STATISTICAL ANALYSIS PLAN

Study Protocol Number:	E7080-J081-116/KEYNOTE 524
Study Protocol Title:	An Open-Label Phase 1b Trial of Lenvatinib Plus Pembrolizumab in Subjects with Hepatocellular Carcinoma
Date:	22 Oct 2019
Version:	Final Version 2.0

Summary of Changes

Revisions to Version 2.0

Date: 22 Oct 2019

Changes	Rationale	Affected Sections
Stated the outline of analysis plan and populations to submit an sNDA/sBLA for accelerated approval (AA) to support the new use of lenvatinib in combination with pembrolizumab for the first-line (1L) treatment of patients with advanced unresectable hepatocellular carcinoma (HCC) not amenable to locoregional treatment.	Updated to clarify the plan for the 1L HCC subjects to support AA applications based on communications with the U.S. Food and Drug Administration (FDA) toward the Type B meeting that was scheduled on 10 Oct 2019.	• Section 3 • Section 5 • Section 5.2.1 • Section 5.4 • Section 5.6
Removed the separate description for primary and secondary endpoints and its analyses by DLT part and Expansion part.	Updated to provide analysis results for primary/secondary endpoints using data from the 1L HCC subjects of this study (regardless of DLT/Expansion part), to support AA applications.	• Section 5.1.1 • Section 5.1.2 • Section 5.4.1
Added the presentation of unconfirmed response, and the summary for depth of tumor shrinkage in target lesions.	Updated to clarify the analyses for AA applications.	• Section 5.4.1
Specified the cutoff date of 31 Oct 2019 for this primary analysis.	Updated to clarify for AA applications.	• Section 3 • Section 5 • Section 5.4.1
Specified the sensitivity analyses based on different populations and the supplementary analysis based on different derivation rule for efficacy endpoints.	Added to support AA applications.	• Section 3 • Section 5 • Section 5.2.1 • Section 5.4.4 • Section 8.2
Added concordance analysis between independent imaging review and investigator.	Added to support AA applications.	• Section 5.4.5
Removed the summary of number of cycles, dose intensity, relative dose intensity, and dose interruption of pembrolizumab.	Updated to align with other pembrolizumab studies.	• Section 5.6.1
Added the summary for clinically significant TEAEs (CSAE) for lenvatinib, and TEAEs of special interest (AEOSI) for pembrolizumab.	Added to support AA applications.	• Section 5.6.3
Defined serious adverse events that develop from 31 days up to 120 days after the subject's last dose of study drug as treatment-emergent SAEs.	Updated to be consistent with protocol amendment 05.	• Section 8.3
Minor editorial changes	To correct errors or clarify text	• Throughout

1 TABLE OF CONTENTS

2 LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AA	accelerated approval
ADA	anti-drug antibody
AE	adverse event
AEOSI	adverse events of special interest
AFP	alpha-fetoprotein
ATC	anatomical therapeutic chemical
BCLC	Barcelona Clinic Liver Cancer
BOR	best overall response
BW	body weight
CI	confidence interval
CR	complete response
CRF	case report form
CSAE	clinically significant adverse event
CSR	clinical study report
CTCAE	common toxicity criteria for adverse events
DCR	disease control rate
DOR	duration of response
DLT	dose-limiting toxicity
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
FDA	Food and Drug Administration
HBcAb	hepatitis B core antibody
HBsAb	hepatitis B surface antibody
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HRQoL	Health Related Quality of Life
IIR	independent imaging review
IV	intravenous
LLT	lower level term
LVEF	left ventricular ejection fraction
MedDRA	Medical Dictionary for Regulatory Activities

Abbreviation	Term
mRECIST	modified RECIST
NE	not evaluable
NYHA	New York Heart Association
ORR	objective response rate
OS	overall survival
PD	pharmacodynamics
PD	progressive disease
PD-L1	PD-1 ligand 1
PFS	progression free survival
PK	pharmacokinetics
PR	partial response
PS	performance status
PT	preferred term
Q3W	every 3 weeks
QD	once daily
QOD	every other day
QT	QT interval
QTc	QT interval corrected for heart rate
RECIST	response evaluation criteria in solid tumor
SAP	statistical analysis plan
SAE	serious adverse event
SD	stable disease
SI	système international
sBLA	supplemental Biologics License Application
sNDA	supplemental New Drug Application
SOC	system organ class
TEAE	treatment-emergent adverse event
TEMAV	treatment-emergent markedly abnormal laboratory value
TLG	tables, listings, and graphs
TNM	tumor-node-metastasis
TTP	time to progression
TTR	time to response
WHO DD	World Health Organization drug dictionary

3 INTRODUCTION

Eisai Protocol E7080-J081-116/KEYNOTE 524 (Study 116/KN524) is a multicenter, open-label Phase 1b study of lenvatinib (E7080) plus pembrolizumab in subjects with advanced/unresectable hepatocellular carcinoma (HCC). As discussed with the U.S. Food and Drug Administration (FDA) toward the Type B meeting (that was scheduled on 10 Oct 2019), Eisai and Merck is planning to submit an sNDA/sBLA for accelerated approval (AA) to support the new use of lenvatinib in combination with pembrolizumab for the first-line (1L) treatment of patients with advanced unresectable hepatocellular carcinoma not amenable to locoregional treatment. Based on communications with the FDA, statistical analyses tables, listings, and graphs (TLGs) will be produced with the data up to the cutoff date of 31 Oct 2019. For the 1L HCC subjects, the first 80 subjects will be used for the primary efficacy evaluation and all 100 subjects will be included in an efficacy sensitivity analysis as well as the safety assessment. A primary Clinical Study Report (CSR) will be prepared for these 1L HCC subjects to support the AA application. The CSR will also include all HCC subjects enrolled into Study 116, including the additional 4 subjects with prior anticancer medication as per the protocol.

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 Study 116/KN524 except for pharmacokinetic, pharmacodynamics, pharmacogenomics, biomarker and Health Related Quality of Life (HRQoL) analysis that will be performed separately from the primary CSR. This SAP is based on the clinical study protocol amendment 05 (dated 17 Jan 2019).

3.1 Study Objectives

3.1.1 Primary Objective

- To evaluate the tolerability and safety for combination of lenvatinib plus pembrolizumab in subjects with hepatocellular carcinoma (HCC).

- (Expansion part) To evaluate objective response rate (ORR) and duration of response (DOR) by modified Response Evaluation Criteria In Solid Tumors for HCC (mRECIST) and RECIST 1.1 based on independent imaging review (IIR)

3.1.2 Secondary Objectives

- (DLT evaluation Part) To evaluate ORR and DOR by mRECIST (based on investigator review and IIR), and by RECIST1.1 based on IIR
- (Expansion Part) To evaluate ORR and DOR by mRECIST based on investigator review
- To evaluate the following efficacy endpoints by mRECIST (based on investigator review and IIR) and RECIST1.1 (based on IIR):
  - Progression-free survival (PFS)
  - Time to Progression (TTP)

- Time to Response (TTR)

- Overall survival (OS)
- To assess the pharmacokinetic (PK) profile of lenvatinib and pembrolizumab
- To detect anti-drug antibodies for pembrolizumab (ADA)

3.1.3 Exploratory Objective

- To evaluate the following efficacy endpoints by mRECIST (based on investigator review and IIR) and RECIST1.1 (based on IIR):
  - Disease control rate (DCR)
  - Clinical benefit rate (CBR)
- To evaluate the impact of treatment on Health Related Quality of Life (HRQoL) for subjects treated using the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30, HCC-specific EORTC QLQ-HCC18 questionnaire and European Quality of Life questionnaire.
  For the subjects in Expansion part added as of the protocol Amendment 03, HRQoL will be evaluated using the EORTC QLQ-C30 and the HCC specific supplement QLQ-HCC18. HRQoL will also be evaluated using the generic EuroQuol five dimension five level (EQ-5D-5L) questionnaire.
- To investigate the relationship between candidate biomarkers and anti-tumor activity of lenvatinib in combination with pembrolizumab:
  - To explore blood and tumor markers (such as PD-L1 expression levels, cytokine and angiogenic factor profiling), and immune cell profiling and evaluate their relationship with clinical outcomes including anti-tumor activity of lenvatinib in combination with pembrolizumab

3.2 Overall Study Design and Plan

This is an open-label Phase 1b study. This study will evaluate the tolerability and safety of lenvatinib in combination with pembrolizumab in subjects with HCC.

This study will begin with lenvatinib 12 mg (body weight [BW] ≥60 kg) or 8 mg (BW <60 kg) /day orally and pembrolizumab 200 mg (every 3 weeks [Q3W], intravenous [IV]) in subjects with HCC. Tolerability of this dose level will be evaluated by dose limiting toxicities (DLTs) during the first cycle (21-day treatment cycle).

