

I4T-MC-JVDE Statistical Analysis Plan

Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study of Ramucirumab and Best Supportive Care (BSC) Versus Placebo and BSC as Second-Line Treatment in Patients With Hepatocellular Carcinoma and Elevated Baseline Alpha-Fetoprotein (AFP) Following First-Line Therapy With Sorafenib

NCT02435433

Approval Date: 21-Dec-2017

1. Statistical Analysis Plan

I4T-MC-JVDE: Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study of Ramucirumab and Best Supportive Care (BSC) Versus Placebo and BSC as Second-Line Treatment in Patients With Hepatocellular Carcinoma and Elevated Baseline Alpha-Fetoprotein (AFP) Following First-Line Therapy With Sorafenib

Confidential Information

The information contained in this Statistical Analysis Plan (SAP) is confidential and the information contained within it may not be reproduced or otherwise disseminated without the approval of Eli Lilly and Company or its subsidiaries. This document and its associated attachments or appendices are subject to United States Freedom of Information Act Exemption 4.

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol I4T-MC-JVDE
Phase 3

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly:
07 January 2016

Statistical Analysis Plan Version 2 electronically signed and approved by Lilly on
electronically signed and approved by Lilly on date provided below.

Approval Date: 21-Dec-2017 GMT

2. Table of Contents

Section	Page
1. Statistical Analysis Plan	1
2. Table of Contents	2
3. Revision History	7
4. Abbreviations	8
5. Study Objectives	10
5.1. Primary Objective	10
5.2. Secondary Objectives	10
5.3. Exploratory Objectives	10
6. A Priori Statistical Methods	11
6.1. Introduction	11
6.2. Study Design	11
6.2.1. Region and Country	11
6.2.2. Determination of Sample Size	12
6.3. Analysis Populations	13
6.3.1. Intent-to-Treat (ITT) Population	13
6.3.2. Safety Population	13
6.3.3. Per-Protocol (PP) Population	13
6.4. Definitions and Data Handling Conventions	14
6.4.1. Safety Definitions and Data Handling	14
6.4.1.1. Treatment-Emergent Adverse Events (TEAE)	14
6.4.1.2. Adverse Events of Special Interest (AESIs)	14
6.4.1.3. Consolidated Adverse Events	14
6.4.1.4. Previous/Concomitant Medications	14
6.4.1.5. Partial Data for Adverse Events and Previous/Concomitant Medications	14
6.4.1.6. Study Drug Exposure	15
6.4.2. Efficacy Definitions and Data Handling	16
6.4.2.1. Overall Survival	16
6.4.2.2. Progression-Free Survival	16
6.4.2.3. Time to Deterioration in FHSI-8	18
6.4.2.4. Time to Deterioration in ECOG PS	18
6.4.2.5. Objective Response Rate and Disease Control Rate	18
6.4.2.6. Time to Radiographic Progression	19
6.4.2.7. Patient-Focused Outcomes (PFOs)	20

6.4.3.	Other General Data Handling	21
6.5.	Data Analysis.....	23
6.5.1.	Disposition of Patients	23
6.5.2.	Demographic and Baseline Characteristics	23
6.5.3.	Analysis of Efficacy Data.....	24
6.5.3.1.	General Considerations.....	24
6.5.3.2.	Primary Efficacy Endpoint - Overall Survival.....	25
6.5.3.3.	Secondary Efficacy Endpoints	25
6.5.3.3.1.	Progression-Free Survival.....	26
6.5.3.3.2.	Time to Deterioration of FHSI-8.....	26
6.5.3.3.3.	Time to Deterioration of ECOG PS.....	26
6.5.3.3.4.	Objective Response Rate and Disease Control Rate (DCR).....	27
6.5.3.3.5.	Time to Radiographic Progression	27
6.5.3.4.	Subgroup Analysis.....	27
6.5.3.5.	Analysis Adjusting for Covariates	28
6.5.3.6.	Restricted Mean Difference Analysis	28
6.5.4.	Postdiscontinuation Anticancer Therapy.....	29
6.5.5.	Patient-Focused Outcomes	29
6.5.5.1.	FHSI-8 Analyses	29
6.5.5.2.	EQ-5D-5L Analyses	30
6.5.6.	Analyses of Safety Data	30
6.5.6.1.	Study Drug Exposure.....	30
6.5.6.2.	Overall Summary of Adverse Events	30
6.5.6.3.	Treatment-Emergent Adverse Events.....	31
6.5.6.4.	Deaths, SAEs, and Other Significant Adverse Events	31
6.5.6.5.	Vital Signs and Other Observations Related to Safety.....	32
6.5.6.5.1.	Laboratory Parameters	32
6.5.6.5.2.	Electrocardiogram (ECG)	32
6.5.6.5.3.	Echocardiogram/Multiple-Gated Acquisition (MUGA).....	33
6.5.6.6.	Hospitalizations	33
6.5.7.	Concomitant Medication and Procedure.....	33
6.5.8.	Pharmacokinetic (PK) Analyses and Immunogenicity Analyses	33
6.5.9.	Exploratory Biomarker Analysis	33
6.6.	Interim Analyses and Data Monitoring Committee (DMC).....	34
7.	Unblinding Plan	35

8. Changes in Planned Analyses from the Protocol..... 36

9. References 37

10. Appendices 38

Redacted Version

Table of Contents

Table		Page
Table 1.	Countries in Each Geographic Region used for Stratification and Data Analyses..	12
Table 2.	Dose Reductions for Ramucirumab	16
Table 3.	Censoring Rules for PFS – Main Analysis.....	17
Table 4.	Censoring Rules for PFS – Sensitivity Analyses.....	18
Table 5.	Rules for Determining Date of Progression or Censor for TTP.....	20

Redacted Version

Table of Contents

Appendix		Page
Appendix 1.	Analyses for Pharmaceuticals and Medical Devices Agency (Japan) (PMDA)	39
Appendix 2.	Clinical Trial Registry Analyses	40

Redacted Version

3. Revision History

Statistical analysis plan (SAP) Version 1 was approved prior to first meeting of the independent data monitoring committee (IDMC).

Version 2 was approved prior to the database lock to assess primary objective to update the following components:

- Added time to deterioration in Functional Assessment of Cancer Therapy (FACT) Hepatobiliary Symptom Index-8 (FHSI-8) total score and Eastern Cooperative Oncology Group performance status (ECOG PS) as secondary projectives
- Updated Region 1 area
- Revised determination of sample size
- Revised Per-protocol population definition and criteria
- Revised the method to summarize dose delay and dose omission
- Removed the last progression-free survival (PFS) sensitivity analysis
- Revised the definition of time to deterioration
- Clarified the criteria of stable disease (SD) as the best overall response
- Revised patient-focused outcomes analysis definition
- Revised missing data handling
- Revised the testing sequence of secondary endpoints
- Added time to deterioration of FHSI-8 individual item analysis
- Revised the study exposure analysis
- Revised overall summary of adverse events
- Revised summary of death, serious adverse events (SAEs), and adverse events of special interest (AESI)
- Added the analysis of treatment-emergent anti-drug antibodies
- Added reference

4. Abbreviations

Term	Definition
AE	adverse event
AESI	adverse event of special interest
AFP	alpha-fetoprotein
BCLC	Barcelona Clinic Liver Cancer (staging)
BOR	best overall response
BSC	best supportive care
CI	confidence interval
CP	Child-Pugh
CR	complete response
CRF/eCRF	case report form (electronic)
DCR	disease control rate
ECG	electrocardiogram
ECOG PS	Eastern Cooperative Oncology Group Performance Status
EQ-5D-5L	EuroQol 5-Dimension 5-Level
FACT	Functional Assessment of Cancer Therapy
FHSI-8	FACT Hepatobiliary Symptom Index-8
HCC	hepatocellular carcinoma
HR	hazard ratio
IC	inclusion criterion
ICF	inform consent form
IDMC	Independent Data Monitoring Committee
ITT	intent-to-treat
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LVEF	left ventricular ejection fraction

