X184-313: Statistical Analysis Plan

A Randomized, Double-Blind, Controlled Phase 3 Study of Cabozantinib in Combination with Nivolumab and Ipilimumab versus Nivolumab and Ipilimumab in Subjects with Previously Untreated Advanced or Metastatic Renal Cell Carcinoma of Intermediate or Poor Risk

Version 3.0

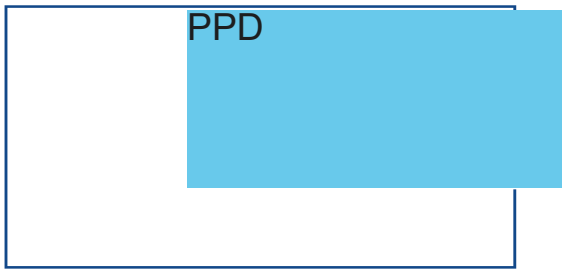
Date Final: 11 April 2022

Prepared by:

PPD

Exelixis Inc.

Approved by:



PPD

Exelixis Inc.

Approved by:

PPD

Exelixis Inc.

CONFIDENTIAL

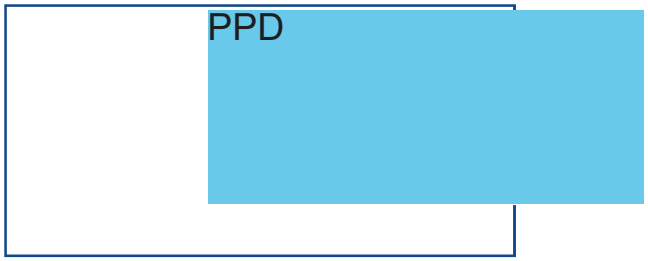
This document (and all copies thereof) is the property of Exelixis, Inc. and contains valuable trade secrets and confidential information, including proprietary technical and business information (hereinafter "Confidential Information"). By accepting this document, you agree that the Confidential Information contained herein will not be disclosed to others without prior written authorization from Exelixis, Inc. and shall only be used for the purposes intended by the document.

Approved by:

PPD

Exelixis Inc.

Approved by:



PPD

Exelixis Inc.

CONFIDENTIAL

This document (and all copies thereof) is the property of Exelixis, Inc. and contains valuable trade secrets and confidential information, including proprietary technical and business information (hereinafter "Confidential Information"). By accepting this document, you agree that the Confidential Information contained herein will not be disclosed to others without prior written authorization from Exelixis, Inc. and shall only be used for the purposes intended by the document.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ADA	Anti-drug antibodies
AE	Adverse event
AESI	Adverse event of special interest
AJCC	American Joint Committee on Cancer
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATA	Adequate tumor assessment
ATC	Anatomical Therapeutic Chemical
BIRC	Blinded independent review committee
BLQ	Below limit of quantitation
BMI	Body mass index
BOR	Best response
BSC	Best supportive care
BUN	Blood urea nitrogen
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CR	Complete response
CRF	Case report form
CRO	Clinical research organization
CSR	Clinical study report
CT	Computerized tomography
CTC	Circulating tumor cell
CTCAE	Common terminology criteria for adverse events
CTMS	Clinical trial management system
DBP	Diastolic blood pressure
DILI	Drug induced liver injury
EBRT	External beam radiation therapy
ECG	Electrocardiogram
ECOG PS	Eastern Cooperative Oncology Group Performance Status
ER	Emergency room visit
ETM	Events to monitor
FDA	Food and Drug Administration

FU	Follow-up visits (post-treatment)
GGT	Gamma-glutamyltransferase
HCRU	Health care resource utilization
HGB	Hemoglobin
HR	Hazard ratio
HRQOL	Health Related Quality of Life
ICH	International Conference on Harmonization
ICI	Immune checkpoint inhibitor
ICU	Intensive care unit
IDMC	Independent Data Monitoring Committee
IF	Information fraction
IMDC	International Metastatic RCC Database
IMM	Immune-modulating medications
INR	Prothrombin international normalized ratio
irAE	Immune-related adverse event
ITT	Intent-to-Treat
IxRS	Interactive Voice/Web Response System
KPS	Karnofsky performance status
LDH	Lactate dehydrogenase
LLN	Lower limit of normal
LLQ	Lower limit of quantitation
MRI	Magnetic resonance imaging
MedDRA	Medical Dictionary for Regulatory Activities
NPACT	Non-protocol anticancer therapy
OS	Overall survival
ORR	Objective response rate
PD	Progressive disease
PFS	Progression-free survival
PITT	PFS Intent-to Treat
PK	Pharmacokinetic
PP	Per-protocol
qd	Once daily
Qod	Every other day
QOL	Quality of Life

RCC	Renal cell carcinoma
rPD	Radiographic progressive disease per RECIST 1.1
SAE	Serious adverse event
SBP	Systolic blood pressure
SAP	Statistical analysis plan
SD	Stable disease
TEAE	Treatment emergent-adverse event
TSH	Thyroid-stimulating hormone
UE	Unable to evaluate
ULN	Upper limit of normal
ULQ	Upper limit of quantitation
UPCR	Urine protein/creatinine ratio
VAS	Visual analogue scale
WHO-DD	World Health Organization drug dictionary

1. ADMINISTRATIVE STRUCTURE AND VERSION HISTORY

Exelixis, Inc. is sponsoring this study. Exelixis is responsible for the statistical design and planning of this study. Statistical programming and analyses are being conducted under contract by Cato-SMS Inc. in conjunction with Exelixis, Inc.

Table 1: Protocol Version History

Date	Version	Primary Reason(s) for Amendment
03 October 2018	Original Protocol	Not Applicable
25 February 2019	Amendment 1.0	Power for secondary endpoint of overall survival (OS) was increased from 80% to 90% . Stratification factor PD-L1 was replaced by Region (No subjects were enrolled prior to this amendment).
09 January 2020	Amendment 2.0	Updated inclusion/exclusion criteria Deleted PD-L1 as an additional stratification factor from primary analyses so that stratification factors used for analyses are those used for randomization. Updated hepatotoxicity, myocarditis and gastrointestinal adverse event (AE) management guidelines.
16 November 2020	Amendment 3.0	Based on four years of follow-up data from the CM214 trial, the assumed median OS in the control arm was changed to 48 months vs. the previously assumed 41 months. As a result, the sample size increased to 840 from 676, and the required number of OS events increased from 342 to 433 to test the underlying null hypothesis for testing OS while preserving a critical delta of approximately 10 months. Safety guidelines for AEs related to nivolumab and ipilimumab were updated per the manufacturer (Bristol Myers Squibb) of these drugs.

Table 2: SAP Version History

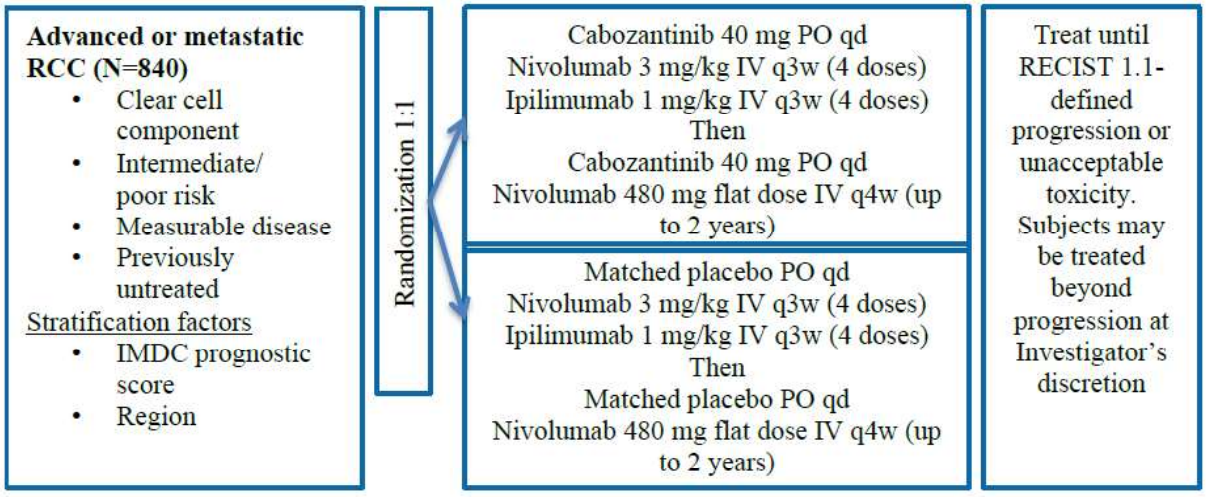
Date	Version	Primary Reason(s) for This Version
21 December 2020	1 (Based on Protocol Amendment 3.0)	Addresses summaries (disposition, demographics/baseline characteristics, progression-free survival (PFS) analysis per BIRC and Investigator assessments, OS analysis, objective response rate (ORR) analysis per BIRC and Investigator assessments, AEs, serious adverse events [SAEs], and related SAEs) needed for futility evaluation of the trial. A future amendment will provide final definitions for all other summaries
18 May 2021	2 (Based on Protocol Amendment 3.0)	This version of the SAP addresses the following: • Additional summaries for disposition to address discontinuation of the treatment components • Additional and sensitivity summaries to support PFS analysis. Estimand text added for analysis of OS endpoint. • AE summaries for summarization of immune-related AEs (irAEs) • Time to onset and Time to resolution of AEs of interest • Challenge/Re-challenge of study treatment • Events to Monitor (ETM) section deleted as it will be part of summary of clinical safety plan
11 April 2022	3	This version of the SAP addresses the following: • While the impact of COVID does not seem to be significant, it appears the number of PFS events will not be reached in a reasonable timeframe in the PITT population. Therefore, the PITT population will be expanded from the first 440 to the first 550 randomized, consistent with the pre-specified provision in the protocol for a similar circumstance due to the COVID-19 pandemic • The section numbers were corrected starting from Section 7 for EFFICACY ANALYSES • Hyperlinks and cross-references were updated throughout the document

Date	Version	Primary Reason(s) for This Version
		• In Section 7.4.2 added language to address situation in which the timing of the first interim analysis of OS transpires at an information fraction much higher than expected (e.g. due to faster-than-expected deaths or slower-than expected PFS events). • Minor Edits thoughout the document

2. STUDY DESCRIPTION

2.1. Study Design

This is a multicenter, randomized, double-blinded, controlled Phase 3 trial of cabozantinib in combination with nivolumab and ipilimumab versus nivolumab and ipilimumab in combination with matched placebo. Approximately 840 eligible subjects with advanced or metastatic renal cell carcinoma (RCC) in the intermediate or poor risk category by International Metastatic RCC Database Criteria (IMDC) will be randomized in a 1:1 ratio (~420 per treatment arm) at approximately 180 sites. The study design schematic for treatment period is presented in the below diagram.



IMDC, international metastatic renal cell carcinoma database consortium; IV, intravenous; PO, oral administration; qd, once daily; q3(4)w once every 3(4) weeks; RECIST, response evaluation criteria in solid tumors; RCC, renal cell carcinoma

Each subject's course of treatment will consist of the following study periods:

Pre-treatment Period: Potential subjects will be screened to determine if they meet the required eligibility criteria. Qualifying screening assessments must be performed within 28 days before randomization unless otherwise specified.

Treatment Period: Subjects who meet all study eligibility criteria will be randomly assigned in a 1:1 ratio to the experimental arm or control arm in a double-blinded manner.

Transition to Unblinded Treatment: The study may transition to unblinded treatment at the discretion of the Sponsor and upon any necessary discussions with regulatory authorities after analyses of PFS and OS have been performed. When the study is declared futile or the null hypothesis for the primary endpoint of PFS is not rejected (negative study), the study will be unblinded and Investigators and subjects will be notified. If the null hypothesis for the primary

endpoint of PFS is rejected (positive study), the study will not be unblinded until the null hypothesis for the secondary endpoint of OS is rejected or the final analysis of OS is performed.

Post-Treatment Period: The post-treatment follow-up (FU) visits FU-1 and FU-2 for safety evaluation will occur 30 (+14) days and 100 (±14) days, respectively, after the date of the decision to permanently discontinue study treatment. In addition, subjects are to be contacted every 12 weeks (± 14 days) after the 100-day FU-2 visit to assess survival status and document receipt of systemic non-protocol anti-cancer therapy (NPACT). See Protocol 'Study design' section for additional details.

Maintenance Phase: When sufficient data have been collected to adequately evaluate all study endpoints, and after treatment assignment has been unblinded, the Sponsor may initiate a Maintenance Phase. Upon initiation of this phase, the Sponsor considers the safety and efficacy profile of the experimental treatment regimen within this study to have been sufficiently established, and data analyses required for regulatory purposes have been completed. Thus, in this phase a subject receiving clinical benefit continues to have long-term access to study drug(s) until protocol-required criteria for discontinuation are met. The nature and frequency of the safety assessments during the Maintenance Phase are to be performed per institutional standard of care and guidance from the Sponsor as necessary.

The study clinical database will be closed upon initiation of the Maintenance Phase. Important safety data (see protocol for details) collected during this phase will be captured in the safety database. Only data collected prior to initiation of this phase will be reported in the clinical study report (CSR).

2.2. Study Treatment

Eligible subjects will be randomly assigned in a 1:1 ratio to the following treatment arms:

Experimental arm (~420 subjects): cabozantinib (40 mg oral, once daily [qd]) + nivolumab (3 mg/kg infusion, q3w) x 4 doses + ipilimumab (1 mg/kg infusion, q3w) x 4 doses, followed by cabozantinib (40 mg oral qd) + nivolumab infusion (480 mg flat dose q4w).

Control arm (~420 subjects): cabozantinib-matched placebo (oral, qd) + nivolumab (3 mg/kg infusion, q3w) x 4 doses + ipilimumab (1 mg/kg infusion, q3w) x 4 doses, followed by cabozantinib-matched placebo (oral, qd) + nivolumab infusion (480 mg flat dose q4w).

Nivolumab will be administered for a maximum of two years.

2.3. Study Objectives and Endpoints

The objective of this study is to evaluate the efficacy and safety of the combination of cabozantinib with nivolumab and ipilimumab versus the combination of nivolumab and ipilimumab in previously untreated (first line) intermediate- and poor-risk advanced or metastatic RCC subjects.

This study has one primary efficacy endpoint (PFS) and one secondary efficacy endpoint (OS). The trial will be declared a success if the null hypothesis is rejected for the primary endpoint.

2.3.1. Primary Efficacy Endpoint

Duration of progression-free survival (PFS) per Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1) by Blinded Independent Review Committee (BIRC).

2.3.2. Secondary Efficacy Endpoint

Duration of overall survival (OS)

2.3.3. Additional Endpoints

- Objective response rate (ORR) per RECIST 1.1 by BIRC
- PFS and ORR per RECIST 1.1 by BIRC according to PD-L1 status
- PFS and ORR per RECIST 1.1 as assessed by the Investigator
- Duration of radiographic response as assessed by the Investigator and by BIRC
- Safety through the evaluation of adverse events (AEs), including immune-related AEs (irAEs), and other safety assessments
- Pharmacokinetics (PK) of cabozantinib given in combination with nivolumab and ipilimumab
- Immunogenicity of nivolumab and ipilimumab given in combination with cabozantinib
- Correlation of biomarker analyses with clinical outcomes
- Health-related quality of life (HRQoL) as assessed by the EuroQol Health questionnaire EQ-5D-5L
- Health care resource utilization

2.4. Power and Sample Size Justification

The study is designed to provide adequate power for analyses of both endpoints of PFS and OS for comparing the experimental arm with the control arm. A larger sample size is needed to provide reasonable power for OS than is required to evaluate PFS. As a result, if PFS were to be

evaluated in the entire study sample, the PFS events may be biased toward shorter progression times. Thus, to allow longer, more robust PFS follow up among a smaller number of subjects, this study employs a "trial within a trial design" (Hessel et al 2016).

The total sample size is 840 subjects, to be randomized in a 1:1 ratio to the experimental and control arms.

The futility analysis will be conducted when at least 100 events have been observed and analysis of primary endpoint of PFS will be conducted when at least 249 events have been observed in the first 440 subjects randomized. This will provide the study with 90% power to reject the null hypothesis of no difference in the survival function of PFS between the two study treatment arms with a 2-sided log-rank test at a 5% level of significance and assuming a true hazard ratio (HR) of 0.66. Assuming an exponential distribution of PFS, this corresponds with a 52% increase in median PFS from 11.6 months to 17.6 months. The minimum observed effect that would result in statistical significance for PFS is an HR of 0.78, a 28% improvement in median from 11.6 to 14.9 months.

As per a provision pre-specified in the protocol, the sample size for determining the primary endpoint (PFS Intent-to-Treat [PITT] population) may be expanded by 25% from the first 440 to first 550 randomized if a review of the accumulating PFS events data suggests that the 249 events required for the analysis will not be reached due to censoring caused by a higher than expected study drop-out or non-compliance stemming from the COVID-19 pandemic. The number of events targeted to perform this event-driven analysis will not change, thus such an increase would have no impact on study power, Type 1 error, criteria for statistical significance or the minimum effect size that rejects the null hypothesis. Version 3.0 of the SAP invoked this provision as described in Section 4.8.1.