In this dose level, 3 subjects will be enrolled first. If 0 or 1 of 3 subjects in a given dose level cohort experiences a DLT, then 3 more subjects will be enrolled into that dose level. If 0 or 1 of 6 subjects in a given dose level cohort experiences a DLT, the dose level will be considered tolerable. Additional 4 subjects can be enrolled without DLT evaluation if further

safety information is considered necessary based on the discussions between the sponsor and investigators.

Enrollment will be interrupted if 2 or more DLTs are observed in this dose level, and after sponsor and investigators' review, enrollment may continue for up to 6 subjects based on the nature and severity of the DLTs. Once 6 subjects are enrolled, after sponsor and investigators' review regarding the nature and severity of the DLTs, an additional 4 subjects (10 subjects in total for DLT evaluation) will be enrolled, and that dose level will be considered to be tolerable if DLT is observed in 3 or less of the 10 subjects in total. An independent medical advisor as third party may be consulted for the review as needed.

In this dose level, at least 3 subjects treated with lenvatinib 12 mg once daily (QD) and pembrolizumab 200 mg Q3W IV (BW≥60 kg) need to be included in the 6 subjects for DLT evaluation. (In case of the 10 subjects for DLT evaluation, at least 5 subjects (BW≥60 kg) treated with lenvatinib 12 mg QD and pembrolizumab 200 mg Q3W IV need to be included.)

Cohort of reduction to lower dose level (lenvatinib) or study discontinuation will be considered, if this dose level (12 mg [BW ≥60 kg] or 8 mg [BW <60 kg] lenvatinib plus 200 mg pembrolizumab) is not tolerable, upon discussions between the sponsor and investigators, and the protocol will be amended as necessary. An independent medical advisor as third party maybe consulted for the consideration as needed.

If there is a potential subject who is not evaluable for DLT (eg, subject who fails to administer ≥75% of the planned dosage of lenvatinib due to a reason other than treatment related toxicity during Cycle 1), the investigator and sponsor will discuss whether or not to include the subject in the DLT Analysis Set. If subject is not evaluable for DLT, then the subject will be replaced.

For determination of the recommended phase 2 dose, all episodes of Grade 3 or 4 thrombocytopenia and neutropenia beyond Cycle 1 will be taken into consideration.

If the dose level is confirmed to be tolerable, an additional (approximately) 20 subjects will be enrolled for consolidation of PK data and safety and efficacy as the **Expansion part**. As of Amendment 03, the Expansion part may be further expanded up to approximately 94 subjects. The decision to expand enrollment will be based on the results of 2 interim analyses that will take place when 20 (6 subjects for DLT evaluation part plus 14 subjects for Expansion part) and 56 subjects (6 subjects for DLT evaluation part plus 50 subjects for Expansion part) have sufficient follow-up to be evaluated for response. The decision to expand enrollment will be assessed by mRECIST based on investigator review. At least 5 subjects (BW≥60 kg) treated with lenvatinib 12 mg QD and at least 5 subjects (BW<60 kg) treated with lenvatinib 8 mg QD will be included in Expansion part.

The study will be conducted in 3 phases: a Pretreatment Phase, a Treatment Phase, and an Extension Phase.

- The **Pretreatment Phase** will last no longer than 28 days and will include a Screening Period and a Baseline Period.
- The **Treatment Phase** consists of the first cycle (21 days) for each subject. The Treatment Phase for each subject ends after completing Cycle 1 of treatment or if they discontinue early. Those subjects who discontinue study treatment in Cycle 1 transition to the Off Treatment (Off-Tx) Visit of the Follow-up Period of the Extension Phase. Those who complete Cycle 1 transition to the Treatment Period of the Extension Phase.
- The **Extension Phase** consists of 2 periods, the Treatment Period and the Follow-up Period. Subjects still receiving study treatment at the end of the Treatment Phase will continue to receive the same study treatment in the Treatment Period of the Extension Phase. Subjects will continue to receive study treatment until disease progression, development of unacceptable toxicity, withdrawal of consent, or sponsor termination of the study. Those subjects who discontinue study treatment transition to the Off-Tx Visit of the Follow-up Period of the Extension Phase.

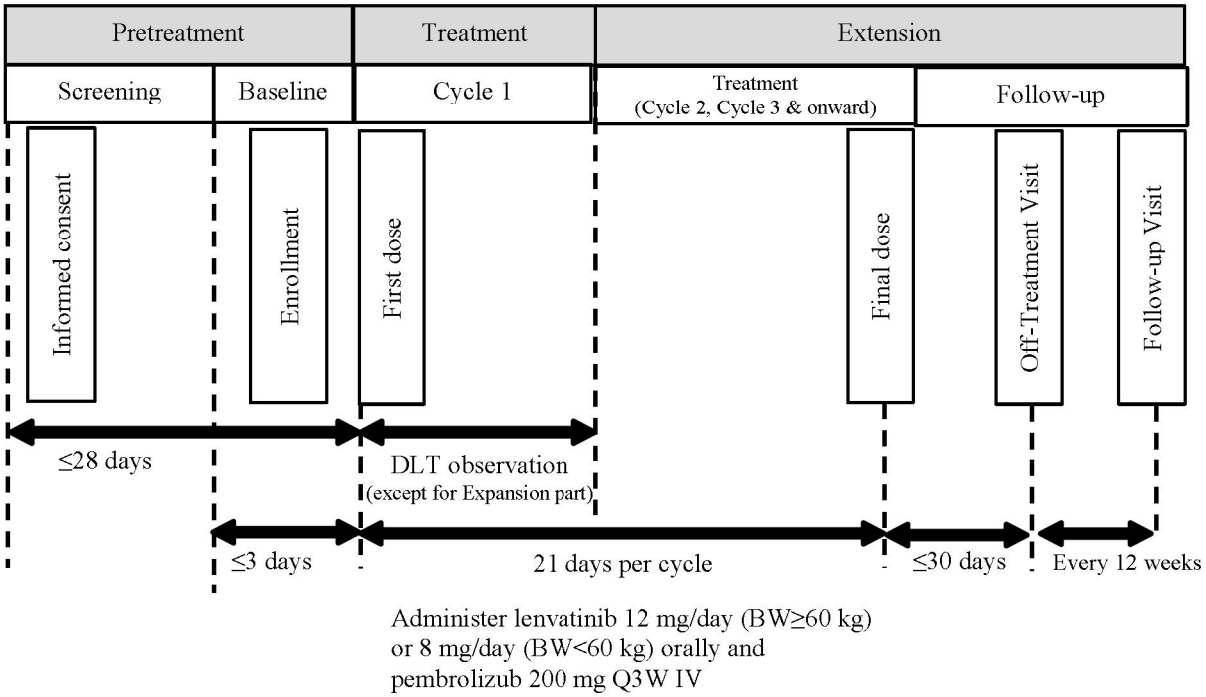

Figure 1 Study Design

4 DETERMINATION OF SAMPLE SIZE

The original sample size in this study was approximately 30 subjects (N=6 to 10 for DLT evaluation part and N=20 for Expansion part). This was not based on statistical power considerations.

The sample size in the Expansion part was 20 evaluable subjects. The associated 2-sided 95% CIs for the ORR of 10% to 90% for 20 subjects are provided in Table 1.

Table 1 2-sided 95% Confidence Interval for the ORR of 10% to 90% (20 subjects)

ORR (N=20)	95% CI
10%	(0.012, 0.317)
20%	(0.057, 0.437)
30%	(0.119, 0.543)
40%	(0.191, 0.639)
50%	(0.272, 0.728)
60%	(0.361, 0.809)
70%	(0.457, 0.881)
80%	(0.563, 0.943)
90%	(0.683, 0.988)

CI = confidence interval, ORR = objective response rate.
95% CI is estimated using Clopper-Pearson method.

As of Amendment 03, the Expansion part may be further expanded up to approximately 94 evaluable subjects. For this expansion decision, 2 interim analyses will take place when 20 (6 subjects for DLT evaluation part plus 14 subjects for Expansion part) and 56 subjects (6 subjects for DLT evaluation part plus 50 subjects for Expansion part) have sufficient follow-up to be evaluated for response. The decision to expand enrollment will be based on the results of 2 interim analyses, which will spend β=0.012 and β=0.024 at the first and second interim analyses, respectively. The decision to expand enrollment will be assessed by mRECIST based on investigator review.

