I4T-CR-JVCR Statistical Analysis Plan Version 5.0

A Randomized, Multicenter, Double-Blind, Placebo-Controlled, Phase 3 Study of Weekly Paclitaxel with or without Ramucirumab (IMC-1121B) in Patients with Advanced Gastric or Gastroesophageal Junction Adenocarcinoma, Refractory to or Progressive after First-Line Therapy with Platinum and Fluoropyrimidine Study

NCT02898077

Approval Date: 30-Jul-2020

1. Statistical Analysis Plan: I4T-CR-JVCR: A Randomized, Multicenter, Double-Blind, Placebo-Controlled, Phase 3 Study of Weekly Paclitaxel with or without Ramucirumab (IMC-1121B) in Patients with Advanced Gastric or Gastroesophageal Junction Adenocarcinoma, Refractory to or Progressive after First-Line Therapy with Platinum and Fluoropyrimidine Study

Ramucirumab (LY3009806) Gastric or Gastroesophageal Junction Adenocarcinoma

This is a randomized, placebo-controlled, double-blind, multicenter, Phase 3 study of patients with advanced gastric or gastroesophageal junction (GEJ) adenocarcinoma, refractory to or progressive after first-line therapy with platinum and fluoropyrimidine. Patients will be randomized to receive either paclitaxel plus ramucirumab or paclitaxel plus placebo. Treatment will continue until disease progression, unacceptable toxicity, or meeting other treatment discontinuation criteria.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I4T-CR-JVCR
Phase 3

SAP Version 1 electronically signed and approved by Lilly 31-Jul-2019
SAP Version 2 electronically signed and approved by Lilly 02-Aug-2019
SAP Version 3 electronically signed and approved by Lilly 05-Aug-2019
SAP Version 4 electronically signed and approved by Lilly 29-Jul-2020
SAP Version 5 electronically signed and approved by Lilly on the date provided below

Approval Date: 30-Jul-2020 GMT

2. Table of Contents

Section **Page**

Figure

Table

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to the first unblinding, to allow execution of activities related to the interim analysis.

SAP Version 2 was approved prior to the first unblinding. AE reporting process was revised in this version.

SAP Version 3 was approved prior to the first unblinding. Revised definition of baseline disease characteristics "time to progression on first-line therapy"

SAP Version 4 was approved prior to the final database lock. The changes made in this version include:

1) Added additional sensitivity analyses for OS and PFS in Section 6.9.3.
2) Adjusted language for major protocol deviations

SAP Version 5 was approved prior to the final database lock. Definition for "time to progression during first-line therapy" was revised.

4. Study Objectives

4.1. Primary Objective

The co-primary objective of this study is to evaluate PFS and OS in patients treated with paclitaxel plus ramucirumab (IMC-1121B, LY3009806) versus paclitaxel plus placebo as second-line treatment in patients with advanced gastric or GEJ adenocarcinoma after failure of any platinum and fluoropyrimidine doublet with or without anthracycline (epirubicin or doxorubicin).

4.2. Secondary Objectives

Secondary objectives are to evaluate:

- Objective response rate (ORR)
- Time to progression (TTP)
- Duration of objective response (DOR)
- Safety profile
- Patient-reported outcome (PRO) measures (EORTC QLQ-C30 and EQ-5D-3L)

4.3. Exploratory Objectives

The exploratory objectives of the study include:

- To explore biomarkers relevant to gastric or GEJ adenocarcinoma or safety, efficacy, and mechanism of action of ramucirumab.

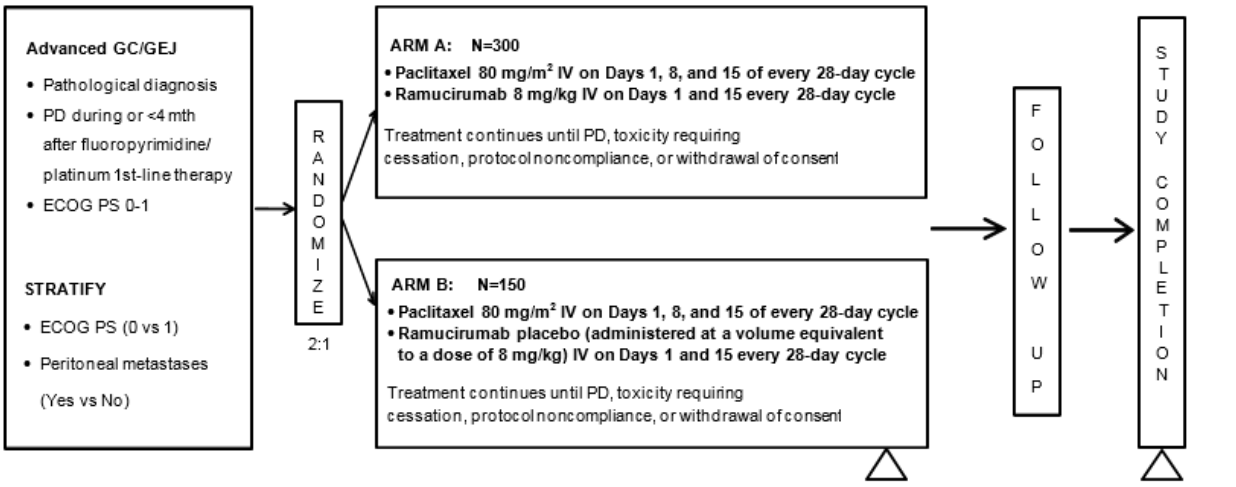
5. Study Design

5.1. Summary of Study Design

Study I4T-CR-JVCR (JVCR) is a multinational, randomized, multicenter, double-blind, placebo-controlled, Phase 3 study in patients with histologically or cytologically confirmed advanced gastric or GEJ adenocarcinoma, refractory to or progressive after first-line therapy with platinum and fluoropyrimidine.

Approximately 450 patients will be randomized on a 2:1 ratio to Arm A (paclitaxel plus ramucirumab) and Arm B (paclitaxel plus placebo) to observe approximately 336 deaths. Randomization will be stratified by ECOG PS (0 versus 1) and peritoneal metastasis (yes versus no).

Figure 5.1 illustrates the study design.



Abbreviations: ECOG PS = Eastern Cooperative Oncology Group performance status; GC = gastric cancer; GEJ = gastroesophageal junction adenocarcinoma; IV = intravenous; mth = months; N = number of patients; PD = progressive disease.

Figure 5.1. Illustration of study design for Study I4T-CR-JVCR

Terms used to describe the periods during the study are defined below:

- **Baseline:** begins when the ICF is signed and ends at the first study treatment (or discontinuation, if no treatment is given)
- **Study Period:** begins at the first study treatment and ends at study completion. The study period does not include the extension period.
    - **Study Treatment Period:** begins at the first study treatment and ends when the patient and the investigator agree that the patient will no longer continue study treatment.

- **Post-discontinuation Follow-Up:** begins the day after the patient and the investigator agree that the patient will no longer continue study treatment.

  Short-term follow-up begins 1 day after the patient and the investigator agree that the patient will no longer continue study treatment and lasts approximately 30 days (no more than 37 days).

  Long-term follow-up begins 1 day after short-term follow-up is completed and continues until the patient's death or overall study completion, whichever comes first.

- **Extension Period:** begins after study completion and ends at the end of trial. During the extension period, patients on study therapy who continue to experience clinical benefit may continue to receive study therapy (except placebo) until one of the criteria for discontinuation is met. The extension period includes extension period follow-up.
  - **Extension Period Follow-Up:** begins 1 day after the patient and the investigator agree that the patient will no longer continue treatment in the extension period and lasts approximately 30 days (no more than 37 days).

