STATISTICAL ANALYSIS PLAN

STUDY TITLE:	**A PHASE III, MULTICENTER, RANDOMIZED, PLACEBO-CONTROLLED, DOUBLE-BLIND STUDY OF ATEZOLIZUMAB (ANTI-PD-L1 ANTIBODY) AS ADJUVANT THERAPY IN PATIENTS WITH RENAL CELL CARCINOMA AT HIGH RISK OF DEVELOPING METASTASIS FOLLOWING NEPHRECTOMY**
STUDY NUMBER:	WO39210
STUDY NAME:	IMmotion010
VERSION NUMBER:	1
ROCHE COMPOUND(S):	Atezolizumab (RO5541267)
EUDRACT NUMBER:	2016-001881-27
IND NUMBER:	119039
NCT NUMBER:	NCT03024996
PLAN PREPARED BY:	[redacted], Ph.D.

STATISTICAL ANALYSIS PLAN APPROVAL

Date and Time(UTC)	Reason for Signing	Name
15-Nov-2021 15:43:28	Company Signatory	[redacted]

SPONSOR:	F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS:	Grenzacherstrasse 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp above

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This Statistical Analysis Plan (SAP) was developed based on Roche SAP enhanced model document (dated 26 October 2020).

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
1	see electronic date stamp on title page	V 10, 12 November 2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

ADA	anti-drug antibody
AE	adverse event
CHMP	Committee for Medicinal Products for Human Use
COVID-19	Corona virus disease 2019
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DFS	Disease-free survival
DMFS	distant metastasis-free survival
DSS	disease-specific survival
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
FWB	Function/well-being
FKSI-19	Functional Assessment of Cancer Therapy Kidney Symptom Index 19
GP5	GP5 item from the FKSI-19
HR	Hazard ratio
iDMC	independent Data Monitoring Committee
IHC	Immunohistochemistry
IND	Investigational New Drug (application)
IRF	Independent Review Facility
ITT	Intent to treat
IV	Intravenous
IxRS	interactive voice/Web response system
MedDRA	Medical Dictionary for Regulatory Activities
MDD	minimally detectable difference
NCI	National Cancer Institute
NGS	Next generation sequencing
OS	Overall survival
PCR	polymerase chain reaction
PD	pharmacodynamic
PK	Pharmacokinetic
PRO	patient reported outcomes
RCC	Renal Cell Carcinoma
RECIST	Response Evaluation Criteria in Solid Tumours
SAE	serious adverse event
SAP	Statistical Analysis Plan

1. INTRODUCTION

This Statistical Analysis Plan (SAP) describes the analyses that are planned to be performed for the Clinical Study Report (CSR) of WO39210 (IMmotion010). Detailed background information of the study can be found in the study protocol. The analyses described in this SAP will supersede those specified in Protocol WO39210 for the purposes of a regulatory filing.

1.1 OBJECTIVES AND ENDPOINTS

This is a Phase III, multicenter, randomized, placebo–controlled, double–blind study (IMmotion010) to evaluate the efficacy and safety of atezolizumab versus placebo in patients with Renal Cell Carcinoma (RCC) who are at high risk of disease recurrence following resection. Specific objectives and corresponding endpoints for the study are outlined in Table 1.

Table 1 Objectives and Corresponding Endpoints

Objectives	Corresponding Endpoints
Primary Efficacy Objective	
• To evaluate the efficacy of adjuvant treatment with atezolizumab	• Investigator-assessed DFS, defined as the time from randomization to death from any cause or the first documented recurrence assessed by investigator, whichever occurs first. Recurrence is defined as any of the following: Local recurrence of RCC New primary RCC Distant metastasis of RCC
Secondary Efficacy Objective	
• To evaluate the efficacy of adjuvant treatment with atezolizumab	• Overall survival, defined as the time from randomization to death from any cause • Investigator–assessed DFS in patients with PD-L1 expression status IC1/2/3* • IRF-assessed DFS, defined as the time from randomization to death from any cause or the first documented recurrence assessed by IRF, whichever occurs first • IRF-assessed DFS in patients with PD-L1 expression status IC1/2/3* • IRF-assessed event-free survival (EFS), defined as the time from randomization to death from any cause, or the first

Objectives	Corresponding Endpoints
	documented recurrence in patients without baseline disease by IRF or the first documented disease progression in patients identified as having baseline disease by IRF, whichever occurs first. Disease progression is defined as either unequivocal progression of baseline disease or new unequivocal lesion • Disease-specific survival, defined as the time from randomization to death from RCC • Distant metastasis-free survival, defined as the time from randomization to death from any cause or the date of diagnosis of distant (i.e., non locoregional) metastases assessed by investigator, whichever occurs first • 1–, 2–, and 3–year investigator–assessed DFS rate, defined as the probability of patients being alive and free of recurrence assessed by investigator at Year 1, 2, and 3 after randomization • 1–, 2–, and 3–year IRF–assessed DFS rate, defined as the probability of patients being alive and free of recurrence assessed by IRF at Year 1, 2, and 3 after randomization
Safety Objective	
• To evaluate the safety and tolerability of atezolizumab in the adjuvant setting	• Incidence, nature, and severity of adverse events graded according to NCI CTCAE v4.0 • Changes in vital signs and clinical laboratory results
Pharmacokinetic Objective	
• To characterize the PK profile of atezolizumab	• Serum concentration of atezolizumab at specified timepoints
Immunogenicity Objectives	
• To evaluate the immune response to atezolizumab	• Prevalence of anti-drug antibodies (ADAs) to atezolizumab at baseline and incidence of ADAs to atezolizumab after treatment