MedDRA®	Medical Dictionary for Regulatory Activities
MUGA	Multiple Gated Acquisition Scan
NE	not evaluable
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PH	proportional hazards
PK	pharmacokinetic
PMDA	Pharmaceuticals and Medical Devices Agency (Japan)
PP	per-protocol
PR	partial response
PRO	patient-reported outcome
PFO	Patient-focused outcome
PT	preferred term
QoL	quality of life
RECIST	Response Evaluation Criteria In Solid Tumors
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	stable disease
SE	standard error
SOC	System Organ Class
TEAE	treatment-emergent adverse event
TTP	time to radiographic progression
VAS	Visual Analogue Scale
VEGF	vascular endothelial growth factor

5. Study Objectives

5.1. Primary Objective

The primary objective is to compare overall survival (OS) for ramucirumab versus placebo in patients with advanced hepatocellular carcinoma (HCC) and elevated baseline alpha-fetoprotein (AFP) after intolerance or progression on prior sorafenib treatment.

5.2. Secondary Objectives

The secondary objectives of the study are to evaluate:

- Progression-free survival (PFS)
- Time to radiographic progression (TTP)
- Objective response rate (ORR)
- Safety profile of ramucirumab
- Ramucirumab pharmacokinetics (PK)
- Immunogenicity of ramucirumab
- Time to deterioration in Functional Assessment of Cancer Therapy (FACT) Hepatobiliary Symptom Index-8 (FHSI-8)
- Time to deterioration in Eastern Cooperative Oncology Group performance status (ECOG PS)
- Patient-reported outcome (PRO) measures of disease-specific symptoms and health-related quality of life

5.3. Exploratory Objectives

The exploratory objective of the study is to explore biomarkers relevant to ramucirumab, angiogenesis, and the disease state and to correlate these markers to clinical outcome.

6. A Priori Statistical Methods

6.1. Introduction

This SAP describes the detailed procedures for executing the planned statistical analyses in order to assess the efficacy and safety profile of ramucirumab when administered to patients with HCC after first-line therapy with sorafenib. The statistical analyses will be conducted by Eli Lilly or a contract research organization using SAS® software Version 9.2 or higher.

6.2. Study Design

Study I4T-MC-JVDE (JVDE) is a global, multicenter, randomized, double-blind, placebo-controlled Phase 3 study evaluating the efficacy and safety of ramucirumab as a single agent for the treatment of patients with advanced HCC and elevated baseline AFP after intolerance or progression on prior sorafenib therapy. Patients with a baseline AFP ≥ 400 ng/mL, based on local laboratory results, who meet all other inclusion/ exclusion criteria will be randomized. The study will randomize (in a 2:1 ratio) approximately 279 patients to receive ramucirumab 8 mg/kg or placebo administered intravenously once every 14 days in an outpatient setting. Randomization will be stratified by the following factors:

- geographic region (Region 1 versus Region 2 versus Region 3)
- macrovascular invasion (yes versus no)
- Eastern Cooperative Oncology Group (ECOG) performance status (PS) (0 versus 1)

All patients will be offered best supportive care (BSC), as determined appropriate by the investigator. Study treatment will be continued until disease progression, unacceptable toxicity, withdrawal of consent, or until other withdrawal criterion is met.

During study, tumor assessment and imaging will be performed as follows:

- every 6 weeks (± 3 days) for the first 6 months after randomization, and
- every 9 weeks (± 3 days) thereafter until the patient has radiographic progressive disease (PD) or dies, whichever occurs first.

During long-term follow-up, the patient will be followed up approximately every 60 days (± 7 days) until death or study completion, whichever occurs first.

The final analysis will be performed after approximately 221 deaths have been observed.

Adverse event (AE) information will be collected until approximately 30 days after the patient and the investigator agree that the patient will no longer continue study treatment and until all SAEs related to protocol procedures or study treatment have resolved or been explained.

6.2.1. Region and Country

The general plan for assigning participating countries to a geographic region is as follows:

1. Region 1 will include countries in the Americas, Europe, Israel, and Australia.

2. Region 2 will include countries in Asia, except for Japan.
3. Region 3 will include Japan only.

Table 1 below shows the planned countries (country with at least 1 randomized patient) and corresponding regions under which they will be grouped for data analyses:

Table 1. Countries in Each Geographic Region used for Stratification and Data Analyses

Region Number	Country
Region 1	Brazil
	Canada
	United States
	Australia
	Austria
	Belgium
	Czech Republic
	Poland
	France
	Germany
	Israel
	Italy
	Spain
	United Kingdom
	Switzerland
Region 2	Korea
	China
	Hong Kong
	Taiwan
Region 3	Japan

6.2.2. Determination of Sample Size

The primary objective of this study is to compare ramucirumab versus placebo in terms of OS in patients with advanced HCC after intolerance or progression on prior sorafenib. The sample size was determined based on the following assumptions:

- hazard ratio (treatment/control) of 0.67, with median OS of 4.5 months in the placebo arm and 6.7 months in the ramucirumab arm
- the randomization ratio is 2:1 (ramucirumab:placebo)
- the overall significance level will be controlled at 1-sided 0.025 (2-sided 0.05)
- the Type II error rate is 20%; that is, the statistical power of the trial is set to 80%

Under the assumptions above, the final analysis will be performed when at least 221 deaths have been observed. Therefore, the study will randomize approximately 279 patients (that is, 20% censoring rate including dropouts, with approximately 186 patients randomized to the ramucirumab arm and 93 patients randomized to the placebo arm).

6.3. Analysis Populations

6.3.1. *Intent-to-Treat (ITT) Population*

The ITT population will comprise all randomized patients. This population will be assessed for baseline characteristics and efficacy including quality of life (QoL). The ITT analysis of these data will consider allocation of patients to treatment groups as randomized regardless what actual treatment they receive.

6.3.2. *Safety Population*

The Safety population will include all patients who received at least 1 dose (including a partial dose) of an investigational product (ramucirumab or placebo). Analysis of safety data will be based on the Safety population and will be performed based on the actual treatment a patient received on the first study treatment administration regardless of which treatment they were randomized to receive (“as treated”).

6.3.3. *Per-Protocol (PP) Population*

Per-protocol population will include all patients who are randomized and treated and do not have any major protocol violations that could potentially affect the efficacy conclusions of the study. Efficacy analyses will also be performed in the PP population. The PP population is a subset of the ITT population that will be defined and finalized prior to database lock for the final analysis. Patients who meet any of the following criteria will be excluded from the PP population:

- Randomized but never received study medication
- All AFP < 400 ng/mL prior to first dose, as determined by local and/or central laboratory testing [per Inclusion Criterion (IC) 9]
- Baseline Child-Pugh (CP) score was not 5 or 6 (CP class A) [per IC 7]
- Patient does not have HCC that is Barcelona Clinic Liver Cancer (BCLC) Stage B or C at randomization [per IC 8]
- Did not have a diagnosis of HCC as defined in the protocol [per IC 1 and EC 17]
- Had not been treated with sorafenib
- Had an ECOG PS score of 2 or above at enrollment [per IC 10]
- Received any previous systemic therapy with vascular endothelial growth factor (VEGF) inhibitors or VEGF receptor inhibitors (including investigational agents) other than sorafenib for treatment of HCC [per Exclusion Criterion 23]
- Received additional concurrent chemotherapy, biological response modifiers, other investigational agents, and radiation therapy (except for palliation to symptomatic sites of disease) during the study treatment
- The patient has a history of or current hepatic encephalopathy (any grade) or current ascites (Grade ≥ 2)
- Patient received incorrect study medication at least 25% of the total number of infusions (that is, Number of incorrect medication infusions/Total number of infusions patient received $\geq 25\%$)

6.4. Definitions and Data Handling Conventions

6.4.1. Safety Definitions and Data Handling

6.4.1.1. Treatment-Emergent Adverse Events (TEAE)

Treatment-emergent adverse event (TEAE) is defined as any untoward medical occurrence that either occurs or worsens at any time after treatment baseline and that does not necessarily have to have a causal relationship with this treatment.

