffmann-La Roche Ltd.

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This Statistical Analysis Plan (SAP) was developed based on Roche SAP model document version (V) 2.0.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
1	see electronic date stamp on the last page	V8.0, 22 March 2024

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Key changes to the SAP as implemented in V2, along with the rationale(s) for each change, are summarized below.

Section	Description of Change	Rationale for Change
4.7.2	The interim analysis of OS (i.e., at the time of the primary analysis of DFS) will be formally tested regardless of the number of OS events observed; p-value boundaries for OS analyses are specified when the number of events is fewer than 91.	Change is being made to accommodate the FDA's request to formally test OS at the interim analysis of OS

TABLE OF CONTENTS

1.	INTRODUCTION.....	9
1.1	Objectives and Endpoints	9
1.2	Study Design	11
1.2.1	Treatment Assignment and Blinding	12
1.2.1.1	Treatment Assignment.....	12
1.2.1.2	Blinding	13
1.2.2	Independent Review Facility	13
1.2.3	Data Monitoring	13
2.	STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATION	14
2.1	Statistical Hypotheses	14
2.2	Type 1 ERROR CONTROL	14
2.3	Sample Size Determination	15
2.4	Investigator-Assessed Disease-Free Survival	15
2.5	Overall survival	15
3.	ANALYSIS SETS	16
3.1	Enrolled POPulation.....	16
3.2	Efficacy Analysis Population	16
3.3	ctDNA-negative Population.....	16
3.4	Safety Evaluable Population	16
3.5	Pharmacokinetic Evaluable Population.....	16
3.6	ADA-Evaluable Population.....	17
4.	STATISTICAL ANALYSES	17
4.1	General Considerations	17
4.2	Primary Endpoint Analysis	17
4.2.1	Definition of Primary Endpoint	17
4.2.2	Main Analytical Approach for Primary Endpoint.....	17
4.2.3	Sensitivity Analyses	18
4.2.3.1	Sensitivity Analyses for Patients in any Treatment Arm Receiving Protocol Prohibited Anti-cancer Therapy.....	18

4.2.3.2	Sensitivity Analyses for Patients Missing Two or More Consecutive Assessment Prior to the DFS Event.....	18
4.2.4	Supplementary Analyses	18
4.2.4.1	Subgroup Analyses for Primary Endpoint	18
4.2.4.2	Analyses at Selected Timepoints	19
4.3	Secondary Endpoint Analysis	19
4.3.1	Confirmatory Secondary Endpoint	19
4.3.2	Supportive Secondary Endpoints.....	19
4.3.2.1	Independent Review Facility-Assessed Disease-Free Survival	19
4.3.2.2	Disease-Specific Survival	19
4.3.2.3	Distant Metastasis-Free Survival	20
4.3.2.4	Time to Confirmed Deterioration of Function and Quality of Life – EORTC Data	20
4.3.2.5	ctDNA Clearance	21
4.4	Exploratory Analysis	21
4.4.1	Descriptive EORTC QLQ-C30 Summaries	21
4.4.2	Efficacy in ctDNA-Negative Population	21
4.5	Safety Analyses	22
4.5.1	Extent of Exposure	22
4.5.2	Adverse Events.....	22
4.5.3	Additional Safety Assessments.....	23
4.6	Other Analyses	23
4.6.1	Summaries of Conduct of Study	23
4.6.2	Summaries of Demographics and Baseline Characteristics	23
4.6.3	Pharmacokinetic Analyses.....	24
4.6.4	Immunogenicity Analyses	24
4.6.5	Biomarker Analyses.....	24
4.7	Interim Analyses	25
4.7.1	Disease-Free Survival	25
4.7.2	Overall Survival.....	25
5.	REFERENCES.....	25

LIST OF TABLES

Table 1	Primary and Secondary Objectives	9
Table 2	Safety Objectives	10
Table 3	Other Objectives and Endpoints	10

LIST OF FIGURES

Figure 1	Study Schema.....	11
Figure 2	Type I Error Control and Analysis Hierarchy	14

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
ADA	anti-drug antibody
AE	adverse event
AESI	adverse events of special interest
CDF	cumulative distribution function
CI	confidence interval
CSR	Clinical Study Report
ctDNA	circulating tumor deoxyribonucleic acid
DFS	disease-free survival
DSS	disease-specific survival
DMFS	distant metastasis-free survival
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC QLQ C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30
GHS/QoL	global health status/quality of life
HR	hazard ratio
IA	interim analysis
iDMC	independent Data Monitoring Committee
IHC	immunohistochemistry
IRC	Independent Review Committee
IRF	Independent Review Facility
IV	intravenous
IxRS	interactive voice/web-based response system
MDD	minimally detectable difference
MedDRA	Medical Dictionary for Regulatory Activities
MIBC	muscle-invasive bladder cancer
NAC	neoadjuvant chemotherapy
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OS	overall survival
PD-L1	programmed death-ligand 1
PK	pharmacokinetic
PRO	patient-reported outcomes
Q4W	every 4 weeks
RF	role function
SAE	serious adverse event

SAP Statistical Analysis Plan
UC urothelial carcinoma

1. **INTRODUCTION**

The analyses described in this Statistical Analysis Plan (SAP) will supersede those specified in Protocol BO42843 for the purposes of a regulatory filing. It is anticipated that positive results from Study BO42843 will support the submission of marketing applications globally for the use of atezolizumab as adjuvant therapy in patients with high-risk muscle-invasive bladder cancer (MIBC) who are circulating tumor deoxyribonucleic acid (ctDNA) positive following cystectomy.

