STATISTICAL ANALYSIS PLAN

Study Protocol Number:	E7080-M001-221
Study Protocol Title:	A single-arm, multicenter, Phase 2 trial to evaluate efficacy and safety of lenvatinib in combination with everolimus in subjects with unresectable advanced or metastatic non clear cell renal cell carcinoma (nccRCC) who have not received any chemotherapy for advanced disease
Date:	11 Oct 2019
Version:	Final Version 2.0

1 TABLE OF CONTENTS

2 LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AE	adverse event
ATC	anatomical therapeutic class
BOR	best overall response
BMI	body mass index
CBR	clinical benefit rate
CI	confidence interval
CR	complete response
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DCR	disease control rate
DOR	duration of response
EAS	evaluable analysis set
ECG	electrocardiograms
ECOG	Eastern Cooperative Oncology Group
FAS	full analysis set
FDA	Food and Drug Administration
IIR	Independent imaging review
LLT	lower level term
MedDRA	Medical Dictionary for Regulatory Activities
nccRCC	non clear cell renal cell carcinoma
NE	not evaluable
NA	not applicable
NYHA	New York Heart Association
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival

Abbreviation	Term
PK	pharmacokinetic
PR	partial response
PT	preferred term
Q1	25th percentile
Q3	75th percentile
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SI	Système International
SOC	system organ class
TEAE	treatment-emergent adverse event
TLG	tables, listings, and graphs
TNM	tumor-node-metastasis
UNK	unknown
WHO	World Health Organization
WHO DD	World Health Organization Drug Dictionary

3 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for Eisai Protocol E7080-M001-221, A single-arm, multicenter, Phase 2 trial to evaluate efficacy and safety of lenvatinib in combination with everolimus in subjects with unresectable advanced or metastatic non clear cell renal cell carcinoma (nccRCC) who have not received any chemotherapy for advanced disease. The focus of this SAP is on the planned primary, secondary and exploratory analyses at the final stage of the study.

3.1 Study Objectives

3.1.1 Primary Objective

The primary objective of the study is to evaluate objective response rate (ORR), as assessed by the investigator, of lenvatinib in combination with everolimus in subjects with unresectable advanced or metastatic nccRCC who have not received any chemotherapy for advanced disease.

3.1.2 Secondary Objectives

The secondary objectives of the study are:

- To assess safety and tolerability of lenvatinib in combination with everolimus
- To evaluate progression-free survival (PFS) as assessed by the investigator
- To evaluate overall survival (OS)
- To assess the pharmacokinetic (PK) profiles of lenvatinib and everolimus during combination therapy in subjects with nccRCC

3.1.3 Exploratory Objectives

The exploratory objectives of the study are:

- To explore ORR as assessed by independent imaging review (IIR)
- To explore PFS as assessed by IIR
- To explore clinical benefit rate (CBR) as assessed by the investigator and by IIR
- To explore disease control rate (DCR) as assessed by the investigator and by IIR
- To explore duration of response (DOR) as assessed by the investigator and by IIR
- To identify and explore tumor and blood biomarkers that correlate with clinical outcomes, including efficacy

- To explore the relationship of population PK derived exposure parameters to biomarker, safety, and efficacy data using a model-based approach

3.2 Overall Study Design and Plan

This is a single-arm, multicenter, Phase 2 study of lenvatinib in combination with everolimus (lenvatinib 18 mg/day + everolimus 5 mg/day) in subjects with unresectable advanced or metastatic nccRCC who have not received any chemotherapy for advanced disease. A Simon's Two-Stage Design has been implemented in 31 enrolled subjects. Sixteen subjects were enrolled in Stage 1 with more than 2 responders as assessed by ORR, so the study proceeded to Stage 2 in which 15 more subjects were enrolled. In the final analysis of 31 subjects, at least 6 responders are required to show a statistically significant improvement of ORR over historical control in the same patient population.

The study consisted of 3 Phases: a Pretreatment Phase (Screening and Baseline Periods), a Treatment Phase (Starting on Cycle 1, Day 1), and a Post-treatment Phase (End of Treatment Visit and survival Follow Up). The study's Schedules of Procedures and Assessments for all study phases are explained in Table 5 of Protocol Amendment 01.

Subjects who discontinued one of the study drugs due to its toxicities might continue to receive the other study drug as long as they demonstrated clinical benefit. Subjects would discontinue both study drugs at the time of documented disease progression, development of unacceptable toxicity, withdrawal of consent, or study termination by the sponsor.