Objectives	Corresponding Endpoints
• To explore the potential relationship of the immunogenic response	• Relationship between ADA status and PK (i.e., minimum serum concentration), safety, and efficacy endpoints
Exploratory Objectives	
• To assess predictive, prognostic, and pharmacodynamic exploratory biomarkers in archival and/or fresh tumor tissue and blood and their association with disease recurrence and overall survival	• Tumor and circulating biomarkers (including, but not limited to PD-L1, PD–1, prevalence of immune subsets, circulating factors, genomic mutations, gene expression, gene expression signatures of tumor and immune biology, and molecular subtypes), as defined by IHC or qRT-PCR, NGS, and/or other methods will be correlated with efficacy measures
• To document patients' perspective regarding treatment tolerability and health–related quality of life	• The Functional Assessment of Cancer Therapy Kidney Symptom Index 19 (FKSI-19)
• To measure health status for health economic modeling	• EQ–5D–5L questionnaire as a measure of patient reported health status to derive utilities
• To evaluate the efficacy of adjuvant treatment with atezolizumab among patients with tumor Fuhrman Grade 4 or sarcomatoid histology (defined by investigator-assessed conventional histopathology)	• Investigator-assessed DFS in patients with tumor Fuhrman Grade 4 or sarcomatoid histology (defined by the investigator-assessed conventional histopathology)

ADA = anti-drug antibody; DFS = disease-free survival; EFS=event-free survival; EQ-5D-5L = EuroQoL 5-Dimension, 5-level; FKSI-19 = Functional Assessment of Cancer Therapy Kidney Symptom Index 19; IC = tumor-infiltrating immune cell; IHC = immunohistochemistry; IRF = Independent Review Facility; NCI CTCAE v4.0= National Cancer Institute Common Terminology Criteria for Adverse Events Version 4.0; NGS = next generation sequencing; PD-1 = programmed death–1; PD-L1 = programmed death ligand-1; PK = pharmacokinetic; qRT-PCR = quantitative reverse transcriptase–polymerase chain reaction; RCC = renal cell carcinoma.

* PD-L1 IC0 is defined as <1% and IC1/2/3 is defined as ≥1% of tumor-infiltrating ICs expressing PD-L1 as assessed by immunohistochemistry using SP142 assay.

1.2 STUDY DESIGN

Approximately 764 patients at high risk for RCC recurrence (see below for high risk definition) will be randomized to one of the following arms in a 1:1 ratio:

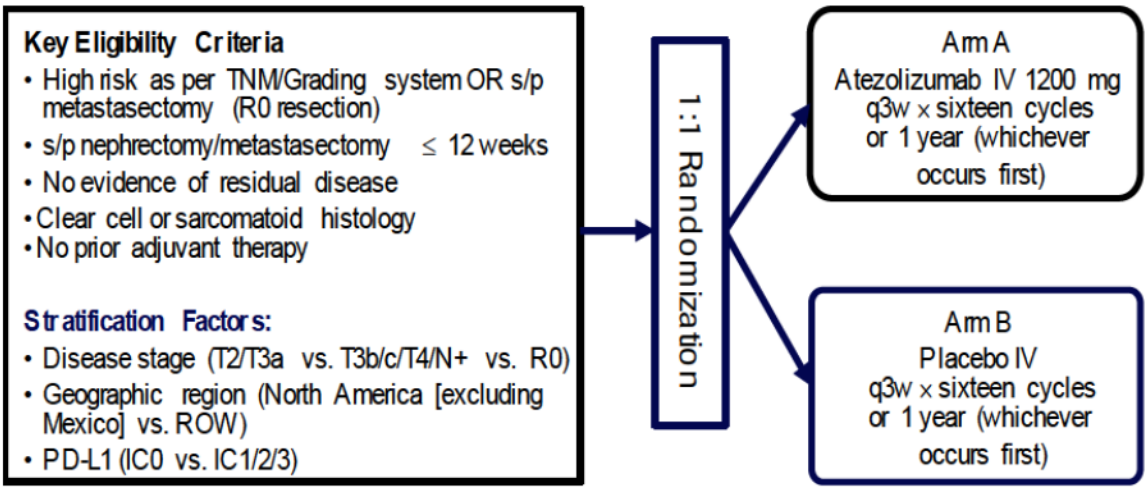
- Arm A (experimental arm): Atezolizumab 1200 mg every 3 weeks (q3w) for 16 cycles or 1 year (whichever occurs first)
- Arm B (control arm): Placebo q3w for 16 cycles or 1 year (whichever occurs first)

High-risk patients are defined as patients with any of the following:

- T2 Grade 4, T3a Grade 3− 4, T3b/c any Grade, T4 any Grade, TxN+ any Grade
- Limited metachronous/synchronous recurrence in patients who undergo complete R0 resection of all sites of disease. These patients will be eligible regardless of their original TNM/grade.

The study schema is shown in Figure 1. For more details of study design, including entry criteria, please refer to the study protocol.

Figure 1 Study Schema



IC=tumor–infiltrating immune cell; IV=intravenous; PD-L1=programmed death ligand–1; q3w=every 3 weeks; s/p=status post; ROW=rest of the world.
Note: TNM/grading: High risk=T2 Grade 4, T3a Grade 3–4, T3b/c any Grade, T4 any Grade, TxN+ any Grade.
PD-L1 IC0 is defined as <1% and IC1/2/3 is defined as ≥1% of ICs expressing PD-L1 as assessed by immunohistochemistry using SP142 assay.

1.2.1 Treatment and Tumor Assessment

This is a randomized, placebo–controlled, double–blind study. The investigator, patient, and Sponsor will be blinded to treatment assignment.

After written informed consent has been obtained and eligibility has been established (including determination of tumor programmed death ligand-1 (PD–L1) expression status by central testing), the study site will enter demographic and baseline characteristics into the interactive voice/web-based response system (IxRS). For those patients who are eligible for enrollment, the study site will obtain the patient's identification number and treatment assignment from the IxRS.