Based on an assumption of H0: 25% ORR and H1: 45% ORR, the 100-subject design with two futility analyses has approximately 96% statistical power at 2-sided α = 0.02 (that corresponds to 1-sided α=0.01). At the first interim analysis (N=20), if there are more than 5 responses, then approximately 36 additional subjects will be enrolled. At the second interim analysis (N=56), if there are more than 16 responses, approximately 44 additional subjects will be enrolled. If there are 5 or fewer responses at the first interim analysis (N=20) or 16 or fewer responses at the second interim analysis (N=56), the sponsor may decide whether to expand enrollment based on clinical outcome (eg, DOR).

The boundaries for the decision to expand enrollment in Expansion part are presented in Table 2.

Table 2 Boundaries for the Decision to Expand Enrollment in the Expansion Part

Analysis Number	Cumulative β Spent	Objective Response Rate	P-value
Interim Analysis 1 (N=20)	0.012	0.258	0.933
Interim Analysis 2 (N=56)	0.024	0.301	0.376
Final Analysis (N=100)	0.041	0.351	0.020

5 STATISTICAL METHODS

Based on communications with the FDA toward the Type B meeting (that was scheduled on 10 Oct 2019), efficacy data up to the cutoff date of 31 Oct 2019 from the first 80 1L HCC subjects (with projected median study follow-up of approximately 12 months, based on 30 Jun 2019 data cutoff) will be the primary data set supporting the AA applications of the combination of lenvatinib plus pembrolizumab for the 1L treatment of patients with advanced unresectable HCC not amenable to locoregional treatment. The data from all 100 1L HCC subjects (with projected median study follow-up of approximately 10.5 months, based on 30 Jun 2019 data cutoff) will be analyzed as a sensitivity analysis for key efficacy endpoints as well as for the overall safety assessment. In addition, the data from the first 60 1L HCC subjects (with projected median study follow-up of approximately 15.5 months, based on 30 Jun 2019 data cutoff) will also be presented for key efficacy endpoints as an efficacy sensitivity analysis to assess the outcome in the subjects with longer duration of follow-up.

Therefore, statistical analyses for the CSR described hereafter will be performed on the 1L HCC subjects to support AA applications, as well as on all HCC subjects including the four subjects with prior anticancer medication as per the protocol, unless otherwise specified. Data from 1L HCC subjects will be presented in a column labeled as "HCC-1L", and data from all HCC subjects (including subjects with prior anticancer medication) will be presented in a column labeled as "All HCC". The analysis tables will minimally have the following headers.

Study subjects (eg, subject disposition, baseline characteristics, etc.) tables

HCC-1L		All HCC		
First 80 (N=80)	All (N=100)	DLT Part (N=6)	Expansion Part (N=98)	Overall (N=104)

Efficacy/Safety tables for the 1L HCC subjects to support AA applications

First 80 HCC-1L (N=80)	All HCC-1L (N=100)	First 60 HCC-1L (N=60)

Efficacy/Safety tables for all HCC subjects (including subjects with prior anticancer medication) as per the protocol

DLT Part (N=6)	Expansion Part (N=98)	Overall (N=104)

The analysis plan for pharmacokinetic, pharmacodynamics, pharmacogenomics, biomarker and HRQoL analysis are described briefly in the corresponding section of this SAP as per the protocol, but the detailed analysis plan and results will be provided in a separate report or a subsequent CSR, once they become available.

All statistical analyses will be performed by the sponsor or designee after the data cutoff for the primary analysis or the study is completed and the database is locked. Statistical analyses will be performed using SAS software or other validated statistical software as required.

All descriptive statistics for continuous variables will be reported using mean, standard deviation (SD), median, 25 percentile (Q1), 75 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 objective is to evaluate the tolerability and safety for combination of lenvatinib plus pembrolizumab in subjects with HCC. Thus, the primary endpoints will be safety related endpoints including DLT.

To evaluate ORR and DOR by mRECIST and RECIST 1.1 based on IIR analysis is also primary objective and primary endpoint.

5.1.2 Secondary Endpoints

The secondary endpoints related to the efficacy endpoints will be ORR and DOR by mRECIST based on investigator review.

The following efficacy endpoints by mRECIST (based on investigator review and IIR) and RECIST1.1 (based on IIR) are also the secondary endpoints in this study:

- PFS
- TTP
- TTR

OS is also the secondary endpoint in this study.

These efficacy endpoints are defined as follows.

- **ORR** is defined as the proportion of subjects who have best overall response (BOR) of complete response (CR) or partial response (PR) at the time of data cutoff.
- **DOR** is defined as the time from the first documentation of CR or PR to the date of first documentation of disease progression or death (whichever occurs first), in subjects with confirmed CR or PR.
- **PFS** is defined as the time from the first study dose date to the date of first documentation of disease progression or death (whichever occurs first).
- **TTP** is defined as the time from the first study dose date to the date of first documentation of disease progression.

- **TTR** is defined as the time from the date of first study dose to the date of first documentation of CR or PR, in subjects with confirmed CR or PR.
- **OS** is measured from the date of first study dose until date of death from any cause. Subjects who are lost to follow-up and the subjects who are alive at the date of data cutoff will be censored at the date the subject was last known alive or the cut-off date, whichever comes earlier.

Responses of PR or CR should be confirmed no less than 4 weeks after the initial response in this study. In order for stable disease (SD) to be considered the BOR, it must occur ≥5 weeks following the first dose of study drug.

Determination of the PK profile of lenvatinib and pembrolizumab while subjects are receiving combination therapy.

Serum ADA will be also measured for pembrolizumab.

5.1.2.1 Pharmacokinetic (PK) Endpoints

PK parameters derived by non-compartmental analysis using plasma concentrations of lenvatinib which include, but are not limited to, are shown as below:

C_{max}	maximum observed concentration
t_{max}	time at which the highest drug concentration occurs
$AUC_{(0\text{-}t)}$	area under the concentration-time curve from zero time to time of last quantifiable concentration
$AUC_{(0\text{-}ti)}$	area under the concentration-time curve from zero (pre-dose) to a given sampling time (t_i)
$AUC_{(0\text{-}inf)}$	area under the concentration-time curve from zero time extrapolated to infinite time
$t_{1/2}$	terminal elimination phase half-life
CL/F	apparent total clearance following oral dosing
V_z/F	apparent volume of distribution at terminal phase
$C_{ss,max}$	maximum observed concentration at steady state
$C_{ss,min}$	minimum observed concentration at steady state
$t_{ss,max}$	time at which the highest drug concentration occurs at steady state
$AUC_{(0\text{-}\tau)}$	area under the concentration-time curve over the dosing interval
CL_{ss}/F	apparent total clearance following oral administration at steady state
$C_{ss,av}$	average steady-state concentration
$R_{ac}(C_{max})$, $R_{ac}(AUC)$	accumulation index

PTF ratio	peak-trough fluctuation ratio

5.1.3 Exploratory Endpoints

The exploratory endpoints related to the efficacy endpoints will be DCR and CBR. These efficacy endpoints are defined as follows.

- **DCR** is defined as the proportion of subjects who have BOR of CR or PR or SD (minimum duration from Cycle 1 Day 1 [C1D1] to SD ≥5 weeks).
- **CBR** is defined as the proportion of subjects who have BOR of CR or PR or durable SD (duration of SD ≥23 weeks).

For the subjects in Expansion part added as of Protocol Amendment 03, HRQoL will be assessed using EORTC QLQ-C30, the HCC-specific questionnaire (HCC-18), and a generic instrument EQ-5D-5L.

It is also the exploratory objective to investigate the relationship between candidate biomarkers and clinical outcomes including anti-tumor activity of lenvatinib in combination with pembrolizumab. Exploratory endpoints will be blood and tumor markers (such as PD-L1 expression levels, cytokine and angiogenic factor profiling), and immune cell profiling.

5.2 Study Subjects

Summary tables will be formatted with the column labels as outlined above in the beginning of Section 5.

5.2.1 Definitions of Analysis Sets

DLT Analysis Set will include all subjects (except for the Expansion part) who have completed Cycle 1 without major protocol deviation with at least 75% of study drug compliance and are assessed for DLT, and subjects who have experienced DLT during Cycle 1. This will be the analysis set to determine tolerability.

Safety Analysis Set and **Efficacy Analysis Set** will include all subjects who received at least 1 dose of study drug. Safety Analysis Set will be used for safety evaluations. Efficacy Analysis Set will be used for efficacy evaluations, as well as for demographic and baseline characteristics, etc.

These analysis sets will be applied to the following populations depending on the purpose, as explained above in the beginning of Section 5.