4 DETERMINATION OF SAMPLE SIZE

The sample size was calculated using Simon's Two-Stage Design for the primary endpoint ORR assuming an ORR of 25% from this study versus an historical control of 8% (Hudes et el., 2007). A total of approximately 31 subjects, including 16 in Stage 1, were planned to be enrolled in the study. The actual number of subjects accrued depended on the results from Stage 1. If $\leq$1 responders were observed in the first 16 subjects treated (Stage 1), the enrollment should be stopped. The total sample size would be the actual number of subjects enrolled at the time of termination of the trial. If $\geq$2 responders were observed in Stage 1, the study proceeded to Stage 2 and a total of 31 subjects would be enrolled. At least 6 responders in 31 subjects are required to claim superiority of the study treatment over the historical control. This design yielded a one-sided α, type I error, of 0.0319 and power of 0.8053 in Stages 1 and 2 combined.

5 STATISTICAL METHODS

All descriptive statistics for continuous variables will be reported using mean, standard deviation (SD), median, 25th percentile (Q1), 75th percentile (Q3), minimum and maximum. Categorical variables will be summarized as number (percentage) of subjects.

5.1 Study Endpoints

5.1.1 Primary Endpoint

The Primary efficacy endpoint is **ORR** based on investigator assessment, defined as the proportion of subjects who have a best overall response (BOR) of complete response (CR) or partial response (PR). Confirmation of CR and PR must have been performed ≥4 weeks after response was first documented. Subjects whose overall response cannot be determined will be considered as not having an objective response.

5.1.2 Secondary Endpoints

The secondary efficacy endpoints are:

- **PFS** based on investigator assessments – defined as the time date of first dose of study drug to date of first documentation of disease progression or death, whichever occurs first. The censoring rules are detailed in Section 8.3.
- **OS** – defined as the time from the date of first dose of study drug until date of death from any cause. In the absence of death before data cutoff, subjects will be censored either at the date last known to be alive or the date of data cutoff, whichever came earlier.

5.1.3 Exploratory Endpoints

The exploratory endpoints are:

- **ORR** based on IIR assessment.
- **PFS** based on IIR assessment.
- **CBR** is the proportion of subjects who have a BOR of CR or PR or durable stable disease (SD). SD must be achieved at ≥ 7 weeks after the first study drug to be considered as a best overall response. Durable SD is a subset of SD with the duration of ≥ 23 weeks after the first study drug. The CBR will be determined based on both investigator and IIR assessments.
- **DCR** is the proportion of subjects who have a BOR of CR, PR, or SD. The DCR will be determined based on both investigator and IIR assessments.
- **DOR** is defined as the time from the date that the criteria are met for CR or PR (whichever is recorded first) to the date that progressive disease (PD) is objectively documented or death, whichever occurs first. The start time is the first recorded instance of response, not the date when the response is confirmed. Only responders will be included in this analysis. The DOR will be determined based on both investigator and IIR assessments.

5.2 Study Subjects

5.2.1 Definitions of Analysis Sets

The Full Analysis Set (FAS) includes subjects who received at least one dose of the study drugs. This will be the analysis set for all efficacy and safety evaluations.

The Evaluable Analysis Set (EAS, a subset of the Full Analysis Set) includes all subjects who have both an evaluable baseline tumor assessment and an evaluable postbaseline tumor assessment per investigator, unless subjects are discontinued because of disease progression or toxicity. This will be used for the sensitivity analyses of efficacy for investigator assessments.

An evaluable post-baseline tumor assessment will be determined by a completed overall tumor timepoint response assessment that is not NE. Similarly, an evaluable baseline assessment will have a diameter recorded for target lesions, or a non-target assessment that is not 'NE' or 'NA'.

The PK Analysis Set will include all subjects who received at least one dose of study drug and have evaluable lenvatinib plasma and/or everolimus whole blood concentration data.

5.2.2 Subject Disposition

The number (percentage) of enrolled subjects and screen failures will be summarized. Reasons for screen failure will also be tabulated.

Subject disposition on study treatment (treatment ongoing at data cutoff date, treatment discontinued, and the primary reason for discontinuation) will be summarized. Study status at data cutoff and reason for discontinue from study will also be summarized using data from the survival follow-ups.

Distribution of subjects by sites in the United States will also be summarized.

5.2.3 Protocol Deviations

All protocol deviations will be determined prior to database lock and will be agreed upon by a review of individual subject data. Major protocol deviations will be summarized and listed by each category.

5.2.4 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics for the FAS will be summarized. Demographic summaries will be repeated for the EAS.

Continuous demographic and baseline variables include age, BMI, weight, and height; categorical variables include sex, age group (≤65 years, >65 years), ethnicity and race. Other variables at study entry include: pregnancy test, ECOG performance status, NYHA, TNM

classification (T-Primary Tumor, N-Regional Lymph Nodes, M-Distant Metastasis), Anatomic Stage (I, II, III, IV), time from original diagnosis to first dose, age at diagnosis, location of primary tumor, subtypes of nccRCC (Papillary, Chromophobe, CDC, RMC, Unclassified, Other), number of target lesions and total sum of target lesion diameters.