6.4.1.2. Adverse Events of Special Interest (AESIs)

Adverse events of special interest (AESIs) are AEs thought to be potentially associated with the study drug or the disease under study. Therefore, aggregate AE terms for AEs typically associated with anti-angiogenesis or possibly associated with the study drug were developed.

Adverse events of special interest may be modified as the understanding of the safety of the investigational drug increases. Please refer to the Investigator's Brochure update for most up-to-date aggregate AE terms. The Medical Dictionary for Regulatory Activities (MedDRA™) preferred terms (PTs) that are grouped under each of the AESI terms will be provided in the compound-level safety document.

6.4.1.3. Consolidated Adverse Events

Given the high level of granularity of the MedDRA dictionary, clinically identical or synonymous PTs reported under different terms in the database, in addition to being reported separately, will also be consolidated in a separate summary. The list of consolidated AE categories will be reported in the clinical study report.

6.4.1.4. Previous/Concomitant Medications

A medication will be regarded as **previous** if:

- the medication was stopped prior to first dose of study treatment (either ramucirumab or placebo) (medication stop date < the date of first dose of study treatment).

A medication will be regarded as **concomitant** if:

- the medication was started on or after the date of first dose of study treatment and before the end of short-term follow-up; or
- the medication was started prior to first dose of study treatment but was ongoing at the time of the first dose of study treatment.

6.4.1.5. Partial Data for Adverse Events and Previous/Concomitant Medications

For the patient data listings, no imputation of incomplete data will be applied. The listings will present the data as entered on electronic case report form (eCRF).

Dates missing the day or both the day and month of the year will adhere to the following conventions:

- If the onset date of an AE or start date of a medication is missing the day, the onset date will be set to:
 - first day of the month that the event occurred, if the onset yyyy-mm is after the yyyy-mm of first study treatment
 - the day of the first study treatment, if the onset yyyy-mm is the same as yyyy-mm of the first study treatment
 - the date of informed consent, if the onset yyyy-mm is before the yyyy-mm of the first study treatment
- If the resolution date of an AE or end date of a medication is missing the day, the resolution date will be set to:
 - the last day of the month of the occurrence
 - if the patient died in the same month, then set the imputed date as the death date
- If the onset date of an AE or start date of a medication is missing both the day and month, the onset date will be set to:
 - January 1 of the year of onset, if the onset year is after the year of the first study treatment
 - the date of the first treatment, if the onset year is the same as the year of the first study treatment
 - the date of informed consent, if the onset year is before the year of the first study treatment
- If the resolution date of an AE or end date of a medication is missing both the day and month, the date will be set to:
 - December 31 of the year of occurrence
 - if the patient died in the same year, then set the imputed date as the death date

If onset date is completely missing, then no imputation will be done and the event will be considered as treatment emergent (for AEs) or concomitant (for medications) unless the end date rules out the possibility.

For all dates prior to signing informed consent form (ICF), the following conventions will be used for imputing partial dates:

- if only the day of the month is missing, use 1st of the month
- if both the day and the month are missing, use 01 January

6.4.1.6. Study Drug Exposure

Exposure analyses will be based on the actual dose administered (in mg) and body weight (in kg) at each treatment cycle.

Ramucirumab /Placebo Treatment:

- Number of infusions

- Duration of treatment (weeks) = [(Date of last dose – date of first dose)+14]÷7
Note: 14 days added to duration of treatment because administration is every 2 weeks [on Day 1 of each 2-week cycle]

Cumulative dose, dose intensity, and relative dose intensity:

- Cumulative dose (mg/kg) = Sum of all (total doses administered [mg] ÷ Last available weight [kg] prior to each dosing)
- Dose intensity (mg/kg/week) = (Cumulative dose) ÷ (Duration of treatment)
- Planned weekly dose intensity (mg/kg/week) = 8 mg/kg / 2 weeks = 4 mg/kg/week
- Relative dose intensity (%) = (Dose intensity) ÷ (Planned weekly dose intensity) x 100.

Number of dose reductions: total number of reduction steps considering the intended dose level before each infusion (as entered in the eCRF) as referenced in [Table 2](#) below.

Table 2. Dose Reductions for Ramucirumab

Dose Level	Ramucirumab
	8 mg/kg
Reduction to first dose level	6 mg/kg
Reduction to second dose level	5 mg/kg

The protocol stipulates that study treatment be administered every 14 days ± 3 days. Hence, the minimum interval between 2 dose administrations is 11 days (14-3 days). For purposes of analysis, the number of dose delays and dose omissions will be defined as follows:

- **Number of Dose delay:** Total number of treatments reported as delayed in CRF that were administered >3 days but ≤ 11 days beyond a scheduled infusion where treatment not given.
- **Number of Dose omitted (Not Administered):** Total number of treatments reported as omitted or delayed in CRF that were administered >11 days beyond a scheduled infusion where treatment not given.

6.4.2. Efficacy Definitions and Data Handling

6.4.2.1. Overall Survival

The OS time is defined as the time from the date of randomization to the date of death from any cause. If a patient is not known to have died on or before the date of data cut-off, OS data will be censored on the last date (on or before the cut-off date) the patient was known to be alive.

6.4.2.2. Progression-Free Survival

Progression-free survival is defined as the time from the date of randomization until the date of objectively determined PD or death from any cause. Censoring rules for the PFS main analysis are described in [Table 3](#).

Table 3. Censoring Rules for PFS – Main Analysis

Situation	Event / Censor	Date of Event or Censor
Tumor progression or death	Event	Earliest date of PD or death
No tumor progression and no death	Censored	Date of last adequate radiological assessment or date of randomization (whichever is later)
<i>unless</i>		
No baseline radiological tumor assessment available	Censored	Date of randomization
No adequate post baseline radiological tumor assessment available <u>and</u> death reported after 2 scan intervals following randomization	Censored	Date of randomization
New anticancer treatment started and no tumor progression or death within 14 days	Censored	Date of adequate radiological assessment prior to (start of new therapy +14 days) or date of randomization (whichever is later)
Tumor progression or death documented <u>immediately after</u> 2 or more scan intervals following last adequate radiological tumor assessment or randomization (whichever is later)	Censored	Date of last adequate radiological assessment or date of randomization (whichever is later)

Abbreviations: CR = complete response; PD = progressive disease; PFS = progression-free survival; PR = partial response; SD = stable disease.

Notes:

1. Clinical progression (that is, symptomatic progressions, which are not radiologically confirmed) will not be considered as progressions.
2. Adequate radiological tumor assessment refers to an assessment with one of the following responses: CR, PR, SD, or PD.
3. If target, non-target and new lesion assessments have different dates within a visit, then the earliest of those dates will be considered as the date of the tumor assessment if the assessment for that visit is PD; otherwise, the latest date will be used.