For OS, the study was originally designed as follows: a total of 342 deaths among 676 randomized subjects are required to provide 90% power to reject the null hypothesis of no difference in OS using the log-rank test, a 2-sided significance level of 5%, three planned interim analyses, and assuming a true HR of 0.70. Assuming an exponential distribution for OS, this corresponds to a 43% improvement in median survival from 41 months to 58.6 months. Under this design, the minimum observed effect that would result in statistical significance for the final analysis of OS is an HR of 0.80, a 25% improvement in median from 41 to 51.1 months corresponding to a minimum difference in observed medians of approximately 10 months.

In response to new data supporting a median OS in the control arm of 48 months instead of the previously assumed 41 months (Albiges et al 2020), the study has been modified per amendment 3.0 as follows: A total of 433 events among 840 randomized subjects (the Intent-to-Treat [ITT]

population) are required to provide 90% power to reject the null hypothesis of no difference in OS using the log-rank test at a 2-sided significance level of 5%, three planned interim analyses, and assuming a true HR of 0.73. Assuming an exponential distribution for OS, this corresponds to a 37% increase in median survival from 48 months to 65.8 months. Under this design the minimum observed effect that would result in statistical significance for OS is an HR of 0.824, a 21% improvement in median from 48 to 58.25 months corresponding to a minimum observed difference in medians of approximately 10 months.

This increase in sample size does not inflate Type 1 error because (a) it was done solely in response to new external data about the expected median OS of the control arm, and (b) the control arm median is a nuisance parameter with respect to the minimum observed difference in medians that results in rejection of the null hypothesis. This change ensures the trial retains the ability to reject the null hypothesis for OS if the observed difference in medians is approximately 10 months.

Three interim analyses of the secondary endpoint of OS are planned at approximately the 27% (~117 death events), 50% (~217 death events) and 75% (~325 death events) information fractions (IFs). Details on the interim analyses of OS and the control of type 1 error are provided in Sections 7.3 and 7.4.

With an assumed constant accrual rate of 40 subjects per month, a 1:1 treatment allocation ratio, and three interim analyses of OS, a total of 840 subjects (420 per treatment arm) will be randomized. The number of events required for the primary analyses of PFS (249 events among the first 440 [or 550] randomized subjects) is expected to be observed approximately 23 months after the first subject is randomized. The number of events required for the final analysis for the secondary endpoint of OS (433 events among 840 subjects) is expected to be observed approximately 69 months after the first subject is randomized. The true intervals required to meet these milestones may be longer or shorter due to divergence from assumptions, including non-constant accrual rate due to the time required for all study sites to become active.

An overview of the endpoints, assumptions, and operating characteristics is shown in Table 3.

Table 3: Summary of Endpoint Operating Characteristics

Accrual per month	**40**	
Randomization allocation	**1:1**	
Endpoint:	**PFS: Primary endpoint**	**OS: Secondary endpoint**
Power	90%	90%
Alpha allocated (2-sided)	0.05	0.05
# of interim analyses (approximate information fraction)	1 (40%) Futility, non-binding (1 -sided)	3 (27%, 50%, 75%)
Assumed median, control (months)	11.6	48
Assumed median, experimental (months)	17.6	65.8
Assumed HR	0.66	0.73
Number of events	249	433
N for analysis Population	440 (or 550) PITT Population	840 ITT Population
Maximum HR to reject H_0 (experimental median in months)	0.78 (14.9)	0.824 (58.25)
Minimum difference in medians (months) to reject H_0	3.3	10.25

Power and sample size estimates were calculated using EAST v6.5 by Cytel Software.

2.5. Randomization and Blinding

Randomization will be stratified by the following factors established at screening:

- IMDC prognostic score (1-2 risk factors [intermediate] vs 3-6 risk factors [poor])
- Region ([US or Canada or Europe or Australia or New Zealand] vs [Latin America or Asia])

The IMDC risk factors are (Heng et al 2009):

- KPS < 80%
- Less than 1 year from initial RCC diagnosis (including original localized disease if applicable) to systemic treatment
- Hemoglobin (g/L) < lower limit of normal (LLN)
- Corrected calcium (mg/dL) > 10 mg/dL
- Absolute neutrophil count (ANC) (GI/L) > upper limit of normal (ULN)
- Platelet count (10^9/L) > ULN

This is a double-blinded study. The cabozantinib-matched placebo will be given in combination with nivolumab and ipilimumab in the control arm to blind (mask) study treatment. The

assignment of study treatment will be unknown to the subjects, Investigators, study centers, Sponsor, and any clinical research organization (CRO) affiliated with the study other than those authorized to access treatment assignment for regulatory safety reporting and submission processes (see Protocol Section 8.2.2), Interactive Voice/Web Response System (IxRS) system administration, and drug supply management.

3. ANALYSIS POPULATIONS

The planned populations for statistical analyses are described in the sections below.

3.1. PFS Intent to Treat Population

PFS Intent-to-Treat (PITT) population is defined as the first 550 (see Section 4.8.1) randomized subjects (based upon Greenwich Mean Time randomization date/time values) regardless of whether any study treatment or the correct study treatment was received.

3.2. Intent to Treat Population

The Intent-to-Treat (ITT) population is defined as all randomized subjects regardless of whether any study treatment or the correct study treatment was received.

3.3. Safety Population

The Safety population will include all subjects who receive any amount of study treatment. Analyses based on Safety population will be performed according to the actual treatment received for the length of the study. Subjects who receive both treatments in error may be excluded from the analyses and their data listed.

3.4. P-Safety Population

The P-Safety population will include subjects in the PITT population who were treated with any amount of study treatment. Select summaries will be presented based on actual treatment received.

3.5. Per Protocol Population

A Per Protocol population may be defined at a later stage to characterize the impact of COVID-19 on the efficacy and safety for the study.

4. GENERAL CONVENTIONS

The statistical principles applied in the design and planned analyses of this study are consistent with International Conference on Harmonization (ICH) E9 and E9(R1) guidelines (ICH 1998 and 2019) and FDA guidance's for Industry on Multiple Endpoints in Clinical Trials (2017) and for handling Clinical Trial Endpoints for Approval of Cancer Drugs and Biologics (2018).

Continuous data will be summarized using descriptive statistics (number of observations, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum). Frequencies and percentages will be used for summarizing categorical (discrete) data.

Confidence intervals (CIs), when presented, will generally be constructed at the 95% level. For binomial variables, the Clopper-Pearson method (Clopper and Pearson 1934) will be employed unless otherwise specified.

A month is operationally defined to be 30.4375 days. Six months is operationally defined to be 183 days.

Efficacy summaries will be presented by treatment arm randomized to and safety summaries by treatment received unless otherwise specified.

4.1. Definition of Baseline

In general, for efficacy endpoints the last observed measurement prior to randomization will be considered the baseline measurement. An exception to this rule is the efficacy biomarkers such as pharmacogenetics blood samples and plasma biomarker samples. These samples per schedule of assessment are to be collected on Week 1 Day 1 (W1D1) (which could be up to 3 days after randomization) but prior to the administration of the first dose; hence baseline for these measurements will be defined with respect to the first dose date. For subjects who did not take any study treatment, any biomarker sample available prior to randomization will be considered as the baseline observation.

For safety endpoints the last observation on or before first dose of study treatment will be considered the baseline measurement unless otherwise specified and will be established for each period defined in Section 4.5.

For assessments on the day of first dose of study treatment where time is not captured, if such procedures are required by the protocol to be conducted before first dose or a nominal pre-dose indicator is available, this will serve as sufficient evidence that the assessment occurred prior to first dose.

The earliest instance of administration of any agent of study treatment will be considered as the time point of first dose. Similarly, the latest instance of administration of any agent of study treatment will be considered as the time point of last dose of study treatment.

4.2. Definition of Study Day

For the purpose of efficacy data summaries, Study Day is defined with respect to the randomization date. For visits (or events) that occur on or after randomization, study day is defined as (date of visit [event] – date of randomization + 1). For visits (or events) that occur prior to randomization, study day is defined as (date of visit [event] – date of randomization). There is no Study Day 0.

For the purposes of safety data summaries, Dose Day is defined with respect to the date of first dose of study treatment received. For visits (or events) that occur on or after first dose, dose day is defined as (date of visit [event] – date of first dose of study treatment + 1). For visits (or events) that occur prior to first dose, dose day is defined as (date of visit [event] – date of first dose of study treatment). There is no Dose Day 0.

For listings (such as for adverse events [AEs]) that include the derivations of "days since last dose," this is defined for each agent as (event date – date of last dose for the agent). Three such fields, one for cabozantinib or placebo, one for nivolumab, and one for ipilimumab will be presented for subjects receiving the combination of these agents. Events that occur on the same day as the last dose of a particular study component will therefore be described as occurring zero days from the last dose of that study component.

4.3. Visit Window Calculation

Analyses will be according to actual visit dates and times. The planned analyses do not require the calculation of visit windows. However, for analyses that require a particular visit as planned, measurements included in that visit must have occurred during the acceptable window defined for the visit in the protocol, unless otherwise specified in this plan.

4.4. Missing and Partial Data

In general, other than for partial dates, missing data will not be imputed and will be treated as missing. The algorithms for imputation of partial dates vary depending upon the parameter.

Data imputation rules are presented in Appendix A.

4.5. Safety Observation Periods

Generally only the safety data (including AEs that start or worsen, laboratory results, vital signs, electrocardiograms [ECGs], Karnofsky Performance Status [KPS], Eastern Cooperative Oncology Group Performance Status [ECOG PS], concomitant medications, etc.) reported during a safety observation period (defined below) will be analyzed and summarized, unless otherwise specified in this plan.

4.5.1. Standard Period

The standard safety observation period is the interval from the first dose date of study treatment to the earliest of:

- Date of last dose of last component of study treatment + 30 days, or
- Date of death, or
- Date subject withdrew consent for non-interventional study assessments, or
- Data cutoff date for analysis under consideration.

4.5.2. Extended Period

To evaluate the incidence of adverse events of interest (AESIs) associated with study treatment, extended observation period is defined as the interval from the first dose date of study treatment to the earliest of the following:

- Date of last dose of last component of study treatment + 100 days, or
- Date of death, or
- Date subject withdrew consent for non-interventional study assessments, or
- Data cutoff date for analysis under consideration.

4.6. Definition of Prior, Concomitant, and Subsequent Therapy

For the purpose of inclusion in summary tables, incomplete start and stop dates for therapies (medications, radiation therapies or surgery) will be imputed as detailed in Appendix A.

Therapies may be summarized as prior, concomitant and/or subsequent.

Based on imputed start and stop dates:

- Prior therapies are defined as those with a start date occurring before the date of first dose of study treatment.
- Concomitant therapies are defined as those that stop or continue on or after the first dose date through the end of the safety observation period.

- Concomitant and subsequent therapies are defined as those that stop or continue on or after the date of randomization.

4.7. Software

All analyses will be conducted using SAS Version 9.4 or higher.

4.8. Changes to Planned Analyses

Substantive changes to the analyses described in the protocol or in approved versions of this plan will be fully documented in a revised version of this plan approved by the Sponsor prior to conducting analyses.

Clarifications, minor corrections, and operational considerations necessary to accurately conduct the analyses that do not materially change the nature of the analysis will be documented in an addendum to this plan that will also be approved by the Sponsor prior to final study analyses.

4.8.1. Expansion of PITT Population

As of January 2022, after more than 17 months minimum of follow up in the first 440 randomized, approximately 209 (84%) of the required 249 PFS events per BIRC have been observed. Notably, the rate of events has slowed since June 2021 with only approximately 12 events occurring in the last 7 months. As the study and Exelixis remain blinded, the reason(s) for the slower event rate are not known. The impact of COVID-19 on the trial does not appear to be significant. However, whatever the cause, it does not appear that the targeted number of events will be reached in a reasonable timeframe.

Version 3.0 of the SAP invokes the protocol provision to expand the PITT population by 25% to include the first 550 subjects randomized to allow the targeted 249 PFS events to be reached without further protracted delay. Although the reason for the slowing event rate may not be due to COVID-19, this change is consistent with the principle and scope of expanding the PITT population pre-specified in the protocol.

The minimum follow up time of the first 550 subjects randomized is 15.6 months, longer than the hypothesized median PFS in the control arm of 11.6 months, a common convention for defining adequate follow up of study samples for time to event (TTE) endpoints.

As the power for TTE endpoints is driven by the total number of events and not the total sample size, this change does not fall into the realm of sample size re-estimation. It does not impact the targeted number of events or other study operating characteristics, only the time required to reach the targeted number of events. It does not change the hypothesized effect size, the criteria for declaring statistical significance, impact the potential clinical significance of trial results, trial

conduct, subject safety, or the statistical or scientific or clinical interpretation of study results. As such, no protocol amendment nor statistical adjustment is required.

5. STUDY POPULATION SUMMARIES

5.1. Enrollment

Subjects are defined to be enrolled at randomization. Enrollment will be summarized by region, country, site, and protocol version for subjects in the PITT and ITT populations.

5.2. Disposition

Subject disposition will be summarized by PITT, ITT, Safety and P-Safety populations.

The following summaries will be presented:

- Number and percentage of subjects in the PITT, ITT, Safety and P-Safety populations
- Subjects treated with cabozantinib/placebo
- Subjects treated with nivolumab
- Subjects treated with ipilimumab
- Subjects still on study treatment component(s) at the time of data cut-off
- Subjects who discontinued any study treatment component
- Subjects who discontinued all study treatment component
- Primary reason for discontinuation from:
    - last study treatment component
    - cabozantinib/placebo
    - nivolumab
    - ipilimumab

  (Note: In the above summaries subjects who discontinued treatment due to AE or SAE that was unrelated to progression of disease under study, relationship to the underlying treatment will also be summarized.)
- Primary reason for discontinuation from radiographic follow-up
- Primary reason for discontinuation from study follow-up
- Number of subject's screen failure and the reasons for screen failure

5.3. Demographic and Baseline Characteristics

Summaries of demographics, stratification factors and baseline characteristics will be presented for subjects in the ITT, PITT, Safety and P-Safety populations.

[A] The demographic characteristics include:

- Age (continuous)
- Age category 1: < 65 years, ≥ 65 years
- Age category 2: < 75 years, ≥ 75 years
- Age category 3: < 65 years, 65 to < 75 years, 75 to < 85 years, ≥ 85 years
- Sex: Male, Female, Not reported
- Ethnicity: Hispanic or Latino, Not Hispanic or Latino, Not Reported
- Race:
    - American Indian or Alaska Native
    - Asian
    - Black or African American
    - Native Hawaiian or Other Pacific Islander
    - White
    - Not Reported
    - Other
- Geographic Region
    - Latin America or Asia
    - US or Canada or Europe or Australia or New Zealand

Note: Birth date is not collected but age in years is collected at informed consent.

[B] Categorical summaries of the stratification factors will be presented as recorded in the (a) IxRS during randomization, (b) case report form (CRF), (c) cross tabulation of the two stratification factors per IxRS, and (d) cross tabulation of the two stratification factors per CRF.

[C] Baseline characteristics include:

- Height in centimeter (cm) – descriptive statistics
- Weight in kilogram (kg) – descriptive statistics

- Body mass index (BMI) in kg/meter^2, calculated as (weight in kg*10000)/(Height in cm)^2 – descriptive summary
- Body surface area (m^2) – descriptive summary
- ECOG PS: 0, 1, Missing
- Karnofsky performance status: < 80%, ≥ 80%, Missing
- Karnofsky performance status: 100-90, 80-70, Missing
- Smoking history: Current, Former, Never
- PD-L1 immunohistochemistry status: ≥ 1%, < 1%, Indeterminate
- Tumor MET immunohistochemistry: High, Low/Negative, Unknown
- Time from initial RCC diagnosis to systemic treatment: < 1 year, ≥ 1 year

[D] Baseline laboratories:

- Hemoglobin (g/L): < LLN, ≥ LLN
- Corrected calcium (mg/dL): ≤ 10, > 10
- Absolute neutrophil count (ANC) (GI/L): ≤ ULN, > ULN
- Platelet count (10^9/L): ≤ ULN, > ULN

5.4. Medical History

General medical history data will be coded per Medical Dictionary for Regulatory Activities (MedDRA).