Baseline radiographic assessment of disease will be performed within 21 days prior to randomization; first treatment will be administered within 7 days following randomization.

Patients in both arms will receive any necessary premedication prior to the first infusion of study therapy at each treatment session.

A planned treatment cycle will be defined as 28 days (4 weeks) in each arm and includes the period of treatment with paclitaxel given on Days 1, 8, and 15, in combination with either ramucirumab or placebo given on Days 1 and 15. The period during which no treatment will be administered is required for patients to recover from toxicities to NCI-CTCAE, Version 4.03 criteria Grade <2 or baseline (with the exception of alopecia), and is expected to last 14 days.

Patients will undergo radiographic assessment of disease status every 6 weeks (±5 days) following the first dose of study therapy for the first 6 months, and every 9 weeks (±5 days) thereafter, until documentation of radiographic PD. Patients in both arms will continue to receive all treatments until there is radiographic or symptomatic progression of disease, toxicity requiring cessation, protocol noncompliance, or withdrawal of consent. In case of treatment discontinuation for any reason other than radiographically confirmed PD, radiographic tumor assessments will continue every 6 weeks (±5 days) following the first dose of study therapy for the first 6 months, and every 9 weeks (±5 days) thereafter, until documentation of radiographic PD.

Follow-up information regarding further anticancer treatment and survival status will be collected at least every 8 weeks (+0 to 7 days). Long-term follow-up will continue as long as the patient is alive, or until study completion, whichever is first.

Adverse event information will be collected up to approximately 30 days (no more than 37 days) after the decision is made to discontinue study treatment, or until all SAEs and all therapy-related

AEs have resolved, stabilized, returned to baseline, or been deemed irreversible, whichever date is later.

Patients with evidence of sustained clinical benefit (that is, no disease progression) may continue to receive study drug(s) for long-term durations during the extension period.

Upon study completion, investigators and patients may be unblinded to study treatment assignment.

5.2. Determination of Sample Size

The study plan is to enroll approximately 450 patients with advanced gastric or GEJ adenocarcinoma after failure of any platinum and fluoropyrimidine doublet with or without anthracycline.

The approximately 450 patients will be randomized in a 2:1 ratio to Arm A (paclitaxel plus Ramucirumab) and Arm B (paclitaxel plus placebo). Patients will be randomized using stratification factors ECOG PS (0 vs 1) and peritoneal metastasis (yes vs no).

Under the following assumptions:

- The randomization ratio is 2:1;
- OS in each arm follows an exponential distribution;
- The OS hazard ratio for Arm A vs Arm B is 0.81;
- Control arm median OS = 10.5 months,

336 deaths (i.e. 25% censoring rate) will provide at least 80% probability to assure that the effect size observed in this study is consistent with the global study, where effect size is defined by $\delta = \frac{1}{HR} - 1$. The above assurance probability is defined by Prob$\lfloor \hat{\delta}_{JVCR} > \rho\hat{\delta}_{JVBE} \rfloor$, where $\hat{\delta}_{JVBE}$ = 0.235 was estimated from the observed HR of 0.81 in the global study JVBE and $\hat{\delta}_{JVCR}$ will be estimated from the observed HR in this study. ρ is a constant set to 0.5, corresponding to 50% effect size retention from the global study.

Assuming further a constant accrual rate of 12 patients per month, it is estimated that the study will take approximately 44 months to obtain the required 336 deaths, with 37 months for full accrual and 7 months for follow-up.

5.3. Method of Assignment to Treatment

Upon completion of all screening evaluations to confirm a patient's eligibility, the site will register the patient via the Interactive Web Response System (IWRS). The IWRS registration consists of assigning the patient a unique study identification number and randomizing the patient to 1 of the 2 treatment arms. Once the patient is registered through the IWRS, he/she is considered to be *enrolled* in the study.

Randomization will be stratified by ECOG PS (0 versus 1) and peritoneal metastasis (yes versus no). To balance the treatment allocation among the stratification factors, a stratified permuted block randomization will be utilized and incorporated into the IWRS.

6. A Priori Statistical Methods

6.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company.

All tests of efficacy will be conducted at a 2-sided significance level of 0.05, unless otherwise stated. All confidence intervals will be given at a 2-sided 95% level, unless otherwise stated.

The assumptions for each statistical method may be evaluated. If there is violation of assumptions, alternative statistical methods may be used.

Additional exploratory analyses of the data may be conducted as deemed appropriate.

6.1.1. Analysis Populations

The **entered** population includes all patients who sign the informed consent document.

The **intention-to-treat (ITT)** population will include all randomized patients. Patients will be grouped according to randomized treatment.

The **per-protocol (PP)** population will include all patients in the ITT population who are compliant with the study protocol without major protocol violations.

The **safety** population will include all randomized patients who received at least 1 dose of study treatment.

Unless otherwise stated, all patient characteristic and efficacy analyses will be performed on the ITT population, and all safety and exposure analyses will be performed on the safety population.

Analysis of duration of response (DOR) will be performed only for ITT patients who achieve a best overall response of CR or PR, which is denoted as **mITT1** population.

Analysis of PROs will include randomized patients that provide a baseline assessment and at least 1 post-baseline assessment, denoted as **mITT2** population; however, compliance will be based on the ITT population.

Unless otherwise stated, all analyses on the ITT population will be performed by **randomized treatment arm**, and all analyses on the safety population will be performed by **actual treatment received on Cycle 1 Day 1**.

Rescreening: For rescreened patients, all analyses including summary of disposition and baseline characteristics, will only consider the data collected during the last screening.

Treatment switching: For patients in the safety population who switched treatment due to any reason, they will still be analyzed based on the first actual treatment received.

6.2. Adjustments for Covariates

As a supportive analysis, PFS and OS will be analyzed after adjusting for selected prognostic factors. Potential prognostic factors include the stratification factors and other factors used for subgroup analysis discussed in Section 6.12.

The hazard ratio for treatment effect will be estimated using an unstratified multivariate Cox proportional hazard 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 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.

6.3. Handling of Dropouts or Missing Data

All analyses and descriptive summaries will be based on the observed data. With the exception of dates, missing data will not be imputed. For the patient data listings, no imputation of incomplete dates will be applied. The listings will present the incomplete dates without any change.

General rules for imputing dates related to AE, concomitant therapy or post discontinuation therapy:

- Onset date of an AE or start date of a concomitant therapy or post discontinuation therapy:
    - If only the day is missing, the 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 treatment.
    - If both the day and month are missing, the complete date will be set to:
        - January 01 of the year of onset, if the onset year is after the year of the first study treatment.
        - The date of the first dose, 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 treatment.
- Resolution date of an AE or end date of a concomitant therapy
    - If only the day is missing, the date will be set to the last day of the month of the occurrence, or to the date of death if the patient died in the same month.
    - If both the day and month are missing, the date will be set to December 31 of the year of occurrence or to the date of death if the patient died in the same year.

If a 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.

General rule for imputing other dates: If a date variable is needed for an analysis, use the following general rule to impute incomplete date:

- If the date has no missing year or month, but the day is missing, then assign 1 to day.
- If the date has no missing year, but has missing month, then assign January to month.

However, after imputation, check if the imputed date is logically consistent with other relevant date variables and make appropriate correction if necessary. For example, if a visit start date was May 10, 2008 and a tumor assessment date was May dd, 2008 (missing day) but it was known that the assessment occurred after that visit, then after imputation, the tumor assessment date would become May 01, 2008. In this case, the imputed tumor assessment date should be compared to the visit start date, and then corrected to be the visit start date, May 10, 2008.