Randomization will be stratified by the following factors:

- Disease stage (T2/T3a vs. T3b/c/T4/N+ vs. patients with resected synchronous/metachronous metastasis [R0])
- Geographic region (North America [excluding Mexico] vs. rest of world)
- PD–L1 IC0 vs IC1/2/3 (<1% vs ≥1% of tumour-infiltrating immune cells (IC) expressing PD-L1 as assessed by immunohistochemistry using SP142 assay)

A stratified, permuted–block randomization will be implemented to balance assignment to each treatment within levels of the stratification factors.

Patients in both treatment arms will receive up to 1 year of treatment with either atezolizumab or placebo. Treatment will be administered intravenously on Day 1 of each 21-day cycle. Treatment will be discontinued in the event of disease recurrence, unacceptable toxicity, consent withdrawal, or study termination by the Sponsor.

Surveillance for tumor recurrence will be performed every 3 months (+/-) 2 weeks for 3 years after randomization. After 3 years, patients will undergo tumor assessment every 6 months (+/-) 4 weeks thereafter until Independent Review Facility (IRF)-assessed disease recurrence, death, consent withdrawal, or study termination by the Sponsor, whichever occurs first. Disease recurrence will be as determined on the basis of radiographic evidence and whenever possible, supported/confirmed by biopsy results. In the absence of IRF-assessed disease recurrence, tumor assessments should continue, regardless of whether patients start new anti-cancer therapy, until death, loss of follow-up, withdrawal of consent, or study termination by the Sponsor, whichever occurs first. Follow-up data capture, including subsequent anti-cancer therapies (including targeted therapies and immunotherapies) and subsequent disease progression, will continue for each patient until death, loss of follow-up, withdrawal of consent, or study termination by the Sponsor, whichever occurs first.

1.2.2 Independent Data Monitoring Committee

An independent Data Monitoring Committee (iDMC) will be convened to evaluate safety data during the study. Members of the iDMC will be external to the Sponsor and will follow an iDMC charter that outlines their roles and responsibilities. The iDMC will evaluate study safety data on a periodic basis, approximately every 6 months, until the time of the analysis of the primary efficacy endpoint. The Sponsor will remain blinded to treatment assignment information until the primary analysis has occurred.

All summaries and analyses by treatment arm for the iDMC review will be prepared by an external independent Data Coordinating Center. The safety data will include demographic data, adverse events, serious adverse events, and relevant laboratory data.

Following their data review, the iDMC will provide a recommendation to the Sponsor in accordance with the iDMC Charter. The final decision will rest with the Sponsor.

Any outcomes of these safety reviews that affect study conduct will be communicated in a timely manner to Health Authorities and to Investigators for notification to their Institutional Review Boards/Ethics Committees.

1.2.3 Independent Review Facility

To support the secondary endpoints, disease recurrence will be as determined by IRF assessment on the basis of radiographic evidence and whenever possible, supported/confirmed by biopsy results. Cases for which biopsy results definitively rule out recurrence of RCC will not be considered as disease recurrence for this study.

All scans should be submitted to the central review facility for independent review. It is important to the integrity of the study that all imaging studies, clinical information (including photographs) and pathology reports for on-study biopsies are forwarded to the IRF before each patient enrolls and as each patient progresses through the study.

1.2.3.1 IRF Tumor Assessment Methodology

An IRF Charter (1st Stream) was implemented to support the independent review of the eligibility and efficacy assessments in the intent to treat (ITT) population. At the time of charter creation in 2016, there were no established imaging guidelines in adjuvant RCC trials like the well-established Response Evaluation Criteria in Solid Tumours (RECIST) guidelines for metastatic disease. Hence, this charter was based on RECIST guidelines where applicable. In 2019, the study team learnt that the efficacy assessment methodology followed by the IRF was not aligned with the intended imaging guidance in the protocol. The main issue was in relation to the global retrospective review process that impacted the accurate assessment of disease-free survival (DFS). Varying degrees of retrospective backdating of recurrence were seen, including backdating the appearance of lesions associated with recurrent disease to a scan in which the lesions had not been prospectively identified and documented. In extreme cases, disease recurrence was backdated to the baseline scans, although almost all of these patients had previously been confirmed by the IRF as being eligible (i.e. disease free) at screening using the same baseline scans. This backdating could potentially underestimate the disease-free survival period and make the results of the study uninterpretable.

After consultations with U.S. FDA and CHMP, a revised IRF Charter (2nd Stream) was developed. In the revised charter, retrospective backdating of recurrence can only be applied to a scan in which a lesion associated with recurrent disease was prospectively identified and documented as an equivocal lesion, as recommended by RECIST guidelines. Moreover, during the efficacy assessment per the revised charter, a patient's baseline scan will be read using the eligibility criteria specified in the protocol to identify baseline disease to confirm the eligibility assessments per the original charter. Lesions in patients who were retrospectively identified by IRF as having baseline disease will be followed with additional lesion growth criteria specified in the Imaging Note To File associated the with 2nd Stream Charter.

The assessment of the secondary endpoint of IRF-assessed DFS will be based on the revised Charter. All scans taken on the study are to be re-read following the methodology specified in the revised charter by a new set of radiology reviewers (i.e. 2nd Stream read). Another secondary endpoint of IRF-assessed event-free survival (EFS) was added to evaluate the treatment benefit for both patients with and without identified baseline disease in the 2nd Stream efficacy review.

All tumor scans will also be read using the original imaging methodology as per the original IRF charter. The efficacy results as per the original charter assessment methodology will be analyzed as part of sensitivity analyses to assess the impact of the revisions to the IRF tumor assessment methodology on IRF-assessed DFS. The detailed analyses are described in Section 5.4.3.1.