MEDICAL HISTORY

The number (percentage) of subjects reporting a history of any medical condition and current medical condition, as recorded on the CRF, will be summarized. A subject data listing will also be provided. Medical History will be coded using MedDRA (version 22.0), and summarized on the FAS by System Organ Class (SOC) and Preferred term (PT).

5.2.5 Prior and Concomitant Therapy

All investigator terms for medications recorded in the CRF will be coded to an 11-digit code using the World Health Organization Drug Dictionary (WHO DD) (version WHODDMAR18_HD_B2).

Prior medications will be defined as medications that started before the first dose of study drug. Concomitant medications will be defined as medications that (1) started before the first dose of study drug and were continuing at the time of the first dose of study drug, or (2) started on or after the date of the first dose of study drug up to 28 days after the subject's last dose. Medications started after 28 days of the last dose of study drug will be considered as post-treatment medications.

The number (percentage) of subjects who took prior and concomitant medications will be summarized on the FAS by Anatomical Therapeutic Chemical (ATC) class and WHO DD PT. All medications will be presented in subject data listings.

Frequencies [n (%)] of previous anticancer medication will be displayed on the FAS by ATC class and WHO DD preferred term. In addition, previous anticancer medication will be summarized on the FAS in terms of number of regimens, duration of last regimen, best response for last regimen, time from end of last regimen to first dose, and therapeutic setting (adjuvant / locally advanced / neoadjuvant / metastatic / maintenance / unknown).

The following data on previous radiotherapy will be summarized on the FAS: time from last radiotherapy to first dose, site of previous radiotherapy, progression of tumor lesion at the site since radiotherapy (yes/no), and number of all previous radiotherapy treatments.

The non-pharmacological procedures and palliative radiotherapy will be only presented in subject data listings.

New anticancer medication and procedures during survival follow-up will be summarized on the FAS respectively by ATC class and preferred term, and by SOC and preferred term.

5.2.6 Treatment Compliance

Records of treatment compliance for each subject will be kept during the study. Clinical research associates will review treatment compliance during investigational site visits and at the completion of the study. Treatment compliance will not be summarized since the data will not be entered into the clinical database. However, dose modifications (i.e., reduction/ interruption) will be summarized in Section 5.6.1.

5.3 Data Analysis General Considerations

5.3.1 Pooling of Centers

This study is a multicenter study. Subjects from all centers will be pooled together for all the analyses.

5.3.2 Adjustments for Covariates

Not applicable.

5.3.3 Multiple Comparisons/Multiplicity

The overall type I error is controlled at 0.05 in the hypothesis tests of the primary endpoint. No multiplicity adjustments for the tests of secondary endpoints will be made in this exploratory trial. A two-sided alpha of 0.05 will be used in the tests unless otherwise specified.

5.3.4 Examination of Subgroups

No subgroup analyses are planned for this study.

5.3.5 Handling of Missing Data, Dropouts, and Outliers

Rules for missing dates for AE and concomitant medications are set out in Section 8.5. For PFS analysis, drop-outs will be censored according to the rules described in Section 8.3.

Potential outlier values will be investigated. They will be analyzed as originally reported in the locked database.

5.3.6 Other Considerations

Not applicable.

5.4 Efficacy Analyses

Tumor response data utilized in the efficacy analyses will be obtained from both investigator's assessment of the imaging scans and independent image review. Tumor response will be assessed by RECIST 1.1 criteria (Eisenhauer et al 2009).

5.4.1 Primary Efficacy Analyses

The primary analysis is the comparison of ORR assessed by investigator in the FAS to the rate from historical data. Subjects who do not have a tumor assessment for any reason will be considered non-responders and included in the denominator when calculating the response rate. Confirmation of CR and PR must have been performed ≥4 weeks after response was first documented; full details are described in Section 8.4.

ORR from historical control data is assumed to be 8%. An ORR equal to 25% constitutes a clinically meaningful improvement in ORR for this trial. Hence, the null and alternative hypotheses are set as follows:

H_0: ORR = 8% vs. Ha: ORR ≥ 25%

Simon's Two-Stage Design is used in hypothesis testing for the primary endpoint, ORR. The interim futility analysis after Stage 1 allowed for an early evaluation of efficacy results in order to stop the trial early for futility, when the interim results suggested the trial was unlikely to achieve statistical significance. As detailed in Table 1, the trial would stop early at the interim stage if no more than 1 responder out of 16 subjects was observed. Otherwise, the trial proceeded to the second stage and with a total of 31 subjects recruited. Six responses for the primary efficacy endpoint in the final analysis are required to conclude statistical significance at the second stage.