The censoring rules used for sensitivity analyses on PFS will be the same as the PFS main analysis above except the one alternative rule as listed below in [Table 4](#) for each analysis.

Table 4. Censoring Rules for PFS – Sensitivity Analyses

Sensitivity Analysis	Situation	Date of Progression or Censor	Outcome
Count clinical progression as progression	Documented progression or clinical progression	Date of documented progression (New Lesion, Unequivocal Progression of Non-Target Lesion, or Progression of Target Lesion). If a tumor assessment was performed on multiple days, use the earliest date for that visit. Or date of clinical progression, whichever occurred first.	Progressed
Ignore new anticancer treatment	New therapeutic anticancer treatment started prior to documented progression or death	Date of progression or death based on subsequent tumor assessments	Progressed
Ignore missing tumor assessments	Death or documented progression after \geq two consecutively missed tumor assessment visits	Date of first scheduled tumor assessment within the time period of the consecutively missed visits	Progressed
Treat lost to follow up as progression (worst case scenario)	Patient lost to follow-up without documented progression or death	Date of next scheduled tumor assessment after last tumor assessment	Progressed

Abbreviation: PFS = progression-free survival.

6.4.2.3. Time to Deterioration in FHSI-8

Time to deterioration in FHSI-8 total score is defined as the time from the date of randomization to the first date observing a decrease of ≥ 3 points from baseline. In case of no FHSI-8 deterioration, the subject will be censored at the time of the last FHSI recording. Additionally, the deterioration threshold and clinically meaningful difference will be evaluated in the trial population as described in the Psychometric Statistical Analysis Plan.

The deterioration criteria that FHSI-8 total score decrease of ≥ 2 or 4 points will be used as sensitivity analyses.

Other patient-focused outcome analysis is defined in Section 6.4.2.7.

6.4.2.4. Time to Deterioration in ECOG PS

Time to deterioration in ECOG PS is defined as the time from the date of randomization to the first date observing ECOG PS ≥ 2 (that is, deterioration from baseline status of 0 or 1). In case of no PS greater than 2, the subject will be censored at the time of the last PS recording.

The deterioration criteria that the PS change from baseline ≥ 2 or PS change to ≥ 3 will be used as sensitivity analysis.

6.4.2.5. Objective Response Rate and Disease Control Rate

The ORR will be calculated as the number of patients who achieve a best response of complete response (CR) or partial response (PR) using investigator response assessments divided by the total number of patients randomized to that study arm. Disease control rate (DCR) will be calculated as the number of patients who achieve a best response of CR, PR, or SD using

investigator response assessments divided by the total number of patients randomized to that study arm. The Response Evaluation Criteria In Solid Tumors Version 1.1 (RECIST 1.1) guideline will be followed in this clinical trial. Patients who do not have a tumor response assessment for any reason are considered non-responders and are included in the denominator when calculating the ORR or DCR.

A PR or CR will not require confirmation to be considered to be a response. In the case of SD, follow-up measurements must have met the SD criteria at least once after study entry at a minimum interval of 6 weeks (that is, 6 weeks – 3 days = 39 days as allowed per protocol) from randomization to assign a best response of SD. If there are no postbaseline scans available for a patient, then the best response will be not evaluable (NE).

6.4.2.6. Time to Radiographic Progression

Time to radiographic progression (TTP) is defined as the time from the date of randomization until the date of radiographic progression according to RECIST 1.1. Censoring rules are described in [Table 5](#).

Table 5. Rules for Determining Date of Progression or Censor for TTP

Situation	Event / Censor	Date of Event or Censor
Tumor progression	Event	Earliest date of PD
No tumor progression (regardless of death)	Censored	Date of last adequate radiological assessment or date of randomization (whichever is later)
<i>unless</i>		
No baseline radiological tumor assessment available	Censored	Date of randomization
No adequate post baseline radiological tumor assessment available	Censored	Date of randomization
New anticancer treatment started and no tumor progression within 14 days	Censored	Date of adequate radiological assessment prior to (start of new therapy +14 days) or date of randomization (whichever is later)
Tumor progression documented <u>immediately after 2</u> or more scan intervals following last adequate radiological tumor assessment or randomization (whichever is later)	Censored	Date of last adequate radiological assessment or date of randomization (whichever is later)

Abbreviations: CR = complete response; PD = progressive disease; PFS = progression-free survival; PR = partial response; SD = stable disease.

Notes:

1. Clinical progression (that is, symptomatic progressions, which are not radiologically confirmed) will not be considered as progressions.
2. Adequate radiological tumor assessment refers to an assessment with one of the following responses: CR, PR, SD, or PD.
3. If target, non-target and new lesion assessments have different dates within a visit, then the earliest of those dates will be considered as the date of the tumor assessment if the assessment for that visit is PD; otherwise the latest date will be used.

6.4.2.7. Patient-Focused Outcomes (PFOs)

Patient-reported symptoms will be assessed using the Functional Assessment of Cancer Therapy (FACT) Hepatobiliary Symptom Index-8 (FHSI-8). The FHSI-8 is a 5-point, 0-4 assessment (a higher score representing high symptom burden) for 8 symptoms including some more specific to hepatobiliary cancer (jaundice, stomach pain/discomfort) and some that are associated with generalized advanced/metastatic malignancy (weight loss, fatigue) (Yount et al. 2002).

The FHSI-8 is scored to obtain a single FHSI-8 total score of 0 to 32, with a higher score representing low symptom burden. The FHSI-8 will be scored according to the developer's guidelines, including imputation by prorating if >50% of items (5 items or more out of the 8) are completed.

The prespecified minimal clinically important difference from baseline in FHSI-8 total score is approximately 10% of the range of the scale, or 3 points (Steel et al. 2005).

Scores for each postbaseline time point will be compared to the baseline value and categorized as follows:

Deterioration: Defined as a decrease of 3 points or more

Improvement: Defined as an increase of 3 points or more

Stable: Defined as no change or increase/decrease <3 points

The best response across all time points will be obtained by categorizing the best total score as deterioration, improvement, or stable. The worst response will be similarly obtained by categorizing the worst total score.

Additional meaningful change thresholds will also be derived from the trial population as described in the Psychometric Statistical Analysis Plan. The time to deterioration analysis will be performed using thresholds derived from these analyses in addition to the prespecified 3-point decrease.

Time to deterioration for individual item of FSHI-8 is the time from the date of randomization to the first date observing deterioration of each item, with the deterioration threshold defined as a decrease in one categorical response from baseline of each item. In case of no deterioration, the subject will be censored at the time of the last assessable FSHI-8 individual item recording.

EuroQol – 5 Dimensions 5 Level (EQ-5D-5L)

Additionally, patients' health status will be assessed using EQ-5D-5L. EuroQol – 5 Dimensions index score is calculated from a set of item weights to derive a scale of 0 to 1, with 0 representing death and 1 representing perfect health, with the possibility of negative scores. The index score will be based on the value set for England, where the lowest possible value is -0.281 (Devlin et al. 2017). The visual analog scale (VAS) is scored from 0 (*worst imaginable health state*) through 100 (*best imaginable health state*) to represent the patient health condition.

For FSHI-8 and EQ-5D-5L, other than baseline visit and end of treatment visit, a visit window is applied. If the actual assessment date falls in the window of study days $[(X-1)*7-20, (X-1)*7 + 21]$, where $X = 7, 13, 19, \dots$, the visit will be presented as Week X. If there are multiple assessments within 1 visit window, the last assessment will be used in the summary by visit.

For percentage compliance of the FSHI-8 and EQ-5D-5L, instruments with at least 1 item completed will be considered as having been completed. No other adjustment or imputation for missing data will be performed. To evaluate the impact of missing data on results, sensitivity analyses will be performed from the following imputations: last observation carried forward, worst observation carried forward, and multiple imputation using Markov Chain Monte Carlo methodology.