Patient-reported outcome analysis: For percentage compliance of the EORTC QLQ-C30 and EQ-5D, instruments with at least one item completed will be considered as having been completed. Multi-item QLQ-C30 scales with ≥50% of items completed will be scored using imputation rules per the scoring manual (Fayers et al. 2001). No other adjustment or imputation for missing data will be performed.

The QLQ-C30 scoring manual imputes missing items as follows:

- Have at least half of the items from the scale been answered?
- If *Yes*, use all the items that were completed, and apply the standard equations for calculating the scale scores; ignore any items with missing values when making the calculations.
- If *No*, set scale score to missing.
- For single-item measures, set score to missing.

Safety analysis: The following rule for missing data processing will apply for safety analysis:

- Missing classifications concerning study medication relationship will be considered as related to study medication.
- If the AE onset date is missing or partial, the date will be compared as far as possible with the date of first dose of study medication when determining whether or not the AE is present at baseline. In this case, the AE will be assumed to be treatment emergent, unless there is clear evidence (through comparison of partial dates) to suggest that the AE started prior to the first dose of study medication.

Time-to-event analysis: All censored data will be accounted for using appropriate statistical methods.

6.4. General Data Handling

- Age will be computed from July 1 of the year of birth (since only year of birth is recorded in eCRF) to the date of Informed Consent, as

  (Date of Informed Consent - July 1 of the year of Birth +1) / 365.25.

  Age will be rounded to the integer part of the computed age using the above formula.

- Baseline Measurements:

  - Efficacy: 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: The last measurement prior to the first study drug administration will be used as the baseline assessment.
  - Other baseline characteristics (including vital sign assessments): 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 (time to event) analysis: 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.

- Time-to-event: The event or censoring time (days) is calculated as:
  - Date of event/censoring – Date of randomization + 1

- Duration: Duration (except for 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; (Days in months = 30.4375, which is average number of days in a year /12)
- Duration (years): (End Date – Start Date + 1) / 365.25; (Average days in a year = 365.25, reflecting the Julian year of three years with 365 days each and one leap year of 366 days)

6.5. Multiple Comparisons/Multiplicity

No adjustment for multiplicity is planned for this study. All *p*-values are of descriptive nature.

6.6. Patient Disposition

A detailed description of patient disposition will be provided. It will include:

- Summary of patients entered and randomized by country
- Total number of patients entered
- Total number of patients randomized, by treatment arm
- Summary of reasons for patients entered, but not randomized
- Total number of patients treated, by treatment arm
- Summary of reasons for patients randomized but not treated

A detailed summary of reasons for patient discontinuation from study treatment will be provided. A listing of AEs leading to treatment discontinuation will be provided.

The number and percentage of patients included/excluded in ITT population, Safety population and PP population as well as exclusion reasons will be summarized for all subjects who signed the Informed Consent.

Moreover, the number and percentage of patients still receiving treatment (i.e., have not completed the End of Treatment visit) and those still on study evaluation (i.e., have not completed the End of Study Evaluation visit) at data cut-off date will be summarized by treatment group. Patients having completed their End of Treatment and End of Study Evaluation visits will be presented by reason for discontinuation of study treatment and study, respectively.

6.7. Patient Characteristics

6.7.1. Demographics and Performance Status

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

- Sex
- Race
- Age (years)

- Age group (<65 vs ⩾65 years)
- Height (cm)
- Weight (kg)
- Body mass index (BMI) (kg/m^2)
- Body surface area (BSA) (m^2)
- Baseline ECOG PS [per eCRF]

Height and weight will be based on "Vital Signs: Baseline" eCRF. BMI is defined as weight divided by the square of height.

6.7.2. Baseline Disease Characteristics

The following baseline disease characteristics will be summarized using the ITT population:

- Duration of disease
- Peritoneal metastasis [per eCRF]
- TNM classification at initial diagnosis
- Extent of disease at study entry (locally advanced vs. metastatic)
- Primary tumor location (gastric vs. gastroesophageal junction tumor)
- Primary tumor present (yes vs. no)
- Disease measurability (measurable vs. non measurable)
- Time to progression on first-line therapy (<6 months vs. ⩾6 months)
- Histopathological diagnosis grade at initial diagnosis (well differentiated, moderately differentiated, poorly differentiated, undifferentiated, unknown)
- Pathological diagnosis HER2 status at initial diagnosis (negative finding, positive finding, not done, unknown)
- Sites of metastatic disease (Liver, Lung, Lymph nodes, etc.)
- Number of organs involved (0, 1, 2, ⩾ 3, derived from sites of metastatic disease)
- Weight loss (⩾ 10% over the 3 months prior to study entry vs. < 10%)
- Presence of ascites (yes vs. no)

Duration of disease is defined as time from date of diagnosis confirmation to date of randomization. Date of diagnosis confirmation will be derived from "Pathological Diagnosis" eCRF. If only the month and year of the date of diagnosis confirmation is known, the day will be imputed as 15.

Tumor location will be reported directly from "Pathological Diagnosis (MH_IPD)" eCRF.

Time to progression on first-line therapy is defined as time from start of first-line therapy to disease progression. It will be calculated as (progressive disease date – start date of first-line therapy + 1).

Pathological diagnosis HER2 status and histopathological diagnosis grade are directly available on "Pathological Diagnosis" eCRF, where the grades are coded according to Table 6.1.

Table 6.1. Coding for Tumor Grades

Tumor Grade	eCRF Entry
Well differentiated	G1 Grade 1
Moderately differentiated	G2 Grade 2
Poorly differentiated	G3 Grade 3
Undifferentiated	G4 Grade 4
Unknown	GX – grade cannot be assessed

Disease measurability will be derived from "Target Tumor: RECIST 1.1" and "Non Target Tumor: RECIST 1.1" eCRFs at baseline. All patients with at least one lesion on the target lesion form will be counted as having measurable disease.

The number of organs involved and sites of metastatic disease will be summarized based on data in "Nature of Disease" eCRF.

Weight loss and presence of ascites are documented in "Stratification and Subgroup Criteria" eCRF. Extent of disease at study entry can be derived from "Pathological Diagnosis" eCRF: If TNM metastasis is "M1" then the extent of disease is "metastatic", otherwise, it's "locally advanced".

6.7.3. Medical History/Pre-existing Conditions

A summary of medical history will be presented by System Organ Class (SOC) and Preferred Term (PT) of the historical/pre-existing conditions using version 21.1 of the Medical Dictionary for Regulatory Activities (MedDRA). The number and percentage of subjects in the ITT population that experienced each condition will be summarized by treatment arm.

6.7.4. Prior Therapies

Prior radiotherapy, surgery, and systemic and locoregional therapy will be summarized. Prior radiotherapy will be categorized by reason for regimen (adjuvant, neoadjuvant, neoadjuvant plus adjuvant, advanced/metastatic). Surgery will be categorized by surgical intent (palliative, curative). Prior systemic and locoregional therapy will be categorized by setting of regimen (adjuvant, locally advanced, metastatic, neoadjuvant) and type of therapy (chemotherapy, monoclonal antibody, etc.). Patients will also be categorized based on the number of regimens

received during first-line therapy. Frequency of each specific therapy will be tabulated within each setting of regimen.