1.1 **OBJECTIVES AND ENDPOINTS**

This study will evaluate the efficacy, safety, and pharmacokinetics of atezolizumab compared with placebo in patients with high-risk MIBC who are ctDNA positive, as determined by the investigational Signatera™ (Natera, Inc.) assay (RD0827369), following cystectomy. The Signatera test received Breakthrough Device designation (Q202123/S002) and will be validated as a companion diagnostic using the data from this study. Specific objectives and corresponding endpoints for the study are outlined below.

Table 1 Primary and Secondary Objectives

Primary Objective	Corresponding Endpoint
<u>Efficacy Objective:</u> Evaluate the efficacy of atezolizumab compared with placebo	- Investigator-assessed DFS, defined as the time from randomization to the first occurrence of a DFS event, defined as any of the following: - Local (pelvic) recurrence of UC (including soft tissue and regional lymph nodes) - Urinary tract recurrence of UC (including all pathological stages and grades) - Distant metastasis of UC - Death from any cause
Secondary Objectives	Corresponding Endpoints
<u>Efficacy Objective:</u> Evaluate the efficacy of atezolizumab compared with placebo	- OS, defined as the time from randomization to death from any cause - Independent Review Facility (IRF)-assessed DFS - Investigator-assessed DSS), defined as the time from randomization to death from UC per investigator assessment of cause of death - Investigator-assessed DMFS), defined as the time from randomization to the diagnosis of distant (i.e., non-locoregional) metastases or death from any cause - Time to confirmed deterioration of function and health-related quality of life (HRQoL), defined as the time from randomization to the date of a patient's first score decrease of ≥ 10 points from baseline on the EORTC QLQ-C30 physical function scale, role

	function scale, and the global health status/quality of life (GHS/QoL) scale (separately), held for at least two consecutive time points or followed by death ctDNA clearance, defined as the proportion of patients who are ctDNA positive at baseline and ctDNA negative at Cycle 3, Day 1 or Cycle 5, Day 1
--	---

ctDNA = circulating tumor deoxyribonucleic acid; DFS = Disease-free survival; DMFS = distant metastasis-free survival; DSS = disease-specific survival; EORTC QLQ30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; OS = overall survival; UC = urothelial carcinoma.

Table 2 Safety Objectives

<u>Safety Objective:</u> Evaluate the safety of atezolizumab compared with placebo	- Incidence and severity of AE, with severity determined according to National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events, Version 5.0 (NCI CTCAE v5.0) - Change from baseline in targeted vital signs - Change from baseline in targeted clinical laboratory test results
<u>Exploratory Safety Objective:</u> Evaluate the safety of atezolizumab compared with placebo	Overall patient-reported AE burden over time, as assessed by the patient global impression of cancer treatment side effects burden item from the EORTC Item Library 46 (IL46)

AE = adverse event; DFS = Disease-free survival; DMFS = distant metastasis-free survival; DSS = disease-specific survival; EORTC = European Organisation for Research and Treatment of Cancer.

Table 3 Other Objectives and Endpoints

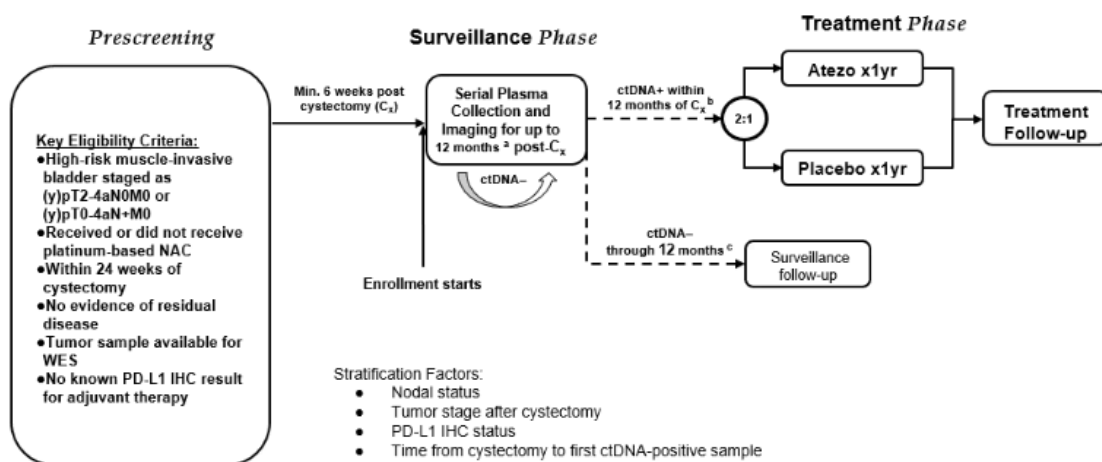
Other Objectives	Corresponding Endpoints
<u>Exploratory Objective:</u> Evaluate the efficacy of atezolizumab compared with placebo	- Cumulative distribution function (CDF) curves of score change from baseline to Cycle 7 for the EORTC QLQ-C30 Physical Function, Role Function, GHS/QoL scales - Descriptive score and score change from baseline summaries for each EORTC QLQ-C30 scale - OS and investigator-assessed DFS in the enrolled population that remains ctDNA negative
<u>Pharmacokinetic Objective:</u> Characterize the pharmacokinetic (PK) profile of atezolizumab	Serum concentration of atezolizumab at specified timepoints
<u>Immunogenicity Objective:</u> Evaluate the immune response to atezolizumab	Prevalence of anti-drug antibody (ADAs) to atezolizumab at baseline (baseline prevalence) and incidence of ADAs to Atezolizumab after initiation of study treatment (postbaseline incidence)

ctDNA=Circulating tumor deoxyribonucleic acid; DFS=Disease-free survival; EORTC QLQ30=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; GHS/QoL=global health status/quality of life.