6.4.3. Other General Data Handling

- **Age:** Age will be computed from July 1 of the year of birth to the date of Informed Consent, as:

$$(\text{Date of Informed Consent} - \text{July 1 of the year of Birth} + 1) / 365.25.$$
- **Missing Data:**

- All analyses and descriptive summaries will be based on the observed data. Unless otherwise specified, missing data will not be imputed or “carried forward.”
- If the causality (relationship to study drug) of an adverse event is missing, then it will be considered as related to study drug.
- **Baseline Measurements:**
 - **Efficacy analysis including QoL:** The last measurement on or prior to the date of randomization will serve as the baseline measurement. In the event such a value is missing, the first assessment completed prior to the first study drug administration will be used as the baseline assessment so long as this assessment was taken within 7 days of randomization.
 - **Safety analysis:** The last non-missing measurement prior to the first study drug administration will be used as the baseline assessment.
 - **Baseline characteristics:** The last measurement on or prior to the date of randomization will serve as the baseline measurement. In the event such a value is missing, the first assessment completed on or prior to the date of first study drug administration will be used as the baseline assessment.
- **Study Day:**
 - For safety analysis: Study day is calculated as:
 - assessment date – first dose date + 1; if the assessment was performed on or after the first dose day.
 - assessment date – first dose date; if the assessment was performed prior to the first dose date
 - For efficacy analysis including QoL: Study day is calculated as:
 - assessment date – randomization date + 1; if the assessment was performed on or after the randomization date.
 - assessment date – randomization date; if the assessment was performed prior to the randomization date.
- **Duration:** Duration (except duration of study treatment) is calculated as:
 - Duration (days): (End Date – Start Date + 1)
 - Duration (weeks): (End Date – Start Date + 1) / 7
 - Duration (months): (End Date – Start Date + 1) / 30.4375
 - (Average number of days in a month = average number of days in a year /12)
 - Duration (years): (End Date – Start Date + 1) / 365.25
 - (Average number of days in a year = 365.25, reflecting the Julian Year of 3 years with 365 days each and 1 leap year of 366 days)

- **Time-to-event:** The efficacy and QoL event or censoring time (days) is calculated as:
Date of event/censoring – Date of randomization + 1
Investigator’s assessments will be used to derive PFS, TTP, and best tumor response.
- **Data Presentation:** Categorical variables will be summarized in frequency tables, with the counts and percentage of patients in each category. Percentages given in the summary tables will be rounded and thus may not always add up to exactly 100 percent. For continuous variables, summary statistics will include N, mean, standard deviation, median, Q1-Q3, and minimum and maximum values (range).

6.5. Data Analysis

6.5.1. Disposition of Patients

The number of patients enrolled (that is, signed informed consent), randomized, treated, and discontinued (that is, discontinued the study treatment for reason(s) such as disease progression, AE, lost to follow-up, or withdrawal of consent) will be presented by site and treatment group. Patients enrolled but not randomized will be summarized by reason for not being randomized.

The number and percentage of patients who were in the ITT population, in the Safety population, and in the PP population will be summarized. Moreover, the number and percentage of patients still receiving treatment at data cut-off date will be summarized by treatment group. Patients having completed their End of Treatment visit will be presented by reason for discontinuation of study treatment. Patients excluded from the PP population will be summarized by reason for exclusion.

Besides summary tables, by-patient listings will also be generated.

6.5.2. Demographic and Baseline Characteristics

The following demographic and baseline characteristics will be summarized based on the ITT population:

- Sex
- Race
- Ethnicity
- Age (years)
- Age group (<65 vs. ≥65 years; <75 vs. ≥75 years)
- Height (cm)
- Weight (kg)

In addition, the stratification factors (per Interactive Web Response System [IWRS] and eCRF) and the discrepancy between 2 data sources will be summarized based on the ITT population:

- Geographical region (Region 1 vs. Region 2 vs. Region 3)
- Macrovascular invasion presence
- ECOG PS

The following disease characteristics and prior therapy information will be summarized based on the ITT population:

- Duration of Disease (defined as time from diagnosis confirmation to randomization; if the day of diagnosis confirmation is unknown it will be replaced by 15MMMYYYY)
- BCLC stage
- Child-Pugh score
- Etiology of liver disease (hepatitis B vs. hepatitis C vs. other)
- Basis of HCC diagnosis
- Tumor differentiation (well differentiated, moderately differentiated, poorly differentiated, unable to determine, not applicable)
- Number of metastatic sites Involved (0, 1, 2, ≥ 3)
- Extrahepatic spread presence
- Duration of prior sorafenib treatment (<5 months vs. ≥ 5 months)
- Reason for discontinuation of sorafenib (PD, toxicity)
- Time from last sorafenib treatment to randomization (<1 month vs. ≥ 1 month)
- Had any prior locoregional therapy (yes vs. no)
- Macrovascular invasion presence (CRF) (yes vs. no)
- Presence of ascites Grade ≥ 2 (yes vs. no)
- Required medical or procedural management for ascites even if Grade ≤ 1 at the time of study entry

Prior anticancer therapy, preexisting medical condition, previous surgery for HCC, prior locoregional therapy, and prior radiotherapy for HCC will all be presented.

6.5.3. Analysis of Efficacy Data

6.5.3.1. General Considerations

The primary statistical analysis will be conducted in the ITT population following the intent-to-treat principle, including eligible and ineligible patients and based on the treatment group to which patients was randomized (irrespective of which treatment was received). Control of the type I error for multiple testing will be maintained by the sequential gatekeeping testing for the primary and secondary efficacy variables (Section 6.5.3.3). Sensitivity and subgroup analyses of the primary and secondary endpoints will be performed.

The stratification categories used for the primary and secondary analyses are:

- geographic region (Region 1 versus Region 2 versus Region 3)
- macrovascular invasion (yes versus no)
- ECOG PS (0 versus 1)

Unless otherwise specified, all stratified analyses of primary and secondary endpoints will be based on the stratification factors per IWRS.

The final analysis will be conducted once approximately 221 deaths have been observed in the ITT population using a 2-sided nominal significance level of 0.05.

6.5.3.2. Primary Efficacy Endpoint - Overall Survival

The primary analysis will compare the observed OS between the 2 treatment groups (ramucirumab plus BSC versus placebo plus BSC).

The primary OS analysis will be conducted in the ITT population. The comparison will use the log-rank test, stratified by the randomization stratification factors.

Hazard ratio (HR) and its two-sided 95% confidence interval (CI) will be estimated using Cox proportional hazards (PH) model stratified by the randomization strata. Additionally, OS curves will be presented using the Kaplan-Meier method by treatment arm, together with a summary of associated statistics (that is, median OS, 6-months and 1-year OS rates including corresponding two-sided 95% CI). The 95% CI for the median survival time will be calculated according to Brookmeyer and Crowley (1982). The 95% CI for the OS rate difference will also be calculated. The total number of events and censored observations will be presented for each treatment group along with reason for censoring (for example, patient still alive at data cut-off date, lost to follow-up, etc.).

The primary analysis will be performed using the stratification variables as captured by the IWRS.

The following sensitivity analyses are planned:

- An analysis of OS based on the PP population
- An analysis of OS in the ITT population with an unstratified log-rank test
- A stratified analysis of OS in the ITT population using the stratification factors as reported in the eCRF
- An analysis of OS in the ITT population with baseline AFP ≥ 400 ng/mL based on central lab result
- An analysis of OS adjusting the treatment effect for significant prognostic factors (see Section 6.5.3.5)

A listing will be prepared presenting relevant information on the OS (including date of randomization, date of death, date last known to be alive, censoring status, and reason for censoring).

