Statistical Analysis Plan I5B-MC-JGDJ

A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Trial of Doxorubicin plus Olaratumab versus Doxorubicin plus Placebo in Patients with Advanced or Metastatic Soft Tissue Sarcoma

NCT02451943

Approval Date: 23-Aug-2018

1. Statistical Analysis Plan for Clinical Study: I5B-MC-JGDJ: A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Trial of Doxorubicin plus Olaratumab versus Doxorubicin plus Placebo in Patients with Advanced or Metastatic Soft Tissue Sarcoma (ANNOUNCE)

Olaratumab (LY3012207) Soft Tissue Sarcoma

I5B-MC-JGDJ is a global, multicenter, randomized, double-blind, placebo-controlled, parallel-group Phase 3 trial that will compare the safety and efficacy in patients with advanced or metastatic STS after treatment with doxorubicin (75 mg/m^2 on Day 1) plus olaratumab (loading dose of 20 mg/kg on Days 1 and 8 in Cycle 1, followed by 15 mg/kg on Days 1 and 8 in subsequent cycles) versus doxorubicin (75 mg/m^2 on Day 1) plus placebo (on Days 1 and 8) in a 21-day cycle. Patients will receive combination treatment for 8 cycles, followed by monotherapy olaratumab or placebo until evidence of progressive disease (PD), unacceptable toxicity, death, or other withdrawal criteria are met.

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol I5B-MC-JGDJ
Phase 3

Version 1 electronically signed and approved by Lilly on 12-Oct-2015
Version 2 electronically signed and approved by Lilly on 01-Aug-2017
Version 3 electronically signed and approved by Lilly on

Approval Date: 23-Aug-2018 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

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 first unblinding of the Sponsor to the aggregate (by-treatment-arm) data. Additional analyses were added after first patient visit but before the Sponsor unblinding, to reflect changes made to the study protocol. The overall changes and rationale for the changes incorporated in Version 2 are as follows:

- The language for the efficacy interim analysis has been updated to state that only the IDMC and regulatory authorities will have access to the unblinded results. Therefore, there will be no alpha-spending associated with this interim analysis.
- Details for progression-free survival 2 analysis were added.
- Details for patient-reported outcomes analyses were added.
- Other edits for clarity and consistency were made.

SAP Version 3 was approved after enrollment had completed but prior to any unblinding of the Sponsor to the aggregate (by-treatment-arm) data. Key changes incorporated in Version 3 are as follows:

- Additional sensitivity and subgroup analyses were added.
- A section was added providing more details, rationale, and hypotheses for the planned efficacy analyses with respect to the presence or absence of baseline lung lesions.
- Other edits for clarity and consistency were made.

4. Study Objectives

4.1. Primary Objective

The primary objective is to compare doxorubicin plus olaratumab versus doxorubicin plus placebo with respect to overall survival (OS) in 2 populations:

(1) Patients with advanced or metastatic soft tissue sarcoma (STS) not amenable to treatment with surgery or radiotherapy with curative intent

(2) Patients with advanced or metastatic leiomyosarcoma (LMS) not amenable to treatment with surgery or radiotherapy with curative intent

4.2. Secondary Objectives

The secondary objectives of the study are to compare doxorubicin plus olaratumab versus doxorubicin plus placebo as follows:

- Progression-free survival (PFS)
- Objective response rate (ORR) (complete response [CR] + partial response [PR])
- Disease control rate (DCR; CR + PR + stable disease [SD])
- Patient-reported outcomes (PROs): pain, health-related quality of life, and health status
- Duration of response (DoR)
- Duration of disease control (DDC)
- Safety and tolerability
- Pharmacokinetics (PK) and immunogenicity

4.3. Exploratory Objectives

The exploratory objectives of this study are to examine clinical variables, such as histological subtypes, and clinical outcomes, and the association between biomarkers and clinical outcomes.

5. A Priori Statistical Methods

5.1. Sample Size

The primary objective of this trial is to compare doxorubicin plus olaratumab versus doxorubicin plus placebo with respect to OS in 2 populations:

(1) Patients with advanced or metastatic STS not amenable to treatment with surgery or radiotherapy with curative intent

(2) Patients with advanced or metastatic LMS not amenable to treatment with surgery or radiotherapy with curative intent

Population (1) will be analyzed for efficacy among all randomized study patients (the intent-to-treat [ITT] population). Population (2) will be analyzed as the subset of randomized patients with LMS. Therefore, this document will consider Population (1) as synonymous with the ITT study population, and refer to Population (2) as the "LMS population". The study will be considered a positive study if either the ITT or LMS populations (or both) show a statistically significant improvement in OS with the regimen of doxorubicin plus olaratumab.

The study will enroll 460 patients in 1:1 randomization (230 patients in the investigational arm and 230 patients in the control arm). Enrollment will be conducted so that approximately 200 patients with LMS and 260 patients with other (non-LMS) histology will be randomized. The final analysis will occur only when both a minimum of 131 OS events have been observed in randomized patients with LMS and a minimum of 322 OS events have been observed in randomized patients overall.

See Section 5.11.1 for a detailed description of the statistical testing plan according to the graphical approach of Maurer and Bretz (2013).