Table 1 Simon's Two-Stage Design for Interim Futility and Final Analyses of ORR

	Threshold
Interim Analysis	Continue if >=2 responders when n_1=16
Final Analysis	Reject H_0 if >=6 responders when n=31

H_0 = null hypothesis, n = sample size for Stages 1 and 2 subjects combined, n1 = sample size for Stage 1 subjects, ORR = objective response rate

Final Analysis

Since there were more than 1 responder out of the first 16 treated subjects, study proceeded to Stage 2 to include 31 total treated subjects. The final analysis will take place after all subjects have been enrolled and on study treatment for at least 6 months or have discontinued due to disease progression, death or toxicity.

A 2-sided Clopper-Pearson 95% confidence interval (CI) will be constructed for ORR. The percentages for each BOR category (CR, PR, SD, PD, NE/UNK) will be displayed.

SENSITIVITY ANALYSIS

ORR per investigator tumor assessments will also be summarized in EAS at the final analysis.

5.4.2 Secondary Efficacy Analyses

All analyses of secondary efficacy endpoints will be performed in FAS. PFS assessed by investigator will also be summarized in EAS.

The PFS assessed by the investigator and OS will be analyzed using Kaplan–Meier product limit estimates. Median PFS and OS will be presented with two-sided 95% Brookmeyer-Crowley confidence intervals (CIs) if estimable. In addition, the cumulative probability of PFS at 3, 6, and 12 months and the cumulative probability of OS at 6, 12, and 18 months will be reported along with the 95% CIs using Kaplan-Meier product limit method and Greenwood log-log formula. The censoring rules for PFS are detailed in Section 8.3.

The cumulative PFS and OS will be plotted over time.

5.4.3 Other Efficacy Analyses

ORR and PFS based on IIR assessment will be analyzed the same as ORR and PFS per investigator assessments, separately. DCR and CBR for both investigator and IIR assessments will be evaluated using the same method as used for ORR.

DOR for both investigator and IIR assessments will be summarized for subjects with BOR of CR or PR. The censoring rule is the same as that of PFS censoring when applicable. KM estimated quantiles and their 95% CIs, using the generalized Brookmeyer and Crowley method will be provided.

In addition, waterfall plots will be presented on the percentage changes from baseline to post-baseline nadir in sum of diameters of the target lesions per investigator and IIR assessments.

5.5 Pharmacokinetic, Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

5.5.1 Pharmacokinetic Analyses

Individual listings with a summary of plasma concentrations of lenvatinib and a figure containing lenvatinib plasma concentrations versus PK sampling time will be provided by subject in the Pharmacokinetic Analysis Set. Other PK analyses will be described in a separate analysis plan.

5.5.2 Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

Not Applicable.

5.6 Safety Analyses

All safety analyses will be performed based on Full Analysis Set. Safety data will be summarized on an “as treated” basis using descriptive statistics (i.e., n, mean, standard

deviation, median, minimum, maximum for continuous variables; n [%] for categorical variables). Safety variables include treatment-emergent adverse events (TEAEs), clinical laboratory parameters, vital signs, 12-lead ECG results, ECOG. Abnormal values will be flagged.

Study Day 1 for all safety analyses is defined as the date of the first dose of study drug.

5.6.1 Extent of Exposure

The duration of treatment (months), calculated as (Date of last dose – Date of first dose +1)/(365.25/12) will be summarized with descriptive statistics for lenvatinib and everolimus combined and separately.

Following variables will be summarized separately for lenvatinib and everolimus:

- Total cumulative dose (mg) per subject is defined as the sum of all doses actually taken by subject.
- Dose intensity (mg/day) per subject is defined as the total cumulative dose (mg) per subject divided by the total duration of treatment in days.
- Relative dose intensity per subject is defined as 100 times the dose intensity divided by the starting planned dose.
- Number of subjects with dose reductions, dose interruptions and treatment discontinuation.
- Time to first dose reduction (weeks) for applicable subjects.

Subject data listings will be provided for all dosing records, and for the above calculated summary statistics.

5.6.2 Adverse Events

AEs will be graded by the investigators using CTCAE v4.03. The AE verbatim descriptions (investigator terms from the CRF) will be classified into standardized medical terminology using the Medical Dictionary for Regulatory Activities (MedDRA). Adverse events will be coded to the MedDRA (version 22.0) lower level term (LLT) closest to the verbatim term. The linked MedDRA preferred term (PT) and primary system organ class (SOC) are also captured in the database.

A treatment-emergent AE (TEAE) is defined as an AE that emerges during treatment or up to 28 days following last dose of study drug, having been absent at pretreatment (Baseline) or

- Reemerges during treatment or up to 28 days following last dose of study drug, having been present at pretreatment (Baseline) but stopped before treatment, or

- Worsens in severity during treatment or up to 28 days following last dose of study drug relative to the pretreatment state, when the AE is continuous.