1.2 STUDY DESIGN

This is a global Phase III, randomized, placebo-controlled, double-blind study designed to evaluate the efficacy and safety of adjuvant treatment with atezolizumab compared with placebo in patients with MIBC who are ctDNA positive, as determined by the investigational Signatera (Natera, Inc.) assay, and are at high risk for recurrence following cystectomy. [Figure 1](#) presents an overview of the study design.

Figure 1 Study Schema



atezo=atezolizumab; ctDNA=circulating tumor DNA; ctDNA+ = ctDNA positive; ctDNA- = ctDNA negative; Cx=cystectomy; IHC = immunohistochemistry; Min.=minimum; NAC=neoadjuvant chemotherapy; SOC=standard of care; WES=whole exome sequencing; yr=year.

^a 1 month=1 calendar month; thus 12 months would be equal to 1 calendar year.

^b Patients enters the treatment phase of the study and are randomized to treatment at the first plasma sample that is ctDNA positive provided that they have fully recovered from cystectomy and that there is no evidence of disease recurrence on imaging as confirmed by investigator and Independent Review Facility within 28 days of randomization.

^c Patients who remain ctDNA negative at 12 months from the date of cystectomy will not be randomized to treatment and will enter follow-up.

After screening and 6 weeks (minimum) post-cystectomy, all enrolled patients enter the surveillance phase for imaging and ctDNA testing. Disease-free patients who are ctDNA-positive are randomized to receive 12 cycles or up to 1 year (whichever occurs first) of treatment with either atezolizumab (fixed dose of 1680 mg) or matching placebo. Treatment is administered by intravenous (IV) infusion on Day 1 of each 28-day cycle.

Atezolizumab/placebo will be discontinued in the event of disease recurrence, unacceptable toxicity, withdrawal of consent, or study termination by the Sponsor.

All patients entered in the treatment phase and underwent scheduled assessments for tumor recurrence at baseline and every 9 weeks (± 7 days) in the first year following randomization. Upon completion of the treatment/placebo phase, patients enter treatment follow-up. Disease status assessments for tumor recurrence are performed every 9 weeks (± 7 days) for Year 2; every 12 weeks (± 10 days) for Year 3; every 24 weeks (± 10 days) for Years 4-5; and at Year 6 (approximately 48 weeks after the last assessment in Year 5).

In the absence of an investigator-assessed disease-free survival (DFS) event (defined as any of the following: local [pelvic] recurrence of urothelial carcinoma (UC); urinary tract recurrence of UC; distant metastasis of UC; or death from any cause), monitoring of tumor recurrence during treatment follow-up should continue regardless of whether patients start new anti-cancer therapy, until withdrawal of consent, loss to follow-up, or study termination by the Sponsor, whichever occurs first. Monitoring for tumor recurrence includes physical examination, laboratory evaluation, and imaging studies of the chest, abdomen, upper urinary tracts, and pelvis. Disease recurrence in the treatment phase is as determined by investigators based on radiographic evidence and independently assessed by Independent Review Facility (IRF). A confirmatory tumor biopsy is mandatory if clinically feasible at the time of disease recurrence. Cases for which biopsy results definitively rule out recurrence of UC are not considered as disease recurrence for this study.

1.2.1 Treatment Assignment and Blinding

1.2.1.1 Treatment Assignment

This is a randomized, double-blind study. After written informed consent has been obtained, all screening procedures and assessments have been completed, and eligibility has been established for a patient, the study site will obtain the patient's identification number. Blinded treatment assignment is provided by the interactive voice/web-based response system (IxRS) for patients eligible to be randomized in the treatment phase of the study.

Patients are randomized to one of the following arms:

- Arm A (experimental arm): atezolizumab 1680 mg IV infusion every 4 weeks (Q4W) on Day 1 of each 28-day cycle

- Arm B (control arm): placebo IV infusion Q4W on Day 1 of each 28-day cycle

Randomization occurs in a 2:1 ratio through the use of a stratified permuted-block randomization method. Randomization is stratified by the following factors:

- Nodal status (positive vs. negative)
- Tumor stage after cystectomy ($\leq$ [y]pT2 vs. [y]pT3/pT4)
- Programmed death-ligand 1 (PD-L1) immunohistochemistry (IHC) status (IHC score of IC0/1 vs. IC2/3)
- Time from cystectomy to first ctDNA-positive sample ($\leq$ 20 weeks vs. $>$ 20 weeks)

Patients who remain ctDNA negative through 12 months after cystectomy are not randomized to treatment and enter surveillance follow-up.