2. STATISTICAL HYPOTHESES

The null hypothesis of no difference in investigator-assessed DFS between the two treatment arms in the ITT population will be tested using the stratified log-rank test with type I error as 0.05. The null and alternative hypotheses in terms of the investigator-assessed DFS functions $S_{DFS\_A}(t)$ and $S_{DFS\_B}(t)$ in Arm A (atezolizumab) and Arm B (placebo) are phrased as below, respectively:

$$H0: S_{DFS\_A}(t) = S_{DFS\_B}(t) \text{ versus } H1: S_{DFS\_A}(t) \neq S_{DFS\_B}(t)$$

If statistical significance is achieved in investigator-assessed DFS, then the null hypothesis of no difference in overall survival (OS) between the two arms in the ITT population will be tested using the stratified log-rank test with the allocated type I error at the interim and final analyses (Section 3.1). The null and alternative hypotheses in terms of the survival functions $S_{OS\_A}(t)$ and $S_{OS\_B}(t)$ in Arm A and Arm B are phrased as below, respectively:

$$H_0: S_{OS\_A}(t) = S_{OS\_B}(t) \text{ versus } H_1: S_{OS\_A}(t) \neq S_{OS\_B}(t)$$

3. SAMPLE SIZE DETERMINATION

3.1 TYPE I ERROR CONTROL

Type I error rate for this study is 0.05 (two–sided). To control the type I error rate at alpha = 0.05 (two–sided) for the primary endpoint of investigator-assessed DFS and the key secondary endpoint of OS, the treatment arms will be compared in a hierarchical fashion. The analysis hierarchy will be investigator-assessed DFS followed by OS. One planned analysis of investigator-assessed DFS and a total of three planned analyses of OS (two interim analyses and one final analysis; see Section 5.8 for details) will be performed by the Sponsor. The analysis hierarchy will be implemented as follows:

Step 1: Investigator-assessed DFS will be evaluated at alpha = 0.05 (two–sided). The final analysis will be conducted when approximately 334 DFS events (44% of

764 patients) have occurred. The investigator-assessed DFS endpoint will be considered positive in the ITT population if statistical significance is achieved at the DFS final analysis.

Step 2: If the investigator-assessed DFS results are statistically significant at the DFS final analysis, alpha = 0.05 will be passed to the analysis of OS. At the time of the investigator-assessed DFS final analysis, the first OS interim analysis will be performed and it is projected that approximately 190 deaths will have occurred (25% of 764 patients). The second interim, and final analyses of OS will be conducted when approximately 222, and 254 deaths (29%, and 33% of 764 patients) have occurred, respectively. For control of the familywise error rate at level 0.05, OS will be evaluated on the basis of the generalized Haybittle–Peto boundary (Haybittle 1971) for statistical significance, with alpha boundaries for the two interim and final analyses specified as the following: 0.036, 0.036, and 0.011. Specifically, if at the DFS final analysis, the observed number of OS events is significantly less than the projected 190 events (e.g., < 170 events), a nominal alpha value (1E–05) will be spent at the first OS interim analysis and the team will conduct an additional interim analysis of OS when 190 events have occurred. If the investigator-assessed DFS results are not statistically significant at the DFS final analysis, formal treatment comparison of OS will not be performed.

3.2 INVESTIGATOR-ASSESSED DISEASE-FREE SURVIVAL

The final analysis of the primary endpoint of investigator-assessed DFS will take place when approximately 334 DFS events have occurred (44% of 764 patients) on the basis of the following assumptions:

- Log–rank test (two–sided)
- Type I error rate (alpha) of 0.05 (two–sided)
- 1:1 randomization ratio
- 5% loss to follow–up over 24 months
- Median DFS for the control (placebo) arm of 47 months and estimated median DFS in the atezolizumab arm of 67 months (corresponding to a hazard ratio (HR) of 0.70, under the proportional hazard assumption)
- 90% power

Accrual of the planned 764 patients is projected to occur over 33 months, assuming a ramp–up period of 8 months. On the basis of these assumptions, the required number of DFS events for final analysis is projected to occur at Month 67 after the first patient is randomized; minimum follow–up at the time of the investigator-assessed DFS final analysis will be 34 months. Also on the basis of these assumptions, it is projected that an observed HR of ≤ 0.80 at the final analysis will result in a statistically significant difference (i.e., minimally detectable difference) between treatment arms. An HR of 0.80 corresponds to an improvement of 12 months in median DFS, from 47 months in the control (placebo) arm to 59 months in the atezolizumab arm.

3.3 OVERALL SURVIVAL

The final analysis of the secondary endpoint of OS will take place when approximately 254 deaths have occurred (33% of 764 patients), on the basis of the following assumptions:

- Log–rank test (two–sided)
- Type I error rate (alpha) of 0.05 (two–sided)
- 1:1 randomization ratio
- 2% loss to follow–up over 24 months for OS
- Two interim analyses for OS will take place when approximately 190 and 222 deaths have occurred, respectively
- Alpha boundary is set to be 0.036, 0.036, and 0.011 for the three analyses of OS based on the generalized Haybittle–Peto boundary
- Median OS for the control (placebo) arm of 100 months and estimated median OS in the atezolizumab arm of 143 months (corresponding to HR of 0.70)
- 75% power

On the basis of these assumptions and projected accrual, the required number of OS events for the final analysis of OS is projected to occur at Month 88 from the time the first patient is randomized. Also on the basis of these assumptions, it is projected that an observed HR of $\leq$ 0.73 and 0.75 at the first and second interim analyses, and an HR of $\leq$ 0.72 at the final analysis of OS, respectively, will result in a statistically significant difference (i.e., minimally detectable difference) between treatment arms.