Only those AEs that are treatment emergent will be included in summary tables. All AEs, treatment emergent or otherwise, will be presented in subject data listings.

An overview table, including the incidence of and the number of subjects with TEAEs, TEAEs with grade 3 or above, treatment-related TEAEs, serious adverse events (SAEs), deaths, and TEAEs that led to study drug withdrawal, dose reduction, or dose interruption will be provided. Treatment-related events are considered by the investigator to be related to study treatment or with missing assessment of the causal relationship.

The incidence of TEAEs will be reported as the number (percentage) of subjects with TEAEs by SOC, PT and CTCAE grade. A subject will be counted only once within a SOC and PT, even if the subject experienced more than one TEAE within a specific SOC and PT, and the highest grade will be chosen. This table will be repeated by PT and CTCAE grade in descending order of frequency.

The number (percentage) of subjects with treatment-related TEAEs will be summarized by SOC, PT and CTCAE grade. The number (percentage) of subjects with SAEs, treatment-related SAEs, TEAEs leading to discontinuation from study drug, and TEAEs leading to dose modification will be summarized by SOC, PT and CTCAE grade in separate tables.

Separate subject data listings of all AEs, fatal AEs, grade 3 or above AEs, SAEs, AEs leading to discontinuation from study drug, and AEs leading to dose modification will be provided.

A listing of all deaths will be provided, and an overall summary of deaths will be computed: deaths within 28 days of the last dose and death after 28 days of the last dose.

In addition, clinically significant TEAEs will be generated and summary of these clinically significant TEAEs will be provided. Appendix 13.1 includes the list of clinically significant TEAEs.

5.6.3 Laboratory Values

Clinical laboratory tests to be performed include hematology, chemistry and urinalysis. Laboratory results will be summarized as appropriate, using the International System of units (SI). For all quantitative hematology and chemistry parameters listed in

Table 2, the actual value and the change from baseline to each post-baseline visit (with at least 10% of subjects) will be summarized by visit using descriptive statistics.

Laboratory parameters will be categorized according to CTCAE v4.03 grades and shifts from baseline CTCAE grades to worst post-baseline values (high/low separately where applicable) will be assessed.

Table 2 Clinical Laboratory Tests

Category	Parameters
Hematology	Hematocrit, hemoglobin, platelets, RBC count, WBC count with differential (bands, basophils eosinophils, lymphocytes, monocytes, neutrophils), and ANC
Chemistry	
Electrolytes	Bicarbonate, chloride, magnesium, potassium, sodium
Liver function tests	Alanine aminotransferase, alkaline phosphatase, aspartate aminotransferase, direct bilirubin, total bilirubin
Renal function tests	Blood urea/blood urea nitrogen, creatinine
Thyroid function tests[a]	TSH and free T4
Pregnancy test	Serum β-hCG (if urine not tested)
Fasting blood tests	Glucose, lipids (total cholesterol, LDL-C, HDL-C, triglycerides)
Other	Albumin, amylase, calcium, lactate dehydrogenase, lipase, phosphorus, total protein, uric acid
Urinalysis	Glucose, ketones, pH, protein, RBCs, specific gravity
Pregnancy test	Urine β-hCG(if serum not tested)

ANC = absolute neutrophil count, β hCG=beta-human chorionic gonadotropin, HDL-C=high density lipoprotein-C, LDL-C=low density lipoprotein-C, RBC=red blood cells, T4=thyroxine, TSH=thyroid stimulating hormone, WBC=white blood cells.

a: Thyroid function will be assessed at the Screening Visit, at Baseline/C1D1, Day 1 of every cycle (starting at C2) throughout the study, and at the End of Treatment Visit.

5.6.4 Vital Signs

Descriptive statistics for vital signs parameters (i.e., diastolic blood pressure [DBP] and systolic blood pressure [SBP], pulse rate, respiratory rate, temperature, weight) and changes from baseline will be presented for each visit (with at least 10% of subjects). Vital signs will be listed by subject.

Blood pressure will also be summarized using a shift table by visit by categories defined based on CTCAE grades (version 4.03): Grade 1 (Pre-Hypertension: SBP 120—139 mmHg or DBP 80 – 89 mmHg), Grade 2 (Stage 1 Hypertension: SBP 140—159 mmHg or DBP 90—99 mmHg), Grade 3 (Stage 2 Hypertension: SBP ≥ 160 mmHg or DBP ≥ 100 mmHg).

5.6.5 Electrocardiograms

Descriptive statistics for ECG parameters and change from baseline will be presented by visit (with at least 10% of subjects). ECG findings (categorized as normal; abnormal, not clinically significant; and abnormal, clinically significant) will be summarized, and the related shift table (from Baseline to Worst Post-Baseline) will be presented.