Randomization to the treatment phase will end after the projected number of DFS events is expected to be reached. After this timepoint, patients who are still enrolled in the surveillance phase will not be randomized but may continue with ctDNA testing for the remainder of the 12-month surveillance period and may receive treatment outside the study as per the investigator's discretion.

1.2.1.2 Blinding

Study site personnel and patients are blinded to treatment assignment during the study. The Sponsor and its agents are also blinded to treatment assignment, with the exception of individuals who require access to patient treatment assignments to fulfill their job roles during a clinical trial. These roles include the unblinding group responsible, clinical supply chain managers, sample handling staff, operational assay group personnel, IxRS service provider, and independent Data Monitoring Committee (iDMC) members.

1.2.2 Independent Review Facility

An IRF is used for this study. DFS based on the IRF assessments is the secondary endpoint. Specific details, including responsibilities and structure of the Independent Review Committee (IRC) are specified in the IRF Charter.

1.2.3 Data Monitoring

An iDMC evaluates safety data during the study. Sponsor affiliates are excluded from iDMC membership. The iDMC follows a charter that outlines the iDMC roles and responsibilities.

Unblinded safety data are reviewed on a periodic basis as described in the iDMC Charter. All summaries and analyses for the iDMC review are prepared by an independent Data Coordinating Center (iDCC). After reviewing the data, the iDMC provides a recommendation to the Sponsor as described in the iDMC Charter. Final decisions rest with the Sponsor.

2. STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATION

2.1 STATISTICAL HYPOTHESES

The primary efficacy endpoint is investigator-assessed DFS, defined as the time from randomization to the first occurrence of a DFS event. DFS will be compared between treatment arms using the stratified log-rank test. The null and alternative hypotheses can be phrased in terms of the survival functions $S_A(t)$ and $S_B(t)$ in Arm A (atezolizumab) and Arm B (placebo), respectively:

$$H_0: S_A(t) = S_B(t) \text{ versus } H_1: S_A(t) \neq S_B(t)$$

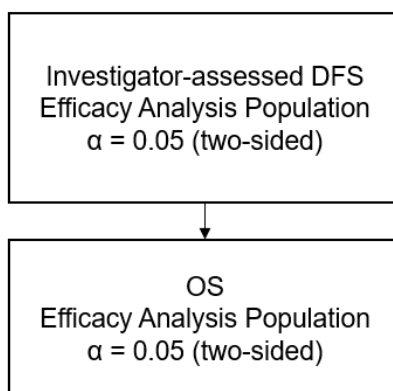
The secondary efficacy endpoint is overall survival (OS), defined as the time from randomization to death from any cause. OS will be compared between treatment arms using the same stratified log-rank test.

2.2 TYPE 1 ERROR CONTROL

The type 1 error (α) for this study is 0.05 (two-sided). Type 1 error will be controlled for the primary endpoint of investigator-assessed DFS and the key secondary endpoint of OS. To control the type 1 error at $\alpha = 0.05$ (two-sided) for the investigator-assessed DFS and OS endpoints, the treatment arms will be compared in a hierarchical fashion as follows (see also [Figure 2](#)):

- Step 1: Investigator-assessed DFS for the efficacy analysis population will be evaluated at $\alpha = 0.05$ (two-sided).
- Step 2: If the investigator-assessed DFS analysis results are statistically significant, $\alpha = 0.05$ will be passed to the analysis of OS (see [Section 4.7.2](#) for more details). If the investigator-assessed DFS results are not statistically significant, formal treatment comparison of OS will not be performed.

Figure 2 Type I Error Control and Analysis Hierarchy



DFS = disease-free survival; OS = overall survival.

2.3 SAMPLE SIZE DETERMINATION

It is projected that up to 800 patients with MIBC will be enrolled in surveillance in order to identify 178 events in approximately 240 ctDNA-positive patients at the time of DFS primary analysis.

The sample size of up to 800 patients for the surveillance population is based on the assumption that approximately 30%-46% of patients will either be ctDNA-positive at the time of enrollment or will become ctDNA positive following cystectomy.

Randomization to the treatment phase will end after the projected number of DFS events is expected to be reached.

2.4 INVESTIGATOR-ASSESSED DISEASE-FREE SURVIVAL

The analysis of the primary endpoint of investigator-assessed DFS in the efficacy analysis population will take place when approximately 178 DFS events have occurred (approximately 74% of the expected 240 patients).

The estimated number of events required for the analysis is based on the following assumptions:

- Two-sided log-rank test at the 0.05 significance level
- 80% power
- 2:1 randomization ratio (atezolizumab to control)
- Median DFS for the control arm of 4.4 months and estimated median DFS in the atezolizumab arm of 6.9 months (corresponding to a hazard ratio [HR] of 0.64)
- 10% probability of drop-out by 24 months in each arm

The required number of investigator-assessed DFS events is projected to occur at Month 40 from the time the first patient is randomized. It is projected that an observed HR of 0.73 or lower will result in a statistically significant difference between treatment arms (i.e., the minimally detectable difference [MDD]) at the analysis; this corresponds to an improvement of 1.6 months in median DFS, from 4.4 months in the control arm to 6.0 months in the atezolizumab arm).