FOR THE 1L HCC SUBJECTS TO SUPPORT AA APPLICATIONS

Efficacy Analysis Set in:

- First 80 HCC-1L subjects (N=80; for primary analysis)
- All HCC-1L subjects (N=100; for sensitivity analysis)

- First 60 HCC-1L subjects (N=60; for sensitivity analysis)

Safety Analysis Set in:

- First 80 HCC-1L subjects (N=80; for summary of extent of exposure only)
- All HCC-1L subjects (N=100; for all safety analyses)

FOR ALL HCC SUBJECTS (INCLUDING 4 SUBJECTS WITH PRIOR ANTICANCER MEDICATION) AS PER THE PROTOCOL

Efficacy Analysis Set/Safety Analysis Set in:

- Subjects from DLT part (N=6)
- Subjects from Expansion part (N=98)
- Overall (N=104)

PK Analysis Set includes all subjects who have received at least 1 dose of lenvatinib and pembrolizumab, and have evaluable concentration data.

The number (percentage) of subjects included in each analysis set will be presented. Subject data listings will be provided.

5.2.2 Subject Disposition

For the summary table of screening subjects, the number (percentage) of subjects who signed informed consent, continued in the study after screening, failed screening, and the primary reason for screen failures will be presented.

The number (percentage) of subjects who were treated will be provided by country and site.

For the summary table of the subject disposition on study treatment, the number (percentage) of subjects who were treated, continued or discontinued study treatment at the time of study cutoff date, along with the primary reason for the study treatment discontinuation will be presented. The number (percentage) of subjects who discontinued treatment but were on survival follow-up at the time of data cutoff will also be provided.

For the summary table of the subject disposition on study, the number (percentage) of subjects who were on study or off study (ie, discontinued study) and the primary reason for study discontinuation will be presented.

Subject data listings will be provided.

5.2.3 Protocol Deviations

Major protocol deviation criteria will be established and subjects with major protocol deviations will be identified and documented before the database lock. The major protocol

deviations identified according to the criteria at study entry and during treatment will be listed and summarized.

5.2.4 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics will be summarized using descriptive statistics. This summary table will be generated on the Efficacy Analysis Set.

Continuous demographic and baseline variables include

- Age
- Height
- Weight

Categorical demographic and baseline variables include

- Age group (<65 years, ≥65 to <75 years, ≥ 75 years)
- Weight group (<60kg, ≥60 kg)
- Sex (Male, Female)
- Race (White, Black or African American, Asian [Japanese, Chinese, Other Asian], American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Country (France, Italy, Japan, Russian Federation, Spain, United Kingdom, United States)
- ECOG-PS (0, 1, 2)
- NYHA (I, II, III, IV and NA for Heart failure)
- BCLC stage (0, A, B, C, D)
- Extrahepatic spread (ES) (Yes, No)
- Macroscopic Portal Vein Invasion (MPVI) (Yes, No [Vp0])
  - Portal vein involvement (Vp1, Vp2, Vp3, Vp4)
- Macroscopic vascular invasion (MVI) (Yes, No)
- Other macroscopic vascular invasion than portal vein involvement (Yes, No)
- MPVI, ES, or Both (Yes, No)

Disease history and characteristics at study entry will also be summarized by:

- Viral tests at screening
  - HCV Ab (Positive / Negative)
  - HBs Ag (Positive / Negative)

The following would also be evaluated when HBs Ag was negative.

- HBs Ab (Positive / Negative)
- HBc Ab (Positive / Negative)

The following would also be evaluated when HBs Ag Positive, HBs Ab Positive or HBc Ab Positive.

- HBV DNA

- Ammonia test at screening
- Child-Pugh score (5, 6, 7, 8, 9)
- TNM staging and each classification (The categories are as in CRF)
- Diagnosis of HCC (Histological or cytological diagnosis of HCC, Clinically confirmed diagnosis of HCC)
- Factor of Carcinogenesis (Hepatitis B, Hepatitis C, Alcohol, Unknown, Other)
- HCC type (Fibrolamellar, Scirrhous, Spindle cell variant, Pleomorphic type, Clear cell type, Trabecular, Multinodular Differentiated, Moderately Differentiated, Poorly Differentiated, Well Differentiated, Biopsy Performed – HCC Type Unknown, Biopsy not Performed, Other)
- Underlying cirrhosis based on independent imaging review (Yes, No)
- Involved disease sites (Liver, Lung, Lymph Nodes, Bone, Other)
- Involved disease sites per subject (1, 2, ≥3)
- Baseline serum alpha-fetoprotein (AFP) level
- Baseline serum alpha-fetoprotein (AFP) level group (<200 ng/mL, ≥200 ng/mL; <400 ng/mL, ≥400 ng/mL)
- Time since confirmed diagnosis of HCC
- Time since last disease progression

Prior anticancer medications will be summarized by:

- Number of prior regimen (0, 1, 2, 3)
- Type of previous regimen (Adjuvant, Neoadjuvant, Metastatic, Locally advanced, Maintenance, Unknown)
- Best response for last regimen (CR, PR, SD, PD, NE, NA, unknown)
- Duration of last regimen (months)
- Time from end of last regimen to the date of first dose of study drug (months)

Prior radiotherapy will be summarized by:

- Subjects with any prior radiotherapy
- Site of previous radiotherapy (The category depends on the actual data)
- Time from last radiotherapy to the date of first dose of study drug (months)
- Lesion progressed since last radiotherapy (Yes, No, Not evaluated)

Subject data listings for demographic and other baseline characteristics will be provided.

Prior anticancer procedures will be summarized by:

- Number of previous procedures (0, 1, 2, 3, 4, ≥5)
- Time from end of last procedure to the date of first dose of study drug (months)

MEDICAL HISTORY

The number (percentage) of subjects reporting a history of any medical condition, as recorded on the CRF, will be summarized for each part and overall by Medical Dictionary for Regulatory Activities (MedDRA) preferred term (PT) and primary system organ class (SOC).

A subject data listing of medical history and current medical condition will be provided.

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 2018Mar) drug codes. Prior therapies will be defined as therapies that started before the first dose of study drug regardless if they were either stopped before the first dose of study drug or continued during the study. Concomitant therapies will be defined as therapies 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 30 days after the subject's last dose.

The following summary table will be presented on the Efficacy Analysis Set.

The number (percentage) of subjects with prior anticancer medications will be summarized by Anatomical class (ATC Level 1), Pharmacological class (ATC Level 3) and WHO DD preferred term.

The number (percentage) of subjects with prior anticancer procedures will be summarized by MedDRA SOC and PT.

The number (percentage) of subjects with prior medications will be summarized by Anatomical class (ATC Level 1), Pharmacological class (ATC Level 3) and WHO DD preferred term. The number (percentage) of subjects with concomitant medications (excluding antihypertensive, antidiarrheal medications, and corticosteroids for systemic use), concomitant antihypertensive medications, concomitant antidiarrheal medications, and

concomitant corticosteroids for systemic use will also be summarized in the same manner. Concomitant corticosteroids for systemic use will be presented in a subject data listing.

The number (percentage) of subjects with prior (or concomitant) non-pharmacological procedures will be summarized by MedDRA SOC and PT

The number (percentage) of subjects with anticancer or other therapies (medications or procedures) during survival follow-up will be presented.

The number (percentage) of subjects with anticancer medications during survival follow-up will be summarized by Anatomical class (ATC Level 1), Pharmacological class (ATC Level 3) and WHO DD preferred term. The number of anticancer medications given to a patient and duration of first anticancer medication during survival follow-up will also be summarized.

The number (percentage) of subjects with anticancer procedures during survival follow-up will be summarized by MedDRA SOC and PT. The number of anticancer procedures given to a patient will also be summarized.

All therapies (ie, all medications and procedures) will be presented in subject data listings.

5.2.6 Treatment Compliance

Treatment compliance will not be summarized since the data will not be entered in the clinical database. Treatment related protocol deviations will be presented in CSR as provided in Section 5.2.3. Information on subject exposure to study drug is described in Section 5.6.1.

5.3 Data Analysis General Considerations

5.3.1 Pooling of Centers

Center will not be considered as a factor in the analysis. Subjects from all centers will be pooled for all analyses.

5.3.2 Adjustments for Covariates

No adjustment for covariates will be performed.

5.3.3 Multiple Comparisons/Multiplicity

No statistical comparison is planned in this study.