4. ANALYSIS SETS

The following populations are defined:

Population	Definition
ITT	All randomized patients regardless of whether the assigned study treatment was received.
Safety-evaluable	All randomized patients who received any amount of study treatment (i.e., atezolizumab or placebo), regardless of whether a full or partial dose was received. Specifically, for patients randomized to the placebo arm, if atezolizumab was received by mistake, patients will be grouped under the atezolizumab arm in the safety analyses.
PRO-Evaluable	All randomized patients with a non-missing baseline PRO assessment and at least one post-baseline assessment. The PRO-evaluable population will be defined for each questionnaire (i.e., the FKSI-19 evaluable population and the EQ-5D-5L evaluable population).
Immunogenicity Analysis	All randomized patients with at least one ADA assessment for atezolizumab
Post-baseline ADA-Evaluable	All randomized patients who received at least one dose of atezolizumab and with at least one postdose ADA assessment
Pharmacokinetic-Evaluable	All randomized patients who received any dose of study treatment and who have at least one measurable post-baseline PK sample available

ADA = anti-drug antibody; EQ-5D-5L = EuroQoL 5-Dimension, 5-level version; FKSI-19 = Functional Assessment of Cancer Therapy Kidney Symptom Index 19; ITT = intent-to-treat; PK=pharmacokinetic; PRO = patient reported outcomes.

5. STATISTICAL ANALYSES

5.1 GENERAL CONSIDERATIONS

All efficacy analyses will be performed in the ITT population, unless otherwise specified. Patients will be analyzed according to the treatment assigned at randomization by IxRS.

All IRF-assessed DFS and EFS analyses will use the tumor assessments per the revised charter (2nd Stream), unless otherwise specified in the sensitivity analyses (Section 5.4.3.1).

All safety analyses will be performed in the safety-evaluable population, unless otherwise specified. Patients will be analyzed according to the treatment they actually received.

5.2 PATIENT DISPOSITION

Enrollment and reasons for discontinuation from the study will be summarized by treatment arm for the ITT population. Study treatment administration and reasons for

discontinuation from study treatment will be summarized for the safety-evaluable population.

5.3 PRIMARY ENDPOINT ANALYSIS

5.3.1 Definition of Primary Endpoint

The primary efficacy endpoint is investigator-assessed DFS, defined as the time from randomization to death from any cause or the first documented recurrence assessed by investigator, whichever occurs first. Recurrence is defined as any of the following (as assessed by radiologic recurrence criteria described in the study protocol): local recurrence of RCC, new primary RCC, and distant metastasis of RCC.

Data for patients without an investigator-assessed event will be censored at the last date the patient was assessed to be alive and investigator-assessed recurrence free (or, for patients with no post baseline disease assessment, at the randomization date).

5.3.2 Main Analytical Approach for Primary Endpoint

The statistical hypothesis which will be tested for investigator-assessed DFS is given in Section 2. Investigator-assessed DFS will be compared between treatment arms using the stratified log-rank test. The HR for recurrence or death will be estimated using a stratified Cox proportional hazards model, and the 95% CI for the HR will be provided.

The stratified analysis will use the same factors as the randomization stratification factors. The stratification factors will be obtained from the IxRS. If there is at least one stratum with less than 20 events, the geographic region will be removed as a factor in the model given its limited prognostic significance. After removing the geographic region, if there is still at least one stratum with less than 20 events, the stratification factor that contains the level with the smallest number of patients will be removed from the stratified analysis.

Results from an unstratified analysis will also be provided.

Kaplan-Meier methodology will be used to estimate for each treatment arm the investigator-assessed DFS rate at specific timepoints (e.g., every 6 months or yearly), and to estimate median DFS; Kaplan-Meier curves will be produced. Brookmeyer-Crowley methodology (Brookmeyer and Crowley 1982) will be used to construct the 95% CI for the median DFS for each treatment arm.

5.3.3 Sensitivity Analyses for Primary Endpoint(s)

5.3.3.1 Sensitivity Analyses for Impacts of Missed Visits

For U.S. registration purposes, the impact of missed visits on investigator-assessed DFS will be assessed. The investigator-assessed DFS will be defined as described in Section 5.3.1 with an additional censoring rule for missed visits. Data for patients with an investigator-assessed DFS event who missed two or more scheduled assessments immediately prior to the investigator-assessed DFS event will be censored at the last

tumor assessment prior to the missed visits. Methods for comparison between treatment arms will be the same as the methods for treatment comparison for the primary endpoint of investigator-assessed DFS.

5.4 SECONDARY ENDPOINT ANALYSES

5.4.1 Overall Survival

OS (the key secondary efficacy endpoint) is defined as the time from randomization to death from any cause. Data for patients who have not died will be censored at the last date known to be alive or at the randomization date for patients with no post baseline information. Statistical hypothesis which will be tested for OS is given in Section 2. Methods for comparison of OS between treatment arms will be the same as the methods for treatment comparison for the primary endpoint of investigator-assessed DFS.

5.4.1.1 Sensitivity Analyses for Impact of Patients in Placebo Arm Receiving Atezolizumab or Similar Immunotherapy

The impact on OS of patients in the placebo arm subsequently receiving atezolizumab or another immunotherapy considered similar to atezolizumab in its mechanism of action may be assessed, depending on the number of such patients. If more than 5% of patients in the placebo arm have received atezolizumab or a similar immunotherapy, the following analyses may be performed to compare treatment arms at each of the interim analyses and the OS final analysis, as applicable:

- OS in the placebo arm will be discounted according to a range of possible effects on OS after having received atezolizumab or another immunotherapy (e.g., 10%, 20%, 30%, etc.).
- Additional sensitivity analyses may be conducted if deemed necessary.

5.4.2 Investigator-assessed DFS in Patients with PD-L1 Expression Status IC1/2/3

Investigator-assessed DFS in patients with PD-L1 expression status of IC1/2/3 is defined as a secondary endpoint. The definition of DFS is the same as that of the primary endpoint of investigator–assessed DFS. Investigator–assessed DFS in this population will be analyzed similarly to the analysis of investigator–assessed DFS in the ITT population.