6.5.3.3. Secondary Efficacy Endpoints

A gatekeeping approach to selected secondary endpoints will be applied so as to protect the study-wise type I error rate and to enable inferential statements; each hypothesis is inferentially tested only if each of the preceding hypotheses were rejected. The sequential order of the confirmatory testing after OS in the ITT population will be: 1) PFS in the ITT population; 2) time to deterioration in FHSI-8 total score in the ITT population; and 3) time to deterioration in ECOG PS in the ITT population. Secondary endpoints will be analyzed at the same time as OS and at the same level of significance (1-sided 0.025).

6.5.3.3.1. Progression-Free Survival

The censoring rules in [Table 3](#) and [Table 4](#) will be used for the primary and sensitivity analyses for PFS. PFS will be analyzed using a stratified log-rank test, stratified by the same stratification factors used in the analysis of the primary endpoint, OS. The HR will be generated using a stratified Cox regression model. An additional PFS analysis with an unstratified log-rank test will also be performed. The estimation of survival curves for the 2 treatment groups will be generated using the Kaplan-Meier method. The medians as well as the 3-month and 6-month PFS rates will be provided along with their 95% CI. The total number of events (disease progressions, deaths, total) and censored observations will be presented for each treatment group along with reason for censoring. PFS will be analyzed for the ITT population. Additionally, PFS will be analyzed for the PP population.

A listing will be prepared presenting relevant information on the PFS time (including date of randomization, date of last radiological scan, date of last radiological scan with overall response SD, PR, or CR, date of radiological progression, date of symptomatic progression, date of death, censoring status, and reason for censoring). A separate listing will be prepared using the sensitivity censoring rules.

6.5.3.3.2. Time to Deterioration of FHSI-8

The time to deterioration of FHSI-8 total score will be compared between the treatment groups from stratified (primary) and unstratified log-rank tests. Kaplan-Meier graphs by treatment arm will be produced. If deterioration is not observed, then time to deterioration is censored at the date of the last FHSI-8 assessment for that patient. The HR and 95% CIs will be estimated using Cox's PH regression model, sensitivity analysis may include covariates of assigned treatment and baseline total score.

Additional sensitivity analyses may be performed using definition described in [Section 6.4.2.3](#).

6.5.3.3.3. Time to Deterioration of ECOG PS

The ECOG performance status results will be summarized using frequency distributions for each scheduled visit, including also the best and worst postbaseline value.

Time to deterioration will be analyzed for the ITT population using the Kaplan-Meier method, using an stratified (primary) and unstratified log-rank test. Additional sensitivity analyses may be performed using definition described in [Section 6.4.2.4](#).

Event/Censoring for the time to deterioration analysis:

- Event date is the first date observing the required PS value or deterioration
- Censor at time of the last PS value if no event

In order to investigate the relationship between objective progression and deterioration in ECOG PS, an analysis of changes from baseline in ECOG PS by objective progression (yes vs. no) by tumor measurement period will be performed.

6.5.3.3.4. Objective Response Rate and Disease Control Rate (DCR)

The ORR will be calculated as the number of patients who achieve a best response of CR or PR using the investigator response assessments, divided by the total number of patients randomized to that study arm in the ITT population. Patients who do not have a tumor response assessment for any reason are considered as non-responders and are included in the denominator when calculating the response rate. The RECIST 1.1 will be followed in this clinical trial.

Frequencies for Best Overall Response (BOR) will be presented by treatment group, as well as ORR and DCR observed in each treatment group together with 95% CI.

The ORRs and DCR observed in each treatment group will be compared using the exact Cochran-Mantel-Haenszel test adjusting for the stratification variables as captured by IWRS.

A listing will be prepared presenting relevant information on response (including date of randomization, date of radiological scans, target lesion percentage change, response on target-lesions, response on non-target lesions, new lesions, and overall response for each tumor assessment, BOR).

6.5.3.3.5. Time to Radiographic Progression

The censoring rules presented in [Table 5](#) will be used for the analysis for TTP. The treatment groups will be compared using the stratified log-rank test, as well as unstratified log-rank test. The estimation of survival curves for the 2 treatment groups will be generated using the Kaplan-Meier method. Median TTP will be provided with a two-sided 95% CI. CIs will be computed in the same way as for OS. The HR will be generated using the stratified as well as unstratified Cox regression model.

Time to radiographic progression will be analyzed for the ITT population.

6.5.3.4. Subgroup Analysis

Subgroup analyses will be performed for OS, the primary endpoint, and PFS. Only unstratified analysis will be performed on each subgroup. Each analysis will use the similar methodology as for the primary analysis. A Forest plot of the estimated HRs with 95% CIs will be provided for subgroups with sufficient number of events (≥ 15).

The following is the list of subgroups:

- Gender (Males vs. Females)
- Age (< 65 years vs. ≥ 65 years)
- Race (white vs. Asian vs. all others)
- Geographical region (Region 1 vs. Region 2 vs. Region 3)
- Etiology of liver disease (hepatitis B vs. hepatitis C vs. other)
- Presence/absence of extra-hepatic metastases
- Presence/absence of macrovascular invasion
- BCLC score (B vs. C)
- Baseline ECOG PS (0 vs. 1)
- Had any prior locoregional therapy

- Reason for discontinuation of sorafenib (PD vs. toxicity)

Tests within each subgroup and tests for subgroup-by-treatment interaction terms will use an unstratified test and unstratified Cox PH model.

The goal of subgroup analyses is to assess internal consistency of study results and whether there is significant treatment heterogeneity across any of the subgroups. Since, even if all patient 'groups' benefit to exactly same extent, smaller or larger estimated effects, even negative effects, may be seen for some subgroups simply by chance alone. Without appropriate interpretation, this can lead to erroneous conclusion in one or more subgroups, in particular where differential treatment effects are not expected across any of the factors assessed. In order to assist with interpretation of the subgroup results, the methodology of Fleming (1995) may be followed to provide background information on the extent of variability that might be expected by chance alone.

6.5.3.5. Analysis Adjusting for Covariates

As supportive analysis, the primary (OS) and secondary (PFS) efficacy endpoints will also be analyzed after adjusting for selected prognostic factors. Potential prognostic factors include the stratification factors and other factors as listed in the list of subgroup analyses in Section 6.5.3.4.

Hazard ratio for treatment effect will be estimated using an unstratified multivariable Cox PH model to be constructed by selecting variables among all the potential variables using stepwise selection method with entry p-value 0.05 and exit p-value 0.1. The covariate selection process will be based on the entire ITT population. The treatment factor will be kept out of the model throughout the covariate selection process and only added into the final model. HR for treatment effect and corresponding 95% CI will be estimated from the final model.

6.5.3.6. Restricted Mean Difference Analysis

The common method for describing benefit on the time scale is to calculate the difference in median event time between the two treatment arms. An alternative method for describing benefit on the time scale is to estimate the average difference between the Kaplan-Meier (KM) curves. This corresponds to calculating the difference in the average time to event for the 2 treatment arms (Irwin 1949; Karrison 1997; Meier et al. 2004). Similar to the HR, this method uses all of the available information across the KM curves, but has the additional advantage of assessing benefit on the time scale.

To estimate an improvement in OS with ramucirumab, we will follow the method of Irwin (1949) detailed in Karrison (1997) and Meier (2004) for estimating the 'difference in average OS', which we will refer to more formally as the restricted mean difference in OS. The area under each KM curve will be calculated using numerical integration (trapezium rule) per Karrison and implemented in SAS using PROC LIFETEST. The difference between treatment arms and a CI for the difference will be formed.