5.5. Cancer History and Current Disease Status

Cancer history and current disease characteristics data collected on the cancer history CRF will be summarized categorically or with descriptive statistics as appropriate for subjects in ITT and PITT populations. The following summaries are planned:

- Diagnosis of RCC by histology or cytology (Yes, No)
- Time to randomization since initial diagnosis of RCC as identified by histology/cytology in years (Note: Incomplete diagnosis dates will be imputed as detailed in Appendix A.) (Descriptive summaries)
- Disease stage at initial diagnosis (Stage I, Stage II, Stage III, Stage IV, Unknown)
- Current disease stage per American Joint Committee on Cancer (AJCC) criteria (Stage III, Stage IV, Other)
- Nephrectomy performed (Yes, No)

- Histology type based on nephrectomy (Clear cell, Papillary, Chromophobe, Sarcomatoid, Other)
- Histology types based on archival tissue or fresh biopsy obtained for this study (Clear cell, Papillary, Chromophobe, Sarcomatoid, Other)
- Time to randomization since most recent documented recurrence/progression (Descriptive summaries)
- Extent of disease at baseline (sites of metastases) per Investigator and per BIRC assessed as target and non-target lesions at baseline, for the following organs:
    - Brain
    - Lymph node
    - Bone
    - Visceral
        - Liver
        - Lung
        - Adrenal gland
        - Kidney
        - Spleen
        - Pancreas
    - Other sites
- Number of organs with target and non-target lesions at baseline per BIRC (1, 2-3, 4-5 ≥ 6)
- Number of organs with target and non-target lesions at baseline per Investigator (1, 2-3, 4-5 ≥ 6)
- Has measurable disease at baseline identified by BIRC (Yes, No)
- Has measurable disease at baseline identified by the investigator (Yes, No)
- Descriptive summary for baseline sum of target-lesion diameters (mm) per BIRC
- Descriptive summary for baseline sum of target-lesion diameters (mm) per Investigator

6. TREATMENTS AND MEDICATIONS

6.1. Prior Non-Radiation Anticancer Therapy

All prior non-radiation anticancer agents will be summarized categorically by Anatomical Therapeutic Chemical (ATC) Class Text and World Health Organization Drug Dictionary

(WHO-DD) substance preferred name by study treatment arm, type of therapy (systemic, local, unknown) and context of therapy (adjuvant, neoadjuvant, other) for ITT and PITT populations.

The following will be summarized categorically or with descriptive statistics as appropriate for all subjects in the ITT and PITT populations:

- Therapy type: Systemic (Neoadjuvant, Adjuvant, Other), Local, Unknown
- Number of prior local non-radiation anticancer agents for RCC per subject (1, 2, ≥ 3) and descriptive statistics
- Radiographic or clinical progression after initiation of the most recent agent (Yes, No)
- The time from the end of most-recent local non-radiation anticancer agent for RCC to randomization
- The time from the end of most-recent systemic adjuvant non-radiation anticancer agent for RCC to randomization

6.2. Prior Radiation Therapy, Surgery and Procedures

Data obtained from the history of radiation therapy CRF will be summarized categorically or with descriptive statistics as appropriate for all subjects in the PITT and ITT populations:

- Subject incidence of radiation therapy by indication (RCC and Other)
- Number of prior radiation therapies for RCC per subject (1, 2, ≥ 3) and descriptive statistics
- Subject incidence of type of radiation therapy received for RCC or other indications
- Subject incidence of radiation therapy by site (Bone, Soft-tissue, Systemic, Unknown) for RCC or other indications
- Number and percent of subjects that received radiation by anatomic sites
- Number of subjects with prior surgery/procedure (Yes, No)
- Subject incidence of surgery type (Partial nephrectomy, Simple nephrectomy, Radical nephrectomy, Other)
- Time from the date of most recent surgery for RCC to randomization in months

6.3. Prior and Concomitant Medications (Excluding Anticancer Therapy)

Prior and concomitant medications, other than prior and subsequent anticancer therapies, will be summarized by treatment group in the ITT and Safety populations by ATC and WHO-DD substance preferred name. Anticancer therapies are addressed in Section 6.9 of this plan.

6.4. Prior and Concomitant Immune Modulating Medications

Immune-modulating medications (IMMs) are a select list of medications that were used to manage AEs due to study treatment. A list of the verbatim terms for these medications will be identified and provided along with the ATC and WHO-DD base substance preferred name. These will be summarized by treatment group in the Safety population by ATC and WHO-DD base substance preferred name. In addition, the following summaries will be provided:

- Subjects with prior use of IMMs
- Subjects with concomitant use of IMMs
- Subjects with concomitant use of IMMs for management of AEs
- Subjects with concomitant use of high dose corticosteroids for management of AEs where high dose is defined as ≥ 40 mg/day of prednisone or equivalent

6.5. Study Treatment Exposure

Study treatment exposure will be summarized descriptively and will include subjects in the Safety and P-Safety populations. Exposure summaries for the oral components and the infusion components will be presented separately.

The algorithms for calculating the various exposure parameters are described in Table 4 for overall dosing summaries, Table 5 for weight-based dosing summaries in cycle 1 – 4, and

Table 6 for dosing summaries based on a flat dose from cycle 5 onwards.

Table 4: Study Therapy Definitions (Overall summary)

	Nivolumab	Ipilimumab	Cabozantinib/Placebo
Cumulative dose (mg)	Sum of doses administered to a subject	Sum of doses administered to a subject	Sum of doses administered to a subject
Number of infusions	Sum of all infusions administered	Sum of all infusions administered	N/A
Average daily dose (mg/day)	NA	NA	Cumulative dose (mg) / (Duration of exposure in days)
Relative dose intensity (%)	(Relative dose intensity (%) during weight-based infusions × number of weight-based infusions received + Relative dose intensity (%) of flat dose infusions × number of flat dose infusions received) / total infusions received	Sum of all relative cycle intensities / N	100 × average daily dose / (40mg/day)
Duration of exposure in days (including holds)	Last non-zero dose date of agent – first dose date of the agent +1	Last non-zero dose date of agent - first dose date of agent +1	Date of decision to discontinue or cut-off date – first dose date of agent +1
Duration of exposure in days (excluding dose holds)	Last non-zero dose date of agent – first dose date of the agent -total days infusion was delayed +1	Last non-zero dose date of agent – first dose date of the agent -total days infusion was delayed +1	Date of decision to discontinue or cut-off date – first dose date of agent – total days with dose received of 0 mg +1
Overall duration of study treatment (including holds)	(Maximum (date of last non-zero dose of nivolumab or ipilimumab or date of decision to discontinue cabozantinib/placebo or cut-off date) – date of first dose of study treatment + 1) / 30.4375		
Overall duration of study treatment (excluding holds)	Overall duration of study treatment (including holds) – total number of days dose was held, or infusion was delayed		

Table 5: Study Therapy Definitions (Weight Based Cycles 1-4)

	Nivolumab	Ipilimumab	Cabozantinib/Placebo
Intended dose for each cycle (i) for infusion components and in mgs for oral components	3 mg/kg/3wk for i=1-3; 3 mg/kg/4wk for i=4	1 mg/kg/3wk for i=1-4	40 mg once daily
$Dose_{(i)}$	Total dose administered in mgs for each cycle is collected on the CRF	Total dose administered in mgs for each cycle is collected on the CRF	Total dose for the interval is collected on the CRF
Cycle $duration_{(i)}$ (wk)	$(Dose\ date_{(i+1)} - Dose\ date_{(i)})/7$ for i=1-3 i=4, 28/7=4 wks	$(Dose\ date_{(i+1)} - Dose\ date_{(i)})/7$ for i=1-3 i=4, 21/7=3 wks	NA
Cycle $intensity_{(i)}$ (mg/kg/wk)	$Dose_{(i)}$/Cycle $duration_{(i)}$ i=1-4	$Dose_{(i)}$/Cycle $duration_{(i)}$ i=1-4	NA
Relative cycle $intensity_{(i)}$ (%)	(Cycle $intensity_{(i)}$ / intended $dose_{(i)}$) × 100 for i=1-4	(Cycle $intensity_{(i)}$ / intended dose(i)) × 100 for i=1-4	NA
Relative dose intensity (%)	Sum of all relative cycle intensities divided by number of weight-based infusions received	Sum of all relative cycle intensities divided by number of weight – based infusions received	See Table 4

Table 6: Study Therapy Definitions (Flat Dose Cycles From 5 Onwards)

	Nivolumab	Cabozantinib or Placebo
Dosing schedule per Protocol	480 mg every 4 weeks	40 mg once daily
Relative dose intensity (%)	Cumulative dose (mg)/[(Last non-zero dose date – First dose date of flat doses + 28) × 480/28)] × 100	See Table 4

6.5.1. Overall Exposure

Overall study treatment exposure will be summarized as follows:

- Descriptive summary for duration of exposure (including dose holds) of study treatment in months calculated as [maximum of (date of last non-zero dose of nivolumab or ipilimumab or date of decision to discontinue cabozantinib/placebo or cutoff date – date of first dose of study treatment+1)/30.4375]. Categorical summary presented in months: $\leq 3, > 3, > 6, > 12, > 24$

- Descriptive summary for duration of exposure (excluding dose holds) of study treatment in months calculated as duration of exposure (including dose holds) – total number of days dose was held or infusions were delayed
- Descriptive summary for duration of exposure (including dose holds) of the first study treatment component discontinued in months calculated as [minimum of (date of last non-zero dose of nivolumab or ipilimumab or date of decision to discontinue cabozantinib/placebo or cutoff date – date of first dose of study treatment+1)/30.4375]. Categorical summary presented in months: ≤ 3, > 3, > 6, > 12, > 24
- Descriptive summary for duration of exposure (including dose holds) of the first component discontinued for treatment combination cabozantinib/placebo+nivolumab in months calculated as [minimum of (date of last non-zero dose of nivolumab or date of decision to discontinue cabozantinib/placebo or cutoff date – date of first dose of cabozantinib/placebo+nivolumab treatment+1)/30.4375]. Categorical summary presented in months: ≤ 3, > 3, > 6, > 12, > 24

6.5.2. Exposure for Cabozantinib/Placebo Component

Exposure for cabozantinib/placebo treatment will be summarized as follows:

- Duration of exposure (including dose holds) calculated as (earlier of date of decision to discontinue oral component or cutoff date – date of first dose of oral component + 1) /30.4375
- Duration of exposure (excluding dose holds) calculated as (earlier of date of decision to discontinue oral component or cutoff date – date of first dose of oral component - total days with oral dose received of 0 mg + 1) /30.4375
- Average daily dose (mg/day) calculated as (total mg received/duration of exposure in days)
- Percent dose intensity, calculated as 100 × (average dose in mg/day)/(40 mg/day)

6.5.3. Exposure for Nivolumab Treatment Component

Exposure for nivolumab will be summarized as follows:

By descriptive statistics:

- Duration (in months) of exposure including delays calculated as (date of last infusion of nivolumab – date of first dose of nivolumab)/30.4375
- Number of infusions administered per subject
- Number of expected infusions per subject
- Cumulative dose

- Average dose (mg/infusion)
- Overall relative dose intensity (%)
- Relative dose intensity (%) for infusions 1-4 (weight based)
- Relative dose intensity (%) for infusions 5 onwards (flat dose received)

By frequency count and percent:

- Duration (in months) of exposure including delays categorized as 0-3m, > 3m-6m, > 6m-12m, > 12m-24m, > 24m
- Number of infusions administered per subject categorized as 0, 1-4, 5-10, 11-16, 17-22, ≥ 23
- Number of subjects with any incomplete nivolumab infusions
- Number of subjects with interrupted nivolumab infusions

6.5.4. Exposure for Ipilimumab Treatment Component

Exposure for ipilimumab will be summarized as follows:

By descriptive statistics:

- Duration (in months) of exposure including delays calculated as (date of last infusion of ipilimumab – date of first dose of ipilimumab)/30.4375
- Number of infusions administered per subject
- Number of expected infusions per subject
- Cumulative dose
- Average dose (mg/infusion)
- Overall relative dose intensity (%)

By frequency count and percentage for:

- Number of subjects completing ipilimumab treatment
- Number of subjects discontinuing ipilimumab early. Amongst these subject's categorical summary for number of doses received (1, 2, 3)
- Number of subjects continuing ipilimumab. Amongst these subject's categorical summary for number of doses received (1, 2, 3)
- Number of ipilimumab infusions received: 0, 1, 2, 3, 4
- Number of subjects by cycles of ipilimumab administered: 0, 1, 2, 3, 4
- Number of subjects with any incomplete ipilimumab infusions

- Number of subjects with interrupted ipilimumab infusions

6.6. Study Treatment Modifications

Treatment modifications will be summarized separately for each component and will include subjects in the Safety and P-Safety populations. For cabozantinib/placebo treatment components treatment modifications includes holds and reductions, for nivolumab/ipilimumab treatment components they include infusion delays. Treatment modifications only due to AE will be summarized. Note that incomplete or interrupted infusions will be summarized separately and will not be considered as treatment modifications.

6.6.1. Dose Reduction due to Adverse Events

Dose reductions are permitted for cabozantinib or placebo but not for nivolumab or ipilimumab. Two dose reduction levels for the oral study medication are 20 mg qd and 20 mg every other day [qod]. The following summaries for dose reductions due to AE by treatment arm will be presented:

Categorical summaries for:

- Subjects with any dose reduction
- Dose level received by a subject (40 mg, 20 mg qd, 20 mg qod)
- Lowest non-zero dose level received
- Last dose level received (excluding dose holds)
- Last dose level received (including dose holds)

Descriptive statistics for:

- Duration of treatment in months for each dose level received per subject
- Time from first dose to first dose level reduction (days)
- Time from first dose to second dose level reduction (days)

6.6.2. Dose Holds (Delays) due to Adverse Events

Nivolumab or ipilimumab infusion may be delayed. An infusion was considered delayed if the delay exceeded by 3 days for nivolumab/ipilimumab. For cabozantinib/placebo dose was considered as held if the subject did not take the dose of the medication. The following hold/delay summaries due to AE for study treatment, both infusion components due to the same AE and each treatment component will be presented:

- Subjects with any dose delay/hold
- Frequency counts and percentages for subjects with any dose hold, hold ≥ 7 days, ≥ 14 days, ≥ 21 days
- Descriptive statistics for number of dose holds (0 mg dose level) due to an AE per subject and classified as 1, 2, 3, and >3
- Descriptive statistics for total duration of dose holds per subject due to an AE, where duration of each dose hold is calculated as (stop date of hold – start date of hold + 1). Categorical summaries for ≥ 7 days, ≥ 14 days, and ≥ 21 days will also be presented
- Descriptive statistics for duration of each dose holds due to an AE, calculated as (stop date of hold – start date of hold + 1). Categorical summaries for ≥ 7 days, ≥ 14 days, and ≥ 21 days will also be presented
- Descriptive statistics for time to first dose hold (also presented for study treatment), time to first dose hold that was ≥ 7 days, ≥ 14 days, and ≥ 21 days. The time to dose hold is calculated as (start date of the hold – first dose date + 1)
- Descriptive statistics for time to second dose hold, time to second dose hold that was ≥ 7 days, ≥ 14 days, and ≥ 21 days

6.6.3. Dose Modifications due to Adverse Events

For cabozantinib a subject may experience three types of dose modifications (dose reductions to 20 mgs qd or 20 mgs qod, and dose holds) due to an AE. For nivolumab/ipilimumab dose modifications include infusion delays due to AE. Thus, the maximum number of dose modification types a subject may experience is four. The following summaries will be presented per treatment arm:

- Frequency counts and percentages for subjects with any dose modifications
- Descriptive statistics for number of dose modification types (0-4)
- Descriptive statistics for time to the first dose modification
- Descriptive statistics for time to the second dose modification

6.7. Incomplete or Interrupted Infusions

The following summaries will be presented for infusion components and will include subjects in the Safety arm:

- Number of subjects with any incomplete ipilimumab infusions
- Number of subjects with interrupted ipilimumab infusions

- Number of subjects with any incomplete nivolumab infusions
- Number of subjects with interrupted nivolumab infusions

6.8. Study Treatment Non-Compliance and Dosing Errors

Treatment non-compliance and dosing errors for reasons other than AE will be summarized in the Safety population. Frequency counts and percentages will be presented by treatment groups for:

- Subjects with dose hold/delay (0 mg) due to non-compliance
- Subjects who received dose > maximum allowed dose level at any time (40 mg qd for cabozantinib,480 mg per infusion for nivolumab, > 1mg/kg mg per infusion for Ipilimumab)
- Subjects who received non-protocol specified dose level (≤ maximum allowed dose level) at any time due to non-compliance, or due to site/logistic error or other reason

6.9. Non-Protocol Anticancer Therapy (NPACT)

For the purpose of supporting safety evaluations:

Concomitant and subsequent (see definition in Section 4.6) non-radiation NPACT will be summarized by ATC text and WHO Drug base substance preferred name in the Safety population.

For the purpose of supporting efficacy evaluations:

Concomitant and subsequent NPACT will be summarized by treatment group in the PITT and ITT population as follows:

- NPACTS will be summarized by ATC text and WHO Drug based substance preferred name by the following categories: systemic, local, other, and unknown
- Time to first systemic NPACT from randomization date will be summarized by descriptive statistics

6.10. Concomitant and Subsequent Surgeries/Procedures/Radiations/Transfusions

Concomitant and subsequent surgery/procedures that impacted the target lesion(s) (Yes, No, Unknown) will be summarized by treatment group in the ITT, PITT, and Safety populations.