5.4.3 IRF-assessed DFS

IRF–assessed DFS is defined as the time from randomization to death from any cause or the first documented recurrence assessed by IRF using the methodology in the revised IRF Charter, whichever occurs first. Data for patients without an IRF-assessed DFS event will be censored at the last date the patient was assessed to be alive and IRF-assessed recurrence free (or, for patients with no post-baseline disease assessment, at the randomization date). Patients who were determined to have baseline

disease in the 2nd Stream read will be censored at the randomization date. IRF-assessed DFS will be analyzed similarly to the analysis of investigator–assessed DFS.

5.4.3.1 Sensitivity Analyses for Impact of IRF Tumor Assessment Methodology Revision

The impact of the IRF tumor assessment methodology revision on the IRF-assessed DFS (see Section 1.2.3.1 for details) will be assessed by analyzing IRF-assessed DFS per the original charter and the concordance between the results assessed following the two charters.

The IRF-assessed DFS per the original charter is defined as the time from randomization to death from any cause or the first documented recurrence assessed by IRF per the original charter, whichever occurs first.

Data for patients without an IRF-assessed DFS event per the original charter will be censored at the last date the patient was assessed to be alive and IRF-assessed recurrence free (or, for patients with no post baseline disease assessment, at the randomization date). Patients who were determined to have baseline exclusionary lesions per protocol or meet recurrence criteria at baseline per the original charter will be censored at the randomization date.

The following analyses will be performed at the final analysis of investigator-assessed DFS:

- The IRF-assessed DFS per the original charter will be compared between treatment arms using similar analysis methods to the ones for the IRF-assessed DFS per the revised charter.
- The concordance between DFS assessed by the two IRF tumor assessment methodologies will be summarized by treatment arm (e.g., the concordance of event numbers and the concordance of IRF-assessed DFS dates between the two IRF tumor assessment methodologies).
- Additional sensitivity analyses may be conducted if deemed necessary.

5.4.4 IRF-assessed DFS in Patients with PD-L1 Expression Status IC1/2/3

IRF-assessed DFS in patients with PD-L1 expression status of IC1/2/3 is defined as a secondary endpoint. The definition of DFS is the same as that of the secondary endpoint of IRF-assessed DFS. IRF-assessed DFS in this population will be analyzed similarly to the analysis of IRF-assessed DFS in ITT population.

5.4.5 IRF-assessed EFS

IRF–assessed EFS is defined as the time from randomization to death from any cause, or the first documented recurrence in patients without baseline disease by IRF or the first documented disease progression in patients identified as having baseline disease by IRF, whichever occurs first. Data for patients without an IRF-assessed EFS event will be

censored at the last date the patient was assessed to be alive and IRF-assessed event free (or, for patients with no post-baseline disease assessment, at the randomization date). IRF–assessed EFS will be analyzed similarly to the analysis of IRF–assessed DFS.

5.4.6 Disease-specific Survival

Disease–specific survival (DSS) is defined as the time from randomization to death from RCC. Data for patients who have not died will be censored at the last date known to be alive. Data for patients who died from causes other than RCC will be censored at the date of death. Data for patients with no post baseline information will be censored at the randomization date. DSS will be analyzed similarly to the analysis of investigator-assessed DFS.

5.4.7 Distant Metastasis−free Survival

Distant metastasis–free survival (DMFS) is defined as the time from randomization to death from any cause or the date of a DMFS event, defined as diagnosis of distant (i.e., non–locoregional) metastases assessed by investigator, whichever occurs first. 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 (or, for patients with no post-baseline disease assessment, at the randomization date). Patients who develop a local (e.g., renal bed or new tumor in ipsilateral kidney) recurrence will continue to be followed until the occurrence of a distant metastasis. DMFS will be analyzed similarly to the analysis of investigator-assessed DFS.

5.4.8 One, Two, and Three-year Investigator-assessed DFS Rate

One, two, and three–year investigator–assessed DFS rate is defined as the probability of patients being alive and free of recurrence assessed by investigator at Year 1, 2, and 3 after randomization. The DFS rates will be estimated by Kaplan-Meier methodology, and the 95% CI will be estimated using Greenwood's formula. The 95% CI for the difference in one, two, and three–year DFS rates between the two arms will be estimated using the normal approximation to the binomial distribution.

5.4.9 One, Two, and Three-year IRF-assessed DFS Rate

One, two, and three–year IRF–assessed DFS rate is defined as the probability of patients being alive and free of recurrence assessed by IRF at Year 1, 2, and 3 after randomization. One, two, and three-year IRF–assessed DFS rate will be analyzed similarly to the analysis of one, two, and three-year investigator-assessed DFS rate.

5.5 EXPLORATORY ENDPOINTS ANALYSIS

Exploratory analyses of the following primary and secondary endpoints will be performed at various timepoints (e.g., every 6 months after randomization): investigator–assessed DFS, OS, IRF–assessed DFS, IRF–assessed EFS, DSS and DMFS. Event–free rates will be estimated by Kaplan-Meier methodology, and the 95% CI will be estimated using

Greenwood's formula. The 95% CI for the difference in event–free rates between the two arms will be estimated using the normal approximation to the binomial distribution.

To assess the consistency of study results in subgroups defined by demographic and baseline characteristics, efficacy outcomes of investigator-assessed DFS, OS, IRF–assessed DFS and IRF–assessed EFS in patient subgroups will be examined. Summaries of these endpoints, including unstratified HRs estimated from the Cox proportional hazards model and Kaplan–Meier estimates of the median will be produced separately for each level of the subgroup.

Investigator–assessed DFS will be summarized in patients with tumor Fuhrman Grade 4 or sarcomatoid histology (defined by investigator-assessed conventional histopathology). The definition of DFS is the same as that of the primary endpoint of investigator–assessed DFS. Investigator–assessed in this population will be analyzed similarly to the analysis of investigator–assessed DFS in the ITT population.

Exploratory biomarker analyses will be performed in an effort to understand the association of these markers with study drug response and mechanism of action, including efficacy and/or adverse events. Biomarker analyses may be described in a separate analysis plan and summarized in a separate report.