5.3.4 Examination of Subgroups

ORR and DOR by mRECIST based on IIR and investigator review, and by RECIST 1.1 based on IIR will be summarized in the first 80 1L HCC subjects for the following subgroups:

- Age group (<65 years, ≥65 years)
- Sex (Male, Female)
- Race (White, Asian, All Others)
- Country (US, Non-US)
- Geographic region (Western, Asia-Pacific)
- ECOG PS (0, 1)
- Macroscopic Portal Vein Invasion (MPVI) (Yes, No)
- Extrahepatic Spread (ES) (Yes, No)
- MPVI, ES or Both (Yes, No)
- Etiology: HBV (Yes, No)
- Etiology: HCV (Yes, No)
- Etiology: Alcohol (Yes, No)
- Weight group (<60kg, ≥60 kg)
- AFP group (<200 ng/mL, ≥200 ng/mL)
- AFP group (<400 ng/mL, ≥400 ng/mL)
- BCLC stage (Stage B, Stage C)

Additional subgroup analyses may also be performed, if deemed appropriate. The definition of subgroup will be determined and documented in the SAP before database lock.

5.3.5 Handling of Missing Data, Dropouts, and Outliers

The details of handling of missing data will be described in Section 8.

5.3.6 Other Considerations

Not applicable.

5.4 Efficacy Analyses

All efficacy analyses will be summarized with available data in the clinical database at the data cutoff. The efficacy analysis will be performed based on the Efficacy Analysis Set. Efficacy data will be formatted as discussed above in the beginning of Section 5 and 5.2.1.

5.4.1 Primary Efficacy Analyses

The primary analyses for ORR and DOR will be performed based on mRECIST and RECIST 1.1 by IIR in the first 80 HCC-1L subjects.

BOR will be summarized. If a subject had a best overall response of non-CR/non-PD, the subject's BOR will be grouped with the SD category. BOR of CR or PR should be confirmed no less than 4 weeks after the initial response in this study. If a subject who had a timepoint response of CR or PR, but the subsequent response was not confirmed no less than 4 weeks after the response, the subject's BOR will be presented as unconfirmed response (ie, unconfirmed CR, or unconfirmed PR) grouped with the SD category. ORR and its corresponding exact 2-sided 95% confidence interval (CI) will be calculated. ORR including confirmed responses and unconfirmed responses will also be presented with its 95% CI.

The analysis for DOR will be performed in the subjects whose response of PR or CR is confirmed. The DOR will be summarized using the Kaplan-Meier method. The median, Q1, and Q3 for DOR with their 95% CIs will be provided. The probability of DOR with its 95% CI at 3, 6, 9 and 12 months will also be calculated and presented. The 95% CIs for the probability of DOR will be calculated based on Greenwood formula using log-log transformation, and 95% CIs for median, Q1, Q3 for DOR will be calculated using a generalized Brookmeyer and Crowley method. The number (percentage) of subjects with event/censored will be summarized. DOR will also be plotted over time by the Kaplan-Meier method.

The calculation of DOR is defined as below:

DOR (days) = Date of first documented PD or death – date of first evaluation of

PR or CR + 1.

For subjects who did not have an event, DOR will be censored per censoring rules shown in Section 8.2.

The number (percentage) of subjects with DOR≥6 months in the subjects whose response of PR or CR is confirmed will also be provided.

The listing of subjects with a confirmed PR or CR will include time to response, type of response, duration of response, and response ongoing status.

A waterfall plot will illustrate the maximum post-baseline percent changes from baseline in the sum of the diameters of target lesions (ie, maximum tumor shrinkage). The number (percentage) of subjects with maximum tumor shrinkage >0%, ≥50%, and ≥75% will be summarized, using the number of subjects with both baseline and postbaseline sum of diameters of target lesions as the denominator. A spider plot will illustrate the percent changes from baseline in the sum of the diameters of target lesions over time. Swimmers' plots showing duration of treatment, the initial response timepoint, disease progression timepoint, death timepoint, and treatment ongoing status will be provided.

Data cutoff for the primary analysis will be done after all subjects in the Expansion part finish at least Cycle 8 assessment and have a tumor assessment of at least Week 24, or discontinue if before Cycle 8. Based on communications with the FDA, this primary analysis will be conducted with the data cutoff date of 31 Oct 2019.

Subject data listings will be provided.

5.4.2 Secondary Efficacy Analyses

ORR and DOR by mRECIST based on investigator review will be analyzed in same manner as primary efficacy analysis.

The calculation of PFS, TTP, TTR and OS is as follows:

- PFS (days) = Date of first documented PD or death – date of first study dose + 1.
- TTP (days) = Date of first documented PD – date of first study dose + 1.
- TTR (days) = Date of first documented CR or PR – date of first study dose + 1.
- OS (days) = Date of death from any cause – date of first study dose + 1.

For a subject without an event of any of the above endpoints, the corresponding censoring date for that endpoint will be used in the above related calculation. The censoring rules for PFS, TTP and OS are detailed in Section 8.2.

The analysis for PFS, and TTP by mRECIST (based on investigator review and IIR) and RECIST1.1 (based on IIR), and OS will be performed in all subjects and summarized in same manner as DOR. TTR will be summarized using descriptive statistics only for the subjects with confirmed CR or PR.

Subject data listings will be provided.

5.4.3 Exploratory Efficacy Analyses

DCR, CBR and their corresponding exact 2-sided 95% CIs will be calculated. For CBR analysis, durable SD is defined as the duration of SD ≥23 weeks. Duration of SD is defined as the time from the date of first study dose to the first documented PD or death, whichever occurs first (ie, same definition as PFS). It is calculated only for subjects who have BOR of SD.

Subject data listings will be provided.

5.4.4 Sensitivity/Supplementary Analyses for Efficacy Endpoints

SENSITIVITY ANALYSES BASED ON DIFFERENT POPULATIONS

Sensitivity analyses of tumor response will be performed using the data from all HCC-1L subjects (N=100; referred to as "Sensitivity Analysis 1"), as well as the data from the first 60

HCC-1L subjects (N=60; referred to as "Sensitivity Analysis 2"), as mentioned above in the beginning of Section 5 and 5.2.1.

SUPPLEMENTARY ANALYSIS BASED ON DIFFERENT DERIVATION RULE

Supplementary analysis of tumor response using the different derivation rule for the situation with more than one missed visit/tumor assessment will also be provided for the first 80 HCC-1L subjects (referred to as "Supplementary Analysis"), following the censoring rule for supplementary analysis described in Section 8.2.

5.4.5 Analyses of Concordance between Independent Imaging Review and Investigator per mRECIST

Concordance on BOR per mRECIST for each subject between assessment by IIR and by investigator will be presented for the first 80 HCC-1L subjects using cross-tabulation table.

5.5 Pharmacokinetic, Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

The details of the pharmacokinetic analyses plan and results will be provided in a separate report or subsequent CSR, once they become available.

5.5.1 Pharmacokinetic Analyses

Pharmacokinetic analyses will be performed in the DLT evaluation part and the expansion part. The Safety Analysis Set will be used for individual lenvatinib plasma concentrations and pembrolizumab serum concentrations listings. The PK Analysis Set will be used for the summaries of lenvatinib plasma concentrations and pembrolizumab serum concentrations, for summaries and listings of PK parameters of lenvatinib and for listings of ADA levels.

5.5.1.1 Plasma or Serum Concentration and its PK Parameter Analysis

Concentration

Plasma concentration values for lenvatinib and serum concentration values for pembrolizumab will be summarized using descriptive statistics (n, mean, standard deviation [SD], median, minimum [min] and maximum [max]) by nominal time point by lenvatinib starting dose group.

Plasma concentrations of lenvatinib, serum concentrations of pembrolizumab and ADA levels will be listed for each subject by actual sampling time by lenvatinib starting dose group.

PK Parameter

PK parameters will be derived by non-compartmental analysis using WinNonlin software (version 6.2.1 or later) according to 302-104.00-MNL.

The following pharmacokinetic parameters for lenvatinib will also be calculated using noncompartmental analysis: C_{max}, t_{max}, $AUC_{(0\text{-}t)}$, $AUC_{(0\text{-}ti)}$, $AUC_{(0\text{-}inf)}$, $t_{1/2}$, CL/F, V_z/F, $C_{ss,max}$, $C_{ss,min}$, $t_{ss,max}$, $AUC_{(0\text{-}\tau)}$, CL_{ss}/F, $C_{ss,av}$, $R_{ac}(C_{max})$, $R_{ac}(AUC)$, PTF ratio.

Other PK parameters may be calculated as appropriate.

Descriptive statistics will be tabulated for the PK parameters of lenvatinib by lenvatinib starting dose group. Descriptive statistics (n, mean, SD, median, min, and max) will be presented for all parameters (apart from t_{max} where mean and SD are not required). In addition, geometric mean and %CV will also be presented for all parameters apart from t_{max}.

PK parameters of lenvatinib for each subject will be listed.

5.5.1.2 Pharmacokinetic Data Figures

Linear and semi-log plots of each subject's lenvatinib plasma and pembrolizumab serum concentration data vs actual time will be displayed by lenvatinib starting dose group. The actual time will be plotted on the X-axis and the concentrations of lenvatinib and pembrolizumab will be plotted on the Y-axis.