Frequency counts and percentages will be presented for concomitant and subsequent radiation therapy indication, type and site by treatment group in the ITT, PITT, and Safety populations.

Concomitant transfusions will be summarized by transfusion type and treatment group for subjects in the Safety population.

7. EFFICACY ANALYSES

Efficacy summary details for primary, secondary and additional endpoints are described in the below sections.

7.1. Primary Efficacy Endpoint (PFS)

The primary efficacy endpoint is duration of PFS as determined by the BIRC per RECIST 1.1.

7.1.1. Primary Estimand

The primary estimand for PFS is the difference in survival functions between treatment conditions in the duration of radiographic PFS in the targeted population:

- irrespective of whether the assigned study treatment was given
- irrespective of clinical deterioration
- irrespective of whether local radiation was given to bone
- irrespective of surgical resection of non-target lesions
- had systemic non-radiation NPACT not been given
- had surgery to resect target tumor lesions not occurred
- had local radiation to soft tissue not been given
- had not missed 2 or more adequate tumor assessments

Estimands are derived as follows:

Table 7: Estimand Definition and Rationale for Primary Estimand

Estimand Attribute[1]	**Primary Definition**		**Rationale (as needed)**
Population	Subjects randomized into the study intended to include subjects with untreated advanced or metastatic RCC of intermediate or poor risk per IMDC (Section 2.1)		The initial clause "Subjects randomized into the study intended to include…" is included to align the estimand with the application of the ITT principle expected by regulatory reviewers.
Endpoint	Duration of radiographic PFS		
Intercurrent events	**Event**	**Strategy**	**Rationale (as needed)**
	Receipt of non-assigned study treatment	Treatment policy	This strategy applied to align the estimand with the application of the ITT principle expected by regulatory reviewers.

Estimand Attribute[1]	Primary Definition		Rationale (as needed)
	Clinical deterioration	Treatment policy	This strategy applied due to the nature of the primary endpoint, which is radiographic PFS.
	Receipt of local radiation to bone	Treatment policy	This strategy applied as local radiation to bone is typically palliative and does not directly confound radiographic evaluations of soft tissue.
	Surgical resection of non-target tumor lesions	Treatment policy	This strategy applied as target lesions are the primary focus of RECIST 1.1 evaluations.
	Surgical resection of target tumor lesions	Hypothetical	This strategy arises as a consequence of the convention of censoring for these intercurrent events that confound radiographic tumor assessments per RECIST 1.1., and the assumption in the Kaplan-Meier model that censored subjects would have behaved in a manner similar to those not censored, ie. not experienced these intercurrent event(s)
	Receipt of systemic non-protocol anti-cancer medications	Hypothetical	
	Receipt of local radiation to soft tissue for disease under study	Hypothetical	
	Missed 2 or more adequate tumor assessments	Hypothetical	
Population summary	Difference in survival functions between treatment conditions as assessed by stratified log-rank test and characterized by the hazard ratio.		

[1]See Appendix E for estimand terminology

7.1.2. Definition of PFS

For the primary analysis used to characterize the primary estimand, duration of PFS is defined as the time from randomization to the earlier of either the date of radiographic progression per BIRC or the date of death due to any cause:

PFS (months) = (earliest date of progression, death, censoring – date of randomization + 1)/30.4375

7.1.3. Hypothesis

The hypotheses to be evaluated in the analysis of the PFS are:

H_0: $S(t)_{cabozantinib+nivolumab+ipilimumab} = S(t)_{placebo+nivolumab+ipilimumab}$

H_A: $S(t)_{cabozantinib+nivolumab+ipilimumab} \neq S(t)_{placebo+nivolumab+ipilimumab}$

where $S(t)_{cabozantinib+nivolumab+ipilimumab}$ and $S(t)_{placebo+nivolumab+ipilimumab}$ are the survivor functions for PFS for the experimental and control arms, respectively.

7.1.4. Conventions for Analysis

Only adequate tumor assessments (ATAs) will be considered in the determination of radiographic progression and censoring dates. For the purpose of this study, an ATA is defined as one that results in a time point assignment of complete response (CR), partial response (PR), stable disease (SD; non-CR, non-PD), or progression, or not applicable (NA). Note that NA is assigned when a subject does not have any evidence of measurable or non-measurable disease.

The recorded date of radiographic progression is the date of the tumor assessment visit at which progression is declared. If multiple scan dates are associated with a tumor assessment visit, the earliest scan date within the set will be chosen as the progression date.

General censoring rules for the primary analysis of PFS are provided in Table 9 (analysis ID PFS-EP-1).

7.1.5. Primary Analysis

The timing of this analysis is event-driven and will be conducted after at least 249 events have been observed in the PITT population (the first 550 randomized subjects; see Section 4.8.1). It is designed to include progression events as determined by the BIRC per RECIST 1.1 or death whichever occurs earlier. Clinical deterioration determined by the Investigator will not be considered as progression events. If more than 249 events are included in the data cut off for the first 550 subjects, the primary analysis of PFS will include the first 249 events; supportive analysis including all PFS events in this population will also be provided. An additional supportive analysis of all PFS events in the first 440 randomized subjects through the data cutoff will be performed.

The hypothesis testing between experimental arm compared to the control arm will be performed using the stratified log-rank test with α=0.05 level of significance. The stratification factors are as described in Section 2.5, and the values used for analysis will be those recorded in the IxRS.

The median duration of PFS along with the associated 95% CI for each study treatment arm will be estimated using the Kaplan-Meier method. The stratified HR and its 95% CI will be estimated using a Cox proportional-hazard model with treatment group as the independent variable using the same stratification factors as were used for the log-rank test.

If the p-value for the stratified log-rank test is less than the critical value for rejecting the null hypothesis and the HR ($\lambda_{\,cabozantinib + nivolumab + ipilimumab} / \lambda_{\,placebo + nivolumab + ipilimumab}$) is < 1, the null hypothesis of no difference between the two study treatment arms will be rejected and it will be inferred that duration of PFS is greater in the experimental arm compared to the control arm.

The unstratified versions of all the above analyses will also be conducted.

A summary of concordance between the IRC and Investigator determinations of radiographic progressive disease per RECIST 1.1 (rPD) status, and rPD dates will also be provided.

7.1.6. Sensitivity Analyses

All sensitivity analyses will include all subjects in the PITT population. Tabulated summaries of survival times, hazard ratios, and log rank test statistics as well as graphs of survival functions will be presented.

7.1.6.1. Sensitivity Analyses

Two sensitivity analyses (PFS-EP-2 and PFS-EP-3) to support the primary PFS analysis (PFS-EP-1) are shown in Table 8. The two sensitivity analyses evaluate the impact of different assumptions or conditions that potentially influence the estimate of the primary estimand:

- The PFS-EP-2 definition evaluates the influence of potentially inconsistent tumor assessment intervals between arms. For subjects who experience radiographic progression, it assigns the date of the scheduled visit as the event date, rather than the date of recorded progression.
- The PFS-EP-3 definition evaluates the influence of the assessor of radiographic progression and is based upon RECIST 1.1 evaluations by the Investigator rather than the BIRC. This definition is used to assess the additional endpoint PFS per RECIST 1.1 as assessed by the Investigator.

Table 8: Event and Censoring Rules for Primary and Sensitivity Analyses

Estimand	Primary		Primary		Primary	
Analysis type	Primary		Sensitivity		Sensitivity	
Analysis purpose	Primary		Evaluate assessment time bias		Evaluate assessor bias	
Analysis ID	PFS-EP-1		PFS-EP-2		PFS-EP-3	
Analysis name	rPFS per BIRC		Uniform dates		rPFS per Investigator	
Estimand endpoint	Radiographic PD		Radiographic PD		Radiographic PD	
Population	PITT		PITT		PITT	
Situation	**Outcome**	**Date of Outcome**	**Outcome**	**Date of Outcome**	**Outcome**	**Date of Outcome**
Radiographic PD per RECIST 1.1 per BIRC	event	date of recorded PD	event	date of scheduled visit (or next scheduled visit if between visits)	NA	NA
Radiographic PD per RECIST 1.1 per Investigator	NA	NA	NA	NA	event	date of recorded PD
Death	event	date of death	event	date of death	event	date of death
	Intercurrent events (excluding those with Treatment Policy strategy for all estimands)					
Clinical deterioration	NA	NA	NA	NA	NA	NA
Systemic NPACT (medications)	censored	date of last ATA* before first initiation of therapy	censored	date of last ATA* before first initiation of therapy	censored	date of last ATA* before first initiation of therapy
Local NPACT (medications)	NA	NA	NA	NA	NA	NA
Surgical resection of target tumor lesion(s)	censored	date of last ATA* before target lesion resection	censored	date of last ATA* before target lesion resection	censored	date of last ATA* before target lesion resection
Local radiation: to soft tissue for disease under study	censored	date of last ATA* before local radiation of soft tissue for disease under study	censored	date of last ATA* before local radiation of soft tissue for disease under study	censored	date of last ATA* before local radiation of soft tissue for disease under study
	Missing data					
No baseline ATA	censored	date of rand.	censored	date of rand.	censored	date of rand.
≥ 2 consecutive missing scheduled ATA immediately prior to analysis event	censored	date of last ATA* before missing visits	censored	date of last ATA* before missing visits	censored	date of last ATA* before missing visits
	Observation ongoing					
None of the above criteria for event or censoring	censored	date of last ATA*	censored	date of last ATA*	censored	date of last ATA*

ATA, adequate tumor assessment; BIRC, blinded independent review committee; ITT, intent-to-treat; NA, not applicable; PD, progressive disease; NPACT, non-protocol anticancer therapy (medications including radiopharmaceuticals but excluding local radiation); rPD, radiographic progressive disease per RECIST 1.1
For convention for date of recorded PD, see Section 7.1.4.
* or date of randomization if no post-randomization ATA
Blue cells indicate changes from primary analysis.

7.1.6.2. Sensitivity Analyses to Evaluate Impact of Potential Informative Censoring

The four additional analyses (PFS-EP-11, PFS-EP-12, PFS-EP-13, PFS-EP-14) directed at the primary estimand will be conducted to evaluate the impact of potentially informative censoring. These analyses are based on the primary analysis PFS-EP-1 but with selected censored subjects re-classified as events, differentially by treatment arm, as shown in Table 9. These are highly conservative definitions intended to evaluate potential bias. The directions of treatment effects are of key interest rather than the values of inferential statistics.

Table 9: Definitions for Differential Sensitivity Analyses of PFS to Further Evaluate Potentially Informative Censoring (based on PFS-EP-1)

Estimand	Primary							
Analysis type	Sensitivity							
Analysis purpose	Differential analyses to explore potentially informative censoring							
Analysis strategy	Selected censored subjects in primary analysis reclassified as events, differentially by assigned treatment arm, as shown below							
Analysis ID	**PFS-EP-11**		**PFS-EP-12**		**PFS-EP-13**		**PFS-EP-14**	
Censoring category for subjects censored in PFS-EP-1 in PITT Population	**C+N+I**	**N+I**	**C+N+I**	**N+I**	**C+N+I**	**N+I**	**C+N+I**	**N+I**
rPD per INV but not BIRC	At recorded date of first rPD per INV		At recorded date of first rPD per INV		At recorded date of first rPD per INV	At recorded date of first rPD per INV	At recorded date of first rPD per INV	
Discontinued radiographic assessment for reasons other than rPD and no sNPACT	At date of last ATA*		At date of last ATA*		At date of last ATA*		At date of last ATA*	
sNPACT prior to rPD per INV or BIRC			At date of first sNPACT	At date of first sNPACT			At date of first sNPACT	

ATA: adequate tumor assessment; BIRC: blinded independent review committee; Cntrl: control arm; Exp: experimental arm; rPD: radiographic progressive disease per RECIST 1.1; INV: Investigator; sNPACT: systemic non-protocol anti-cancer therapy (medications including radiopharmaceuticals)
* or date of randomization if no post-randomization ATA

7.1.7. Alternative Estimands

Four alternative estimands for PFS arising from changes in strategy for handling some intercurrent events are as defined below:

- Alternative estimand 1 changes the strategy for selected clinical intercurrent events to "composite," resulting in an endpoint that includes as events earlier of either radiographic progression per BIRC, clinical deterioration, receipt of systemic non-protocol anti-cancer medications, receipt of radiation therapy to soft tissue for disease under consideration or surgical resection of target tumor lesions or death.

- Alternative estimand 2 changes the strategy to “composite” only for systemic non-protocol anti-cancer medications, yielding an endpoint that includes as events earlier of either radiographic progression (as well as death) or initiation of systemic NPACT.
- Alternative estimand 3 changes the strategy to “composite” for systemic non-protocol anti-cancer medications and missed tumor assessment yielding an endpoint that includes radiographic progression as well as death despite missed tumor assessments or initiation of systemic NPACT.
- Alternative estimand 4 changes the strategy to “composite” only for missed tumor assessment yielding an endpoint that includes radiographic progression as well as death despite missed tumor assessments.

More details are provided in Table 10.

Table 10: Alternative Estimands

Estimand Attribute [1]	Event	Alternative 1 Definition	Alternative 2 Definition	Alternative 3 Definition	Alternative 4 Definition
Population		Subjects randomized into the study intended to include subjects with untreated advanced or metastatic RCC of intermediate or poor risk per IMDC (see Section 2.1)			
Endpoint		Duration of radiographic and clinical progression-free survival	Duration of radiographic progression or initiation of systemic NPACT	Evaluate potentially informative censoring: Duration of radiographic progression despite missed 2 ATAs or initiation of systemic NPACT	Evaluate potentially informative censoring: Duration of radiographic progression despite missed 2 ATAs
Intercurrent events		**Strategy**	**Strategy**	**Strategy**	**Strategy**
	Receipt of non-assigned study treatment	Treatment policy	Treatment policy	Treatment policy	Treatment policy
	Clinical deterioration	Composite	Treatment policy	Treatment policy	Treatment policy
	Receipt of local radiation to bone	Treatment policy	Treatment policy	Treatment policy	Treatment policy
	Surgical resection of non-target tumor lesions	Treatment policy	Treatment policy	Treatment policy	Treatment policy
	Surgical resection of target tumor lesions	Composite	Hypothetical	Hypothetical	Hypothetical
	Receipt of systemic non-protocol anti-cancer medications	Composite	Composite	Composite	Hypothetical
	Receipt of local radiation to soft tissue for disease under study	Composite	Hypothetical	Hypothetical	Hypothetical
	Missed 2 or more adequate tumor assessments	Hypothetical	Hypothetical	Composite	Composite
Population summary		Difference in survival functions between treatment conditions as assessed by stratified log-rank test and characterized by the hazard ratio.			

Shaded cells differ from primary estimand.

[1]See Appendix E for estimand terminology

7.1.7.1. Supplemental Analyses

Event and censoring definitions are provided in Table 11 for four supplemental analyses PFS-EA1, PFS-EA2, PFS-EA3, PFS-EA4 corresponding to alternate estimands PFS-EA1, PFS-EA2, PFS-EA3, PFS-EA4 for PFS respectively.

Table 11: Event/ Censoring Rules for Supplementary Analyses of Alternative PFS Estimands

Estimand	Alternative 1		Alternative 2		Alternative 3		Alternate 4	
Analysis type	Supplementary		Supplementary		Supplementary		Supplementary	
Analysis purpose	Alternate progression definition		Alternate progression definition		Evaluate potentially informative censoring		Evaluate potentially informative censoring	
Analysis ID	PFS-EA1		PFS-EA2		PFS-EA3		PFS-EA4	
Analysis name	Investigator claims		rPFS or sNPACT		rPFS or sNPACT despite missing ATA		rPFS despite missing ATA	
Estimand endpoint	Radiographic or clinical PD		Radiographic PD or initiation of systemic NPACT		Radiographic PD despite missed 2 ATAs or initiation of systemic NPACT		Radiographic PD despite missed 2 ATAs	
Population	PITT		PITT		PITT		PITT	
Situation	**Outcome**	**Date of Outcome**	**Outcome**	**Date of Outcome**	**Outcome**	**Date of Outcome**	**Outcome**	**Date of Outcome**
Radiographic PD per RECIST 1.1 per BIRC	NA	NA	event	date of recorded PD	event	date of recorded PD	event	date of recorded PD
Radiographic PD per RECIST 1.1 per Investigator	event	date of recorded PD	NA	NA	NA	NA	NA	NA
Death	event	date of death	event	date of death	event	date of death	event	date of death
	Intercurrent events (excluding those with Treatment Policy strategy for all estimands)							
Clinical deterioration	event	date of determination of clinical deterioration	NA	NA	NA	NA	NA	NA
Systemic NPACT (medications)	event	date of first initiation of therapy	event	date of first initiation of therapy	event	date of first initiation of therapy	censored	Date of last ATA before systemic NPACT
Local NPACT (medications)	NA	NA	NA	NA	NA	NA	NA	NA
Surgical resection of target tumor lesion(s)	event	date of last ATA* before target lesion resection	censored	date of last ATA* before target lesion resection	censored	date of last ATA* before target lesion resection	censored	date of last ATA* before target lesion resection
Local radiation: to soft tissue for disease under study	event	date of last ATA* before local radiation	censored	date of last ATA* before local radiation	censored	date of last ATA* before of local radiation	censored	date of last ATA* before of local radiation
	Missing data							
No baseline ATA	censored	date of rand.	censored	date of rand.	censored	date of rand.	censored	date of rand.
≥ 2 consecutive missing scheduled ATA immediately prior to analysis event	censored	date of last ATA* before missing visits	censored	date of last ATA* before missing visits	Event	date of last ATA* before missing visits	Event	date of last ATA* before missing visits
	Observation ongoing							
None of the above criteria for event or censoring	censored	date of last ATA*	censored	date of last ATA*	censored	date of last ATA*	censored	date of last ATA*

ATA, adequate tumor assessment; BIRC, blinded independent review committee; ITT, intent-to-treat; NA, not applicable; PD, progressive disease; NPACT, non-protocol anti-cancer therapy (medications including radiopharmaceuticals but excluding local radiation)
For convention for date of recorded PD, see Section 7.1.4.
* or date of randomization if no post-randomization ATA
Blue cells indicate changes from primary analysis.