Since the KM curve may be ill-determined beyond a certain range, or even undefined (if the longest observation is censored), for evaluation and comparison of means, the area under each KM curve will be calculated between time 0 and restriction time T, which is why this is referred

to as a "restricted mean". Following the suggestion of Karrison, the restriction time T will be chosen as largest time point t such that the standard error (SE) of the survival estimate at time t in each treatment group is no more than 0.075. For this purpose, we will use the simple, albeit conservative, formula proposed by Peto et al. (1977) for calculating the SE of S(t) as:

$$SE(S(t)) = S(t)\sqrt{(1 - S(t))/n(t)},$$

where n(t) is the number of patients still at risk at time t.

Besides for OS, similar restricted mean analysis can be constructed for PFS.

6.5.4. Postdiscontinuation Anticancer Therapy

The number and percentage of patients with any additional postdiscontinuation anticancer treatment (PDT) including systemic, locoregional, and surgery will be presented for the ITT population by the treatment arm as well as by the region. Postdiscontinuation anticancer treatments will be summarized by therapy type and by drug name, and will also be reported in listings.

Imbalances between treatment arms in PDT use can confound the evaluation of the treatment effect for OS. If a notable imbalance in PDT use is observed (either overall or with respect to important agents or classes of agent), a sensitivity analysis for OS may be conducted in which patients will be reweighted in such a way as put more weight on patients with PDT in the arm that had less PDT use, and less weight on patients with PDT in the arm that had more PDT use, and thereby the rate of PDT use will be balanced between arms on a weighted basis. The PDT-weighted analysis will help assess the impact of any observed difference in the rate of PDT use between arms.

Additional analysis may be explored to help interpret OS results, for example, analysis of OS censoring at the start of PDT.

6.5.5. Patient-Focused Outcomes

For each of the instruments, compliance will be summarized for the ITT population. Compliance at an assessment time point is defined as the number of patients who provided data divided by the expected number of patients on study at that time point. A patient who answers at least one item at a time point is considered to have been assessed. Compliance will be summarized by time point, and reasons for non-compliance will be tabulated. Compliance at different time points along with the reasons for noncompliance may be represented graphically using a stacked bar chart.

6.5.5.1. FHSI-8 Analyses

The proportion of patients in each arm with deterioration in scores, improvement in scores, and stable scores at specific time points will be presented both in a summary table and graphically using stacked bar charts. The proportion of patients who deteriorated at each assessment time point may be compared using Fisher's exact test. Patients with no data will be considered deteriorated.

FHSI-8 score will be summarized descriptively at each assessment time by treatment arm, as well as changes from pre-treatment visit, provided at least 30% of the patients have an FHSI-8 score at that time point. Best and worst responses across all the visits will also be summarized in a table.

The time to deterioration of FHSI-8 individual item will be compared between the treatment groups from stratified and unstratified log-rank tests. The HR and 95% CIs will be estimated using Cox's PH model, covariates may include assigned treatment and baseline score.

6.5.5.2. EQ-5D-5L Analyses

The 5-dimension single-item 5-level EQ-5D-5L responses will be summarized using frequency distributions for each dimension and for each assessment time by treatment arm.

The index score and the VAS will be presented using summary statistics for continuous variables for each assessment time by treatment arm.

6.5.6. Analyses of Safety Data

The 'safety population' as defined in the Section 6.3.2 will be used in the analysis of safety.

In addition, selected AE summaries will be provided for subgroups including age, gender, and race; additional subgroups may be evaluated.

6.5.6.1. Study Drug Exposure

The following exposure-related variables will be reported using summary statistics (number of patients, mean, and standard deviation) by treatment group:

- Exposure: duration of treatment; number of cycles received; number of patients completing ≥ 1 cycle, ≥ 2 cycles, ..., ≥ 6 cycles, and mean, standard deviation; number of patients with dose adjustments: dose reduction, dose delay, and dose omission;
- Reasons for dose adjustments.

The following exposure-related variables will be reported using summary statistics (number of patients, mean, standard deviation, median, first and third quartiles, minimum, and maximum) by treatment group:

- Dose intensity: cumulative dose; weekly dose intensity; relative dose intensity, overall weekly dose intensity, overall relative dose intensity.

Details of study drug administration will be included in patient listings.

The number and percentage of patients having received any premedication will be presented by Anatomical Therapeutic Chemical level 2. Selected combinations of premedication will also be summarized.

6.5.6.2. Overall Summary of Adverse Events

All adverse events (AEs) will be summarized by MedDRA System Organ Class (SOC) and PT. The incidence and percentage of patients with at least one occurrence of a preferred term will be

included, according to the most severe National Cancer Institute - Common Terminology Criteria for Adverse Events (NCI-CTCAE) Version 4.0 grade.

An overall summary of AEs will be provided to summarize the following categories using frequency counts and percentages:

- patients with at least 1 TEAE, SAE, Common Terminology Criteria for Adverse Events (CTCAE) Grade ≥ 3 TEAE
- patients with AEs that led to death
- patients with AEs that led to study treatment discontinuation
- patients with SAEs that led to study treatment discontinuation.

The summary will be provided for regardless of study drug causality, and repeated for events deemed by the investigator to be related to study treatment.

6.5.6.3. Treatment-Emergent Adverse Events

Causality (relationship to study drug) as determined by the investigator will be separately summarized. Missing classifications concerning study medication relationship will be considered as related to study medication. Start and end dates of AE will be included in the listings along with action taken, and outcome. If more than one AE is recorded for a patient within any SOC or PT term, the patient will only be counted once on the most severe grade and the closest relationship to treatment.

The following set of summaries will be provided for any TEAE:

- By SOC and PT (all grade and Grade ≥ 3)
- By PT (all grade and Grade ≥ 3) by decreasing frequency on ramucirumab arm
- By consolidated terms
- By maximum Grade 1-5

These summaries will also be produced for TEAEs considered related, by the investigator, to study treatment.

All collected AEs (treatment emergent or non-treatment emergent) will be presented in a listing.

6.5.6.4. Deaths, SAEs, and Other Significant Adverse Events

Reasons for deaths (study disease, AE [any AE, study treatment-related AE]) will be summarized separately for 1) all deaths, 2) death on treatment, 3) death within 30 days of study treatment discontinuation, and 4) deaths after 30 days of study treatment discontinuation.

Serious adverse events will be summarized by SOC and PT, by PT, and repeated for events deemed by the investigator to be possibly related to study medication, with consolidated summary performed if needed. A listing of SAEs will be produced.

In addition, the following analyses will be performed (*repeated for events deemed by the investigator to be possibly related to study medication, †include consolidated summary):

- Adverse events leading to death by PT†
- Adverse events leading to study treatment discontinuations by PT†
- Adverse events leading to study treatment dose modification by PT†
- Adverse events of Special Interests

Listing of AESI will be provided.

6.5.6.5. Vital Signs and Other Observations Related to Safety

For systolic and diastolic blood pressure, shift from baseline to pre and post ramucirumab/placebo administration will be summarized at each assessment time point using summary statistics according to the following categories:

- Systolic blood pressure (<140, 140-<160, ≥160 mmHg)
- Diastolic blood pressure (<90, 90-<100, ≥100 mmHg)

Weight at baseline will be presented using summary statistics. Changes from baseline to on-treatment weight assessments will be presented by time point by considering the frequency of patients with changes falling in the following categories: <-10% (loss), ≥-10% - <10%, ≥10% (gain).

A summary of ECOG PS at each scheduled time point will be provided. Refer to Section [6.5.3.3.3](#) for analysis of ECOG PS.