The overall type I error rate for the study is controlled at 0.025 (1-sided). The associated sample sizes for the hypotheses that form the primary objective are based on the initial allocation of alpha (OS ITT $\alpha = 0.02$ and LMS $\alpha = 0.005$, 1-sided). One interim efficacy analysis will be performed at 194 ITT OS events.

If the ITT OS hazard ratio (HR) is assumed to be 0.723, and assuming 322 ITT OS events, a log-rank comparison of OS between study arms in the ITT population will have 80% power at the nominal alpha of 0.02. Assuming 30% censoring, a total sample size of 460 randomized patients is required.

5.2. General Considerations

This document describes the statistical analyses planned prior to final treatment assignment unblinding of the aggregate database. Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol and the justification for making the change will be described in the clinical study report (CSR). Additional exploratory analyses of the data will be conducted as deemed appropriate.

All tests of treatment effects will be conducted at a two-sided alpha level of 0.05 unless otherwise stated, and all confidence intervals (CIs) will be given at a two-sided 95% level, unless otherwise stated. Statistical analysis will be performed using SAS software (SAS, Version 9.1.2 or higher).

The following general terms will be used globally in the SAP:

- Unless otherwise specified, summary statistics stand for n, mean, standard deviation, median, minimum, and maximum for continuous variables; and frequency and percentage for categorical variables.

- **Study Treatment Period:** begins on the day the first dose of study treatment is administered and ends when the patient and the investigator agree that the patient will no longer continue study treatment. The date of this agreement is to be reported on the electronic case report form (eCRF) as the Date of Discontinuation from study treatment.
- **Postdiscontinuation 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 the day after the patient and the investigator agree that the patient will no longer continue study treatment and lasts approximately 30 days (±7 days) (until the short-term 30-day safety follow-up visit is completed).

  Long-term follow-up begins the day after short-term follow-up is completed and continues until the patient's death or overall study completion (whichever is earlier).

5.2.1. Definitions of Analysis Variables

Definitions of efficacy, safety, and PRO analysis variables are listed in Section 5.2.1.1, Section 5.2.1.2, and Section 5.2.1.3, respectively. Other variables are listed below alphabetically.

- **Age (years):** (Informed Consent Date - Date of Birth + 1)/365.25.
  Note. Average days in a year = 365.25, reflecting the Julian Year of 3 years with 365 days each and 1 leap year of 366 days. Birth month and day are imputed to be 01 July because only birth year is collected through eCRF.

- **Baseline measurement** is the last non-missing measurement prior to first dose for safety analyses, and the last non-missing measurement prior to randomization for demographic and efficacy analyses.

- **Duration** 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
    Note. Days in months = (1/12)*average number of days in a year
  - Duration (years): (End Date – Start Date + 1)/365.25

- **Duration of disease** is defined as months from first diagnosis of cancer to randomization.

- **Measurable disease (Yes/No)** is defined as yes for patients with at least 1 target lesion based on radiographic assessment data collected at baseline. If no target lesions are present, then patients would be categorized as a No.

- **Study Day:** Study day indicates the number of days the patient has been receiving study treatment. It is calculated as assessment date – first dose date + 1 day if the assessment is done on or after the first dose day. If the assessment is done prior to the first dose day, study day will be calculated as assessment date – first dose date. Date of first dose is defined as Study Day 1.

5.2.1.1. Efficacy Analysis Variables

Definition of efficacy analysis variables are listed.

Overall survival (OS) is defined for each patient as the time from the date of randomization to the date of death from any cause. If the patient is alive at the cutoff date for the analysis (or was lost to follow-up without a confirmed date of death), OS will be censored for analysis on the last date the patient was known to be alive.

Progression-free survival (PFS) is defined for each patient as the time from the date of randomization to the first date of radiologic disease progression (as defined by Response Evaluation Criteria In Solid Tumors, Version 1.1 [RECIST v.1.1]) based on investigator assessments or death due to any cause. Table JGDJ.5.1 defines the rules of censoring to be applied to PFS. In addition, sensitivity analyses of PFS will be performed using different rules for censoring (as defined by Table JGDJ.5.2).

Table JGDJ.5.1. Censoring Rule of PFS Primary 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)
unless		
No baseline radiological tumor assessment available	Censored	Date of randomization
No adequate postbaseline radiological tumor assessment available and 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 immediately after 2 or more consecutive missing scan intervals following last adequate radiological tumor assessment or randomization (whichever is later)	Censored	Date of last adequate radiological assessment prior to the missing assessment or date of randomization (whichever is later)

Abbreviations: CR = complete response; PD = progressive disease; PR = partial response; SD = stable disease.

a Symptomatic deteriorations (that is, symptomatic progressions, which are not radiologically confirmed) will not be considered as disease progressions.

b Adequate radiological tumor assessment refers to an assessment with one of the following responses: CR, PR, SD, or PD.

c The 2-scan interval is counted from the date of last adequate tumor assessment to the date of next 2 scheduled tumor assessments plus 14 days (adjusted by tumor assessment window).

d Refer to flow chart in Appendix 1 if a patient meets multiple censoring criteria.

e If there are multiple dates associated with 1 assessment, the assessment date will be set to the first date when the overall response is PD and the last date otherwise.