7.1.8. Exploratory Analyses

Exploratory analyses of the effect of baseline characteristics, stratification factors, and other variables on PFS may be conducted if necessary, using Cox regression models and subgroup analyses performed employing Kaplan-Meier methods. At the time of the primary analysis, PFS-EP-1 will also be repeated in the ITT population.

7.2. Secondary Efficacy Endpoint

The secondary efficacy endpoint for this study is duration of OS. Formal hypothesis testing is planned for this endpoint provided the null hypothesis of no difference between treatment arms for testing Primary PFS endpoint is rejected.

7.2.1. Estimand for Secondary Efficacy Endpoint of Overall Survival

The estimand for OS is the difference in survival functions between treatment conditions in the duration of survival in the targeted population:

- irrespective of whether the assigned study treatment was given
- irrespective of any subsequent therapy given

Estimands are derived as follows:

Table 12: Estimand Definition and Rationale for Overall Survival

Estimand attribute [1]	**Primary definition**		**Rationale (as needed)**
Population	Subjects randomized into the study intended to include subjects with untreated advanced or metastatic RCC of intermediate or poor risk per IMDC (see Section 2.1)		The initial clause "Subjects randomized into the study intended to include…" is included to align the estimand with the application of the ITT principle expected by regulatory reviewers
Endpoint	Duration of overall survival		
Intercurrent events	**Event**	**Strategy**	**Rationale (as needed)**
	Receipt of non-assigned study treatment	Treatment policy	This strategy applied to align the estimand with the application of the ITT principle expected by regulatory reviewers.
	Receipt of any subsequent therapy	Treatment policy	
Population summary	Difference in survival functions between treatment conditions as assessed by stratified log-rank test and characterized by the hazard ratio.		

[1]See Appendix E for estimand terminology

7.2.2. Definition

Duration of OS (months) = (earliest date of death or censoring – date of randomization + 1)/30.4375.

General censoring rules for the analysis of OS are described below:

- For subjects who are alive at the time of data cutoff but are permanently lost to follow-up, duration of OS will be right censored at the date the subject was last known to be alive.
- Those who withdraw consent from survival follow-up and are alive will be right censored at the date the subject withdrew consent from survival follow-up.
- Subjects alive on or after the data cutoff or those who died after the data cutoff will be right censored at the date of data cutoff.

7.2.3. Primary Analysis

The analysis of OS will include all subjects in the ITT population. Statistical methods used for analysis are as described for the primary endpoint of PFS (see details in Section 7.1.5).

The primary analysis of OS is planned to be conducted after a total of 433 deaths have been observed in the study. Three interim analyses of OS are planned at approximately the 27%, 50% and 75% IFs (see details in Section 7.4.2). If null hypothesis of no differences in OS is rejected at a planned interim OS analysis in favor of the experiment arm, no subsequent testing of OS is planned.

The analysis of OS will be performed amongst subjects in the ITT population provided the primary analysis of PFS is positive.

7.2.4. Exploratory Analysis

Exploratory analyses of the effect of baseline characteristics, stratification factors, and other variables on OS may be conducted if necessary, using Cox regression models and subgroup analyses performed employing Kaplan-Meier methods.

7.3. Control of Type I Error and Timing of Analyses

Inflation of Type 1 error associated with testing multiple endpoints (primary and secondary) will be controlled by applying a hierarchical testing procedure: The Primary PFS endpoint will be tested at 2-sided alpha of 5%. The secondary OS endpoint will be conducted at the same 5% alpha level only if the null hypothesis of no difference between arms in Primary PFS is rejected.

Three interim analyses of OS are planned at the 27%, 50%, and 75% IFs. The timing of the first interim OS analysis will coincide with the timing of the primary PFS analysis and will be conducted provided the null hypothesis for PFS is rejected in favor of the experimental arm. Type 1 error associated with interim analyses will be controlled using a Lan-DeMets O'Brien-Fleming alpha-spending function based upon a 5% total alpha allocation for OS (see details in Section 7.4).

All statistical evaluations of efficacy other than the primary or secondary endpoints will not be adjusted for the Type 1 error and the respective p-values will be considered as descriptive.

7.4. Interim Analyses

7.4.1. PFS Futility Analyses

A non-binding futility analysis of PFS is planned to be performed at the 40% IF (approximately 100 events). The futility boundary based on 100 events is 1.059 using a Lan-DeMets O'Brien-Fleming boundary and may have to be recomputed based on the actual number of events observed. The analysis will be based upon radiographic evaluations per the BIRC. Results will be reviewed by the IDMC as described as below:

- If the results of the futility analysis yield an observed HR for PFS per BIRC of < 1.059 (and in the absence of significant safety concerns), the IDMC will recommend the trial continue.
- If the results of the futility analysis yield an observed HR for PFS per BIRC of ≥ 1.059, the IDMC will refer the analysis results and a recommendation to a Sponsor Executive Committee (composed of individuals not involved in day-to-day trial conduct). The Executive Committee will review the analysis results and IDMC recommendation to take a decision about whether the trial will continue or be declared futile.
- If the trial is declared futile, study actions may include, but are not limited to discontinuing enrollment and/or discontinuing study treatment for study subjects.

The stratified HR and its 95% CI will be estimated using a Cox proportional-hazard model with treatment group as the independent variable.

7.4.2. OS Interim Analyses

Three interim analyses of OS are planned at approximately the 27% (estimated), 50% and 75% IFs based on a total alpha of 5% and are to be conducted only if the null hypothesis between treatment arms to test the Primary PFS endpoint is rejected. The timing of the first interim OS analysis is planned to coincide with the timing of the primary PFS analysis (estimated 27%

information fraction for OS) and will be conducted provided the null hypothesis for PFS is rejected in favor of the experimental arm. Details and boundaries for interim and the final analysis are shown in the Table 13.The actual analyses may include more or fewer events than the target IFs, hence the actual critical values employed at the interim and final analyses of OS will depend upon the actual information fraction at the time of the analyses.

At an interim or final analysis, if the actual IF exceeds the planned IF due to faster than expected accrual or shorter than expected event times and/or event ascertainment lag, the analysis conducted per the information fractions shown in Table 13 will be summarized descriptively as supportive analysis.

Table 13: Details and Boundaries for the Interim and Final Analyses

	Critical Values for Testing OS	
OS alpha total	5.0%	
OS analysis information fraction	Critical p-value	Maximum HR to reject
IA1: 27%* (~117 death events)	0.0000324	0.464
IA2: 50% (~217 death events)	0.003	0.668
IA3: 75% (~325 death events)	0.018	0.77
FA: 100% (~433 death events)	0.044	0.824

* The timing of the first interim OS analysis is planned to coincide with the timing of the primary PFS analysis (estimated 27% information fraction for OS) and will be conducted provided the null hypothesis for PFS is rejected in favor of the experimental arm.
IA, interim analysis; FA, final analysis.

Should the timing of the first interim analysis of OS (IA1) transpire at an information fraction much higher than expected (e.g. due to faster-than-expected deaths or slower-than expected PFS events), the interim analyses of OS will be conducted as follows.

Table 14: Interim Analysis Strategy for OS

OS IA1 IF	Proceed as	Rationale
<30%	perform IA1, IA2 and IA3 as planned	similar conditions to original plan
30 to < 40%	• perform IA1 at time of primary PFS • Conduct IA2 at 65% IF • Do not conduct IA3	Preserves two IA as planned with reasonable intervals between them
40 to < 50%	• perform IA1 at time of primary PFS • Conduct IA2 at 75% IF • Do not conduct IA3	Preserves two IA as planned with reasonable intervals between them
50 to 65%	• perform IA1 at time of primary PFS • Conduct IA2 to 80% IF • Do not conduct IA3	Preserves two IA as planned with reasonable intervals between them
≥65%	• Perform a IA1 at time of primary PFS • Do not conduct IA2 and IA3	IF is high enough that additional interims would not be practical

7.5. Additional Endpoints

Each exploratory endpoint will be analyzed using an appropriate two-sided statistical test without adjustment for multiplicity unless specified otherwise. Statistical results for exploratory endpoints will be considered supportive. Exploratory analyses will be performed on the ITT population unless specified otherwise.

7.5.1. Objective Response Rate (ORR) by BIRC

For each subject, best overall response (BOR) is defined as the best tumor assessment category as determined per RECIST 1.1 by BIRC that occurs through the first overall time point response of PD and prior to any of the censoring events defined for the primary analysis of PFS as described in Section 7.1. Tumor assessment categories are ranked as: confirmed CR, PR, SD, progressive disease (PD) and not evaluable (NE). To be classified as confirmed CR or confirmed PR, confirmation must have occurred on a subsequent visit that is ≥ 28 days after the response was first observed. To be classified as SD, at least one overall time point response of SD must be documented ≥ 56 days after randomization. The ORR is defined as the proportion of subjects with a best response (BOR) of confirmed CR or confirmed PR.

Descriptive testing for ORR will be performed using an unstratified two-sided chi-squared test at 5% level of significance. If a sufficient number of responders are observed, analysis using the

Cochran-Mantel-Haenszel method to adjust for the stratification factors (see Section 2.5) will be generated.

Point estimates of ORR for each study treatment arm, the difference in ORR between the study treatment arms, and associated (1-α) % confidence intervals will be provided. The odds ratio and associated (1-α) % confidence intervals will also be presented. The 2-sided CIs for the point estimate of ORR will be calculated using the Clopper Pearson's method. The 2-sided CIs for the difference in ORR between the two study treatment arms and for the odds ratio will be calculated using asymptotic methods. All summaries for ORR will be provided for the ITT and PITT populations.

Time to objective response is an arithmetic summary amongst those with an objective response and is defined as time from randomization to the first CR or PR that is subsequently confirmed. Time to objective response will be summarized with descriptive statistics.

Waterfall plots displaying maximum percent tumor reduction since baseline in target lesions will be generated for tumor assessment data per BIRC.

These plots will include subjects with a tumor assessment at baseline and at least one post-baseline visit. For each subject, data on or after the progression/censoring date of the respective PFS analyses described in Section 7.1 will be excluded from the waterfall plots.

A summary of the concordance between Investigator and IRC determined objective response status and date will also be provided.

7.5.2. ORR by Investigator

The ORR per RECIST 1.1 by Investigator will be analyzed as described in Section 7.5.1 above, respectively. The ITT and PITT populations will be used.

7.5.3. Duration of Objective Response (DOR)

Duration of objective response (DOR) is defined as the time from the first documentation of objective response that is subsequently confirmed at a visit that is ≥ 28 days later to disease progression or death due to any cause.

Duration of response (months) = (earliest date of PD or death due to any cause or censoring – date of first objective response + 1)/30.4375

The DOR will be computed only among subjects who experience an objective response (confirmed CR or confirmed PR). The DOR will be analyzed and presented separately per BIRC and per Investigator for subjects in the ITT and PITT populations. The dates of progression and

censoring for DOR are shown in Table 8 (see column PFS-EP-1 per BIRC and column PFS-EP-3 per Investigator).

The DOR will be analyzed using the same analysis method (Kaplan-Meier) as for the analysis on PFS (see Section 7.1.5).

7.5.4. PFS by Investigator

The analysis of PFS per RECIST 1.1 per Investigator is described in PFS-EP-3 in Section 7.1.6. and will be conducted on subjects in the ITT and PITT populations.

7.5.5. PFS and ORR per RECIST 1.1 by BIRC according to PD-L1 status

Tumor markers such as PD-L1 will be measured at baseline on tumor specimen samples. Analysis of PFS and ORR per RECIST 1.1 by BIRC will be performed by PD-L1 status ($\geq$ 1%, < 1%, Indeterminate) as defined in Section 7.5.8. The analyses will be conducted by methodology described in Section 7.1.5. Similar analyses may be undertaken for MET status (High/Positive or Low/Negative) using a $\geq$ 50% cutoff of tumor cell membrane staining and intensity score of 2+ or greater. The analysis will be conducted on subjects in ITT and PITT populations.

7.5.6. Pharmacokinetics (PK)

Pharmacokinetics (PK) analyses are outside the scope of this plan. A separate PK analysis plan and PK report will be provided.

7.5.7. Immunogenicity Assessments

Blood samples will be obtained from all subjects in both arms for immunogenicity assessment (anti-drug antibodies [ADA] for nivolumab and ipilimumab) by validated immunoassays predose on W1D1, W14D1, W26D1, and FU-1 and FU-2 visits. Immunogenicity assessments for binding antibody will utilize a tiered based approach. First, the screening assay will be performed for all available samples. Positive screening samples will be further confirmed for the ADA status. Positive samples from the confirmatory assay will be subsequently tested for titer. Neutralizing antibody assay may be applied if deemed necessary. The relationship between nivolumab and/or ipilimumab exposure and the status of ADA (binding) will be explored. A separate ADA analysis plan and ADA report will be provided.

7.5.8. Biomarkers

Biomarkers such PD-L1 expression, MET and other select ones will be summarized. For assessment of PD-L1 expression, evaluation of at least 100 viable tumor cells is required.

Positive staining is defined as ≥ 1% and negative is defined as < 1% of cells that exhibit circumferential and/or partial linear plasma membrane staining at any intensity. High cytoplasmic staining can interfere with membrane scoring and will define as indeterminate. Additionally, failure of PD-L1 staining due to suboptimal tissue sample integrity or assay failure will also be defined as indeterminate.

Baseline and percent changes from baseline for select biomarkers may be summarized. Descriptive statistics may be presented by treatment group using all available data from protocol-defined time-points. Appropriate transformations may be applied to normalize the data for presentation or analysis.

A waterfall plot of best percentage change from baseline in respective biomarker will be presented for each treatment group using the ITT population. In all calculations, best change from baseline will be based only on protocol-defined time-points. Best change is defined as either the largest reduction or the largest increase depending on the directionality of the biomarker being summarized.

7.5.9. Health-Related Quality of Life (HRQOL)

Health-related quality of life (HRQOL) will be assessed by the EuroQol Health questionnaire EQ-5D-5L (Reenen et al 2015). The questionnaire will be self-completed by the subjects at various time points regardless of whether study treatment was given, reduced, held, or discontinued until the date of the last tumor imaging assessment or the study met its primary endpoint and will provide a generic measure of health for clinical appraisal (see protocol Section 5.7.7). The EQ-5D-5L questionnaire has two pages: a descriptive page, which assesses on an increasing severity scale of 1 (no problems with function/no symptoms) – 5 (unable to perform function/extreme symptoms) changes in the following five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The second page has a 0-100 visual analogue scale (VAS),), which records the subjects self-rated health between 100 (best health you can imagine) and 0 (worst health you can imagine) and serves as a quantitative measure of health by the subject (see protocol Appendix D and EQ-5D-5L User Guide 2015).