Select prior therapies of interest will also be summarized, as listed below:

- Medication classes of interest: Anthracyclines and related substances, detoxifying agents for antineoplastic treatment, folic acid analogues, monoclonal antibodies, platinum compounds, pyrimidine analogues, and taxanes
- Previous treatment with a targeted agent containing regimen (EGFR only, HER2 only, EGFR and HER2, others) vs. previous treatment without a targeted agent
- Previous treatment with platinum/fluoropyrimidine with anthracycline vs. without anthracycline
- Previous gastrectomy (total, partial, other, no)

Listings of prior systemic therapies, radiotherapies and surgeries will be provided.

6.8. Concomitant Therapy

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

- It started on or after the date of first dose of study treatment and within 30 days after the date of last study treatment; or
- It started prior to date of first dose of study treatment but was ongoing at the time of the first dose of study treatment.

The following concomitant medications will be summarized:

- All concomitant medications, by medication class and preferred term.
- Selected concomitant medications, including growth factors (erythroid growth factors, granulocyte-colony stimulating factors, granulocyte-macrophage colony-stimulating factor), antiemetics, anti-hypertensive agents, and antibiotics.
  Note: Such drugs to be used for programming will be identified through reviewing of the unique drug terms collected in the study.
- Premedication for study drug.

Premedication usage will be summarized overall and by cycle.

The frequency and percentage of patients with any blood transfusions experienced in study treatment period or the 30-day post-discontinuation follow-up period will be summarized by treatment group. Transfusions will be further characterized by transfused blood product (e.g. packed red blood cells, whole blood, platelets, or fresh frozen plasma). Reasons for transfusion will be presented by listing the corresponding AE and MH PTs.

6.8.1. Post-discontinuation Therapies (PDT)

Post-discontinuation anti-cancer therapies, including systemic therapies, radiotherapies and surgeries, will be summarized by treatment arm. Systemic therapies will also be summarized by type of therapy, medication class and specific medication terms. The number of subjects who received 1, 2, 3, … subsequent regimens will be summarized. Patient-level listings of PDT will be provided.

6.9. Efficacy Analyses

Efficacy endpoints will be analyzed using the ITT population and presented by treatment group. The stratification factors for the analyses of primary and secondary endpoints are:

- Eastern Cooperative Oncology Group (ECOG) performance status (0 versus 1)
- Peritoneal metastasis (yes versus no)

The stratification factors are captured in the IWRS and on eCRFs. Unless otherwise specified, all primary efficacy analyses will be based on the stratification factors per IWRS. A cross-tabulation of the frequency of each level of each stratification factor per IWRS and per eCRF will be produced.

6.9.1. Primary Outcome and Methodology

6.9.1.1. Progression-free Survival

The PFS is defined as the duration from the date of randomization to the date of the first radiographically documented PD as defined by RECIST 1.1, or to the date of death due to any cause, whichever is earlier. The primary censoring rules are described in Table 6.1.

The analysis of PFS in this study will estimate the HR and its corresponding CI in a stratified Cox proportional hazards model using treatment arm as a single covariate. The stratification factors included in the Cox model are the same as randomization strata. An additional analysis with an unstratified Cox proportional hazards model will be employed to explore the impact of the stratification factor. Stratified (using the same stratification factor) and unstratified log-rank tests will also be conducted.

The survival curves for each randomized arm will be estimated using the Kaplan-Meier (KM) product-limit method. CI for median PFS will be computed by the Brookmeyer and Crowley method (Brookmeyer and Crowley 1982) using log-log transformation. Survival rates at 6, 12, 18 and 24 months will also be estimated using KM estimates on the PFS curve for each randomized arm. Associated CI for survival rates will also be calculated using normal approximation.

In addition, a multivariate Cox model will be used to adjust for baseline covariates as imbalances in potential prognostic factors could impact the analysis of survival. Covariates that are planned to be included are described in section 6.12.

A patient-level listing of PFS time will be provided.

The primary PFS analysis will be performed when at least 256 PFS events have occurred (see Section 6.13).

Table 6.2. Primary Censoring Rules for PFS

Index	Situation	Date of Event or Censor	Event / Censor
1	**No baseline** radiological tumor assessment available	Date of randomization	Censored
2	**No post baseline** radiological tumor assessment available and **no death** reported within 2 scan intervals following randomization	Date of randomization	Censored
3	**No post baseline** radiological tumor assessment available but **death** reported within 2 scan intervals following randomization	Date of death	Event
4	**No tumor progression** (per RECIST 1.1) and **no death** reported within 2 scan intervals following adequate last radiological tumor assessment (or following randomization if no adequate assessment available)	Date of last adequate radiological tumor assessment (or date of randomization if no adequate assessment available)	Censored
5	**No tumor progression** (per RECIST 1.1) but **death** reported within 2 scan intervals following last adequate radiological tumor assessment	Date of death	Event
6	**Tumor progression** (per RECIST 1.1) documented within 2 scan intervals following adequate previous radiological tumor assessment or following randomization	Earliest of the target, non-target and new tumor assessment dates	Event
7	**Tumor progression** (per RECIST 1.1) documented after ≥2 consecutively missed tumor assessment visits following adequate previous radiological tumor assessment (or following randomization if no adequate assessment available)	Date of last adequate radiological assessment before missed assessments (or date of randomization if no adequate assessment available)	Censored
8	**New anticancer treatment** started	Date of previous adequate radiological assessment prior to start of new therapy (or randomization date if no adequate assessment available)	Censored
9	**No tumor progression** (per RECIST 1.1) and patient lost to follow-up or withdrawal of consent	Date of last adequate radiological assessment (or randomization date if no adequate assessment available)	Censored

Notes: (1) Symptomatic deteriorations (i.e. symptomatic progressions, which are not radiologically confirmed) will not be considered as progressions.
(2) 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 progressive disease (PD); otherwise the latest date will be used.
(3) Adequate radiological tumor assessment refers to an assessment with overall response of CR, PR, SD or PD.
(4) New anticancer treatment includes systemic and locoregional therapies, radiotherapy and surgical procedures.

6.9.1.2. Overall Survival

Overall survival (OS) is measured from the date of randomization to the date of death from any cause. For each subject who is not known to have died as of the data cut-off date, if the subject is lost to follow-up, then OS for that subject will be censored to the last known alive date, as recorded in "Post Discontinuation Follow-up" CRF page; otherwise, OS will be censored at the date of last contact prior to the data cut-off date. Contacts considered in the determination of last contact date include: clinical evaluation dates (AE, vital sign, ECG), laboratory evaluation dates (local and central), tumor assessment dates, visit dates and withdrawal of consent date.

The analysis of OS distributions in the two treatment arms will be conducted via similar methods as for PFS (i.e. HR and CI from stratified and unstratified Cox models, p-value from stratified and unstratified log-rank tests, KM curve, median estimates with CI and survival rate estimates with CI). A patient-level listing of overall survival time will be provided.

An interim OS analysis is planned, and a final OS analysis will be performed if the primary objective is not achieved at interim. Details are discussed in Section 6.13.