Table JGDJ.5.2. Censoring Rules for PFS Sensitivity Analysis Definitions

Sensitivity Analysis Definition #	Situation	Date of Progression or Censor	Censored / Progressed
SA 1: Count symptomatic deterioration as progression	Radiographic documented progression or symptomatic deterioration	Date of documented progression or date of symptomatic deterioration, whichever occurred first	Progressed
SA 2: Ignore new anticancer treatment	New anticancer treatment (systemic therapy) started before radiographic documented progression or death	A) Date of radiographic documentation of progression or death, whichever is earlier B) Last adequate radiological assessment if no radiographic documented progress or death occured	A) Progressed B) Censored
SA 3: Ignore missing tumor assessment	Death or radiographic documented progression after ≥ 2 consecutively missed tumor assessment visits	Date of radiographic documentation of progression or death, whichever is earlier	Progressed
SA 4: Treat lost to follow up as progression	Patient is lost to follow-up without radiographic documented progression or death	Date of next scheduled postbaseline radiological assessment at or after becoming lost to follow-up	Progressed

Abbreviations: PFS = progression-free survival; SA = Sensitivity Analysis.

Progression-free survival 2 (PFS2) is defined as the time from the randomization date to the date of disease progression on next-line treatment, or death due to any cause, whichever occurs first. Table JGDJ.5.3 defines the rules of censoring to be applied to PFS2. Note that disease progression on next-line treatment in this study will be recorded by investigators without details of corresponding radiologic assessment results. In the event that the date of disease progression on next-line treatment is reported only to the nearest month, the date will be imputed for analysis (assumed to have occurred on the 15th day of the reported month).

Table JGDJ.5.3. Censoring Rules of PFS2 Analysis

Situation					
PD on study therapy	PD after end of study therapy but prior to next line of treatment	PD on next line of treatment	Death	**Event / Censor**	**Date of Event or Censor**
Yes	No	Yes	Yes / No	Event	Event at PD date on next line of treatment
Yes	No	No	No	Censored	Censoring at the date PD-free on next line of treatment
Yes	No	No	Yes	1) Event 2) Censored	Event at death date if no further treatment, otherwise censoring at the date PD-free on next line of treatment
No	Yes	Yes	Yes / No	Event	Event at PD date on next line of treatment
No	No	Yes	Yes / No	Censored	Censoring at start of next line of treatment
No	Yes / No	No	No	Censored	Censoring at the date PD-free on next line of treatment
No	Yes	No	Yes	1) Event 2) Censored	Event at death date if no further treatment, otherwise censoring at the date PD-free on next line of treatment
No	No	No	Yes	Censored	Censoring at start of next line of treatment

Time to any progression (censoring for death without progression) is defined identically to PFS, except that the time to any progression will be censored at the date of death if there is no prior or concurrent radiologic disease progression. Otherwise, censoring follows the rules described in Table JGDJ.5.1.

Time to any new metastasis (censoring for death and/or for progressive disease (PD) due to increased sum of target lesions) is defined for each patient as the time from the date of randomization to the first date of radiographic documentation of 1 or more new lesions. Time to any new metastases will be censored at the first date of radiologic disease progression if that progression was based solely on an increased sum of target lesions (without new lesions). If there is no radiologic disease progression, time to any new metastases will be censored following the rules of Table JGDJ.5.1, with the exception that censoring will be applied at the date of death (if no previous event or censoring).

New-metastases-free survival (nMFS) is defined for each patient as the time from the date of randomization to the first date of radiographic documentation of 1 or more new lesions, or to the date of death from any cause (whichever occurs first). nMFS will be censored at the first date of radiologic disease progression if that progression was based solely on an increased sum of target lesions (without new lesions). Otherwise, nMFS will be censored for analysis in a manner analogous to PFS (following the rules of Table JGDJ.5.1).

Time to any progression based solely on increased sum of target lesions is defined as the time from the date of randomization to the first date of radiologic disease progression based solely on an increased sum of target lesions. Time to progression based on an increased sum of

target lesions will be censored at the first date of radiologic disease progression if that progression was based solely on new lesions. If there is no radiologic disease progression, time to any progression based on an increased sum of target lesions will be censored following the rules of Table JGDJ.5.1, with the exception that censoring will be applied at the date of death (if no previous event or censoring).

Objective response rate (ORR) is defined as the proportion of randomized patients achieving a best overall response of PR or CR per RECIST v.1.1. Patients who do not have any postbaseline tumor response assessments are considered non-responders and are included in the denominator when calculating the response rate. Tumor assessments performed after initiation of new anticancer treatment (systemic therapy) will be excluded from evaluating the best overall response.

Disease control rate (DCR) is defined as portion of randomized patients achieving a best overall response of CR, PR, or SD per RECIST v.1.1. Patients who do not have any postbaseline tumor response assessments for any reason are considered non-responders and are included in the denominator when calculating the response rate. Tumor assessments performed after initiation of new anticancer treatment (systemic therapy) will be excluded from evaluating the best overall response.

Duration of response (DoR) is defined for each patient with a best response of CR or PR as the duration from the first date of CR or PR to the first date of radiologic disease progression or death due to any cause. DoR will be censored according to the same rules as PFS.