To compare the two treatment arms for subjects in the ITT population, the following summaries are planned at each time point for each of the 6 questions:

Within each treatment arm:

- Descriptive statistics (number of observations, mean and standard deviation)

- Rate of completion for the questionnaire at each time point. This is defined as total number of subjects who answered all questions on the EQ-5D-5L questionnaire / the expected total number of subjects still on study at the visit
- Mean change from baseline at each time point and the corresponding 95% CI and p-value from one-sample t-test
- Effect size for change from baseline within arm, calculated as (mean of change in score/pooled standard deviation for baseline scores). An effect size greater than ≥ 0.3 will be considered potentially clinically meaningful
- Shift in the severity scale since baseline

Across treatment arms

- The difference in the effect sizes will be presented
- Difference in mean change from baseline for the questions at the time point will be evaluated by a two sample t-test
- Line plots for mean ± standard error and the corresponding mean for change from baseline over time. Data from both treatment arms will be displayed on the same plot. In addition, these plots will also show the average state of the subjects at 3 landmark points, namely, around end of treatment, around progression and around 30 days post-treatment follow-up for the two treatment arms
- Percentage of subjects in Level 1 (no problem) vs. Levels 2-5 (any problems) will be summarized over time
- Percentage of subjects with any problems (Level 2-5) will be compared between the treatment arms using a bar chart

The EQ-5D-5L may be converted into a single index (EQ-Index) value normalized across different countries where the index is validated. See Appendix D for conversion details. For EQ-VAS and EQ-Index, descriptive statistics for change from baseline at each time will be presented. Plots for mean ± standard error and mean change from baseline ± standard error over all time points for the two treatment arms will be generated. In addition, repeated-measures mixed-effects models will be used to explore treatment differences over time. These analyses will include the outcome variable of QOL score change from baseline. The predictors (fixed effects) will be the baseline scores, treatment arms, visit, and randomization strata described in Section 2.5. The individual subject nested within the planned treatment arm will be the random effect. All available data will be included for the analysis. The estimated least squares means for the two treatment arm and their difference, the p-values comparing the 2 treatment arms and the

effect size will be presented. No adjustment will be made for multiple comparisons. An effect size of difference ≥ 0.3 will be considered potentially clinically meaningful. All summaries will be provided for the ITT population.

7.5.10. Health Care Resource Utilization

For this study the following health care resource utilization (HCRU) parameters collected during the study observation period will be summarized for subjects in the ITT population:

- Days of hospitalization due to SAEs
- Days in intensive care unit (ICU) due to SAEs
- Number of emergency room (ER) visits due to SAEs
- Number and type of surgeries
- Number and type of transfusions

The summaries will include:

- Number and percentage of subjects in each category of HCRU
- Descriptive statistics for each HCRU category amongst those subjects who utilized the respective resource
- Total number of days or visits as applicable for each HCRU
- Per person year summary for each HCRU

To calculate the per person year value for a subject for a HCRU parameter, the numerator is the sum of the days or visits for that subject for the parameter; and the denominator is defined as: (safety observation period – date of randomization + 1) / 365.25.

7.6. Subgroups

The following subgroups based on baseline characteristics and stratification factors as reported by the Investigator on the CRF will be explored for primary and secondary efficacy endpoints. Summary tables will be provided for the ITT population. Forest plots will be provided for hazard ratios/odds ratio as applicable.

- Age category
  - < 65 years
  - ≥ 65 years
- Sex
  - Male
  - Female

- Race
  - Asian
  - Black or African American
  - White
  - Rest of the races reported/Not Reported
- Geographic Regions
  - US or Canada or Europe or Australia or New Zealand
  - Latin America or Asia
- ECOG Performance status at baseline:
  - 0
  - 1
  - Missing
- KPS (%) at baseline:
  - 100-90
  - 80-70
  - Missing
- IMDC prognostic score per CRF:
  - Poor
  - Intermediate
- IMDC prognostic score per IxRS:
  - Poor
  - Intermediate
- Current histology types by archival tissue or biopsy (per cancer history CRF):
  - Clear cell
  - Papillary
  - Chromophobe
  - Sarcomatoid
  - Other
- Current disease stage per AJCC criteria (per cancer history CRF):
  - Stage III
  - Stage IV
  - Other
- Prior Nephrectomy
  - Yes
  - No

- Number of organs with target and non-target lesions per BIRC
  - 1
  - 2-3
  - 4-5
  - ≥ 6
- Baseline SoD (mm) of target lesions per BIRC
  - < Median
  - ≥ Median
- Visceral Liver Metastasis at baseline
  - Yes
  - No
- Visceral Metastasis at baseline
  - Yes
  - No
- Bone Metastasis at baseline
  - Yes
  - No
- Tumor PD-L1 status
  - ≥ 1%
  - < 1%
  - Indeterminate
  - Missing

8. SAFETY SUMMARIES

All safety analyses will be performed on data as described in Section 4.5 for subjects in the Safety population. A few summaries may be generated on the P-Safety population. No formal statistical comparison between the two treatments arms is planned.

Safety and tolerability will be assessed by the incidence of treatment emergent-adverse events (TEAEs), changes in laboratory parameters and vital signs from baseline, KPS, and ECOG PS.

An Independent Data Monitoring Committee (IDMC) will monitor safety of the subjects during the study on a regular basis. The committee will operate independently from the Sponsor and the clinical Investigators.

The primary responsibility of the IDMC is to review the accumulating safety data on a regular and an ad hoc basis and make recommendations to the Sponsor regarding the continued conduct

of the study. Safety data will be provided at regular intervals to the IDMC in the form of summary reports or data listings from the Sponsor (blinded) or its designated representative (unblinded).

Details regarding IDMC membership, schedule and format of meetings, format for presentation of data, access to interim data, mode and timing of providing interim reports to the IDMC, and other issues relevant to committee operations are described in the IDMC charter.

The IDMC members will use their expertise, experience, and judgment to evaluate the safety data from the trial and to recommend to the Sponsor whether the trial should continue or be stopped early for safety. No formal statistical rules recommending early stopping for safety are planned.

8.1. Adverse Events

Adverse event terms recorded on the CRF will be mapped to preferred terms and system organ class using the MedDRA. The severity of AEs will be measured by CTCAE version 5.0 (Common Terminology Criteria for Adverse Events [CTCAE] Version 5.0 2009). The Investigator will judge each event to be "not related" or "related" to each study treatment component. Adverse events related to study treatment summaries will include all events that are judged to be related to at least one treatment component in the respective study arms.

A TEAE is defined as any event with an onset date on or after the date of the first dose of study drug or any ongoing event on the date of the first dose of study drug that worsens in severity after the date of the first dose of study drug.

Unless otherwise specified, only TEAEs with an onset date (or date of increase in severity) through the end of the relevant safety observation period (see Section 4.5) will be tabulated in summary tables.

All deaths starting at any time after informed consent, irrespective of when they occur, are recorded in the CRFs as Grade 5 AEs. However, only those associated with an event starting after informed consent through 100 days after the date of the decision to discontinue the last component study treatment (or starting at any time after informed consent and judged to be related to study treatment) require reporting them as an SAE.

For the purpose of calculating treatment emergence and inclusion in summary tables, incomplete onset dates will be imputed as detailed in Appendix A.

An overall summary of adverse events will be provided with the number and percent of subjects who experienced the following types of events during the standard safety observation period

(unless otherwise noted as “at any time” or “during the extended safety observation periods” below) in each study treatment arm. Unless specified explicitly otherwise, the related criteria will be with respect to the study treatment regimen (at least one treatment component in the treatment regimen meeting the criteria) for the below summaries:

- Subjects with a TEAE
- Subjects with a Related TEAE
- Subjects with a Serious TEAE
- Subjects with a Serious Related TEAE
- Subjects with a Worst-Grade 3 or 4 TEAE
- Subjects with a Worst-Grade 3 or 4 Related TEAE
- Subjects with a Worst-Grade 3 TEAE
- Subjects with a Worst-Grade 3 Related TEAE
- Subjects with a Worst-Grade 4 TEAE
- Subjects with a Worst-Grade 4 Related TEAE
- Subjects with a Grade 5 TEAE (all deaths) at any time after first dose date and in the below categories:
    - ❖ ≤ 30 days after last dose of study treatment
    - ❖ ≤ 30 days after last dose of study treatment and not causally related to PD
    - ❖ > 30 days after last dose of study treatment
    - ❖ > 30 days after last dose of study treatment and not causally related to PD
    - ❖ ≤ 100 days after last dose of study treatment
    - ❖ ≤ 100 days after last dose of study treatment and not causally related to PD
    - ❖ > 100 days after last dose of study treatment
- Subjects with a Related Grade 5 TEAE at any time after first dose date and those in the below categories:
    - ❖ ≤ 30 days after last dose of study treatment
    - ❖ ≤ 30 days after last dose of study treatment and not causally related to PD
    - ❖ > 30 days after last dose of study treatment
    - ❖ > 30 days after last dose of study treatment and not causally related to PD
    - ❖ ≤ 100 days after last dose of study treatment
    - ❖ ≤ 100 days after last dose of study treatment and not causally related to PD
    - ❖ > 100 days after last dose of study treatment
- Subjects with a TEAE leading to dose modification (reduction or hold/delay)
- Subjects with a TEAE leading to dose reduction
- Subjects with a TEAE leading to dose hold/delay

- Subjects with TEAE leading to discontinuation:
    - Of at least one treatment component
    - Of all treatment components due to the same AE
    - Of cabozantinib/placebo
    - Of nivolumab
    - Of ipilimumab
- Subjects with Related TEAE leading to discontinuation:
    - Of at least one treatment component
    - Of all treatment components due to the same AE (note: the TEAE has to be related to all treatment components)
    - Of cabozantinib/placebo only
    - Of nivolumab only
    - Of ipilimumab only
- Subjects with TEAE leading to discontinuation and TEAE is not related to treatment and not causally related to disease under study:
    - Of at least one treatment component
    - Of all treatment components due to the same AE
    - Of cabozantinib/placebo only
    - Of nivolumab only
    - Of ipilimumab only
- Subjects with TEAE leading to discontinuation and TEAE is not related to treatment and is causally related to disease under study:
    - Of at least one treatment component
    - Of all treatment components due to the same AE (note: the TEAE has to be causally related to disease under study for all treatment components)
    - Of cabozantinib/placebo only
    - Of nivolumab only
    - Of ipilimumab only
- Subjects with immune-mediated TEAE up to extended safety observation period

The following summaries of AEs will be provided in general for standard observation period for subjects in the Safety population unless indicated otherwise:

Table 15: Adverse Event Summaries

TEAE Included	Displays Sorted By	Format
Subject Incidence per SOC and PT by Severity		
All AEs	SOC and PT (MedDRA standard)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Events with An Incidence of ≥5% (Any Grade) or ≥2% (Grade 3-4) in either arm	SOC and PT (SOC per MedDRA standard, PT Within SOC by Decreasing Difference In Percent For Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Events with An Incidence In The C+N+I Arm of ≥5% (Any Grade) or ≥2% (Grade 3-4) compared to control arm	SOC and PT (SOC per MedDRA standard, PT Within SOC by Decreasing Difference In Percent For Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Subject Incidence of Non-Serious Adverse Event with An Increase of ≥ 5% (Any Grade) in any arm	SOC and PT (SOC per MedDRA standard, PT Within SOC by Decreasing Difference In Percent For Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Related AEs	SOC and PT (MedDRA standard)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
SAEs	SOC and PT (MedDRA standard)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Related SAEs	SOC and PT (MedDRA standard)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Subject Incidence per PT and Severity		
All AEs	PT (Descending Freq. of Any Grade) PT (Descending Frequency of Grade 3-4) PT (Descending Freq. of Difference in Percent Between The Two Arms for Any Grades) PT (Descending Frequency of Difference in Percent Between The Two Arms for Grade3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
All AEs	PT (Descending Frequency of Difference in Any Grade)	Worst severity: Any Grade Odds-ratio, Relative Risk and Risk Difference
All AEs (for label)	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 1, Grade2, Grade 3, Grade 4, Grade 5
Related AEs	PT (Descending Frequency of Any Grade) PT (Descending Frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
SAEs	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Related SAEs	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Leading to Treatment Modifications and Discontinuations		
Leading to Dose Reduction (applicable only for oral components)	PT (Descending Frequency of Any Grade) PT (Descending Frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Leading to Dose Hold/Delay	PT (Descending Frequency of Any Grade) PT (Descending Frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Leading to Dose Modification	PT (Descending Frequency of Any Grade) PT (Descending Frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Leading to Treatment Discontinuation (5 summaries as shown below) • Of At Least One Component • Of All Treatment Components due to The Same AE	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5

TEAE Included	Displays Sorted By	Format
• Of cabo/placebo • Of nivolumab • Of ipilimumab		
Leading to Treatment Discontinuation and Related to Treatment (5 summaries as shown below) • Of At Least One Component • Of All Treatment Components due to The Same AE • Of cabo/placebo • Of nivolumab • Of ipilimumab	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Leading to Treatment Discontinuation and Not Related to Treatment (5 summaries as shown below) • Of At Least One Component • Of All Treatment Components due to The Same AE • Of cabo/placebo • Of nivolumab • Of ipilimumab	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Leading to Treatment Discontinuation and Not Causally Related to Disease Progression (5 summaries as shown below) • Of At Least One Component • Of All Treatment Components due to The Same AE • Of cabo/placebo • Of nivolumab • Of ipilimumab	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Leading to Treatment Discontinuation and Causally Related to Disease Progression (5 summaries as shown below) • Of At Least One Component • Of All Treatment Components due to The Same AE • Of cabo/placebo • Of nivolumab • Of ipilimumab	PT (Descending Frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Grade 5 AEs		
Gr 5 AEs <= 30 days After The Last Dose of Study Treatment	PT (Descending Frequency of Grade 5)	Grade 5, Odds Ratio, Relative Risk, Risk Difference
Gr 5 AEs <= 30 Days After The Last Dose of Study Treatment and Judged Not To Be Causally Related to PD	PT (Descending Frequency of Grade 5)	Grade 5, Odds Ratio, Relative Risk, Risk Difference
Gr 5 AEs <= 100 days After The Last Dose of Study Treatment	PT (Descending Frequency of Grade 5)	Grade 5, Odds Ratio, Relative Risk, Risk Difference
Gr 5 AEs <= 100 Days After The Last Dose of Study Treatment and Judged Not To Be Causally Related to PD	PT (Descending Frequency of Grade 5)	Grade 5, Odds Ratio, Relative Risk, Risk Difference
Gr 5 AEs Related to Study Treatment At Any Time	PT (Descending Frequency of Grade 5)	Grade 5, Odds Ratio, Relative Risk, Risk Difference

TEAE Included	Displays Sorted By	Format
Gr 5 AEs Related to Study Treatment and <=30 Days After The Last Dose of Study Treatment	PT (Descending Frequency of Grade 5)	Grade 5, Odds Ratio, Relative Risk, Risk Difference
Gr 5 AEs Related to Study Treatment and <=100 Days After The Last Dose of Study Treatment	PT (Descending Frequency of Grade 5)	Grade 5, Odds Ratio, Relative Risk, Risk Difference

The following data listings will also be provided with indicators for grade, relationship, seriousness, dose day of the event start/stop, days since last dose, actions taken with study treatment:

- All AEs
- Serious AEs other than Gr 5 AEs
- All AEs that led to treatment discontinuation
- All AESIs

8.2. Deaths

All subject deaths (Grade 5 TEAEs) will be summarized for all subjects in the Safety population.

Deaths will be summarized as follows:

- Number of subjects who died
- Deaths ≤ 30 days after the date of receipt of the last dose of study treatment
- Deaths ≤ 100 days after the date of receipt of the last dose of study treatment

In addition, under each category causality to study disease will also be summarized.

All reported subject deaths will be listed.

8.3. Adverse Events of Special Interest

Adverse events of special interest (AESIs) consist of AEs potentially associated with immune checkpoint inhibitor (ICI) therapies, cases of potential drug induced liver injury (DILI) and suspected transmission of an infectious agent by the study treatment.

Specific AEs may be observed due to treatment with ICIs. Routine and early recognition and management of these events may mitigate severe toxicities. Hence the sponsor has defined a set of preferred terms that are grouped under various AESI categories that will provide a consistent, reproducible, and transparent compilation of safety information over time. The select preferred terms are revisited and updated periodically.

Analysis of potential DILI is discussed in Section 8.7.2 and cases of transmission of infectious agent by study treatment, if any, will be discussed in patient narratives.