6.9.2. Secondary Efficacy Analyses

Table 6.3. Censoring rules for TTP

Index	Situation	Date of Event or Censor	Event / Censor
1	**No baseline** radiological tumor assessment available	Date of randomization	Censored
2	**No post-baseline** radiological tumor assessment available and no death reported within 2 scan intervals following randomization	Date of randomization	Censored
3	**No post baseline** radiological tumor assessment available but death reported within 2 scan intervals following randomization	Date of death	Censored
4	**No tumor progression** (per RECIST 1.1) and no death reported within 2 scan intervals following last adequate radiological tumor assessment	Date of last adequate radiological tumor assessment	Censored
5	**No tumor progression** (per RECIST 1.1)	Date of death	Censored

	but death reported within 2 scan intervals following last adequate radiological tumor assessment		
6	**Tumor progression** (per RECIST 1.1) documented within 2 scan intervals following previous adequate radiological tumor assessment	Earliest of the target, non-target and new tumor assessment dates	Event
7	**Tumor progression** (per RECIST 1.1) documented after ≥2 consecutively missed tumor assessment visits following previous adequate radiological tumor assessment	Date of last adequate radiological assessment before missed assessments	Censored
8	**New anticancer treatment** started	Date of previous adequate radiological assessment prior to start of new therapy	Censored
9	**No tumor progression** (per RECIST 1.1) and patient lost to follow-up or withdrawal of consent	Date of last adequate radiological assessment	Censored

Notes: (1) Symptomatic deteriorations (i.e. symptomatic progressions, which are not radiologically confirmed) will not be considered as progressions.

(2) 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 response for that visit is PD; otherwise the latest date will be used.

(3) Adequate radiological tumor assessment refers to an assessment with overall response of CR, PR, SD or PD

6.9.2.1. Time to Progression

Time to progression (TTP) is defined as the duration from the date of randomization to the date of the first radiographically documented PD as defined by RECIST 1.1. TTP will be analyzed using similar methods as described for PFS (i.e. the stratified and unstratified Cox model and log-rank test, KM curve and median estimates with CI). Censoring rules for this analysis are described in Table 6.3.

6.9.2.2. Objective Response Rate and Disease Control Rate

The Best Overall Response (BOR) will be determined using RECIST (Version 1.1). It is defined as the best response across all time points from randomization until radiologically confirmed tumor progression, if any. Tumor assessments performed after initiation of new anticancer treatment (including systemic therapy, surgical procedures and radiotherapy) will be excluded from evaluating the best overall response.

Objective response rate (ORR) will be calculated as the number of patients who achieve a BOR of CR or PR divided by the total number of patients randomized to the corresponding treatment arm. 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.

Disease control rate (DCR) will be calculated as the proportion of patients with BOR of CR, PR or SD, using the same denominator as ORR.

Frequencies and proportions of BOR will be presented by treatment group, along with two-sided 95% Clopper-Pearson CI. The ORR and DCR observed in each treatment group will be compared using the Cochran-Mantel-Haenszel test adjusting for ECOG PS and peritoneal metastasis as captured by IWRS.

Waterfall plots of the best percent change in tumor size will be provided by treatment arm for subjects in the ITT population who have measurable disease (i.e. who have at least 1 target lesion at baseline tumor assessment).

A patient-level listing of overall tumor response will be provided.

6.9.2.3. Duration of Response

Duration of objective response (DOR) is measured from the time measurement criteria are first met for CR or PR (whichever is first) to the date of the first radiographically documented PD as defined by RECIST 1.1, or to the date of death due to any cause, whichever is first. DOR will be analyzed only for patients who achieve a BOR of CR or PR. Patients alive and without radiographic disease progression at data cut-off date or patients lost to follow-up will be censored at the date of their last adequate radiographic tumor assessment. KM method will be used to analyze DOR to estimate the median and corresponding CI for each treatment arm.

6.9.3. Sensitivity Analyses

6.9.3.1. Sensitivity Analyses for OS

The following sensitivity analyses will be performed for OS:

- Analyses based on PP population
- Censoring OS for users of post-discontinuation therapy (PDT) to the start date of PDT
- Censoring OS for users of targeted PDT (including immunotherapy, anti-angiogenesis, anti-HER2)
- Censoring OS for users of anti-angiogenesis PDT

6.9.3.2. Sensitivity Analyses for PFS

PFS will be analyzed based on PP population. Additional sensitivity analyses will be performed for PFS using alternative censoring rules described in Table 6.4.

Table 6.4. Alternative Censoring Rules for PFS

Index	Sensitivity Analysis	Situation in Table 6.2	Date of Progression or Censor	Outcome
1	Count symptomatic deterioration as progression	Documented progression or symptomatic deterioration	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.	Progressed

Index	Sensitivity Analysis	Situation in Table 6.2	Date of Progression or Censor	Outcome
			Or date of symptomatic deterioration, whichever occurred first.	
2	Ignore new anticancer treatment	New anticancer treatment started prior to documented progression or death	Date of progression or death based on subsequent tumor assessments	Progressed
3	Ignore missing tumor assessments	Death or documented progression after ≥ two consecutively missed tumor assessment visits	Date of first missed tumor assessment	Progressed
4	Treat lost to follow-up as progression	Patient lost to follow-up without documented progression or death	Date of next scheduled visit after last tumor assessment	Progressed
5	Treat lost to follow-up as progression for experimental arm only (worst case scenario)	Patient lost to follow-up without documented progression or death (worst case scenario)	Patients in the active treatment group will be considered to have progressed at the date of the next scheduled visit after last tumor assessment. Patients in the control group will be censored at the date of last adequate tumor assessment	Progressed (experimental) or Censored (control)

6.10. Health Outcomes/Quality-of-Life Analyses

Patient-reported outcome (PRO) measures (EORTC QLQ-C30 and EQ-5D) data will be summarized by treatment arm using the ITT population for patients who had a baseline assessment and at least one post-baseline assessment.

Questionnaire compliance will be summarized for the ITT population. Compliance rates for each PRO instrument will be calculated at each assessment time point. Compliance at an assessment time point is defined as the number of patients who were assessed for that PRO instrument divided by the expected number of patients at that time point. The expected number of patients:

- at baseline is equal to the number of patients randomized
- at any post-baseline visit is equal to the number of patients who are alive and have not progressed

A patient who answers at least one item at a time point is considered to have been assessed. Reasons for non-compliance will be tabulated by treatment arm and by assessment time point.

6.10.1. EORTC QLQ-C30

Following the EORTC guidelines (Fayers et al. 2001), the 30 items of the QLQ-C30 will be transformed into the 15 scales (1 global health status/QoL scale, 5 functional scales, and 9 symptoms scales/items). For each scale, a linear transformation will be used to obtain scores

ranging from 0 to 100, with a high score indicating a high level of functioning for the global and the functional scales and, conversely, a high level of symptomatology/problems for the symptom scales/items. The scoring for EORTC QLQ-C30 is detailed in Appendix 1.

The following variables will be derived for each scale:

- Change from baseline is calculated by subtracting baseline assessment result from current assessment result.
- Best score for each scale is the highest post-baseline score for the functional scales and global health status, and is the lowest score for the symptom scales. The change from baseline corresponding to the best score will be called the **best change from baseline**.
- Worst score for each scale is the lowest post-baseline score for the functional scales and global health status, and is the highest score for the symptom scales. The change from baseline corresponding to the worst score will be called the **worst change from baseline**.

Each scale at every assessment point will be compared to its baseline value and be categorized as follows (Osoba et al. 1998):

- **Deterioration**: Defined as an increase of ≥10 points for the symptom scales or a decrease of ≥10 points for the functional scales and global health status/QoL scale.
- **Improvement**: Defined as a decrease of ≥10 points for the symptom scales or an increase of ≥10 points for the functional scales and global health status/QoL scale.
- **Stable**: Defined as no change or an increase/decrease of <10 points.

Floor/Ceiling effect: The ceiling effect is defined as the percentage of patients who have a baseline score of <10 points on the symptom scales/items, and >90 points on the functional scales and global health status/QoL scale. The floor effect is defined as the percentage of patients who have a baseline score of >90 points on the symptom scales, and <10 points on the functional scales and global health status/QoL scale.