5.6.6 Other Safety Analyses

ECOG Performance Status will be provided in a listing and a shift table of baseline ECOG result versus worst post-baseline value will be presented.

5.7 Other Analyses

No other analyses are planned.

5.8 Exploratory Analyses

Exploratory analyses may be conducted as appropriate. Any exploratory analyses that are performed will be appropriate titled and labeled as exploratory and will be clearly distinguished from planned analyses when results are reported in the Clinical Study Report.

6 INTERIM ANALYSES

An interim analysis was planned to be performed by the sponsor after 16 subjects in Stage 1 have completed at least 2 tumor assessments (e.g., Week 16 tumor assessments), unless discontinued due to disease progression, death or toxicity. However, the formal interim analysis was waived and the study continued to Stage 2 because more than one responders were confirmed before the planned interim analysis and the safety profiles of both drugs were acceptable.

7 CHANGES IN THE PLANNED ANALYSES

Treatment duration will be analyzed in months instead of nominal cycles mentioned in the protocol. The number (percentage) of subjects with TEAEs will be summarized by worst CTCAE grade instead of the maximum severity (mild, moderate, or severe by highest CTCAE grade) mentioned in the protocol. Protocol mentioned that the number (percentage) of subjects with TEAEs will also be summarized by relationship to each study drug and combined, however, only the relationship to combined drug will be summarized based on the data collection in the CRF.

Biomarker related analyses are no longer planned for this study, therefore, the Pharmacodynamic Analysis Set will not be defined in the analysis dataset.

8 DEFINITIONS AND CONVENTIONS FOR DATA HANDLING

8.1 Baseline

Baseline is defined as the most recently non-missing value collected before the first dose.

8.2 Day to Month to Year

1 month = 30.4375 days; 1 year = 365.25 days.

8.3 Censoring Rule

The PFS derivation rules in Table 3 follow the publication by the Food and Drug Administration (FDA), "Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics (2007)".

Table 3 Censoring Rules for Analysis of Progression-Free Survival

No.	Situation	Date of Progression or Censoring	Outcome
1	No baseline or post-baseline tumor assessments	Date of first dose	Censored
2	Progression documented between scheduled visits	Date of first radiological PD assessment	Progressed
3	No progression at the time of data cut-off or discontinuation from study treatment	Date of last adequate radiologic assessment prior to or on date of data cut-off or discontinuation from study treatment, or date of the first dose if no adequate post baseline tumor assessment was conducted.	Censored
4	New anticancer treatment started	Date of last adequate radiologic assessment prior to or on date of new anticancer treatment, or date of the first dose if no adequate post baseline tumor assessment was conducted.	Censored
5	Death before first PD assessment	Date of death	Progressed
6	Death between adequate assessment visits*	Date of death	Progressed
7	Death or progression after more than one missed visits**	Date of last adequate radiologic assessment before missed tumor assessments, or date of the first dose if no adequate post baseline tumor assessment was conducted.	Censored

CR = complete response, PD = progressive disease, PR = partial response, SD = stable disease,

* Adequate tumor assessment is a radiologic assessment of CR, PR, SD, non-CR/non-PD or PD.

New anticancer treatment includes palliative radiotherapy.

** More than one missed visits is defined as having either one of the following two durations being longer than 18 weeks - 1 day, which is 125 days (= ((8+1) x 2 x 7) – 1) for subjects on the every 8 week tumor assessment schedule in this study: duration between two consecutive tumor assessments; or duration between the last adequate tumor assessment and death or PD.

The priority of the censoring rules is as follows:

1. If the subject had PD or death within 4 weeks of study drug discontinuation, the following sequence will be applied:

- If a subject did not have baseline tumor assessment (No. 1), the subject will be censored on date of first dose. However, if the subject died within 125 days after first dose and did not receive new anticancer treatment, the date of death will be the PFS event date (not censored).
- If a subject had new anticancer treatment before PD or death (No. 4), the subject will be censored on the date of the last tumor assessment prior to or on the date of new anticancer treatment.
- If a subject missed more than one assessment before PD or death (No. 7), the subject will be censored on the date of the last tumor assessment before PD or death. Note that if a subject is censored by both this criterion and the anticancer treatment criteria, the earliest censoring date will be used.
- Otherwise, if a subject had an event (No. 2, No. 5, or No. 6), the earliest event date will be used.

2. If a subject did not have PD or death, the censoring date will be the earliest censoring date if the subject met multiple censoring criteria (No. 1, No. 3, No. 4, No. 7).