The linear and semi-log mean (+SD) plots of lenvatinib plasma concentration versus nominal time will be displayed. The nominal time will be plotted on the X-axis and the mean (+SD) will be plotted on the Y-axis on the same graph by lenvatinib starting dose group by visit (Cycle 1 Day 1 and Cycle1 Day 15).

The linear and semi-log mean (+SD) plots of pembrolizumab serum concentration versus nominal time will be displayed by lenvatinib starting dose group. The nominal time will be plotted on the X-axis and the mean (+SD) will be plotted on the Y-axis on the same graph by visit (Cycle 1 Day 1 and Cycle1 Day 15).

Box plots, with individual sparse PK data, will also be displayed by lenvatinib starting dose group for both lenvatinib and pembrolizumab.

5.6 Safety Analyses

All tolerability analyses will be performed on the DLT Analysis Set. Safety analysis will be performed based on the Safety Analysis Set. Safety data will be presented with the format as mentioned above in the beginning of Section 5 and 5.2.1 by using summary statistics (eg, n, mean, standard deviation, median, Q1, Q3, minimum, maximum for continuous variables; number [percentage] for categorical variables). Safety data by lenvatinib starting dose (ie. 8 mg, 12 mg) will also be presented in all HCC-1L subjects, using overview table of AE, and additional tables if deemed necessary.

5.6.1 Extent of Exposure

The parameters for extent of exposure will be computed for combination (lenvatinib plus pembrolizumab), lenvatinib alone and pembrolizumab alone, respectively. The parameters to be derived are defined as follows:

1. Number of administrations of pembrolizumab = Total number of IV administrations.
2. Duration of treatment (days) = Last dosing date – first dosing date + 1
3. Duration of treatment (months) = (Last dosing date – first dosing date + 1) / (365.25/12)
4. Total doses (mg) of lenvatinib = Sum of all the actual dose
5. Dose intensity of lenvatinib (mg / (days)) = Total doses / Duration of treatment
6. Received dose as percentage of planned starting dose per subject (%) = 100 × Dose intensity / Planned starting dose

The last dosing date should be the date received non 0 (zero) mg dose. The following table provide the parameters to be summarized for the study drug.

Lenbatinib plus Pembrolizumab	Lenvatinib	Pembrolizumab
3* *: defined as the duration between the earliest first dose start date of either medication and the latest last dose end date of either medication	3, 4, 5**, 6 **: presented by starting dose (8 mg, 12 mg)	1, 3

Furthermore, the following information of administration will also be summarized for lenvatinib only.

The number (percentage) of subjects who experienced dose reductions and treatment interruptions will be summarized based on actual dose records from study medication data. The number (percentage) of subjects with dose reduction, interruption by period (eg, every 3 weeks) will also be summarized based on their time of first dose reduction, interruption, respectively.

Time (weeks) to first dose reduction and to first dose interruption of lenvatinib will be summarized by descriptive statistics for subjects with a dose reduction/interruption event.

Frequency of dose interruptions will also be summarized by appropriate categories (eg, 1, 2, 3, ≥4).

Frequency of dose reductions will be summarized by categories (1, 2, 3, 4).

- Definition of Dose Reduction of lenvatinib:
  - Dose level reduces from previous dose level after dose interruption period per Protocol Section 9.4.1.3.2 Table 3. For example, 12 mg followed by 0 mg, followed by 8 mg; 8 mg followed by 0 mg and then followed by 4 mg.
- Definition of Dose interruption of lenvatinib:
  - Only include the scenario that the before and after dose 0 mg (interruption period), the dose levels are the same. For example: 12 mg followed by 0 mg and followed by 12 mg; 8 mg followed by 0 mg, then followed by 8 mg.
  - If dose level reduces from previous dose level after dose interruption period (dose 0), it should be counted as dose reduction, but not dose interruption. For example, 12 mg followed by 0 mg followed by 8 mg, the period with 0 mg should not be counted as dose interruption, instead it should be counted as a dose reduction.
    Note that change from 4 mg QD to 4 mg QOD is also considered as dose reduction.
  - If after dose 0 mg, the subject discontinued from treatment permanently, it should be counted as a treatment discontinuation instead of a dose interruption.

Subject data listings will be provided.

5.6.2 Dose Limiting Toxicity

The number (percentage) of subjects who experienced any DLT will be summarized. The category of DLT will also be summarized.

Subject data listings will be provided.

5.6.3 Adverse Events

The AE verbatim descriptions (investigator terms from the CRF) will be classified into standardized medical terminology using the MedDRA (version 22.1). Adverse events will be coded to the MedDRA lower level term (LLT) closest to the verbatim term. The linked MedDRA PT and SOC will also be captured in the database.

Only treatment-emergent adverse events (TEAEs), defined in Section 8.1, will be included in summary tables. All AEs, treatment emergent or otherwise, will be presented in subject data listings.

MedDRA preferred terms 'Neoplasm Progression', 'Malignant Neoplasm Progression' and 'Disease Progression' will be excluded from TEAE summary tables, if not related to the study drug.

A subject will be counted only once within a SOC and PT, even if the subject experienced more than 1 TEAE within a specific SOC and PT.

An overview table, including the incidence of and the number of subjects with TEAEs, treatment-related TEAEs, TEAEs/treatment-related TEAEs with worst CTCAE grade 3 or above, treatment-emergent serious adverse events (SAEs), deaths, and TEAEs that led to treatment discontinuation of lenvatinib, pembrolizumab, or both, dose reduction of lenvatinib, or dose interruption of lenvatinib, pembrolizumab, or both will be provided.

The incidence of the following events will be reported as the number (percentage) of subjects with TEAEs by MedDRA SOC and PT:

- TEAEs
- TEAEs by worst CTCAE grade
- TEAEs with worst CTCAE grade 3 or above
- TEAEs with any CTCAE grade 3 or 4
- Treatment-emergent SAEs
- Fatal TEAE
- Non-fatal treatment-emergent SAEs
- Treatment-related TEAEs
- Treatment-related TEAEs by CTCAE grade
- Treatment-related TEAEs with worst CTCAE grade 3 or above
- Treatment-related TEAEs with any CTCAE grade 3 or 4
- Treatment-related treatment-emergent SAEs
- Treatment-related Fatal TEAE
- TEAEs leading to discontinuation of lenvatinib
- TEAEs leading to dose reduction of lenvatinib
- TEAEs leading to dose interruption of lenvatinib
- TEAEs leading to dose modification (reduction/interruption) of lenvatinib
- TEAEs leading to discontinuation of pembrolizumab
- TEAEs leading to dose interruption of pembrolizumab
- TEAEs leading to discontinuation of both lenvatinib and pembrolizumab
- TEAEs leading to dose interruption of both lenvatinib and pembrolizumab

Subject data listings of all deaths, SAEs, AEs leading to death, treatment discontinuation, dose reduction, or dose interruption will be provided. All AE regardless of treatment-emergent or not will be included in the subject data listings.

The following TEAE tables will be provided for clinically significant TEAEs (CSAE) for lenvatinib, and TEAEs of special interest (AEOSI) for pembrolizumab.

- Overview of CSAE for lenvatinib (for overall and respective CSAE group)
- Overview of AEOSI for pembrolizumab (for overall and respective AEOSI group)
- CSAE for lenvatinib by PT and worst CTCAE grade
- AEOSI for pembrolizumab by PT and worst CTCAE grade

5.6.4 Laboratory Values

Laboratory results will be summarized using Système International (SI) units. For all quantitative parameters listed in Protocol Section 9.5.1.5.3, the actual value and the change from baseline to each postbaseline visit will be summarized by visit using summary statistics. Qualitative parameters listed in Protocol Section 9.5.1.5.3 will be summarized by number and percentage of subjects, and changes from baseline to each postbaseline visit will be reported using shift tables. Percentages will be based on the number of subjects with both nonmissing baseline and relevant postbaseline results.

Box and whisker plots will be used to show the longitudinal change of the parameters across visits.

Laboratory test results will be assigned a low/normal/high (LNH) classification according to whether the value is below (L), within (N), or above (H) the laboratory parameter's reference range. The result of LNH classification will be provided in a subject data listing.

Laboratory parameters will be graded by CTCAE ver. 4.03. Changes from CTCAE grade at baseline to worst postbaseline will be reported using shift tables.

CTCAE ver. 4.03 will be used to identify subjects with treatment-emergent markedly abnormal laboratory values (TEMAV). The number (percentage) of subjects with TEMAV (markedly abnormal high/low) will be summarized using worst CTCAE grade for overall study period. The TEMAV is defined in Section 8.3. The number (percentage) of subjects with an increase of at least 1 CTCAE grade from baseline to the worst postbaseline value of any grade will also be presented.