The presence of a significant ceiling effect suggests that not much improvement is possible for that scale. Likewise, the presence of a floor effect suggests that deterioration is less likely.

6.10.1.1. Descriptive Summaries

The proportion of patients in each arm with deterioration in scores, improvement in scores, stable scores, and missing scores (due to progression, death, or failure to complete) at each assessment time point will be presented for each of the 15 scales both in a summary table and graphically using stacked bar charts.

The denominator should include all patients from the group with a baseline assessment and at least one post-baseline assessment for the given scale. The proportion of patients who deteriorated at each assessment time point will be compared using Pearson's χ^2 test.

Data will be summarized descriptively for all 15 scales at each assessment time by treatment arm, provided at least 30% of the patients have assessments at that assessment time. Shift tables will be presented for the 6 single-item symptom scales.

Floor and ceiling effects will be summarized for each of the 15 scales.

6.10.1.2. Analysis of Best and Worst Post-Baseline Change Score

The best change score from baseline for each of the 15 scales will be compared statistically between treatment arms using analysis of covariance (with baseline score and randomization stratification factors per IWRS as covariates) and the Mann-Whitney-Wilcoxon test. The worst change score from baseline will be analyzed in a similar manner.

6.10.1.3. Time to Deterioration (TTD) from Baseline

For each of the 15 scales, the following analyses will be performed for time to first deterioration from baseline score:

- Obtain p-value from stratified log-rank test
- Survival estimates and curves using Kaplan-Meier product limit method
- Hazard ratio estimate and CI using stratified Cox proportional hazard model (covariates include treatment arm and baseline score)
- Forest plot to display HR and CI

Censoring: If deterioration is observed immediately after a missing value, it will be assumed that the deterioration occurred at the time of the missing value. Patients who do not deteriorate will be censored at the date of last EORTC assessment.

6.10.2. EQ-5D-3L Analysis

By treatment arm and at each assessment time, the frequency and percentage of patient's responses (no problem, some problem, and extreme problem) will be presented for each of the 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The responses will be presented graphically using stacked bar charts by assessment time point.

The index score for the EQ-5D-3L will be calculated from a combination of responses using a weighting algorithm currently in development based on EQ-5D-3L Chinese validation study (EuroQol Group 1990; Brooks et al. 1996; Liu et al. 2014). The model for calculating index scores corresponding to each of 243 health states is given in Appendix 2. The index score and the VAS will be presented using summary statistics, including change from baseline, for each assessment time by treatment arm. These will also be graphically represented using box plots by assessment time point.

6.11. Safety Analyses

6.11.1. Extent of Exposure

Exposure analyses will be based on the actual dose administered (in mg) and body weight (in kg) (for ramucirumab/placebo) or BSA (in m^2) (for paclitaxel). The body weight to be used for calculating each dose of ramucirumab or placebo will be the last available measure of the patient recorded on CRF prior to each infusion. The BSA to be used for calculating each dose of paclitaxel will be calculated using the Dubois & Dubois formula (see below) based on the last available weight and height prior to each infusion.

$$\text{BSA } [m^2] = (Weight\ [kg]^{0.425} * Height\ [cm]^{0.725}) * 0.007184$$

For patients who did not receive any amount of study drug, the dose exposure parameters for that treatment (number of infusions, duration of treatment, cumulative dose, dose intensity, relative dose intensity) will be set to 0. Cumulative dose, dose intensity and relative dose intensity will remain missing if they cannot be derived due to missing weight or BSA.

6.11.1.1. Summary of Exposure

The following variables will be presented with summary statistics for each study drug (ramucirumab DP or placebo and paclitaxel) by treatment arm:

- Number of infusions
- Number of cycles received by patient
- Duration on therapy
- Cumulative dose
- Weekly dose intensity
- Relative dose intensity

Additionally, the following information will be presented:

- Number of patients treated by cycle
- Number and percentage of patients with dose delay, dose reduction, dose omission and infusion interruption

6.11.1.2. Calculation of Exposure Parameters

The calculation of exposure parameters will be based on the following rules.

Ramucirumab DP or Placebo Treatment:

- Duration on therapy (weeks; 14 days added to duration of treatment because administration is every 2 weeks [on day 1, 15 of each 4-week cycle]) = [(Date of last dose – Date of first dose) + 14] ÷ 7
- Cumulative dose, dose intensity, relative dose intensity:

  - Cumulative dose (mg/kg) = Sum of all (total dose administered [mg] ÷ Last available weight [kg])
  - Weekly dose intensity (mg/kg/week) = (Cumulative dose) ÷ (Duration on therapy)
  - Planned weekly dose intensity (mg/kg/week) = 2 × 8 mg/kg ÷ 4 weeks = 4 mg/kg/week
  - Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) × 100
- Number of dose level reductions: Sum of the number of dose level reductions as reported in CRF
- Dose delays: As reported in CRF
- Dose omissions: As reported in CRF

Paclitaxel Treatment:

- Duration on therapy (weeks) = [(Date of last dose – Date of first dose) + 14] ÷ 7
- Cumulative dose, dose intensity, relative dose intensity:
  - Cumulative dose (mg/m^2) = Sum of all (total dose administered [mg] ÷ BSA using last available weight and height [m^2])
  - Weekly dose intensity (mg/m^2/week) = (Cumulative dose) ÷ (Duration on therapy)
  - Planned weekly dose intensity (mg/m^2/week) = 3 × 80 mg/m^2 ÷ 4 weeks = 60 mg/m^2/week
  - Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) × 100
- Number of dose level reductions: Sum of the number of dose level reductions as reported in CRF
- Dose delays: As reported in CRF
- Dose omissions: As reported in CRF

In the case of dose reduction, the correct level of reduction is given in Table 6.5.

Table 6.5. Dose Reduction for Ramucirumab DP and Paclitaxel

Dose Level	Ramucirumab DP	Paclitaxel
Starting dose	8 mg/kg	80 mg/m²
Reduction to first dose level	6 mg/kg	70 mg/m²
Reduction to second dose level	5 mg/kg	60 mg/m²

6.11.2. Adverse Events

All AEs will be summarized by MedDRA System Organ Class (SOC) and Preferred Term (PT). AE verbatim text will be mapped by the sponsor or designee to corresponding terminology within MedDRA, and the resulting SOC and PT will be used for AE reporting. AEs will be graded according to National Cancer Institute – Common Terminology Criteria for Adverse Events Verson 4.0 (NCI-CTCAE v4.0).

Unless otherwise specified, when summarized by PT, AEs will be presented in decreasing frequency of PT across treatment arms; when summarized by SOC and PT, AEs will be presented in decreasing frequency of PT within SOC across treatment arms. 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 CTCAE grade and the closest relationship to treatment.

The investigator should classify all “probably related”, “possibly related”, or “does not know” AEs and SAEs as related to study treatment/procedure and report this relationship in “Adverse Events (AE)” eCRF. Missing classifications concerning study medication relationship will be considered as related to study medication by the investigator.

Consolidated AE: Consolidated terms comprising clinically synonymous MedDRA preferred terms, have been defined in order to assist in identifying relevant safety differences between treatment arms. Summaries of the incidence of these consolidated terms will supplement the summaries by MedDRA PTs specified in this section. The MedDRA preferred terms that are grouped under each of the consolidated terms will be provided in compound-level safety document.