2.5 OVERALL SURVIVAL

The final analysis of the secondary endpoint of OS will take place when approximately 130 deaths have occurred on the basis of the following assumptions:

Two-sided log-rank test at the 0.05 significance level

- 60% power
- 2:1 randomization ratio (atezolizumab to control)
- Median OS for the control arm of 15.4 months and estimated median OS in the atezolizumab arm of 23.3 months (corresponding to a HR of 0.66)

- 10% probability of drop-out by 24 months in each arm
- One interim analysis (IA) for OS (see Section 4.7.2)

On the basis of these assumptions, the required number of OS events is projected to occur at Month 49 from the time the first patient is randomized. It is projected that an observed HR of 0.688 or lower will result in a statistically significant difference between treatment arms (i.e., the MDD at the analysis; this corresponds to an improvement of 7 months in median OS, from 15.4 months in the control arm to 22.4 months in the atezolizumab arm).

3. ANALYSIS SETS

3.1 ENROLLED POPULATION

The enrolled population is defined as patients who enter the surveillance run-in phase of the study, which includes all patients who are randomized to receive study treatment as well as those who are not randomized, including the ctDNA-negative population.

3.2 EFFICACY ANALYSIS POPULATION

The efficacy analysis population includes all patients randomized to receive study treatment regardless of the length of time between cystectomy and ctDNA-positive status, when the primary analysis of DFS is conducted.

3.3 ctDNA-NEGATIVE POPULATION

The ctDNA-negative population is defined as enrolled patients who were

- Disease-free at baseline (i.e., enrolment to surveillance)
- Never had a ctDNA+ result
- Have had at least one ctDNA- result
- Have at least one post-baseline computer tomography scan, biopsy, and/or investigator assessed disease recurrence
- Completed at least 48 weeks of surveillance post-cystectomy or completed / discontinued surveillance prior to 48 weeks without a ctDNA+ result

These patients will not be randomized to treatment. This population is considered to be an exploratory analysis population.

3.4 SAFETY EVALUABLE POPULATION

The safety evaluable population is defined as all randomized patients who received at least one dose of study treatment (atezolizumab or placebo).

3.5 PHARMACOKINETIC EVALUABLE POPULATION

The pharmacokinetic evaluable population is defined as all patients who received any dose of atezolizumab and who have evaluable pharmacokinetic (PK) samples.

3.6 ADA-EVALUABLE POPULATION

The anti-drug antibodies (ADA) evaluable population is defined as all patients treated with atezolizumab who have at least one post-baseline ADA result.

4. STATISTICAL ANALYSES

4.1 GENERAL CONSIDERATIONS

For the efficacy analyses, participants will be analyzed according to the treatment to which they were assigned in the efficacy analysis population; whereas for safety analyses, participants will be included in the analyses according to the treatment they actually received in the safety evaluable population.

4.2 PRIMARY ENDPOINT ANALYSIS

4.2.1 Definition of Primary Endpoint

The primary endpoint is investigator-assessed DFS, as defined in [Table 1](#). Participants without an investigator-assessed DFS event as of the clinical data cutoff date (CCOD) will be censored at the last date the patient was assessed to be alive and recurrence free as determined with radiographic evidence. DFS time will be censored at the time of randomization plus 1 day if no post-randomization disease assessment is performed or if recurrence occurs prior to randomization.

4.2.2 Main Analytical Approach for Primary Endpoint

Investigator-assessed DFS will be analyzed in the efficacy analysis population for the primary analysis. Treatment comparisons will be based on the stratified log-rank test at the two-sided level of significance as described in [Section 2.1](#). The HR, λ_A/λ_B , where λ_A and λ_B represent the hazard of experiencing an DFS event in Arm A and Arm B, respectively, will be estimated using a stratified Cox regression model with the same stratification variables used for the stratified log-rank test, and the 95% confidence interval (CI) for the HR will be provided. Results from an unstratified analysis will also be provided.

Kaplan-Meier methodology will be used to estimate median DFS for each treatment arm and to construct DFS curves for each treatment arm. The Brookmeyer-Crowley methodology will be used to construct the 95% CI for the median DFS for each treatment arm ([Brookmeyer and Crowley 1982](#)).

The stratification factors used for the analysis are described in [Section 1.2.1](#) and the values will be those recorded in the IxRS. Due to the potential risk of over stratification ([Akazawa et al. 1997](#)), only programmed death-ligand 1 (PD-L1) status, nodal status, tumor stage after cystectomy will be used in analysis; and the time from cystectomy to first ctDNA-positive sample will not be used in analyses. This set of stratification factors will be applied to all endpoints where stratified analyses are planned, including stratified log-rank test and stratified Cox regression model, unless otherwise specified.

4.2.3 Sensitivity Analyses

4.2.3.1 Sensitivity Analyses for Patients in any Treatment Arm Receiving Protocol Prohibited Anti-cancer Therapy

The impact of protocol prohibited anti-cancer therapy on DFS may be assessed depending on the number of patients who receive such therapy before a DFS event. If >5% of patients received protocol prohibited anti-cancer therapy before a DFS event in any treatment arm, data for these patients will be censored at the last disease assessment date before they received such therapy in DFS analyses that may be performed for the comparisons between treatment arms.

4.2.3.2 Sensitivity Analyses for Patients Missing Two or More Consecutive Assessment Prior to the DFS Event

Data for patients with a DFS event who missed two or more consecutive scheduled assessments immediately prior to the DFS event will be censored at the last date the patient was assessed to be alive and recurrence free as determined with radiographic evidence prior to the missed visits or at the randomization date if there is no post-baseline tumor assessment prior to the DFS event.