8.4 Best Overall Response (BOR)

Best Overall Response does not consider tumor assessments after 1st documented PD, new anticancer treatment and more than one consecutive missing tumor assessments. BOR is defined according to the following order of operations:

1. If subject does not have a baseline tumor assessment, then BOR=NE / UNK
2. Else confirmed CR will be derived by the following rules:
   - Consecutive CR [CR/NE] CR (>= 28 days between the first CR and last CR)
3. Else if not confirmed CR, the confirmed PR will be determined by the following rules:
   - Consecutive PR [CR/PR/NE] CR/PR (>= 28 days between the first PR and the last CR/PR)
   - Consecutive PR SD CR/PR (>= 28 days between the first PR and the second PR/CR)
4. Else if not confirmed CR or PR, BOR of SD will be determined by the following rules:
   - Consecutive PR SD (>=28 days between PR and SD)
   - Any timepoint with response CR/PR/SD/Non-CR/Non-PD that is 49 days after the first dose date
5. Else if not confirmed CR or PR, or SD, and any timepoint with response PD, then BOR=PD
6. Else if not confirmed CR or PR, or SD or PD, then BOR=NE / UNK

8.5 Missing Date Handling

Missing data will not be imputed unless otherwise stated in the SAP or in the footnote of an output. Missing data will not be included in any analysis unless otherwise specified. When relevant, the number of subjects with missing data will be presented

8.5.1 Adverse Event with Missing Dates

If the missing start and/or end dates of the AE did not indicate that the AE started prior to the study treatment or after the 28th day post last study treatment, it will be classified conservatively as the TEAE (Section 5.6.2). This data handling is just for the determination of TEAEs no imputation will be performed to AE start/end dates in datasets and listings.

8.5.2 Medication with Date Missing or Partially Missing

See Section 5.2.5 for the missing date handling in determination of prior and concomitant medications.

9 PROGRAMMING SPECIFICATIONS

The rules for programming derivations and dataset specifications are provided in a separate document (analysis datasets specification).

10 STATISTICAL SOFTWARE

All statistical analyses will be performed using SAS version 9.3 or higher.

11 MOCK TABLES, LISTINGS, AND GRAPHS

The study TLG shells will be provided in a separate document, which will show the content and format of all tables, listings, and graphs in detail.

12 REFERENCES

Common Terminology Criteria for Adverse Events (CTCAE). Version 4.03. United States Department of Health and Human Services, National Institutes of Health, National Cancer Institute, Washington, DC, USA, June 14, 2010.

Criteria Committee, New York Heart Association. Nomenclature and criteria for diagnosis of diseases of the heart and great vessels. 9th ed. Boston: Little, Brown & Co; 1994. p. 253-6.

Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A, editors. Kidney. In: AJCC Cancer staging manual. 7th ed. New York: Springer, 2010. p. 479-89.

Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.

Oken MM, Creech RH, Torney DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982;5(6):649-55.

13 APPENDICES

13.1 Clinically Significant TEAEs per MedDRA Version 22.0

The following are the Clinically Significant AE groups considered in the analyses.

For Lenvatinib:

1. Arterial TE Events
2. Cardiac Dysfunction
3. Fistula Formation
4. GI Perforation
5. Haemorrhage terms (excl laboratory terms)
6. Hepatotoxicity
7. Hypertension
8. Hypocalcemia
9. Hypothyroidism
10. Palmar plantar erythrodysesthesia
11. Posterior reversible encephalopathy syndrome
12. Proteinuria
13. Torsade de pointes/QT prolongation
14. Renal Events

For Everolimus:

1. Dyslipidaemia
2. Angioedema
3. Interstitial lung disease
4. Rash
5. Stomatitis

13.2 Eastern Cooperative Oncology Group Performance Status

Scale	ECOG Performance Status
0	Fully active, able to carry on all predisease performance without restriction.
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature (eg, light house work, office work).
2	Ambulatory and capable of all self-care, but unable to carry out any work activities. Up and about more than 50% of waking hours.
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours.
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair.
5	Dead.

ECOG = Eastern Cooperative Oncology Group.
Adapted from Oken MM, et al. Am J Clin Oncol. 1982;5:649-55.

13.3 New York Heart Association Cardiac Disease Classification

The New York Heart Association Cardiac Disease Classification provides a functional and therapeutic classification for the prescription of physical activity for heart failure patients based on cardiac functional capacity. Based on NYHA definitions, heart failure subjects are to be classified as follows:

Class	NYHA Status
Class I	Subjects with no limitation of activities; they suffer no symptoms from ordinary activities.
Class II	Subjects with slight, mild limitation of activity; they are comfortable at rest or with mild exertion.
Class III	Subjects with marked limitation of activity; they are comfortable only at rest.
Class IV	Subjects who should be at complete rest, confined to bed or chair; any physical activity brings on discomfort and symptoms occur at rest.

NYHA = New York Heart Association.
Adapted from The Criteria Committee of the New York Heart Association. *Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels*. 9th ed. 1994:253-6.