Duration of Disease Control (DDC) is defined for each patient with a best response of CR, PR, or SD as the time from randomization to the first date of radiologic disease progression or death due to any cause. DDC will be censored according to the same rules as PFS.

Time to first worsening in Eastern Cooperative Oncology Group (ECOG) performance status is defined as the time from the date of randomization to the first date observing a 1-point (or greater) deterioration from baseline. For each patient without a worsening in ECOG performance status, censoring will be applied at the last date in which ECOG performance status was reported.

5.2.1.2. Safety Analysis Variables

Definitions of variables for safety analysis are listed by category and alphabetically within category.

Adverse event (AE)-related variables are listed below:

- **Adverse event (AE)** is defined as any untoward medical occurrence in a patient administered a pharmaceutical product, without regard to the possibility of a causal relationship.

AEs of special interest (AESIs)

AESI for olaratumab:

- Infusion-related reactions (IRRs)

AESI for doxorubicin:

- Cardiac arrhythmias
- Cardiac dysfunction

AESI for the combination of olaratumab and doxorubicin:

- IRRs
- Cardiac arrhythmias
- Cardiac dysfunction

Notes.
Categories of AESI may be modified as the understanding of the safety of the investigational drug increases. The final list of categories will be maintained at both compound and study level and reported in the CSR.

Consolidated AEs are composite AE terms consisting of synonymous PTs to allow meaningful interpretation of the AE data. Consolidated AE categories and PTs will be maintained at compound and/or study level and reported in the CSR.

Serious adverse event (SAE) is any AE that results in one of the following outcomes:

- death
- a life-threatening experience (that is, immediate risk of dying)
- persistent or significant disability/incapacity
- initial or prolonged in-patient hospitalization
- congenital anomaly/birth defect
- considered significant by the investigator for any other reason

Treatment-emergent adverse event (TEAE) is defined as an event that first occurred or worsened in severity after baseline and up to 30-day short-term follow-up visit.

Exposure-related variables are listed below:

- **Number of dose level reductions:** Sum of the number of dose level reductions as reported in the eCRF
- **Dose delays:** As reported in the eCRF
- **Dose withheld/skip (Not Administered):** As reported in the eCRF
- **Dose interruption(intravenous hold due to IRR):** As reported in the eCRF

Olaratumab or placebo treatment:

- Duration of treatment (weeks; 21 days added to duration of treatment because administration is every 3 weeks [on Days 1 and 8 of each 3-week cycle]) = ([Date of last cycle Day 1 − date of first dose] + 21) ÷ 7

- Cumulative dose, dose intensity, and relative dose intensity:
  - Cumulative dose (mg/kg) = Sum of (dose administered at each infustion [mg] ÷ Last available weight [kg])
  - Weekly dose intensity (mg/kg/week) = (Cumulative dose) ÷ (Duration of Treatment [week])
  - Planned weekly dose intensity (mg/kg/week) = 2 x (20 mg/kg + (number of cycles – 1) x 15 mg/kg) / (number of cycles x 3 weeks)
  - Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 100

Doxorubicin treatment:

- Duration of treatment (weeks) = ([Date of last dose − date of first dose] + 21) ÷ 7

- Cumulative dose, dose intensity, relative dose intensity:
  - Cumulative dose (mg/m^2) = Sum of (dose administered at each infusion [mg] ÷ Last available body surface area [BSA] [m^2])
  - Weekly dose intensity (mg/kg/week) = (Cumulative dose) ÷ (Duration of treatment)
  - Planned weekly dose intensity (mg/m^2/week) = $75mg/m^2$/ 3 weeks = 25 mg/m^2/week
  - Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 100

5.2.1.3. Patient-Reported Outcome Analysis Variables

Three scales will be used to assess patient reported Quality of Life (QoL) outcomes: EORTC-QLQ-C30, Modified Brief Pain Inventory – short form (mBPI-sf), and EQ-5D-5L.

EORTC-QLQ-C30

The European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 version 3.0 (EORTC-QLQ-C30), a self-administered, cancer-specific questionnaire consisting of 30 questions with multidimensional scales.

Assessments will be scored according to the EORTC QLQ-C30 Scoring Manual (Fayers et al., 2001). The 30 items (Q1-Q30) of the QLQ-C30 are scored to obtain 15 scales (1 global health status/QoL scale, 5 functional scales, and 9 symptom scales/items). A linear transformation is used to obtain scales ranging from 0 to 100 where:

- A high score for a functional scale represents a high / healthy level of functioning.

- A high score for the global health status / QoL represents a high QoL.
- A high score for a symptom scale / item represents a high level of symptomatology / problems.