4.2.4 Supplementary Analyses

4.2.4.1 Subgroup Analyses for Primary Endpoint

To assess the consistency of study results in subgroups defined by demographic and baseline characteristics and stratification factors, efficacy outcomes in patient subgroups will be examined.

Summaries of these endpoints, including unstratified HRs and Kaplan-Meier estimates of the median will be produced separately for each level of the categorical variables.

The subgroups to be considered include, but are not limited to, the following:

- Age at randomization (<65 years vs. ≥65 years)
- Race
- Sex
- Region
- PD-L1 IHC status (IHC score of IC0/1 vs. IC2/3) (IxRS)
- Number of lymph nodes resected
- Nodal status (positive, negative) (IxRS)
- Tumor stage after surgery ($\leq$ (y)pT2 vs. (y)pT3/pT4) (IxRS)
- Prior neoadjuvant chemotherapy (NAC)
- Eastern Cooperative Oncology Group (ECOG) Performance Status at baseline
- Time from cystectomy to first ctDNA-positive sample ($\leq$ 20 weeks vs. >20 weeks) and ($\leq$ 12 months vs. >12 months)

4.2.4.2 Analyses at Selected Timepoints

Analyses at selected timepoints will be performed. The rates of investigator-DFS at various timepoints (i.e., every 6 months after randomization) will be estimated by the Kaplan-Meier methodology for each arm and the 95% CI will be calculated using Greenwood's formula. The 95% CIs for the difference in rates between the two arms will be estimated using the normal approximation method.

4.3 SECONDARY ENDPOINT ANALYSIS

4.3.1 Confirmatory Secondary Endpoint

Overall survival is defined as the time from randomization to death from any cause. Data for patients who have not 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.

OS will be analyzed in a hierarchical fashion as described in Section 2.2. Methods for comparison of OS between treatment arms will be the same as the methods for treatment comparison for the primary efficacy endpoint of investigator-assessed DFS. A summary of OS by patient subgroups and analyses at selected timepoints will be provided in the same fashion as for investigator-assessed DFS.

Two analyses of OS are planned (one IA and one final analysis; see Section 4.7.2). The IA of OS will occur at the time that the primary endpoint of DFS is analyzed in the efficacy analysis population.

4.3.2 Supportive Secondary Endpoints

4.3.2.1 Independent Review Facility-Assessed Disease-Free Survival

An IRF assessed DFS is defined as the time from randomization to the first occurrence of a DFS event. See Table 1 for the definition of a DFS event. IRF-assessed DFS will be analyzed in efficacy analysis population. Methods of analysis will be the same as for investigator-assessed DFS.

4.3.2.2 Disease-Specific Survival

Investigator-assessed disease-specific survival (DSS) is defined as the time from randomization to death from bladder cancer. Data for patients who have not died will be censored at the last date they were known to be alive. Data for patients who died from causes other than bladder cancer will be censored at the date of death. Methods for comparison of DSS between treatment arms will be the same as the methods for treatment comparison for the secondary efficacy endpoint of OS. For the purposes of evaluating DSS, for each patient death, the cause of death will be assessed as related or not related to bladder cancer by the investigator. DSS will be analyzed in the efficacy analysis population.

4.3.2.3 Distant Metastasis-Free Survival

Investigator-assessed distant metastasis-free survival (DMFS) is defined as the time from randomization to the first occurrence of a DMFS event, defined as diagnosis of distant (i.e., non-locoregional) metastasis or death from any cause. Data for patients without a DMFS event will be censored at the last date the patient was assessed to be alive and free of distant metastasis. Data for patients with no post-baseline disease assessment will be censored at the randomization date. Methods for comparison of DMFS between treatment arms will be the same as the methods for treatment comparison for the primary efficacy endpoint of investigator-assessed DFS. DMFS will be analyzed in efficacy analysis population.

4.3.2.4 Time to Confirmed Deterioration of Function and Quality of Life – EORTC Data

Time to confirmed deterioration of physical function, role function, and health-related quality of life is defined as the time from randomization to the first documented confirmed deterioration. A confirmed deterioration is defined as a patient's first ≥ 10 -point score decrease from baseline on the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ C30) physical function, role function, or global health status/quality of life (GHS/QoL) scale ([Osoba et al. 1998](#)), that must be held for at least two consecutive time points or followed by death attributable to cancer progression (not death from other causes, including adverse events [AEs]) within 8 weeks from the last patient-reported outcomes (PRO) assessment. Patients without a confirmed deterioration at the time of clinical data cutoff are censored at the last non missing PRO assessment date (regardless of whether there was deterioration at that time).

Additional sensitivity analyses will be performed using time to confirmed deterioration alternative score deterioration thresholds ([Cocks et al. 2012](#)). Confirmed physical function deterioration is defined as an initial ≥ 5 -point score decrease from baseline on the physical function (PF) scale that is held for at least two consecutive time points or followed by death within 8 weeks from the last PRO assessment. Confirmed role function (RF) deterioration is defined as an initial ≥ 7 -point score decrease from baseline on the RF scale that is held for at least two consecutive time points or followed by death within 8 weeks from the last PRO assessment. Confirmed health-related quality of life (HRQoL) deterioration is defined as an initial ≥ 5 -point score decrease from baseline on the GHS/QoL scale that is held for at least two consecutive time points or followed by death within 8 weeks from the last PRO assessment.