6.11.2.1. Overview of Treatment-Emergent Adverse Events (TEAE)

An AE will be regarded as **treatment-emergent**, if

- Its onset date occurs any time on or after the date of administration of the first dose of study treatment (either ramucirumab DP or placebo or paclitaxel) up to 30 days after the last dose of study treatment (or up to any time if serious and considered related to study treatment); or
- It occurs prior to first dose date and worsens while on therapy or up to 30 days after the last dose of study treatment (or up to any time if serious and considered related to study treatment).

An overview of TEAEs will be provided to summarize the following categories using frequency counts and percentages:

- Patients with at least one TEAE, SAE, CTCAE Grade ≥3 TEAE
- Patients with TEAEs that led to death
- Patients with TEAEs that led to study treatment discontinuation
- Patients with TE-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.

6.11.2.2. Summary of TEAE

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

- By SOC and PT (all grades, and grade >= 3)
- By PT (all grades, and grade >= 3)
- By consolidated terms (all grades, grades 3, 4 and 5)
- By maximum CTCAE grade

These summaries will be repeated for TEAEs considered related to study treatment by the investigator. TEAEs leading to hospitalization will be summarized by SOC and PT.

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

6.11.2.3. Treatment-Emergent Serious Adverse Events

Treatment-emergent serious adverse events will be summarized:

- By SOC and PT (all grades, and grade >= 3)
- By maximum CTCAE grade

These summaries will be repeated for TE-SAEs considered related to study treatment by the investigator.

A listing of SAEs (regardless of treatment-emergent or not) will be produced.

6.11.2.4. Other Notable Adverse Events

AEs leading to discontinuation of study drug, AEs leading to dose modification, and AEs leading to death will be summarized by SOC and PT, and by maximum CTCAE grade.

Patient-level listings will be provided for the aforementioned AEs.

6.11.2.5. Treatment Emergent Adverse Events of Special Interest (AESI)

The list of AESIs consists of pre-specified selected adverse events that are given special consideration because they have been associated with other agents in a similar class of drugs or

that were observed in preclinical evaluation or earlier clinical studies of ramucirumab. For each AESI, a list of clinically relevant MedDRA preferred terms will be pre-specified based on review of blinded data as well as review of standard MedDRA queries (SMQs).

In order to review AESIs in the most medically meaningful way, selected preferred terms that represent the same phenomenon will be grouped together. The MedDRA preferred terms that are grouped under each of the AESI terms will be provided in compound-level safety document.

Treatment-emergent AESIs will be summarized by AESI category and PT. AESI categories for ramucirumab include the following: arterial and venous thromboembolic events, bleeding/hemorrhage events, GI hemorrhage events, congestive heart failure, GI fistula, GI perforation, hypertension, proteinuria, liver injury/failure, IRR, healing complication, RPLS, thrombotic microangiopathy, pulmonary hemorrhage events, and hepatic bleeding/hemorrhage events.

Liver injury/liver failure AESI will also be summarized by clinical and laboratory terms, as defined in compound-level safety document. Immediate and non-immediate hypersensitivity reactions AESI will be separately summarized.

6.11.3. Deaths

All deaths that occur after the administration of the first dose of any study drug and within 30 days after treatment discontinuation, as well as primary cause of death, will be presented in a listing for the Safety Population. The number and percentage of patients who died during the study and within 30 days after the treatment discontinuation will be summarized by primary cause of death and by treatment arm using the Safety Population.

6.11.4. Clinical Laboratory Evaluation

Laboratory values will be converted to standard (SI) units, and the numerical results will be graded using NCI-CTCAE Version 4.03. Laboratory results not corresponding to a NCI-CTCAE Version 4.03 term will not be graded. The list of lab parameters that can be graded as well as the programming specifications can be found in compound-level safety document.

For the graded laboratory parameters, a summary of lab test severity by cycle and by treatment arm will be provided. Shift from baseline will be provided by cycle and for the worst value on the study. For test results of each lab parameter, side-by-side boxplot for the absolute values and change from baseline will be provided by cycle.

Laboratory results will be presented in listings with a flag for values outside of the laboratory normal range. If the normal range is not available from the institution performing the laboratory test, Harrison's (Braunwald et al. 2005) will be used as a general reference. Retest results will be reported in the listing as well.

6.11.5. Vital Signs and Other Physical Findings

Vital signs and shift from baseline will be summarized by cycle.

For systolic and diastolic blood pressure, shift from baseline will be provided by cycle and for the worst value on the study using the following categories:

- Systolic Blood Pressure (<140, 140-<160, ≥160 mmHg)
- Diastolic Blood Pressure (<90, 90-<100, ≥100 mmHg)

A patient-level listing of vital signs will be provided.

6.11.6. Electrocardiograms and Echocardiogram/MUGA

Baseline ECG and left ventricular ejection fraction (LVEF) will be summarized. Shift of "ejection fraction decreased" from baseline to worst post-baseline will be summarized by treatment arm.

6.11.7. Hospitalizations

The number and percentage of patients who were hospitalized will be summarized by treatment arm. The reasons for hospitalization, number of hospitalizations due to AE and total duration of hospitalization due to AE will also be summarized by treatment arm.

6.12. Subgroup Analyses

Subgroup analyses will be performed for PFS and OS for subgroups identified below. Each analysis will include the following:

- Estimate HR and 95% CI by fitting unstratified Cox proportional hazard regression model with treatment arm as single covariate on subgroup population.
- Obtain *p*-value from unstratified log-rank test on subgroup population.
- Obtain *p*-value for interaction term between treatment arm and subgroup factor by fitting unstratified Cox PH model with treatment arm, subgroup factor and their interaction as covariates on the whole ITT population.

A forest plot containing above information will be provided.

Subgroup factors: In addition to the two stratification factors, i.e. ECOG PS (0 versus 1) and peritoneal metastasis (yes versus no) per IWRS, the following subgroup factors are added:

- Gender (males vs. females)
- Age group (<65 years vs. ≥65 years)
- Country (China vs. non-China)
- Post-discontinuation therapy use (yes vs. no)
- Weight loss over the 3 months prior to study entry (≥10% vs. <10%)
- Primary tumor location (gastric vs. GEJ)
- Prior first-line chemotherapy (doublets vs. triplets)

- Number of metastatic site (≤2 vs. ≥3)
- Liver metastasis (yes vs. no)
- Tumor differentiation at initial diagnosis (well, moderately, poorly, undifferentiated, unknown)
- Time to progression on first-line therapy (<6 months vs. ≥6 months)
- Number of previous treatment lines (1 vs >1 [including 1st line, adjuvant and neoadjuvant])
- Presence of ascites (yes vs. no)
- Prior gastrectomy (yes [full or partial] vs. no)
- Post-discontinuation therapy usage (yes vs. no)

6.13. Protocol Deviation

Important protocol deviations (IPDs) are defined as those deviations that could potentially impact data interpretation, data integrity and patient safety across the study. IPDs will be identified from the clinical database and from site monitoring. Categories of IPDs will be documented in the trial issue management plan (TIMP), which will contain detailed criteria used to identify IPDs.

After further review, a comprehensive listing of patients with major protocol deviations (MPDs) among IPDs that might have significantly affected the interpretation and integrity of data or patient safety will be provided. Patients with MPDs will be excluded from PP population.

A listing of patients with MPDs will be defined prior to database lock, and the PP population will be derived using this list. The PP population may be used for select sensitivity analyses of the efficacy data.

IPDs and MPDs will be summarized by treatment arm and by the category and subcategory of deviation.

6.13.1. Protocol Deviations due to COVID-19

A listing of protocol deviations (both important and non-important) caused by COVID-19 will be provided.