EORTC QLQ-C30 (Version 3) Summary of Fifteen Scales and Scoring

Scale	Raw Score: Mean of items	Score
Global health status/QoL (QL2)	Q29, Q30	{1-(Raw Score -1)/6} x 100
Functional scales		
Physical functioning (PF2)	Q1-Q5	{1-(Raw Score -1)/3} x 100
Role functioning (revised) (RF2)	Q6, Q7	
Emotional functioning (EF)	Q21 - Q24	
Cognitive functioning (CF)	Q20, Q25	
Social functioning (SF)	Q26, Q27	
Symptom Scales		
Fatigue (FA)	Q10, Q12, Q18	{(Raw Score -1)/3} x 100
Nausea and vomiting (NV)	Q14, Q15	
Pain (PA)	Q9, Q19	
Dyspnoea (DY)	Q8	
Insomnia (SL)	Q11	
Appetite loss (AP)	Q13	
Constipation (CO)	Q16	
Diarrhea (DI)	Q17	
Financial difficulties (FI)	Q28	

Worsening will be defined as an increase of at least 10 points for the symptom scales or a decrease of at least 10 points for the functional scales and the global health status/QoL scale. Time to worsening will be calculated as the time from the first study drug dose(baseline date) to the first observation of worsening. If worsening is observed after a missing value, it may be assumed that the worsening occurred at the time of the missing value. Otherwise, the patient may be considered lost to follow-up and censored at the date of lastassessment. Patients with no post-baseline assessment will be censored at the baseline date. Patients who have the worst possible score at baseline or a missing baseline assessment (i.e. no assessment prior to or on the day of the first study drug dose) will not be included in this analysis. In addition, the following variables will be derived for each scale score:

- For each patient, change from baseline will be calculated for every postbaseline assessment by subtracting the baseline assessment result from the current assessment result.
- Maximum improvement and maximum worsening scores (over baseline) will be determined from the set of all postbaseline change scores.

- Maximum improvement and maximum worsening scores (over an assessment at start of monotherapy) will be determined from a set of change scores from 2nd monotherapy cycle onwards.
- A patient first improving over baseline by 10 points or more without prior worsening of 10 points or more will be categorized as having "improved" for that particular scale score during the study. A change of ≥10 points on the 100-point scales is considered clinically meaningful (Osoba et al. 1998)

Modified Brief Pain Inventory – short form (mBPI-sf)

The mBPI-sf assesses the severity of pain and its impact on functioning. The assessment will be analyzed primarily in terms of the "worst pain" score from each assessment. The first assessment will occur at Cycle 1 Day 1. **Time to first worsening of the mBPI-sf (Brief Pain Inventory Short Form Modified) "worst pain" score (TWP)** is defined for each patient as the time from the date of the first study drug dose (baseline date) to the first date of either a "worst pain" score increase of ≥2 points from baseline with no change in analgesic use or "worst pain" score increase of ≥1 point from baseline and an analgesic drug class increase of ≥1 level (Farrar et al., 2001; Rowbotham, 2001). If the patient has not worsened by either of these criteria, TWP will be censored for analysis on the last date the mBPI-sf was administered. Patients with a baseline worst pain score of 8 or more or with a missing baseline score (i.e. no assessment prior to or on the day of the first study drug dose) will not be included in the time to first worsening analysis.
A clinical pain response will be defined as a a "worst pain" score decrease of ≥2 points from baseline with no change in analgesic use or a "worst pain" score decrease of ≥1 point from baseline and an analgesic use decrease of ≥1 level. Patients with a baseline worst pain score of 0 (best possible score) or with a missing baseline score will not be included in the analysis of this endpoint. The cumulative distribution of the percentage of pain responders by treatment arm as a function of time is to be presented graphically.

In addition to the TWP variable defined above, the following variables will be derived for the "worst pain" score:

- For each patient, change from baseline will be calculated for every postbaseline assessment by subtracting the baseline assessment result from the current assessment result.
- Maximum improvement and maximum worsening scores (over baseline) will be determined from the set of all postbaseline change scores.
- Maximum improvement and maximum worsening scores (over an assessment at start of monotherapy) will be determined from a set of change scores from 2nd monotherapy cycle onwards.
- A patient first improving over baseline by 2 points or more without prior worsening of 2 points or more will be categorized as having "improved" his/her score during the study.

EQ-5D-5L

The EQ-5D-5L responses may be incorporated into a cost-utility analyses. The EQ-5D-5L data will be scored as described in literature (van Hout et al. 2012). The index score is calculated

from a set of item weights to derive a score of 0 to 1, with 1 representing the best health status. United Kingdom (UK) weights will be applied for the base case (EuroQol, n.d). Geographic-specific weights will be used as appropriate and when available as part of the cost-utility analysis for that specific geography.

Each patient completing the EQ-5D-5L report the level, or score for each of 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), as well as a visual analog scale (VAS) score. The index score is calculated as a function of individual levels (1=no problem, 2=slight, 3=moderate, 4=severe, and 5=extreme problem) from each of the 5 dimensions. The index score will not be computed for an assessment if the patient has 1 or more missing values among the 5 items. The VAS is a score reported by the patient ranging from 100 (best imaginable health state) to 0 (worst imaginable health state).

The analysis will include all cycles for which at least 25% of patients in each arm have an assessment. For cycles with data available from <25% of participants, analysis will be descriptive only.

The following variables will be derived for the EQ-5D-5L Index and VAS:

- For each patient, change from baseline will be calculated for every postbaseline assessment by subtracting the baseline assessment result from the current assessment result.
- Maximum improvement and maximum worsening scores (over baseline) will be determined from the set of all postbaseline change scores.
- Maximum improvement and maximum worsening scores (over an assessment at start of monotherapy) will be determined from a set of change scores from 2nd monotherapy cycle onwards.