Shifts from baseline to worst postbaseline value for proteinuria determined by dipstick (negative, trace, 1+, 2+, 3+, 4+), and a summary of proteinuria from 24-hr urine collection will be provided.

The frequency and percentage of values $\leq$ the upper limit of normal (ULN) and >ULN of thyroid function test will be summarized at baseline and worst postbaseline value.

Subject data listings will be provided.

5.6.5 Vital Signs

Summary statistics for vital signs parameters (diastolic and systolic blood pressure, pulse, and temperature) and weight, and changes from baseline will be presented by visit.

Hypertension based on vital signs (diastolic and systolic blood pressure) will be presented using shift from baseline to worst postbaseline CTCAE grade. Category of hypertension is defined as Table 3.

Table 3 Hypertension Grades Based on Blood Pressure

Grade	Blood Pressure (mm Hg)
0 (Normal)	Systolic<120 and Diastolic<80
1 (Prehypertension)	Systolic 120 – 139 or Diastolic 80 – 89
2 (Stage 1 Hypertension)	Systolic 140 – 159 or Diastolic 90 – 99
3 (Stage 2 Hypertension)	Systolic ≥ 160 or Diastolic ≥ 100

Box and whisker plots will be used to show the longitudinal change of the parameters by visit.

Subject data listings will be provided.

5.6.6 Electrocardiograms

The results of ECG assessments performed at each visit will be evaluated. Summary statistics for ECG parameters (Heart Rate, RR, PR, QRS, QT, and QTcF) and changes from baseline will be presented by visit.

Shift tables will present changes from baseline in ECG interpretation (categorized as normal; abnormal, not clinically significant; and abnormal, clinically significant) to each visit.

In addition, the number (percentage) of subjects who met below criteria at least once in QTcF will be presented:

Absolute QTcF interval prolongation:
- QTcF interval >450 ms
- QTcF interval >480 ms
- QTcF interval >500 ms

Change from baseline in QTcF interval:
- QTcF interval increases from baseline >30 ms
- QTcF interval increases from baseline >60 ms

Subject data listings will be provided.

5.6.7 Other Safety Analyses

LVEF

LVEF and change from baseline to lowest postbaseline value will be summarized by visit. Change from baseline to lowest postbaseline category (no reduction, >0 - ≤10%, >10 - ≤15%, >15%), change from baseline to lowest postbaseline value <50% (frequency and percentage) will be presented. Subjects will be included in the change summary only if they had both a baseline value and at least one postbaseline value.

A table summarizing the shift from baseline to the worst postbaseline LVEF category will be presented. LVEF values will be categorized as:

- Hyperdynamic/Normal: LVEF greater than 70%, or LVEF 50% to 70% (midpoint 60%)
- Mild dysfunction: LVEF 40% to 49% (midpoint 45%)
- Moderate dysfunction: LVEF 30% to 39% (midpoint 35%)
- Severe dysfunction: LVEF less than 30%.

Subject data listings will be provided.

ECOG PS

The number (percentage) of subjects for each category of ECOG PS will be summarized by shift from baseline to the highest postbaseline scale of ECOG PS.

Subject data listings will be provided.

5.7 Other Analyses

A separate pre-specified HRQoL analysis following FDA and EMEA PRO Guidelines will be performed and detailed in a separate SAP and HRQoL report. Scoring of EQ-5D-5L and derivation of utility for health economic analysis will also be accomplished in a separate analysis and described in a separate HRQoL report.

5.8 Exploratory Analyses

Exploratory analyses may be conducted as appropriate. Any performed exploratory analyses will be appropriately 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

As of Amendment 03, Expansion part may be further expanded up to approximately 94 evaluable subjects. For this expansion decision, 2 interim analyses will take place when 20 (6 subjects for DLT evaluation part plus 14 subjects for Expansion part) and 56 subjects (6 subjects for DLT evaluation part plus 50 subjects for Expansion part) have sufficient follow-up to be evaluated for response. The decision to expand enrollment will be based on the results of 2 interim analyses, which will spend β= 0.012 and β= 0.024 at the first and second interim analyses, respectively. The decision to expand enrollment will be assessed by mRECIST based on investigator review.

Based on an assumption of H0: 25% ORR and H1: 45% ORR, the 100-subject design with two futility analyses has approximately 96% statistical power at 2-sided α = 0.02 (that corresponds to 1-sided α = 0.01). At the first interim analysis (N = 20), if there are more than 5 responses, then approximately 36 additional subjects will be enrolled. At the second interim analysis (N = 56), if there are more than 16 responses, approximately 44 additional subjects will be enrolled. If there are 5 or fewer responses at the first interim analysis (N=20) or 16 or fewer responses at the second interim analysis (N=56), the sponsor may decide whether to expand enrollment based on clinical outcome (eg, DOR).

The boundaries for the decision to expand enrollment in Expansion part are presented in Table 2.

7 CHANGES IN THE PLANNED ANALYSES

There were no major changes and deletions in the planned analyses in the protocol. Based on communications with the FDA toward the Type B meeting (that was scheduled on 10 Oct 2019), statistical analysis plan for the 1L HCC subjects to support AA applications were added, in addition to the planned analyses in the protocol on all HCC subjects including the subjects with prior anticancer medication.

8 DEFINITIONS AND CONVENTIONS FOR DATA HANDLING

The data will be handled as follows. The sponsor will determine how to handle all data prior to data base lock.

8.1 General Data Handling

Definition of Baseline data

Baseline is defined as the last non-missing value observed on or prior to the first dose of study treatment for a given parameter.

Definition of Change from Baseline, Percent Change from Baseline

Change from baseline is defined as post-baseline value minus baseline value.
Percent change from baseline is defined as follows:

% Change from baseline = (Change from baseline / Baseline) * 100%

For any Baseline value of 0, the subject's corresponding percent change from baseline will not be included in the summary statistics table.

Handling of data not within specified periods or within the follow-up period

All the safety parameters will be used for summary statistic tabulation, except when it is irrelevant such as the situation where tests scheduled to be conducted before administration of the investigational drug were actually conducted after its administration.

8.2 Efficacy Data Handling

Handling of missing data

For the analysis of tumor response related endpoints (ORR, DCR, and CBR), subjects with missing response status (subjects whose baseline or post-baseline tumor assessment is missing) will be coded as non-responders on Efficacy Analysis Set.

Censoring rules for primary PFS analysis

The following table provides the censoring rules for primary PFS analysis.

No.	Situation	Date of Event (Progression/Death) or Censoring	Outcome
1	No baseline tumor assessments	Date of first dose	Censored
2	Progression documented between scheduled visits	Date of first radiologic PD assessment	Event
3	No progression at the time of data cut-off	Date of last adequate radiologic assessment before or on date of data cut-off	Censored
4	New anticancer treatment started	Date of last adequate radiologic assessment prior to or on date of new anticancer treatment	Censored
5	Death before first PD assessment	Date of death	Event
6	Death between adequate assessment visits*	Date of death	Event
7	Death or progression after more than one missed visit/tumor assessment**	Date of last adequate radiologic assessment before missed tumor assessments	Censored

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

* Adequate tumor assessment is radiologic assessment of CR, PR, SD, non-CR/non-PD, or PD at regular interval as defined in the protocol.

** More than one missed visit/adequate tumor assessment is defined as having either one of the following two durations being longer than 12 weeks + 2weeks (ie, the date of death or PD – the date of last adequate tumor assessment >97 days) for subjects on the every 6 week tumor assessment schedule in the first 8 cycles of treatment and 18 weeks + 2 weeks (ie, the date of death or PD – the date of last adequate tumor assessment >139 days) for subjects on the every 9 week tumor assessment schedule after Cycle 8 in this study.

The priority of the censoring rules is as follows:

1. If the subject had PD or death, the following sequence will be applied:

- If a subject did not have a baseline tumor assessment (No. 1), the subject will be censored on the date of first dose. However, if the subject died within 97 days (14 weeks -1 day) after first dose and did not receive a new anticancer treatment, it will be counted as PFS event at the date of death.
- If a subject had new anticancer treatment before PD or death (No. 4), the subject will be censored on the date of the last adequate tumor assessment prior to or on the date of new anticancer treatment.
- If a subject missed two or more tumor assessments before PD or death (No. 7), the subject will be censored on the date of the last adequate tumor assessment before PD or death. Note that if a subject is censored by both this criterion and the anticancer treatment criterion, 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, or No. 7).