The following summaries will be provided for potential adverse events associated with treatment due to ICI therapies for subjects in the Safety population. These summaries will be for such AESIs collected up to 100 days since last dose of study treatment:

Table 16: Adverse Event of Special Interest Summaries

TEAE included	Row-levels (sorted by)	Columns
Subject Incidence by Group, Preferred Term and Severity		
All AESIs	AESI and PT (AESI, PT within AESI by descending frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
AESIs leading to dose reduction/delay: • Of at least one component • Of all components due to the same AESI • Oral components • Of any infusion component	AESI and PT (AESI, PT within AESI by descending frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
AESIs leading to discontinuation • Of at least one component • Of all components due to the same AESI • Oral components • Of any infusion component	AESI and PT (AESI, PT within AESI by descending frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
AESIs leading to initiation of immune-modulating medication	AESI and PT (AESI, PT within AESI by descending frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
AESIs leading to initiation of high dose (defined in Section 6.4) of immune-modulating medication	AESI and PT (AESI, PT within AESI by descending frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
AESIs leading to initiation of immune-modulating medication and any infusion treatment discontinuation	AESI and PT (AESI, PT within AESI by descending frequency Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
All serious AESIs	AESI and PT (AESI , PT within AESI by descending frequency of Any Grade)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Serious AESIs leading to initiation of immune-modulating medication	AESI and PT (AESI, PT within AESI by descending frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Serious AESIs leading to initiation of high dose (defined in Section 6.4) of corticosteroids	AESI and PT (AESI, PT within AESI by descending frequency of Grade 3-4)	Worst severity: Any Grade, Grade 3-4, Grade 4, Grade 5
Time to Onset of Event		
AESIs	Time to first onset of AESIs by Group Term	Any Grade, Grade 3-5
AESIs leading to initiation of immune-modulating medication	Time to first onset of AESIs by Group Term	Any Grade, Grade 3-5

TEAE included	Row-levels (sorted by)	Columns
AESIs leading to initiation of high-dose corticosteroids	Time to first onset of AESIs by Group Term	Any Grade, Grade 3-5
Time to Resolution		
AESIs	Time to Resolution by Group Term	Any Grade, Grade 3-5
AESIs leading to initiation of immune-modulating medication (choose the interval for the longest PT)	Time to Resolution by Group Term	Any Grade, Grade 3-5
AESIs leading to initiation of high-dose corticosteroids (choose the interval for the longest select PT)	Time to Resolution of AESIs by Group Term	Any Grade, Grade 3-5
Recur-after-Rechallenge		
AESIs	Group Term	Any Grade, Grade 3-5
AESIs leading to initiation of immune-modulating medication	Group Term	Any Grade, Grade 3-5
AESIs leading to initiation of high-dose corticosteroids	Group Term	Any Grade, Grade 3-5
Liver test abnormalities (laboratory data) ALT or AST > 3 x ULN ALT or AST > 10 x ULN		

8.4. Time-to Onset for Adverse Events of Special Interest

Time to onset of AESI PTs grouped in special categories since start of the study treatment will be summarized for events observed amongst subjects who experienced at least one AESI. Descriptive statistics will be presented by any grade and grade 3-5 for each treatment arm for summaries shown in Table 15. See algorithm details in Appendix G.

8.5. Time-to Resolution of Adverse Events of Special Interest

Time to resolution for AESIs will be summarized by any grade and grade 3-5 by treatment arms for summaries shown in Table 15 amongst subjects who experienced at least one AESI.

Analysis of time-to resolution will be by Kaplan-Meier method as described in Section 7.1.5. Number, percentage, median and 95% CI of the select PTs will be reported. See algorithm details in Appendix H.

8.6. Challenge/Re-Challenge for Select Adverse Events of Special Interest

Subjects who experience at least one AESI or liver test abnormalities as characterized by ALT or AST and were re-challenged with either the study treatment, cabozantinib, nivolumab or

ipilimumab after its resolution will be summarized by treatment arm for summaries shown in Table 15.

8.7. Laboratory Assessments

8.7.1. Variables

The following treatment-emergent laboratory abnormalities will be summarized (Table 17). A treatment emergent laboratory abnormality is defined as any laboratory abnormality with an onset date on or after the date of the first dose of study drug.

Table 17: Summary of Laboratory Tests

Category	Abnormality	SDTM LBTESTCD	Grading System
Hematology	WBC decreased	WBC	CTCAE
	ANC decreased	NEUT	CTCAE
	Lymphocytes increased Lymphocytes decreased	LYM	CTCAE
	Platelets decreased	PLAT	CTCAE
	Hemoglobin increased Hemoglobin decreased	HGB	CTCAE
Serum chemistry	Albumin decreased	ALB	CTCAE
	ALP increased	ALP	CTCAE
	Amylase increased	AMYLASE	CTCAE
	ALT increased	ALT	CTCAE
	AST increased	AST	CTCAE
	Calcium, corr increased Calcium, corr decreased	CACORR	CTCAE
	Creatinine increased	CREAT	CTCAE
	GGT increased	GGT	CTCAE
	Glucose increased Glucose decreased	GLUC	CTCAE
	LDH increased	LDH	Sponsor
	Lipase increased	LIPASE	CTCAE
	Magnesium increased Magnesium decreased	MG	CTCAE
	Phosphate decrease	PHOS	CTCAE
	Potassium increased Potassium decreased	K	CTCAE
	Sodium increased Sodium decreased	NA	CTCAE

Category	Abnormality	SDTM LBTESTCD	Grading System
	Total bilirubin increased	BILI	CTCAE
Urine chemistry	UPCR increased	PROTCRT	Sponsor
Endocrinology[1]	Thyroid Stimulating Hormone increased Thyroid Stimulating Hormone decreased	TSH	HLN

[1] TSH is held in the SDTM "chemistry" laboratory category; will use HLN = high, low, normal classification based on normal range

Sponsor-defined grades are to be applied to the following analytes:

LDH:

- Grade 1 if > ULN to ≤ 2 × ULN
- Grade 2 if > 2 × ULN to ≤ 3 × ULN
- Grade 3 if > 3 × ULN

UPCR:

- Grade 1 if ≥ 17.0 to ≤ 121.0 mg/mmol (≥ 0.15 to ≤ 1.0 mg/mg)
- Grade 2 if > 121.0 to ≤ 396.0 mg/mmol (>1.0 to < 3.5 mg/mg)
- Grade 3 if > 396.0 mg/mmol (> 3.5 mg/mg)

8.7.2. Analysis

All laboratory data parameters and visits will include flags for values above or below laboratory reference ranges. Toxicity grades will be assigned programmatically by applying the CTCAE v5.0 guidelines. Only results with assessment dates through the end of the standard safety observation period (see Section 4.5) will be tabulated in summary tables.

Laboratory summaries will be presented in SI units. Continuous laboratory test results will be summarized by treatment group using descriptive statistics for actual values and for changes from baseline by scheduled visit. Plots showing mean and standard error at each scheduled visit (with visits shown on x-axis) may also be presented for some laboratory parameters. For test results which are below or above the quantification level, the imputed values as described in Appendix C will be used for deriving the grade and then summarized.

Tables summarizing baseline laboratory values and the incidence of laboratory abnormalities by baseline and maximum post-baseline CTCAE grade over all records will be presented. In addition, the following summaries will also be presented:

A. Liver function abnormalities will be assessed as follows:
   - Shift from baseline based on normal ranges

- Summaries of subjects meeting Hy's Law laboratory screening criteria as shown below:
    - > 3 × ULN (ALT or AST), > 2 × ULN Total Bilirubin, and < 2 × ULN ALP
    - > 3 × ULN (ALT or AST), > 2 × ULN Total Bilirubin, and ≥ 2 × ULN ALP
    - > 3 × ULN (ALT or AST), > 2 × ULN Total Bilirubin

B. For renal failure surveillance, a summary of subjects meeting renal failure laboratory screening criteria as shown below will be provided:
- Serum creatinine ≥ 3.0 × ULN and ≥ 2.0 × baseline value or
- Estimated Glomerular Filtration Rate (eGFR) ≤ 50% of the baseline value or
- eGFR < 30 mL/min/1.73 m^2 and ≥ 25% reduction from the baseline value

  eGFR = 186 x (Creatinine in mmol per L / 88.4)$^{-1.154}$ x (Age)$^{-0.203}$ × (0.742 if female) × (1.210 if black) [from website: https://ukidney.com/nephrology-resources/egfr-calculator]

C. Categorical summaries for baseline status for TSH and Free T4 will be presented. In addition, categorical summaries for post-baseline TSH and Free T4 status among subjects with TSH and Free T4 in the normal range at baseline will also be presented.

For descriptive summaries of change from baseline in continuous laboratory variables and analyses of shift in grade from baseline or worst grade after baseline, all available results will be considered regardless of whether from the central or local lab. See Table 18 for the list of laboratory tests which will be included in the analysis.

Table 18: Laboratory Data to Be Included

Summary Type	Include central lab results?	Include local lab results?
Subject-Incidence of Treatment Emergent Laboratory Abnormalities in Selected Laboratory Tests by CTCAE Grade	Y	Y
Change from Baseline in Laboratory Values	Y	Y
Shift from Baseline in Laboratory Values by CTCAE Grade	Y	Y
Shift from Baseline in Laboratory Values by High/Low/Normal	Y	Y
Shift from Baseline in Laboratory Values by Sponsor-defined Grades	Y	Y
Laboratory Abnormalities with a subject-incidence of ≥5% (All Grades) or ≥2% (Grades 3-4) in either arm	Y	Y
Laboratory Abnormalities with a subject-incidence of ≥5% (All Grades) or ≥2% (Grades 3-4) in the experimental arm	Y	Y
Laboratory Abnormalities with an increase in subject-incidence of ≥5% (All Grades) or ≥2% (Grades 3-4) in the experimental arm	Y	Y

8.8. Vital Signs

8.8.1. Variables

The following vital signs will be summarized:

- Weight
- Systolic blood pressure (mmHg)
- Diastolic blood pressure (mmHg)

8.8.2. Analysis

Subject incidence of clinically meaningful changes since baseline for weight and blood pressure will be presented as shown below:

- Proportion of subjects with weight loss ≥ 10% after first dose
- Subjects with at least 2 post-baseline assessments and who got worse since baseline and met the following blood pressure criteria on 2 or more visits (need not be consecutive) after first dose (JNC criteria were modified to include single measurement per time point when triplicate assessments were unavailable; Chobanian et al 2003):
    - Pre-hypertension: SBP 120-139 mmHg or DBP 80-89 mmHg
    - Stage 1: SBP 140-159 mmHg or DBP 90-99 mmHg
    - Stage 2: (SBP ≥ 160 mmHg and DBP <120 mmHg) or DBP 100-119 mmHg
    - Stage 3: DBP ≥ 120 mmHg

Only results with assessment dates through the end of the standard safety observation period (see Section 4.5) will be considered for the summaries.

8.9. Karnofsky/ECOG performance status

The Karnofsky/ECOG performance status of the subject will be assessed during screening and at each scheduled safety assessment starting on W1D1 (see Appendix F for KPS/ECOG-PS criteria). For the purpose of evaluating safety, Karnofsky/ECOG will be summarized as shift from baseline tables for, at minimum, the worst value recorded after the initiation of study treatment. Only results with assessment dates through the end of the standard safety observation period (see Section 4.5) will be tabulated in summary tables.

Frequencies of ECOG worsening of ≥ +1 and +2 change from baseline to worst value after first dose will also be summarized.

8.10. Electrocardiogram (ECG)

Only results with assessment dates through the end of the standard safety observation period (see Section 4.5) will be considered for summaries. The following categorical summaries will be presented per Investigator and per independent review:

- Number of subjects with triplicate average QTc > 500 ms or increase in triplicate average QTc from baseline of > 60 ms after first dose
    - Number of subjects with triplicate average QTc > 500 ms after first dose
        - Number of cases referred for independent review
        - Number of cases confirmed as QTc > 500 ms by independent review
    - Number of subjects with increase in triplicate average QTc from baseline of > 60 ms after first dose
- Number of subjects with increase in triplicate average QTc from baseline of > 30 ms after first dose

For the above summaries, the most-recent average value from triplicate measurements taken before first dose will be used as baseline. If no triplicate measurements were taken before first dose, the most-recent single or duplicate value taken before first dose will be used as baseline. If > 3 measurements are taken for an assessment, all will be included in the average.

8.11. Impact of COVID-19 Pandemic

Missed visits due to COVID-19 are captured on the CRFs and/or as protocol deviations. Summaries and analyses to describe and/or assess the impact of the COVID-19 pandemic will be included in the clinical study report. These may include tabulations of COVID-19 related protocol deviations and patterns of missing data, summaries of COVID-19 AEs, the addition of a per-protocol population, and comparative analyses of selected endpoint before, during and after the pandemic.

9. IDENTIFICATION AND SUMMARY OF PROTOCOL DEVIATIONS

Important protocol deviations are pre-specified in the study Protocol Deviation Management Plan. In accordance with ICH E3 (1995), important eligibility deviations per the inclusion/exclusion criteria (documented on study CRFs) and important post-randomization protocol deviations (tracked in study clinical trial management system, CTMS) will be identified and listed separately by study center and subject. Important deviations will be summarized for

the ITT population by deviation code (a standardized description e.g. “did not satisfy eligibility criteria” or “received prohibited medication”) and deviation category (a standardized severity classification: “important” or “not important”).

10. DATA QUALITY ASSURANCE

The Clinical, Data Management, Biostatistics, and Medical Writing departments at Exelixis and designees will work diligently and collaboratively to ensure that the data collected and analyzed for this study are of the highest quality. In addition to electronic evaluation of the data and verification of data from source documents at the respective sites, a data review meeting will be held to review the data and correct significant data anomalies before the study database is locked or data are extracted for the purpose of analysis.

11. REFERENCES

Albiges L, Tannir NM, Burotto M, McDermott DF, Plimack ER, Barthélémy P, et al. Nivolumab plus ipilimumab versus sunitinib for first-line treatment of advanced renal cell carcinoma in CheckMate 214: 4-year follow-up and subgroup analysis of patients without nephrectomy. Ann Oncol. 2020;31 (Suppl 4):S559–S560. Abstract 711P.

Chobanian AV, Bakris GL, Black HR, Cushman C, Green LA, Izzo JL, et al. The seventh report of the joint National Committee On Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: The JNC 7 Report. JAMA 2003 May 21; 289(19):2560-72.

Clopper CJ and Pearson ES, The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934, 26(4):404-413.

Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, USDHHS, NIH, NCI; publish date May 28, 2009 (v4.03: June 14, 2010).

FDA Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drug and Biologics (December 2018).

FDA Guidance for Industry: Multiple Endpoints in Clinical Trials (draft, January 2017).

Heng DY, Xie W, Regan MM, Warren MA, Golshayan AR, Sahi C, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. J Clin Oncol. 2009;27(34):5794-9.

Hessel C, Mangeshkar M, Motzer RJ, Escudier B, Powles TB, Schwab G, et al. Evaluation of the novel “trial within a trial” design of METEOR, a randomized phase 3 trial of cabozantinib versus everolimus in patients (pts) with advanced RCC. Ann Oncol. 2016;27(6):266-95. Abstract 8158.

International Conference on Harmonization ICH E3: Structure and Content of Clinical Study Reports (30 Nov 1995).

International Conference on Harmonization ICH E9: Statistical principles for clinical trials (05 February 1998 [Europe], September 1998 [FDA]).

International Conference on Harmonization ICH E9(R1): Statistical principles for clinical trials (20 November 2019).

Reenen MV, Janssen B et al. EQ-5D-5L User Guide, Version 2.1, April 2015.

Appendix A: Date Imputation Rules

Incomplete Cancer Diagnosis Date

If *year* is missing (or completely missing): do not impute

If only *day* is missing: set to 15th of the month.

If *day* and *month* are missing: set to July 1st.

If either imputation rule above results in a diagnosis date > informed consent:

set diagnosis date to the date of informed consent - 1.

Incomplete Adverse Event Onset Date

Assumption: For on-study Adverse Events.

If *year* is missing (or completely missing): set to the date of first dose.

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

If year = year of first dose: set the date to the first dose date.

If *year* < year of first dose: set *month* and *day* to December 31st.

If *year* > year of first dose: set *month* and *day* to January 1st.

If *month* and *year* are present and *day* is missing:

If *year* = year of first dose, and:

If *month* = month of first dose: set *day* to day of first dose.

If *month* < month of first dose: set *day* to last day of *month*.

If *month* > month of first dose: set *day* to 1st day of *month*.

If *year* < year of first dose: set *day* to last day of month.

If *year* > year of first dose: set *day* to 1st day of month.

For all other cases: set to date of first dose.

Do not impute start/stop dates for incomplete prior radiation, prior surgery/procedures, prior transfusions, and prior concomitant medications.

Incomplete Start Date for prior and concomitant medication and prior anticancer therapy.

If *year* is missing (or completely missing): do not impute.

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

Set *month* and day to January 1st.

If *year* and *month* are present and *day* is missing:

Set *day* to 1st day of month.

Incomplete Concomitant Medication End Date

If *year* is missing (or completely missing): do not impute.

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

Set *month* and day to December 31st.

If *year* and *month* are present and *day* is missing:

Set *day* to last day of the month.

Incomplete Concomitant and Subsequent Anticancer Therapy Start Date,

Assumption: Anticancer therapies reported on the Subsequent Anticancer Therapy CRF.

If *year* is missing (or completely missing): set to date of last dose of study treatment + 1

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

If *year* > year of the last dose: Set *month* and *day* to January 1st.