5.3. Adjustments for Covariates

Analyses of all efficacy variables and patient-reported outcome variables defined in Sections 5.2.1.1 and 5.2.1.3 will be stratified using the randomization stratification factors listed below (defined based on IWRS data). Sensitivity analyses may also be performed to include additional stratification factors, non-stratified, or covariate adjusted analyses (for example, Cox models with covariates).

- Number of prior systemic therapies for advanced/metastatic disease (0 versus ≥1)
  NOTE: Any therapy administered in the adjuvant/neoadjuvant setting will not be considered as a prior line of therapy here.
- Histological tumor type (LMS versus liposarcoma versus pleomorphic sarcoma versus other STS types)
- ECOG performance status (0 versus 1)

Prospectively planned sensitivity, subgroup, and multivariate analyses are described in more detailed in Section 5.11.4.

Other baseline covariates that may be of interest include (but may not be limited to) the following:

- liver lesions (presence at baseline vs. absence at baseline)

- lung lesions (presence at baseline vs. absence at baseline)
- sex (females vs. males)
- age
- weight
- duration of disease since diagnosis
- grade of STS at diagnosis (1/low vs. 2/intermediate vs. 3/high)
- albumin level
- prior systemic therapy in the adjuvant or neo-adjuvant setting (none vs. any)
- prior systemic anti-cancer therapy (none vs. any)
- alanine aminotransferase (ALT)
- bone lesions (presence at baseline vs. absence at baseline)
- prior radiation therapy (none vs. any)
- duration of most recent prior systemic therapy
- hemoglobin
- platelets
- leukocytes

5.4. Handling of Dropouts or Missing Data

Rules for handling dropouts or missing data are listed by type of analysis alphabetically. Unless otherwise specified, observed data will be used and missing data will not be imputed or carried forward.

In the event that the date of second disease progression (disease progression occurring during post-study systemic anticancer therapy) is reported only to the nearest month, the date will be imputed for analysis (assumed to have occurred on the 15th day of the reported month).

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

- Onset date of an AE or start date of a concomitant 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 year and month 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 before/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 an onset date for an AE is missing, then the AE will be considered treatment emergent with unknown onset date.

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 only the day is missing, then assign Day 15 of the month, or the date of death if the patient died prior to 15th of the same month to the day.
- If month is missing, then the date will be set to July 1st of the year, or the date of death if the patient died prior to July 1st of the same year.

However, after imputation, check if the imputed date is logically consistent with other relevant date variable(s) and make appropriate correction if necessary. For example, if a visit start date was 16 May 2008 and a tumor assessment date was xx May 2008 (missing day) but it was known that it occurred on or after that visit, then after imputation, the tumor assessment date became 16 May 2008. In this case, the imputed tumor assessment date should be compared to the visit start date and then corrected to be the later of the 15th day of the month and the visit start date.

Patient-reported outcome analysis: For percentage compliance of the mBPI-sf, EORTC QLQ-C30, and EQ-5D-5L, instruments with at least 1 item completed will be considered as having been completed. Please refer to Section 5.13 for additional details.

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 (both components).
- 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. See Sections 5.2.1 and 5.11 for details.

5.5. Multicenter Studies

This is a multicenter, randomized, double-blind study. Due to the expected large number of investigative centers used for this study, investigative center was not used as a stratification factor and will not be used for covariate adjustment or subgroup analysis. Retrospective exploratory analyses of center-specific data or region-based subgroup analysis may be conducted as deemed appropriate to support global regulatory requests.

5.6. Multiple Comparisons/Multiplicity

The independent data monitoring committee (IDMC) will perform unblinded interim safety and efficacy analyses. Family-wise type I error is controlled for this study's key efficacy outcomes (OS, PFS, and ORR) as described in Section 12.1 of the study protocol and Section 5.11 of this SAP. Regarding all other study analyses, multiplicity of statistical error is not controlled or adjusted for in any way.

5.7. Study Patients

The following summaries (frequency and percentage) and listings for patient disposition will be performed:

- Patient disposition by investigator site and country and overall: patients entered (that is, signed informed consent), entered but not randomized, randomized (that is, ITT population), randomized but not treated, treated (that is, safety population), in Per-Protocol population, and eligible for PRO analysis (refer to Section 5.7.1 for analysis population definitions)
- The primary reasons for discontinuation from study treatment and patients still receiving treatment will be summarized by study treatment arm using frequency and percentage. The following discontinuation reasons will be presented: AE, PD (radiologically documented objective deterioration, symptomatic deterioration), death (due to AE, PD, or other), and other.
- Listings of:
    - primary reason for discontinuation from study regimen
    - date of randomization, first dose administration, last dose administration, and discontinuation from study regimen

5.7.1. Analysis Populations

Unless otherwise specified, all analyses will be based on the population of patients enrolled in the main protocol (Cohort 1). Patients enrolled under the extended enrollment addendum in Taiwan will be referred to as Cohort 2. Patients enrolled in Cohort 2 will not be included in the primary analysis of overall survival. As necessary, efficacy and safety analyses will be summarized for patients enrolled in Taiwan by pooling Cohort 1 and Cohort 2. All pooled analyses will be for descriptive purposes only.