6.14. Interim Analyses and Data Monitoring

An Independent Data Monitoring Committee (IDMC) will be established as an oversight mechanism to monitor overall study conduct. The membership, roles and responsibilities for the IDMC during the conduct of the study are defined in the IDMC Charter. All interim analyses will include complete assessments of safety-related data. The IDMC will evaluate unblinded data and make recommendations to the designated representative of the sponsor. Only the IDMC is authorized to evaluate unblinded interim efficacy analyses. All study team members will remain

blinded to all interim results, except in case that IDMC concludes that criteria for early efficacy have been demonstrated.

Since this study is a bridging study in nature, no study-wise Type I error control or multiplicity adjustments will be done in the testing/estimation procedures. All analyses (including *p*-values) are descriptive.

One interim OS analysis and the primary PFS analysis will be conducted to provide early efficacy information and could potentially result in early communication with regulatory agencies if the primary objective is claimed to have been achieved based on these results.

These analyses will be performed after at least 256 PFS events have occurred. At this time, approximately 160 OS events would have occurred assuming the same median OS values in Section 5.2.

Assuming a true underlying HR of 0.81, with 160 OS events, the probability of obtaining a point estimate of HR less than 1 is at least 90%. Also assuming an exponential PFS distribution for each treatment arm and median PFS durations of 4.35 months and 3 months (HR = 0.69) for Arm A and Arm B, respectively, 256 PFS events will provide at least 80% probability of obtaining a one-sided *p*-value less than 0.025 for testing the hypothesis HR = 1 versus HR < 1 using log-rank test (as calculated with EAST with up-to-date accrual status).

The primary objective will be claimed achieved if both of the following conditions are satisfied at interim:

(1) The point estimate of stratified HR for investigator-assessed OS is less than 1.

(2) The analysis of investigator-assessed PFS shows a one-sided, stratified *p*-value favoring Ramucirumab+paclitaxel arm less than 0.025 (or a two-sided *p*-value less than 0.05).

If either condition is not met at interim, the blinded study will continue until approximately 336 OS events have occurred, at which time the final OS analysis will be performed.

6.15. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

Summary of adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Adverse Events are summarized: by treatment group, by MedDRA preferred term.

- An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (TEAE).
- An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:

  - the number of participants at risk of an event
  - the number of participants who experienced each event term
  - the number of events experienced.

- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

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 taken 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 interim and final analyses. Dummy treatment assignment will be used in the reporting database until the database lock for the final analysis. During this time, analyses using unblinded treatment codes will be performed only at the interim analysis points specified in the protocol/SAP. For those safety and efficacy analyses assigned to the IDMC, only the designated Statistical Analysis Center (SAC), who is independent of the sponsor, will perform analyses on unblinded data.

Data sets will be created for the purpose of aggregate data review in which treatment assignment is scrambled so that personnel involved in the day-to-day conduct of the trial and development and validation of analysis programs will be blinded to patient treatment assignment.

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. References

Braunwald E.; Fauci A.; Hauser S.; Jameson J.; Kasper D.; Longo D.; Harrison's Principles of Internal Medicine. 16th Edition. McGraw-Hill Medical Publishing Division, New York, 2005. Appendix A, Reference Values For Laboratory Tests, pages A1-A9.

Brookmeyer R.; Crowley J. (1982). "A Confidence Interval for the Median Survival Time". *Biometrics*. 38(1): 29-41.

Brooks R.; Group E. (1996-7-31). "EuroQol: the current state of play". *Health policy*. 37(1): 53-72.

Clopper C.; Pearson E. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". *Biometrika*. 26: 404-413.

EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy*. 1990 Dec;16(3): 199-208.

Liu G. G.; Wu H.; Li M; Gao C.; Luo N. (2014). "Chinese Time Trade-Off Values for EQ-5D Health States". *Value in Health*. 17: 597-604.

Meier P.; Karrison T.; Chappell R.; Xie H. (2004). "The Price of Kaplan-Meier", *Journal of the American Statistical Association*, 99(467): 890-896.

Wilson E. B. (1927). "Probable inference, the law of succession, and statistical inference", Journal of the American Statistical Association, 22: 209-212.

9. Appendices

Appendix 1. Scoring for the EORTC QLQ-C30 Questionnaire Version 3.0

EORTC QLQ-C30 will be scored according to the scoring manual including transforming the raw score to the final score and handling of missing item responses as follows:

- Raw Scores (RS):
    a) Global Health Status:
        • Global Health Status (QL2) = (Q29+Q30)/2
    b) Functional Scales:
        • Physical Functioning (PF2) = (Q1+Q2+Q3+Q4+Q5)/5
        • Role Functioning (RF2) = (Q6+Q7)/2
        • Emotional Functioning (EF) = (Q21+Q22+Q23+Q24)/4
        • Cognitive Functioning (CF) = (Q20+Q25)/2
        • Social Functioning (SF) = (Q26+Q27)/2
    c) Symptoms/Items:
        • Fatigue (FA) = (Q10+Q12+Q18)/3
        • Nausea and Vomiting (NV) = (Q14+Q15)/2
        • Pain (PA) = (Q9 +Q19)/2
        • Dyspnea (DY) = Q8
        • Insomnia (SL) = Q11
        • Appetite Loss (AP) = Q13
        • Constipation (CO) = Q16
        • Diarrhea (DI) = Q17
        • Financial Difficulties (FI) = Q28

- Transformed Scores (TS):
    a) For Global Health Status: TS = (RS-1) /6 *100
    b) For Symptom Scales / Items: TS = (RS-1) /3 *100
    c) For Functioning Scales: TS = [1-(RS-1) /3]*100

- Missing Data:

If there are missing items, raw scores derived from more than one question can be prorated so long as the respondent completed >50% of the items on a given subscale:

a) For status/scales/items derived from an odd total number of questions:

[(Number of questions answered+1)/2] questions must be answered

b) For status/scales/items derived from an even total number of questions:

[Number of questions answered/2] questions must be answered

Prorated raw score = [Sum of answered item scores] ÷[N of items answered]

If any subscale has > 50% missing items then the subscale total score will be missing.

Appendix 2. Model for Calculating China Population-Based Predicted Index Scores

An ordinary least squares multiple regression model was proposed and fitted by Liu et al (2014) to calculate the EQ-5D-3L index (utility) scores based on Chinese population. To recapitulate, the five dimensions are mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD) and anxiety/depression (AD). The independent variables include indicator variables corresponding to levels 2 (some problem) and 3 (extreme problem) of each of the 5 dimensions, plus an indicator variable for the presence of any level 3 problems in the state. This model was referred to as the N3 model in the article.

The fitted N3 model can be written as follows:

$$\begin{aligned}\text{Index score} = 1 - (&0.039 + 0.099\text{MO2} + 0.246\text{MO3} + 0.105\text{SC2} + 0.208\text{SC3} + 0.074\text{UA2} \\ &+ 0.193\text{UA3} + 0.092\text{PD2} + 0.236\text{PD3} + 0.086\text{AD2} + 0.205\text{AD3} + 0.022\text{N3})\end{aligned}$$

In the above model, N3 = 1 if any of the 5 responses is 3 and N3 = 0 otherwise. Moreover, MO2 = 1 if mobility = 2 and MO2 = 0 otherwise; other indicator variables are defined similarly. An important exception is that when all responses are 1, the index score is set to 1.

If answers to one or more dimensions are missing, no imputation will be done, and the index score will be set to missing.

Leo Document ID = a2e74790-6b39-4f71-9edf-4e77bdbe1912

Approver: PPD

Approval Date & Time: 30-Jul-2020 02:00:16 GMT

Signature meaning: Approved