13.4 Common Terminology Criteria for Adverse Events (v4.03)

The National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE) v4.0 [published 28 May 2009 (v4.03: 14 June 2010)] provides descriptive terminology to be used for adverse event reporting in clinical trials. A brief definition is provided to clarify the meaning of each AE term. To increase the accuracy of AE reporting, all adverse event terms in CTCAE version 4.0 have been correlated with single-concept, Medical Dictionary for Regulatory Activities (MedDRA®) terms.

CTCAE v4.0 grading refers to the severity of the AE. CTCAE grades 1 through 5, with unique clinical descriptions of severity for each AE, are based on this general guideline:

Grade	CTCAE Status
1	Mild: asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.
2	Moderate: minimal, local, or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living (ADL).[a]
3	Severe or medically significant but not immediately life-threatening: hospitalization or prolongation of hospitalization indicated; disabling, limiting self-care ADL.[b]
4	Life-threatening consequences: urgent intervention indicated.
5	Death related to adverse event.

CTCAE = Common Terminology Criteria for Adverse Events.
a: Instrumental ADL refers to preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.
b: Self-care ADL refers to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden.
Adapted from the Cancer Therapy Evaluation Program, NCI. CTCAE v4.0. Available from: http://evs.nci.nih.gov/ftp1/CTCAE/About.html (Accessed 25 Jun 2015).

For further details regarding MedDRA, refer to the MedDRA website at: http://www.meddra.org/. CTCAE v4.03 is available online at: http://evs.nci.nih.gov/ftp1/CTCAE/About.html (Accessed 25 Jun 2015).

13.5 Stage Information for Renal Cell Cancer

http://www.cancer.gov/types/kidney/hp/kidney-treatment-pdq

Definitions of TNM

The TNM Staging System (tumor-node-metastasis) is the most widely used system for cancer staging in the world. The staging system for renal cell cancer is based on the degree of tumor spread beyond the kidney (Bassil B et al 1985; Golimbu M et al, 1986; Robson CJ et al 1969). Involvement of blood vessels may not be a poor prognostic sign if the tumor is otherwise confined to the substance of the kidney. Abnormal liver function test results may be caused by a paraneoplastic syndrome that is reversible with tumor removal, and these types of results do not necessarily represent metastatic disease. Except when computed tomography (CT) examination is equivocal or when iodinated contrast material is contraindicated, CT scanning is as good as or better than magnetic resonance imaging for detecting renal masses (Magnetic Resonance Imaging JAMA 1988)

Definitions of TNM

The American Joint Committee on Cancer has designated staging by TNM classification to define renal cell cancer (Edge SB et al 2010)

Table 1. Primary Tumor (T)[a]

TX	Primary tumor cannot be assessed.
T0	No evidence of primary tumor.
T1	Tumor ≤7 cm in greatest dimension, limited to the kidney.
T1a	Tumor ≤4 cm in greatest dimension, limited to the kidney.
T1b	Tumor >4 cm but not >7 cm in greatest dimension, limited to the kidney.
T2	Tumor >7 cm in greatest dimension, limited to the kidney.
T2a	Tumor >7 cm but ≤10 cm in greatest dimension, limited to the kidney.
T2b	Tumor >10 cm, limited to the kidney.
T3	Tumor extends into major veins or perinephric tissues but not into the ipsilateral adrenal gland and not beyond Gerota fascia.
T3a	Tumor grossly extends into the renal vein or its segmental (muscle containing) branches, or tumor invades perirenal and/or renal sinus fat but not beyond Gerota fascia.TX Primary tumor cannot be assessed.
T3b	Tumor grossly extends into the vena cava below the diaphragm.
T3c	Tumor grossly extends into the vena cava above the diaphragm or invades the wall of the vena cava.
T4	Tumor invades beyond Gerota fascia (including contiguous extension into the ipsilateral adrenal gland).

Table 2. Regional Lymph Nodes (N)[a]

NX	Regional lymph nodes cannot be assessed.
N0	No regional lymph node metastasis.
N1	Metastases in regional lymph node(s).

Table 3. Distant Metastasis (M)[a]

MX	No distant metastasis.
M0	Distant metastasis

Table 4. Anatomic Stage/Prognostic Groups[a]

Stage	T	N	M
I	T1	N0	M0
II	T2	N0	M0
III	T1 or T2	N1	M0
	T3	N0 or N1	M0
IV	T4	Any N	M0
	Any T	Any N	M1

[a]Reprinted with permission from AJCC: Kidney. In: Edge SB, Byrd DR, Compton CC, et al., eds.: AJCC Cancer Staging Manual. 7th ed. New York, NY: Springer, 2010, pp 479-89.

SIGNATURE PAGE

Author:

PPD

11 Oct 2019
Date

Approval:

Date

19 Oct 2019
Date

PPD

Oncology Business Group