5.6 SAFETY ANALYSES

Unless specified otherwise, the safety analyses described below will be conducted for the safety-evaluable population (see Section 4), with patients grouped according to whether or not any atezolizumab treatment was received (i.e., patients who received any dose of atezolilzumab will be included in the atezolizumab arm).

5.6.1 Extent of Exposure

Summaries of exposure to study treatment will include treatment duration, number of doses, and dose intensity.

5.6.2 Adverse Events

Verbatim description of adverse events will be summarized by mapped thesaurus term using Medical Dictionary for Regulatory Activities (MedDRA), and graded according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v4.0. All adverse events occurring during or after the first study−drug dose until the clinical cutoff date will be summarized by treatment arm and NCI CTCAE grade. In addition, serious adverse events, Grade ≥3 adverse events, adverse events of special interest, adverse events leading to treatment discontinuation, and adverse events leading to treatment interruption will be summarized. Multiple occurrences of the same event will be counted once at the maximum grade.

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

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

5.7 OTHER ANALYSES

5.7.1 Summaries of Conduct of Study

Major protocol deviations, including major deviations of inclusion and/or exclusion criteria, will be summarized by treatment arm.

5.7.2 Summaries of Treatment Group Comparability

Demographic characteristics (e.g., age, sex, race/ethnicity), stratification factors (disease stage [T2/T3a vs. T3b/c/T4/N+ vs. patients with resected synchronous/metachronous metastasis], geographic region [North America (excluding Mexico) vs. rest of world], PD-L1 expression [IC0 vs. IC1/2/3]), and baseline disease characteristics (e.g., Eastern Cooperative Oncology Group (ECOG) performance status) will be summarized by treatment arm for the ITT population.

Continuous variables will be summarized using means, standard deviations, medians, and ranges. Categorical variables will be summarized by frequencies and percentages. Baseline measurements are the last available data obtained prior to the patient receiving the first dose of study treatment.

5.7.3 Pharmacokinetic Analyses

Pharmacokinetic (PK) analyses will be performed in the pharmacokinetic-evaluable population (see Section 4). Atezolizumab serum concentration data (minimum [C_{min}] and maximum [C_{max}]) will be reported and summarized (e.g., mean, SD, coefficient of variation, median, range, geometric mean, geometric mean coefficient of variation) for each cycle where collected as appropriate.

Additional PK analyses will be conducted as appropriate based on the available data.

5.7.4 Immunogenicity Analyses

The immunogenicity analysis population will consist of all patients with at least one ADA assessment for atezolizumab. The post-baseline ADA-evaluable population will be grouped according to treatment received (see Section 4). 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 considering post-baseline incidence, patients are considered to be ADA positive if they are ADA negative or with missing data at baseline but develop an ADA response following study treatment administration (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 4–fold greater (i.e., $\geq$ 0.60 titer units) than the titer of the baseline sample (treatment–enhanced ADA response). Patients are considered to be ADA negative if

they are ADA negative or with 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 units greater than the titer of the baseline sample (treatment unaffected).

The relationship between ADA status and safety, efficacy, and PK endpoints may be analyzed and reported descriptively via subgroup analyses.

5.7.5 Patient-Reported Outcomes Analyses

5.7.5.1 Functional Assessment of Cancer Therapy Kidney Symptom Index 19 (FKSI-19)

PRO analyses will be performed in the FKSI-19–evaluable population (see Section 4), unless otherwise specified.

Compliance rates for the FKSI-19 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 in the ITT population. Reasons for missing assessments, if available, will be summarized with frequencies and percentages.

Instruments will be scored per their user manual; subscales with less than 50% of the items completed will be considered missing. Missing data will not be imputed.

The FKSI-19 (version 2) can be scored as a Total scale and as five subscales:

- Total scale (Total, 19 items; score range of 0-76)
- Physical disease-related symptoms (DRS-P, 12 items; score range of 0-48)
- Emotional disease-related symptoms (DRS-E, 1 item; score range of 0-4)
- Treatment side effects (TSE, 3 items; scored individually 0-4)
- Function/well-being (FWB, 3 items; score range of 0-12)
- Disease-related symptoms (DRS, 9 items; score range of 0-36)

Higher scores indicate lower symptom or disease burden. The scoring of the FKSI-19 and it subscales will be done in accordance with FKSI-19 Scoring guidelines (Version 2; FKSI-19).

Descriptive analyses will include summary statistics (mean, SD, median, interquartile range [IQR], minimum, maximum) of each score and score changes from baseline at each assessment timepoint by treatment arm. Descriptive summaries will be reported for each FKSI-19 subscale (DRS-P, DRS-E, TSE, FWB, DRS) as well as the FKSI-19 Total

scale. Graphs of mean scores and/or score changes from baseline along with two-sided 95% CIs may be presented.

Time to confirmed deterioration (TTCD) analyses will be performed for each FKSI-19 subscale and Total scale in the ITT population. TTCD is defined as the time from the date of randomization to the date of a clinically meaningful score change from baseline that is held for at least two consecutive timepoints, or followed by death within 6 weeks from the last PRO assessment. Patients who do not experience a confirmed deterioration event at the time of clinical data cutoff will be censored at the last timepoint when they completed an assessment. Patients with no baseline or no post-baseline assessment will be censored at the randomization date. TTCD using the FSKI-19 scale will be analyzed similarly to the analysis of investigator-assessed DFS.

A clinically meaningful score change is defined as a 3-point or more score decrease on the FKSI-19 Total scale and DRS subscale (Cella et al. 2007, Yost and Eton 2005); a 4 point or more score decrease on the DRS-P subscale; and a 1-point or more score decrease on the DRS-E, TSE, and FWB subscales (Cella et al. 2018, Yost and Eton 2005).