Censoring rules for primary DOR and TTP analysis

The censoring rules for DOR and TTP will be the same as those for primary PFS analysis, except that:

No baseline tumor assessments (No. 1) will not apply for DOR,

Death before first PD assessment (No. 5) and death between adequate assessment visits (No. 6) will not apply to TTP. Death without new anticancer treatment will be censored at the last adequate tumor assessment for TTP.

Censoring rules for supplementary analysis

The censoring rules for the supplementary analysis described in Section 5.4.4 will follow the censoring rule for the primary PFS analysis specified above with the exception that the following rule will be applied instead of the primary rule for situation No. 7.

No.	Situation	Date of Event (Progression/Death) or Censoring	Outcome
7	Death or progression **immediately** after more than one missed visit/tumor assessment**	Date of last adequate radiologic assessment before missed tumor assessments	Censored

- If a subject missed two or more tumor assessments **right** before PD or death (No. 7), the subject will be censored on the date of the last adequate tumor assessment before PD or death. Note that if a subject is censored by both this criterion and the anticancer treatment criterion, the earliest censoring date will be used.

For example, following this rule, supplementary analysis will be handled as below for each scenario:

- PR, PR, two missing scans, PR, PR, then DOR is censored at the fourth PR date (instead of the censoring at the second PR based on the primary analysis)
- PR, PR, two missing scans, PD, then DOR is censored at the second PR date (same as the primary analysis)
- PR, PR, two missing scans, PR, PD, then DOR has an event at the PD date (instead of the censoring at the second PR based on the primary analysis)

etc.

Censoring rules for OS

No.	Situation	Date of Event (Death) or Censoring	Outcome
1	Death before or on data cut-off	Date of death	Death
2	Death after data cut-off	Date of data cut-off	Censor
3	Alive at data cut-off	Date of data cut-off	Censor
4	Lost to follow-up or withdrawal of consent before data cut-off	Date last known to be alive	Censor

8.3 Safety Data Handling

Definition of derived variables for extent of exposure

The derivation rule will be presented in the Section 5.6.1.

Treatment-emergent adverse event

A treatment-emergent adverse event (TEAE) is defined as an AE that emerges during the time from the first dose of study drug (either lenvatinib or pembrolizumab, whichever earlier) to 30 days following the last dose of study drug (either lenvatinib or pembrolizumab, whichever later), having been absent at pretreatment (Baseline) or

- Reemerged during treatment or up to 30 days following the last dose of study drug, having been present at pretreatment but stopped before treatment, or
- Worsened in severity during treatment or up to 30 days following the last dose of study drug relative to the pretreatment state, when the AE was continuous.

Serious adverse events (SAEs) that develop from 31 days up to 120 days after the subject's last dose of study drug are considered as treatment-emergent SAEs.

TEMAV (Treatment-Emergent Markedly Abnormal Value)

The treatment-emergent markedly abnormal laboratory values (TEMAV) will be defined as the postbaseline value of Grade 3 or 4 with an increase of at least 1 CTCAE grade from baseline. TEMAV is derived as follows based on CTCAE grade.

If the baseline CTCAE grade is 0 (normal), 1 or 2, and the postbaseline CTCAE grade is 3 or 4 then the postbaseline value is considered a TEMAV. Or, if the baseline CTCAE grade is 3 or 4 and the postbaseline CTCAE grade is higher than the baseline CTCAE grade then the postbaseline value is considered a TEMAV.

Handling of below lower quantification values in laboratory results

In the cases where a laboratory result contains a value of below the limit of quantification (BLQ), it will be replaced with the lower limit value of quantification (LLOQ) for summary tables.

The priority of use for blood pressure

For systolic and diastolic blood pressures, an additional confirmatory assessment may be done more than 60 minutes after initial assessment, if necessary. In the case where there are both the initial and confirmatory assessment results in a day, the confirmatory (latest) one will be used.

Visit windows for safety analyses

In the calculation of descriptive statistics for safety parameters (eg, laboratory values and vital signs, etc.) per scheduled visit, and change from baseline per visit, if multiple observations fall in the same scheduled visit, the first record will be used for summary tables. Other safety analyses (eg, worst grade laboratory results) will include all observations until the date of 30 days following the last dose of study drug.

8.4 Pharmacokinetic Data Handling

8.4.1 Lower Limit of Quantification of lenvatinib Plasma Concentration and Everolimus Blood and Concentration

The LLOQ of lenvatinib plasma concentrations is 0.250 ng/mL

The LLOQ of pembrolizumab serum concentrations is 25 ng/mL

8.4.2 BLQ Handling for Calculation of PK Parameters

While calculating PK parameters in WinNonlin, BLQ values will be handled according to 302-104.00-MNL, for non-compartmental pharmacokinetic analysis.

8.4.3 BLQ Handling for Developing Concentration-Time Profiles

When developing individual concentration-time profiles, BLQ values will be handled according to 302-104.00-MNL for non-compartmental pharmacokinetic analysis.

8.4.4 Handling of Anomalous Concentration Values

The handling of anomalous concentration values will follow the guidance in the SWP for non-compartmental pharmacokinetic analysis (302-104.00-MNL).

8.4.5 General Rules for Presentation of Drug Concentrations and PK Parameters

When presenting individual/raw (raw, hereafter) values and summary statistics, the following rule will be applied: for drug concentrations and concentration-dependent pharmacokinetic parameters, all summary statistics (mean, median, geometric mean, SD and coefficient variation (CV)) will have 3 significant digits. For t_{max}, raw values and their median are shown in fixed 2 decimal places.

Variable	Unit	N	Digit rule	Raw/ Minimum/ Maximum	Mean Median	SD	Geometric Mean	CV (%)
drug concentration	ng/mL	X	Significant digits	3	3	3	-	-
C_{max}, $C_{ss,max}$, C_{min}, $C_{ss,av}$	ng/mL	X	Significant digits	3	3	3	3	3
t_{max}, $t_{ss,max}$	h	X	Fixed decimal places	2	2	-	-	-
λ_z(C1D1&D15)	1/h	X	Significant digits	3 (Listing only)	-	-	-	-
$t_{1/2}$(C1D1&D15)	h	X	Significant digits	3	3	3	3	3
$AUC_{(0-t)}$, $AUC_{(0-inf)}$, $AUC_{(0-ti)}$, $AUC_{(0-\tau)}$	ng•h/mL	X	Significant digits	3	3	3	3	3
CL/F, CL_{ss}/F	L/h	X	Significant digits	3	3	3	3	3
V_z/F(C1D1&D15)	L	X	Significant digits	3	3	3	3	3
R_{ac}		X	Significant digits	3	3	3	3	3
PTF ratio		X	Significant digits	3	3	3	3	3

Mean, SD, geometric mean and CV will not be calculated for t_{max}, $t_{ss,max}$.

CV(%)= sqrt(exp[SD**2 of log transformed data]-1)*100

NOTE

1. The following parameters are reported in the CSR, but appear in Listings only. They are important information to confirm that individual $t_{1/2}$ and its related parameters such as $AUC_{(0-inf)}$ are appropriately derived and allow those PK parameters to be reproduced when necessary.
   a. Time points used for estimation of λ_z (lower and upper)
   b. Number of the time points used for λ_z
   c. Adjusted regression coefficient (R_{2adj})

In Listings, a) are shown in same digits as actual sampling time after dosing used for calculation of PK parameters. For b), integer number is used in Listings. For c), 3 significant digits are used in Listing.

9 PROGRAMMING SPECIFICATIONS

The rules for programming derivations and dataset specifications are provided in separate documents.

10 STATISTICAL SOFTWARE

All statistical analyses will be conducted by Eisai or Takumi Information Technology, using validated standard programs or double programming. For analyses needed in data review, single programming will be used.

All statistical analyses will be performed using SAS (version 9.3 or later). As necessary, other validated statistical software will also be used.

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

- FDA Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, May 2007. A revision of the guidance in Dec 2018 is available online at:
  https://wayback.archive-it.org/7993/20190916121529/https://www.fda.gov/media/71195/download
- FDA Guidance for Industry Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics, Apr 2015 is available online at:
  https://wayback.archive-it.org/7993/20190916121636/https://www.fda.gov/media/116860/download

13 APPENDICES

13.1 National Institute for Health: Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03

National Cancer Institute (NCI) Cancer therapy evaluation program Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 May 2009 (v4.03 June 2010) is available online at:

https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf

SIGNATURE PAGE

Authors:

PPD

23 Oct 2019
Date

PPD MDC

PPD

22 Oct 2019
Date

PPD OBG

Approval:

PPD

24 Oct 2019
Date

PPD OBG

PPD

PPD

24 Oct 2019
Date

OBG

PPD

23 OCT 2019
Date

PPD MDC