Analysis methods for the time to confirmed deterioration endpoints will be the same as those performed for the primary efficacy endpoint. The PRO data will be analyzed in the efficacy analysis population.

4.3.2.5 ctDNA Clearance

ctDNA clearance, defined as the proportion of patients ctDNA positive at baseline and ctDNA negative at Cycle 3, Day 1 or Cycle 5, Day 1, will be analyzed in the efficacy analysis population. An estimate of the proportion of patients with ctDNA clearance and its 95% CI will be calculated using the Clopper-Pearson method for each treatment arm. The CI for the difference in the proportion between the two arms will be determined using the normal approximation to the binomial distribution. The proportions will be compared between the two arms with the use of the stratified Cochran-Mantel-Haenszel test. The stratification factors of this analysis will be the same as those described in Section 4.2.2.

4.4 EXPLORATORY ANALYSIS

4.4.1 Descriptive EORTC QLQ-C30 Summaries

Compliance rates for the EORTC QLQ-C30 in the analysis populations will be calculated as the number of patients who completed the assessment divided by the number of patients expected to complete the assessment at each timepoint for each treatment arm. Reasons for missing assessments, if available, will be summarized using frequencies and percentages.

Descriptive analyses of each EORTC QLQ-C30 scale will include summary statistics (mean, 95% CIs, median, interquartile range, minimum, maximum) of scores and score changes from baseline at each assessment timepoint by treatment arm. Graphs of mean scores and/or score changes from baseline along with two-sided 95% CIs may be presented.

The treatment phase for this study is one year. PRO instruments were administered on Day 1 of each odd-numbered cycle (i.e., Cycles 1, 3, 5, 7, 9, and 11) during the treatment/placebo period. Cycle 7 is after 6 cycles of treatment and is the midpoint. Thus, Cycle 7 is selected for the change from baseline cumulative distribution function (CDF) analysis. CDF plots on the change scores from baseline to Cycle 7 on the physical function, role function, and GHS/QoL scales in reference to participants from each treatment arm will be provided.

4.4.2 Efficacy in ctDNA-Negative Population

Exploratory analyses of investigator-assessed DFS and OS will be performed for descriptive purpose for the ctDNA-negative population, and the relevant subgroups similar to those defined for the primary endpoint (Section 4.2.4.1). DFS/OS is time from cystectomy in these analyses. The statistical details for the analysis will be documented in a separate Biomarker Analysis Plan. These results may be reported separately from the Clinical Study Report (CSR).

4.5 SAFETY ANALYSES

Unless specified otherwise, safety analyses will be conducted for the safety evaluable population, with patients grouped according to whether any atezolizumab treatment was received (i.e., patients who received any dose of atezolizumab will be included in the atezolizumab arm).

4.5.1 Extent of Exposure

Exposure to study treatments, including treatment duration, dose intensity, number of cycles received, total cumulative dose will be summarized with descriptive statistics by treatment group.

4.5.2 Adverse Events

Verbatim descriptions of AEs will be mapped to Medical Dictionary for Regulatory Activities (MedDRA) terms and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 5.0. Adverse events will be summarized by MedDRA term, appropriate MedDRA levels (system organ class [SOC] and preferred term [PT]), and by NCI CTCAE grade if applicable. For each patient, if multiple incidences of the same AE occur, the maximum severity reported will be used in the summaries. Only AEs starting after the first dose of any study treatment will be included in the summary tables of AEs, regardless of relationship to study drug, as assessed by the investigator.

Summary tables by treatment group will include but are not limited to the following:

- AEs
- AEs by highest NCI CTCAE grade
- Severe AEs (Grade $\geq$ 3)
- AEs related to study treatment
- Serious adverse events (SAEs)
- SAEs by highest NCI CTCAE grade
- SAEs related to study treatment
- AEs leading to discontinuation of study treatment
- AEs leading to dose interruption of study treatment
- AEs resulting in death
- All protocol defined adverse events of special interest (AESI)

For AESIs, a set of comprehensive definitions comprising Sponsor-defined, Standardized MedDRA Queries (SMQ), high level terms [HLTs, and Sponsor-defined AE Grouped Terms [AEGTs]) will be used to identify and summarize AESI by medical concept. The medical concepts include atezolizumab-associated important identified risks and potential risks.

Deaths and cause of death will be summarized.

4.5.3 Additional Safety Assessments

Laboratory data (hematology, coagulation, and biochemistry) will be summarized descriptively over time including change from baseline by treatment group.

Laboratory data will be classified according to NCI CTCAE 5 as applicable. The highest NCI CTCAE grade post-baseline will be reported and shift tables from baseline to worst post-baseline will be presented by treatment group. Clinically significant abnormal laboratory data will be reported as AEs and summarized in the AE tables.

For vital signs, summary of abnormalities will be summarized by treatment group, including change from baseline.

The baseline electrocardiogram (ECG) of patients will be summarized and results of on-study ECGs will be listed.

Patient-reported overall AE burden, as measured by the EORTC Item Library (IL) 46 item, will be summarized descriptively at each assessment timepoint. The number and percentage of patients reporting each response option ("Not at all" to "Very much") to the item and the change from baseline by response option will be summarized at each assessment timepoint by treatment arm. The distribution of responses and/or change from baseline will be illustrated by stacked bar charts for each arm.