Overall AE burden, as measured by the FKSI-19 GP5 item, will be summarized descriptively (numbers and proportions of responses) at each assessment timepoint in the safety-evaluable population. The distribution of responses and/or change from baseline will be illustrated by stacked bar charts for each arm.

To assess the consistency of the FKSI-19 results, the TTCD analysis (as defined above) will be performed on the subgroups defined by demographic and baseline characteristics. The outputs will include unstratified HRs estimated from the Cox proportional hazards model with associated 95% confidence intervals and, Kaplan-Meier estimates of the median will be produced separately for each level of the subgroup.

5.7.5.2 EuroQoL 5-Dimension, 5-level (EQ–5D–5L)

Health economic data, as assessed by the EQ–5D–5L, will be used to derive utilities for pharmacoeconomic models. The results from the health economic data analysis will be reported separately from the CSR.

5.8 INTERIM ANALYSES

5.8.1 Planned Interim Analyses

There is no planned interim efficacy analysis of the primary endpoint of investigator–assessed DFS for this study. A total of three analyses of OS are planned (two interim analyses and one final analysis). OS will be evaluated on the basis of the generalized Haybittle–Peto boundary (Haybittle 1971) for statistical significance, with p–value boundaries at each interim or final analysis specified as the following: 0.036, 0.036, and 0.011.

The first interim analysis of OS will be performed at the time of the investigator–assessed DFS final analysis (projected to occur at Month 67 after the first patient is randomized), if DFS is statistically significant at the final analysis. On the basis of the projected median OS for each treatment arm, the projected number of deaths observed at the DFS final analysis is approximately 190 deaths (25% of 764 patients), which corresponds to approximately 75% of the number of deaths required for the final analysis of OS. The observed HR of OS that is projected to result in a statistically significant difference between treatment arms at the DFS final analysis is less than or equal to 0.73.

The second interim analysis of OS will be performed when approximately 222 deaths have occurred (29% of 764 patients), corresponding to approximately 87.5% of the number of deaths required for the final analysis of OS. The required number of OS events for the second interim analysis of OS is projected to occur at Month 77. The observed HR that is projected to result in a statistically significant difference between treatment arms at this analysis is OS HR 0.75 or lower.

The boundary for statistical significance at the OS interim and final analyses will be adjusted based on the actual number of events observed. Specifically, if at the DFS final analysis, the observed number of OS events is significantly less than the projected 190 events (e.g., < 170 events), a nominal alpha value (1E–05) will be spent at the first OS interim analysis and the team will conduct an additional interim analysis of OS when 190 events have occurred.

Table 2 Projected Interim and Final OS Analysis Characteristics

Analysis	No. of events	% information	Event to Patient Ratio	Projected Cutoff Date[a]	Projected MDD[b]	Projected Boundary (p-value)[c]
First interim	190	75%	25%	Month 67	0.73	p ≤0.036
Second interim	222	87.5%	29%	Month 77	0.75	p ≤0.036
Final	254	100%	33%	Month 88	0.72	p ≤0.011

MDD=minimally detectable difference; OS=overall survival.

Note: Assumes 2% dropout rate over 24 months for OS analyses.

[a] Study month at which required number of events are projected to occur, where Study Month 1 is the month the first patient is enrolled.

[b] The largest observed hazard ratio that is projected to be statistically significant.

[c] The projected boundary for statistical significance for the number of events shown (actual boundary to be calculated at time of analysis based on actual number of events).

6. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document.

APPENDIX 1: CHANGES TO PROTOCOL-PLANNED ANALYSES

None.

7. REFERENCES

Sponsor Reports

WO39210: A phase III, multicenter, randomized, placebo-controlled, double-blind study of atezolizumab (anti-PD-L1 antibody) as adjuvant therapy in patients with renal cell carcinoma at high risk of developing metastasis following nephrectomy. Clinical study protocol Version 10, dated 12 November 2021.

[IRF Charter 1st Stream] WO39210: A phase III, multicenter, randomized, placebo-controlled, double-blind study of atezolizumab (anti-PD-L1 antibody) as adjuvant therapy in patients with renal cell carcinoma at high risk of developing metastasis following nephrectomy. Independent Review Charter 1st Stream, Version 1, dated 29 March 2017.

[IRF Charter 2nd Stream] WO39210: A phase III, multicenter, randomized, placebo-controlled, double-blind study of atezolizumab (anti-PD-L1 antibody) as adjuvant therapy in patients with renal cell carcinoma at high risk of developing metastasis following nephrectomy. Independent Review Charter 2nd Stream, Version 1, dated 4 December 2020.

[iDMC Charter] WO39210: A phase III, multicenter, randomized, placebo-controlled, double-blind study of atezolizumab (anti-PD-L1 antibody) as adjuvant therapy in patients with renal cell carcinoma at high risk of developing metastasis following nephrectomy. Charter for the independent data monitoring committee, Version 3.0, dated 31 January 2021.

Literature References

Brookmayer R and Crowley J. A Confidence Interval for the Median Survival Time. Biometrics 1982;38:29–41.

Cella D, Yount S, Brucker PS, et al. Development and validation of a scale to measure disease-related symptoms of kidney cancer. Value in Health 2007;10(4):285-293.

Cella D, Grünwald V, Escudier B, et al. Patient-reported outcomes of patients with advanced renal cell carcinoma treated with nivolumab plus ipilimumab versus sunitinib (CheckMate 214): a randomised, phase 3 trial. Lancet Oncol;20(2):297-310.

Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. Br J Radiol 1971;44:793–7.

National Comprehensive Cancer Network/Functional Assessment of Cancer Therapy Kidney Cancer Symptom Index - 19 Item Version (FKSI-19). Available from [cited 02 November 2021]: https://www.facit.org/measures-scoring-downloads/nfksi-19-scoring-downloads

Yost, KJ, Eton DT. Combining distribution-and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005;28(2):172-191.