If *year* = year of the last dose: Set *month* and *day* to date of last dose of study treatment + 1

If *year* and *month* are present and *day* is missing:

Set *day* to 1st day of month if the resulting imputed date is greater than date of last dose or if the month is before the month of last dose date and year is same or before the year of the last dose date. Otherwise set the imputed date to date of last dose + 1

Incomplete Death Date

Identify date of last known alive (LA) prior to death from the following:

1. Date of decision to discontinue study treatment from End of Treatment CRF
2. Date of last radiographic assessment from End of Radiographic Follow Up CRF
3. Date last known alive from Survival Follow Up CRF
4. Date of last lab assessment from the Labs dataset

If *year* is missing (or completely missing): set to date of LA + 1

If only *day* is missing: set to the maximum of the first of month or LA + 1

If *month* and *day* are missing:

If *year* of LA = year of death

Set death date to date of LA + 1

If *year* of most-recent contact < year of death

Set *month* and *day* to Jan 1st.

Incomplete Study Treatment Start Date

Define previous sequential dosing "milestone" as the latest of previous dose stop date, previous dose hold stop date, date of first dose or randomization date.

If *year* is missing (or completely missing): set to date of previous sequential dosing "milestone" + 1

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing): set to January 1st

If *year* and *month* are present and *day* is missing: set to the first day of the month

If the imputed date is before the previous sequential dosing "milestone": set to the date of previous sequential dosing "milestone" + 1

Incomplete Study Treatment Stop Date

Define next sequential dosing "milestone" as the earliest of next dose start date, next dose hold start date, date of last dose from EOT CRF or the cutoff date.

If *year* is missing (or completely missing): set to date of next sequential dosing "milestone" - 1

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing): set to December 31st

If *year* and *month* are present and *day* is missing: set to the last day of the month

If the imputed date is after the next sequential dosing "milestone": set to the date of next sequential dosing "milestone" – 1

Appendix B: Rounding Rules for Reported Percentages

For percentages ≥ 10%:

- Values ≥ X.5 or above round to X+1.
- Values > X but < X.5 round to X.

For percentages <10%:

- Values ≥ X.Y5 or above round to X.Y+0.1.
- Values >X.Y but < X.Y5 round to X.Y.

Appendix C: Imputation Rules for Laboratory Values Outside of Quantification Range

- Lab values below the lower level of quantification (LLQ) that are reported as "< LLQ" or "≤ LLQ" in the database will be imputed by LLQ × 0.99 for analysis purposes. However, the original value will also be maintained.

- Lab values above the upper level of quantification (ULQ) that are reported as "> ULQ" or "≥ ULQ" in the database will be imputed by ULQ × 1.01 for analysis purposes. However, the original value will also be maintained.

Appendix D: EQ-5D-5L Index Value Conversion Guidelines

The EQ-index conversion algorithm (EQ-5D-5L User Guide 2.1, April 2015. Available from: http://www.euroqol.org/about-eq-5d/publications/user-guide.html):

- Calculate ***health state***
    - Each of the 5 dimensions comprising the EQ-5D descriptive system is divided into 5 levels of perceived problems:
        - Level 1: indicating no problem
        - Level 2: indicating slight problems
        - Level 3: indicating moderate problems
        - Level 4: indicating severe problems
        - Level 5: indicating extreme problems
    - A unique health state is defined by combining 1 level from each of the 5 dimensions.
      For example, state 11111 indicates no problems on any of the 5 dimensions, while state 12345 indicates no problems with mobility, slight problems with washing or dressing, moderate problems with doing usual activities, severe pain or discomfort and extreme anxiety or depression. Note that missing values will be coded as '9'. Ambiguous values will be treated as missing values.
- ***EQ-index values*** for each country = ***health state * the country specific conversion factors*** for each dimension (EQ-5D-5L Index Value Calculator, version 1)

Appendix E: Estimands Terminology

Source: ICH E9 R1

Estimand Attributes:

#	Estimand Attribute
1	The population, that is, the patients targeted by the scientific question.
2	The variable (or endpoint), to be obtained for each patient, that is required to address the scientific question.
3	The specification of how to account for intercurrent events to reflect the scientific question of interest.
4	The population-level summary for the variable which provides, as required, a basis for a comparison between treatment conditions.

Strategies for Addressing Intercurrent Event(s):

Strategy	Description
Treatment policy	The occurrence of the intercurrent event is irrelevant: the value for the variable of interest is used regardless of whether or not the intercurrent event occurs.
Composite	The occurrence of the intercurrent event is taken to be a component of the variable, i.e. the intercurrent event is integrated with one or more other measures of clinical outcome as the variable of interest.
Hypothetical	A scenario is envisaged in which the intercurrent event would not occur: the value to reflect that scientific question of interest is that which the variable would have taken in the hypothetical scenario defined.
Principal stratum	The target population might be taken to be the principal stratum (see Glossary) in which an intercurrent event would not occur.
While on treatment	Response to treatment prior to the occurrence of the intercurrent event is of interest.

Appendix F: Performance Scales: Karnofsky and ECOG Grades

Karnofsky Status	Karnofsky Grade	ECOG Grade	ECOG Status
Normal, no complaints	100	0	Fully active, able to carry on all pre-disease performance without restriction
Able to carry on normal activities. Minor signs or symptoms of disease	90	0	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light housework, office work
Normal activity with effort	80	1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light housework, office work
Care for self. Unable to carry on normal activity or to do active work	70	1	Ambulatory and capable of all selfcare but unable to carry out any work activities. Up and about more than 50% of waking hours
Requires occasional assistance, but able to care for most of his needs	60	2	Ambulatory and capable of all selfcare but unable to carry out any work activities. Up and about more than 50% of waking hours
Requires considerable assistance and frequent medical care	50	2	Capable of only limited selfcare, confined to bed or chair more than 50% of waking hours
Disabled. Requires special care and assistance	40	3	Capable of only limited selfcare, confined to bed or chair more than 50% of waking hours
Severely disabled. Hospitalization indicated though death non-imminent	30	3	Completely disabled. Cannot carry on any selfcare. Totally confined to bed or chair
Very sick. Hospitalization necessary. Active supportive treatment necessary	20	4	Completely disabled. Cannot carry on any selfcare. Totally confined to bed or chair
Moribund	10	4	Completely disabled. Cannot carry on any selfcare. Totally confined to bed or chair
Dead	0	5	Dead

Appendix G: Time to Onset (TTO) Algorithm

1. Analysis unit: episode
2. Analysis method: KM
3. Analysis population: subgroup of those with the specified PT in a Grouped AE term
4. For time to onset for Grouped AEs of any grade
    a. For each subject identify all records with PTs that are members of the Group
    b. Sort records by subject, Group and AE start date
    c. For each subject select the first PT within each grouping
    d. Aval = (AE start date-treatment start date+1)/30.4375
    e. Consider this as an event (CNSR=0)
    f. Subjects who do not have any events for a grouping are considered censored and their Aval = (end of observation period date-treatment start date+1)/30.4375
5. For time to onset of Grouped AEs of grade 3-5 follows above steps (a) to (f) except (a) is modified to include only AEs of grade 3 to 5

Appendix H: Time to Resolution (TTR) Algorithm

1. Analysis unit: episode
2. Analysis method: KM
3. Analysis population: subgroup of those with the specified PT or Grouped AE term
4. For specific AE preferred terms (PT):
    a. Exelixis AE CRF is one-episode-per page and captures all changes in event characteristics (grade, seriousness, relationship) as "rows" within the same record.
    b. An "episode" is thus defined as:
        i. Start = earliest interval start date within the record
            1. For "Grade 3-5" summaries, start = earliest interval start date within the record where Grade ≥ 3
        ii. Stop = stop date value for the last interval within the record
            1. If missing or event ongoing:
                a. Set to end date of safety observation period
                b. Set status to 'censored'
            2. If not missing
                a. If grade ^= 5 set status to 'event'
                b. If grade= 5 then set status to 'censored'
    c. Compute TTR per Kaplan-Meier (KM) analysis for each PT episode
        i. May include > 1 episode per subject (with same or different verbatim term) with same PT
5. For grouped AEs: Option 1
    a. Identify all records with PTs that are members of the Group; assign each the Grouped AE name.
    b. Compute episode start/stop/status for each PT as above
    c. Consolidate contiguous (overlapping, bookended, or 1-day gap) records into a common episode [easier said than done]
            1. For "Grade 3-5" summaries, start = earliest interval start date within the record where Grade ≥ 3; after consolidation
    d. Compute TTR per KM for each Grouped AE value episode
        i. May include > 1 episode per subject (with same or different PT) with same Grouped AE name

DocuSign

Certificate Of Completion

Envelope Id: PPD
Subject: Please DocuSign: XL184-313 SAP v3.0.docx
Source Envelope:
Document Pages: 84 Signatures: 5
Certificate Pages: 6 Initials: 0
AutoNav: Enabled
EnvelopeId Stamping: Disabled
Time Zone: (UTC-08:00) Pacific Time (US & Canada)

Status: Completed

Envelope Originator:
PPD
PPD
IP Address: PPD

Record Tracking

Status: Original
11-Apr-2022 | 11:33

Holder: PPD
PPD

Location: DocuSign

Signer Events	Signature	Timestamp
PPD Exelixis, Inc. - Part 11 Security Level: Email, Account Authentication (Required)	PPD Signature Adoption: Pre-selected Style Signed by link sent to PPD Signature ID: PPD Using IP Address: PPD With Signing Authentication via DocuSign password With Signing Reasons (on each tab): I approve this document	Sent: 11-Apr-2022 | 11:36 Resent: 11-Apr-2022 | 13:51 Viewed: 11-Apr-2022 | 13:53 Signed: 11-Apr-2022 | 13:54
Electronic Record and Signature Disclosure: Accepted: 11-Aug-2020 | 13:07 ID: PPD		
PPD Exelixis, Inc. Security Level: Email, Account Authentication (Required)	PPD Signature Adoption: Pre-selected Style Signed by link sent to PPD Signature ID: PPD Using IP Address: PPD With Signing Authentication via DocuSign password With Signing Reasons (on each tab): I approve this document	Sent: 11-Apr-2022 | 11:36 Viewed: 11-Apr-2022 | 13:07 Signed: 11-Apr-2022 | 13:07
Electronic Record and Signature Disclosure: Accepted: 28-Jul-2020 | 14:59 ID: PPD		

Signer Events	Signature	Timestamp
PPD Exelixis - Legal Security Level: Email, Account Authentication (Required) **Electronic Record and Signature Disclosure:** Accepted: 21-Dec-2020 | 14:07 ID:PPD	PPD Signature Adoption: Pre-selected Style Signed by link sent to PPD Signature ID: PPD Using IP Address: PPD With Signing Authentication via DocuSign password With Signing Reasons (on each tab): I approve this document	Sent: 11-Apr-2022 | 11:36 Viewed: 11-Apr-2022 | 11:45 Signed: 11-Apr-2022 | 11:46
PPD Exelixis Security Level: Email, Account Authentication (Required) **Electronic Record and Signature Disclosure:** Accepted: 14-Feb-2022 | 14:43 ID: PPD	PPD Signature Adoption: Pre-selected Style Signed by link sent to PPD Signature ID: PPD Using IP Address: PPD With Signing Authentication via DocuSign password With Signing Reasons (on each tab): I approve this document	Sent: 11-Apr-2022 | 11:36 Viewed: 11-Apr-2022 | 11:38 Signed: 11-Apr-2022 | 11:38
PPD Exelixis - Legal Security Level: Email, Account Authentication (Required) **Electronic Record and Signature Disclosure:** Accepted: 25-Feb-2021 | 13:55 ID:PPD	PPD Signature Adoption: Pre-selected Style Signed by link sent to PPD Signature ID: PPD Using IP Address: PPD With Signing Authentication via DocuSign password With Signing Reasons (on each tab): I approve this document	Sent: 11-Apr-2022 | 11:36 Viewed: 11-Apr-2022 | 11:53 Signed: 11-Apr-2022 | 11:54

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	11-Apr-2022 | 11:36
Certified Delivered	Security Checked	11-Apr-2022 | 11:53
Signing Complete	Security Checked	11-Apr-2022 | 11:54
Completed	Security Checked	11-Apr-2022 | 13:54

Payment Events	Status	Timestamps

Electronic Record and Signature Disclosure

Note: Electronic signatures executed in DocuSign are the legally binding equivalent of your traditional hand written signatures and subject to the electronic requirements of 21 CFR Part 11 and Annex 11.

ELECTRONIC RECORD AND SIGNATURE DISCLOSURE

From time to time, Exelixis, Inc. - Part 11 (we, us or Company) may be required by law to provide to you certain written notices or disclosures. Described below are the terms and conditions for providing to you such notices and disclosures electronically through the DocuSign system. Please read the information below carefully and thoroughly, and if you can access this information electronically to your satisfaction and agree to this Electronic Record and Signature Disclosure (ERSD), please confirm your agreement by selecting the check-box next to 'I agree to use electronic records and signatures' before clicking 'CONTINUE' within the DocuSign system.

Getting paper copies

At any time, you may request from us a paper copy of any record provided or made available electronically to you by us. You will have the ability to download and print documents we send to you through the DocuSign system during and immediately after the signing session and, if you elect to create a DocuSign account, you may access the documents for a limited period of time (usually 30 days) after such documents are first sent to you. After such time, if you wish for us to send you paper copies of any such documents from our office to you, you will be charged a $0.00 per-page fee. You may request delivery of such paper copies from us by following the procedure described below.

Withdrawing your consent

If you decide to receive notices and disclosures from us electronically, you may at any time change your mind and tell us that thereafter you want to receive required notices and disclosures only in paper format. How you must inform us of your decision to receive future notices and disclosure in paper format and withdraw your consent to receive notices and disclosures electronically is described below.

Consequences of changing your mind

If you elect to receive required notices and disclosures only in paper format, it will slow the speed at which we can complete certain steps in transactions with you and delivering services to you because we will need first to send the required notices or disclosures to you in paper format, and then wait until we receive back from you your acknowledgment of your receipt of such paper notices or disclosures. Further, you will no longer be able to use the DocuSign system to

receive required notices and consents electronically from us or to sign electronically documents from us.

All notices and disclosures will be sent to you electronically

Unless you tell us otherwise in accordance with the procedures described herein, we will provide electronically to you through the DocuSign system all required notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you during the course of our relationship with you. To reduce the chance of you inadvertently not receiving any notice or disclosure, we prefer to provide all of the required notices and disclosures to you by the same method and to the same address that you have given us. Thus, you can receive all the disclosures and notices electronically or in paper format through the paper mail delivery system. If you do not agree with this process, please let us know as described below. Please also see the paragraph immediately above that describes the consequences of your electing not to receive delivery of the notices and disclosures electronically from us.

How to contact Exelixis, Inc. - Part 11:

You may contact us to let us know of your changes as to how we may contact you electronically, to request paper copies of certain information from us, and to withdraw your prior consent to receive notices and disclosures electronically as follows:
To contact us by email send messages to:PPD

To advise Exelixis, Inc. - Part 11 of your new email address

To let us know of a change in your email address where we should send notices and disclosures electronically to you, you must send an email message to us at PPD and in the body of such request you must state: your previous email address, your new email address. We do not require any other information from you to change your email address.

If you created a DocuSign account, you may update it with your new email address through your account preferences.

To request paper copies from Exelixis, Inc. - Part 11

To request delivery from us of paper copies of the notices and disclosures previously provided by us to you electronically, you must send us an email to PPD and in the body of such request you must state your email address, full name, mailing address, and telephone number. We will bill you for any fees at that time, if any.

To withdraw your consent with Exelixis, Inc. - Part 11

To inform us that you no longer wish to receive future notices and disclosures in electronic format you may:

i. decline to sign a document from within your signing session, and on the subsequent page, select the check-box indicating you wish to withdraw your consent, or you may;

ii. send us an email to PPD and in the body of such request you must state your email, full name, mailing address, and telephone number. We do not need any other information from you to withdraw consent.. The consequences of your withdrawing consent for online documents will be that transactions may take a longer time to process..

Required hardware and software

The minimum system requirements for using the DocuSign system may change over time. The current system requirements are found here: https://support.docusign.com/guides/signer-guide-signing-system-requirements.

Acknowledging your access and consent to receive and sign documents electronically

To confirm to us that you can access this information electronically, which will be similar to other electronic notices and disclosures that we will provide to you, please confirm that you have read this ERSD, and (i) that you are able to print on paper or electronically save this ERSD for your future reference and access; or (ii) that you are able to email this ERSD to an email address where you will be able to print on paper or save it for your future reference and access. Further, if you consent to receiving notices and disclosures exclusively in electronic format as described herein, then select the check-box next to 'I agree to use electronic records and signatures' before clicking 'CONTINUE' within the DocuSign system.

By selecting the check-box next to 'I agree to use electronic records and signatures', you confirm that:

- You can access and read this Electronic Record and Signature Disclosure; and
- You can print on paper this Electronic Record and Signature Disclosure, or save or send this Electronic Record and Disclosure to a location where you can print it, for future reference and access; and
- Until or unless you notify Exelixis, Inc. - Part 11 as described above, you consent to receive exclusively through electronic means all notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you by Exelixis, Inc. - Part 11 during the course of your relationship with Exelixis, Inc. - Part 11.