4.6 OTHER ANALYSES

4.6.1 Summaries of Conduct of Study

Enrollment, major protocol deviations including major deviations of inclusion/exclusion criteria, and reasons for discontinuation from the study will be summarized by treatment arm for efficacy analysis population.

Study treatment administration and reasons for discontinuation from the study treatment will be summarized by treatment arm for treated patients.

4.6.2 Summaries of Demographics and Baseline Characteristics

Demographic variables such as age, sex, and race/ethnicity will be summarized by treatment arm.

Stratification factors (nodal status, tumor stage after cystectomy, PD-L1 IHC status, and time from cystectomy to first ctDNA-positive sample), baseline and disease characteristics (e.g., time since initial diagnosis, time since cystectomy, surgery-related information, type of urinary diversion, ECOG Performance Status, number of lymph nodes resected, and prior NAC) will be summarized by treatment arm for efficacy analysis population.

Continuous variables will be summarized using means, standard deviations, medians, and ranges. Categorical variables will be summarized by proportions. The baseline value of any variable will be defined as the last available value prior to the first administration of study treatment for efficacy analysis population.

4.6.3 Pharmacokinetic Analyses

Atezolizumab serum concentration data (minimum concentration [C_{min}] and maximum concentration [C_{max}]) will be tabulated and summarized for each cycle at which pharmacokinetics are to be measured. Descriptive statistics will include means, medians, ranges, and standard deviations, as appropriate. Exposure by ADA status will also be summarized.

Additional PK and pharmacodynamic analyses will be conducted as appropriate.

4.6.4 Immunogenicity Analyses

The immunogenicity analysis population will consist of all patients with at least one ADA assessment. Patients will be grouped according to treatment received or, if no treatment is received prior to study discontinuation, according to treatment assigned.

The numbers and proportions of ADA-positive patients and ADA-negative patients at baseline (baseline prevalence) and after drug administration (post-baseline incidence) will be summarized by treatment group. When determining post-baseline incidence, patients are considered to be ADA positive if they are ADA negative or have missing data at baseline but develop an ADA response following study drug exposure (treatment-induced ADA response), or if they are ADA positive at baseline and the titer of one or more post-baseline samples is at least 0.60 titer unit greater than the titer of the baseline sample (treatment-enhanced ADA response). Patients are considered to be ADA negative if they are ADA negative or have missing data at baseline and all post-baseline samples are negative, or if they are ADA positive at baseline but do not have any post-baseline samples with a titer that is at least 0.60 titer unit greater than the titer of the baseline sample (treatment unaffected).

The evaluation of PK, efficacy and safety endpoints by ADA status may be analyzed and reported via descriptive statistics.

4.6.5 Biomarker Analyses

Exploratory biomarker analyses will be performed in an effort to understand the association of these markers with study drug response, including efficacy and/or AEs. Planned analyses of biomarkers in this study will be described separately in a biomarker analysis plan. These results may be reported separately from the CSR. Signatera assay performance will be formally evaluated in a separate analysis plan as part of its companion diagnostic submission.

4.7 INTERIM ANALYSES

4.7.1 Disease-Free Survival

No interim efficacy analyses of investigator-assessed DFS are planned in any analysis population.

4.7.2 Overall Survival

A total of two analyses of OS will be performed by the Sponsor (one IA and one final analysis).

The boundaries for statistical significance at the interim and final OS analyses will be determined based on the Lan-DeMets implementation of the O'Brien-Fleming use function ([Lan and DeMets 1983](#)). The projected p-value boundaries for statistical significance stated below are based on the number of projected events; the actual boundary p-values will be calculated at the time of analysis based on the actual number of events observed.

The final analysis of OS will be performed when approximately 130 deaths have occurred (see Section 2.5). If 130 deaths have occurred at the final analysis, statistical significance will be declared if $p \leq 0.0441$ (two-sided).

One IA of OS will be performed at the time of the planned analysis of investigator-assessed DFS in the efficacy analysis population. Based on the OS assumptions in Section 2.5, it is projected that approximately 98 deaths will have occurred at the interim analysis of OS, which corresponds to approximately 75% of the 130 deaths required for the final analysis of OS. If 98 deaths have occurred at this analysis, statistical significance will be declared if $p \leq 0.0197$ (two-sided); an observed HR of 0.607 or lower will result in a statistically significant difference between treatment arms at this analysis.

However, if there are fewer than 91 deaths observed at the time of the DFS analysis, a user-defined alpha spend function will be used with a p-value boundary of 0.0001 at the interim analysis of OS. A p-value boundary of 0.05 will be used for the final analysis of OS.

An iDMC will be convened to evaluate safety results at periodic intervals until the analysis of the primary endpoint (see Section 1.2.3).

5. REFERENCES

Akazawa K, Nakamura T, and Palesch Y. Power of logrank test and Cox regression model in clinical trials with heterogeneous samples. *Statistics in Medicine*, 1997;16:583-97.

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

- Cocks K, King MT, Velikova G et al. Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *European Journal of Cancer*, 2012;48:1713–21.
- Lan KKG and DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika*, 1983;70:659–63.
- Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology*, 1998;16:139–44.

Signature Page for Statistical Analysis Plan - BO42843 - Published

System identifier: RIM-CLIN-806111

Approval Task	Company Signatory 11-Nov-2024 16:41:05 GMT+0000
---------------	--