Listings of ECOG and vital signs will be provided.

6.5.6.5.1. Laboratory Parameters

Laboratory results will be graded according to the NCI-CTCAE, v 4.0, when applicable. Grading will be purely based on the numeric results and no investigator's assessment will be considered. Laboratory values will be converted to standard (SI) units, as referenced in the NCI-CTCAE v 4.0. Laboratory results not corresponding to a NCI-CTCAE v 4.0 terms will not be graded.

For the graded laboratory parameters, shift tables from baseline to the worst value on the study will be generated.

Laboratory results will be presented in a listing with a flag for values outside of the laboratory normal range. Retest results will be reported in the listings as well.

6.5.6.5.2. Electrocardiogram (ECG)

ECG abnormalities (Yes, No) will be summarized using frequency distributions by visit. For this study, ECG is performed only at baseline and short-term follow up visit.

6.5.6.5.3. Echocardiogram/Multiple-Gated Acquisition (MUGA)

Echocardiogram abnormalities (Yes, No) will be summarized using frequency distributions by visit. For this study, echocardiogram/MUGA is performed only at baseline and short-term follow-up visit.

6.5.6.6. Hospitalizations

The number of patients hospitalized on study or within 30 days of treatment discontinuation drug will be presented by reason for hospitalization. The number of patients with 1, 2, 3, and >3 hospitalizations due to AEs will be summarized. The total duration of hospitalizations will be summarized by treatment group.

6.5.7. Concomitant Medication and Procedure

The number and percentage of patients receiving concomitant medications/procedures of interest will be tabulated. The list may include transfusions, anti-emetic and anti-nauseants, analgesics, anti-infectives, erythropoietic agents, granulocyte colony-stimulating factor or granulocyte macrophage colony-stimulating factor. Final list will be determined prior to database lock.

6.5.8. Pharmacokinetic (PK) Analyses and Immunogenicity Analyses

Serum concentrations of ramucirumab prior to infusion (C_{min}) and at 1 hour post-end of ramucirumab infusion (approximately C_{max}) will be summarized using descriptive statistics. Additional analysis utilizing a population PK approach based on an established population PK model may also be conducted if deemed appropriate. Relationships between ramucirumab exposure and measures of efficacy and safety will be explored. Detailed exposure-response analysis plan is in a separate document.

A subject who is evaluable for treatment-emergent anti-drug antibodies (ADA) is treatment-emergent ADA-positive (TE ADA+) if either of the following holds:

- **Treatment-induced ADA+ subject:** The subject has baseline status of ADA Not Present and at least 1 postbaseline status of ADA Present with titer $\geq 2 \times \text{MRD}$ (see Assay Operating Characteristics).
- **Treatment-boosted ADA+ subject:** The subject has baseline and postbaseline status of ADA Present, with the postbaseline titer being 2 dilutions (4-fold) greater than the baseline titer. That is, the subject has baseline status of ADA Present, with titer 1:B, and at least 1 postbaseline status of ADA Present, with titer 1:P, with $P/B \geq 4$. For immunogenicity, the number and percent of patients with treatment-emergent ramucirumab ADA will be summarized. Listings will be provided of (i) Treatment-emergent adverse events for patients with either at least 1 sample of ramucirumab ADA present or infusion-related reaction or both, and (ii) Antibody to ramucirumab and drug concentration data for patients who have at least 1 sample result of ADA present.

6.5.9. Exploratory Biomarker Analysis

Pharmacodynamic biomarker analyses will be described in a separate SAP.

6.6. Interim Analyses and Data Monitoring Committee (DMC)

An Independent Data Monitoring Committee (IDMC) will be established prior to the inclusion of the first patient in the trial. The IDMC will review unblinded analyses of safety data produced by an independent statistical team. The study team will only review blinded information. Detailed procedure is described in the IDMC charter.

Selected tables and listings will be prepared for the purpose of the interim safety analysis, all based on the safety population.

7. Unblinding Plan

Randomization will occur using an IWRS system. Assignment to treatment groups will be determined by a computer-generated random sequence. Security measures will be implemented so that treatment group code and other variables that can link patients to study arm will be blinded in the database.

In order to maintain the scientific integrity of this double-blind trial, access to study data will be strictly controlled prior to the final analyses. Dummy treatment assignment will be used in the reporting database until the primary database lock for the final analysis of overall survival. While every effort will be made to blind both the patient and the investigator to the identity of the treatment, the inadvertent unblinding of a patient may occur. This unblinding will not be sufficient cause (in and of itself) for that patient to be excluded from any safety or efficacy analyses.

8. Changes in Planned Analyses from the Protocol

No changes have been made from analyses planned in the protocol.

Redacted Version

9. References

- Brookmeyer R, Crowley J. A confidence interval for the median survival time. *Biometrics*. 1982;38(1):29-41.
- Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ*. 2017. Epub ahead of print.
- EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy*. 1990;16(3):199-208.
- Fleming TR. Interpretation of subgroup analyses in clinical trials. *Drug Inf J*. 1995;29:1681S-1687S.
- Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, Bonsel G, Badia X (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*. 2011;20(10):1727-1736.
- Irwin JO. The Standard Error of an Estimate of Expectational Life. *J Hygiene*. 1949;47:188-189.
- Karrison T. Use of Irwin's Restricted Mean as an Index for Comparing Survival in Different Treatment Groups-- Interpretation and Power Considerations, *Controlled Clin Trials*. 1997;18:151-167.
- Meier P, Karrison T, Chappell R, Xie H. The Price of Kaplan-Meier. *J Am Stat Assoc*. 2004;99(467):890-896.
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer*. 1977;35(1):1-39.
- Steel JL, Eton DT, Cella D, Olek MC, Carr B. Clinically meaningful changes in health-related quality of life in patients diagnosed with hepatobiliary carcinoma. *Ann Oncol*. 2005;17(2):304-312.
- Szende A, Oppe M, Devlin N. EQ-5D value sets: Inventory, comparative review and user guide. EuroQol Group Monographs Volume 2. Springer, 2006.
- Yount S, Cella D, Webster K, Heffernan N, Chang C, Odom L, et al. Assessment of patient-reported clinical outcome in pancreatic and other hepatobiliary cancers: the FACT Hepatobiliary Symptom Index. *J Pain Symptom Manage*. 2002;24(1):32-44.

10. Appendices

Redacted Version

Appendix 1. Analyses for Pharmaceuticals and Medical Devices Agency (Japan) (PMDA)

Selected efficacy and safety analyses, appropriate for PMDA submission, will be performed for the following subgroups based on the country of each site: Japan, East Asia (excluding Japan) and other. Table APP.1 describes the actual country grouping for these analyses. A list of analyses is presented in a separate document.

Table APP.1. Grouping of Countries that Randomized Patients for Japan PMDA

Grouping for PMDA (Japan)	Country
Japan	Japan (JPN)
East Asia (excluding Japan)	Hong Kong (HKG)
East Asia (excluding Japan)	Korea (KOR)
East Asia (excluding Japan)	Taiwan
East Asia (excluding Japan)	China
Other	All other countries

Appendix 2. Clinical Trial Registry Analyses

For the purpose of fulfilling the Clinical Trial Registry requirements, summary of SAEs (whether treatment emergent or not) and ‘Other’ AEs (that is, non-serious TEAEs) by PT and treatment group will be performed. For each PT, the number of patients at risk, patients who experienced the event, and the number of events will be presented. In addition, the summary will be provided as a dataset in XML format.

██
██
Approval Date & Time: 21-Dec-2017 17:31:35 GMT
Signature meaning: Approved

Redacted Version