Table JGDJ.5.4 lists analysis population definitions and associated data type for analysis.

Table JGDJ.5.4. Analysis Populations

Population	Definition	Analysis Type / Variable	Note
Intention-to-treat (ITT) Population	All randomized patients	Baseline characteristics, concomitant medication, all efficacy analyses	Patients will be grouped according to randomized treatment.
Safety Population (SP)	All randomized patients who received any quantity of study drug	Safety, e.g. dosing/exposure, AE and resource utilization	Patients will be grouped according to treatment received as defined by the first dose received.
LMS subset of ITT Population	All randomized patients with LMS histologic subtype	Baseline characteristics, concomitant medication, all efficacy analyses	Patients will be grouped according to randomized treatment.

Abbreviations: AE = adverse event; ITT = intent-to-treat; LMS = leiomyosarcoma.

Compliance for the PRO instruments will be reported for the ITT population. All other PRO analyses will be on the ITT population and will include those from whom a completed PRO instrument was obtained at baseline and at least 1 postbaseline (either during study treatment period or 30-day postdiscontinuation follow-up period); thus, the actual patients included for each analysis will depend on the instrument and response variable.

A patient listing of analysis population details will be provided. This listing will be presented by treatment group and will include: investigator site, patient identifier, inclusion/exclusion flag for each population, and reason for exclusion from each population. All patients screened will appear on this listing.

Significant protocol violations that potentially compromise the data integrity and patients' safety will be summarized for the ITT population. These violations will include deviations that can be identified programmatically and those that can only be identified by the clinical research associate during monitoring. Significant protocol violations are described in the Trial Issue Management Plan within the study Trial Master File. The list of significant protocol violations to identify patients to be excluded from the Per-Protocol population will be defined prior to the final database lock.

5.8. Demographic and Other Baseline Characteristics

The following patient demographic and other baseline characteristics will be summarized:

- patient demographics: age (years) and age group (< 65 vs. ≥ 65), gender,ECOG performance status, country, race (White, Black, Asian, All Other), height (cm), weight (kg), and BSA (m^2)
- potential prognostic factors as listed in Section 5.3
- baseline disease characteristics (at study entry): current disease stage, duration of disease (months)
- prior cancer therapies: type of therapy (surgery, radiotherapy, systemic therapy), type of prior surgery, type of prior radiotherapy, and type of prior systemic therapy

- historical illness (no versus at least 1 diagnosis) by Medical Dictionary of Regulatory Activities (MedDRA®) preferred term (PT), presented in decreasing frequency
  Note. Subjects reporting more than 1 condition/diagnosis within a PT will be counted only once for that PT.

- comparison between the eCRF and interactive web response system (IWRS) values of the stratification factors

Patient listings of demographic data and baseline characteristics will be provided. Patient listings of prior cancer therapies (surgery, radiotherapy, and systemic therapy) will be provided.

5.9. Concomitant Medications

The following concomitant medications used in study treatment period or the 30-day postdiscontinuation follow-up period will be summarized by numbers and percentages by treatment group, presented in decreasing frequency of the World Health Organisation drug term across treatment arms:

- all concomitant medications

- premedication for study drug

- growth factors

The proportions of patients reporting use of concomitant medications will be compared between the treatment groups. Patient listing of all concomitant therapies and premedications will be provided. Additional exploratory analyses of impact of premedication use on the rate and severity of IRRs, of growth factors on rates of neutropenia and febrile neutropenia, and of dexrazoxane use on the rate and severity of cardiac dysfunction will be considered.

5.10. Treatment Compliance

Olaratumab/placebo and doxorubicin will be intravenously administered only at the investigational sites. As a result, patient compliance is ensured.

5.11. Efficacy Analyses

5.11.1. Group Sequential Testing Using Graghical Approach

The statistical testings will be conducted according to the graphical method of Maurer and Bretz (2013) so as to control the overall type I error rate at 0.025 (one-sided) or equivalently, 0.05 (two-sided). The graphical approach can be characterized by first defining a set of prespecified null hypotheses that are organized graphically by providing initial alphas for each hypothesis and weights for each edge of the graph that will determine the propagation of α through the entire hypothesis-testing scheme.

The hypotheses objectives are:

Primary objectives: H_{11}: OS in the LMS population

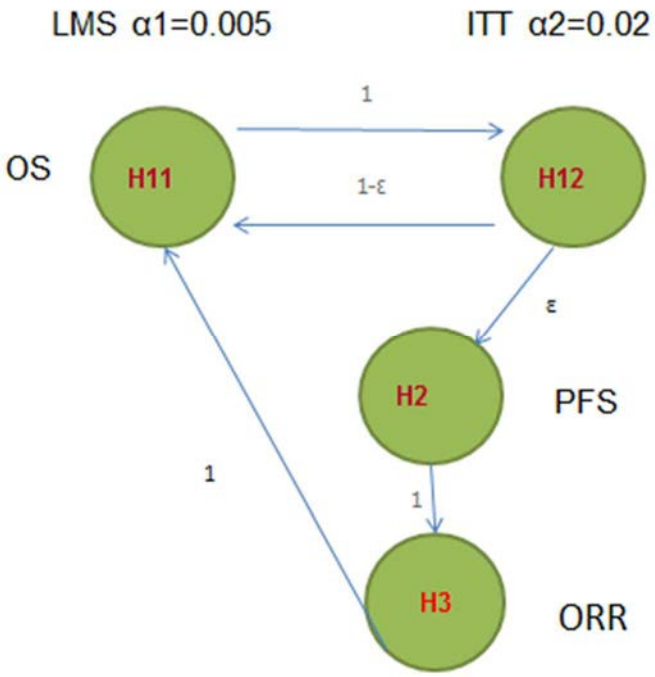
H_{12}: OS in the ITT population

Secondary objectives: H_2: PFS in the ITT population

H_3: ORR in the ITT population

Initially, the overall one-sided alpha of 0.025 is split between the primary objectives of OS in the LMS population (H_{11}) and OS in the ITT population (H_{12}), with H_{11} tested at a one-sided α_1=0.005 and H_{12} tested at a one-sided α_2=0.02. Zero alpha is initially assigned to the other hypotheses.

The initial graph with alpha at each node and weight for each edge is shown in Figure JGDJ.5.1.

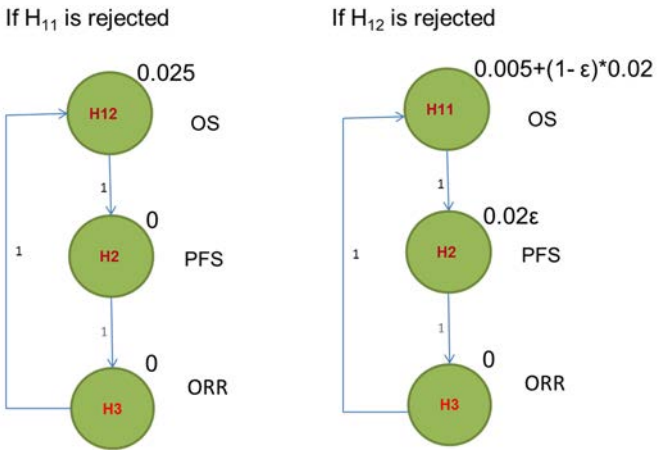


Abbreviations: ITT = intent-to-treat; LMS = leiomyosarcoma; ORR = overall response rate; OS = overall survival; PFS = progression-free survival. ε is set equal to 0.001.

Figure JGDJ.5.1. Initial graphical representation of testing sequence.

After rejection of any hypothesis, the graph will be updated on alphas at each node and weight at each edge, as described in Maurer and Bretz (2013).

It is valuable to note that the weight of the line from H_{11} to H_{12} is 1, indicating that H_{12} receives all of the alpha allotted to H_{11} if H_{11} is rejected, and similarly, the weight of the line from H_{12} to H_{11} is 1- ε, where ε is set equal to 0.001, indicating that H_{11} receives virtually all of the alpha allotted to H_{12} if H_{12} is rejected. Consequentially, after rejecting H_{11} and/or H_{12}, further testing essentially becomes a gatekeeping strategy to assess PFS and ORR in fixed sequence (except with an additional loop back from H_3), as illustrated below in Figure JGDJ.5.2.



Abbreviations: ORR = overall response rate; OS = overall survival; PFS = progression-free survival. ε is set equal to 0.001.

Figure JGDJ.5.2. Graphical representation of testing sequence conditional on rejection of either H_{11} or H_{12} to illustrate updating of alphas at each node and weight at each edge.

Assigning ε=0.001 initially focuses virtually all the alpha on the primary endpoints. The iterative process of subsequent updating of the graph and redistribution of α is described in Table JGDJ.5.5 below and is repeated until no further hypotheses can be rejected.

Table JGDJ.5.5. Planned Alpha Spending Using Graphical Approach

Test	Node α at Analysis
1. Initial Graph (Figure JGDJ.5.1)	
LMS OS	0.005
ITT OS	0.02
2. Graph when H_{11} LMS OS is rejected (Figure JGDJ.5.2; left)	
ITT OS	0.025
2.1 PFS* (If ITT OS is further rejected)	0.025
2.1.1 ORR**(If PFS is further rejected)	0.025
3. Graph when H_{12} ITT OS is rejected (Figure JGDJ.5.2; right)	
LMS OS	0.02498
3.1 PFS* (if LMS OS is further rejected)	0.025
3.1.1 ORR** (if PFS is further rejected)	0.025
3.2. PFS* (if LMS OS is not rejected)	0.00002
3.2.1 ORR** (if PFS is further rejected)	0.00002

Abbreviations: ITT = intent-to-treat; LMS = leiomyosarcoma; ORR = objective response rate; OS = overall survival; PFS = progression-free survival.

* Information fraction for PFS events at 60% OS events interim is 82% based on design assumptions that PFS median for control arm is 4.5 months, and experimental arm is 7.15 months.

** Information fraction for ORR response at 60% OS events interim is 85% based on design assumptions that ORR in response time is 6 months for both arms, and response rate for control arm is 12% and for experimental arm is 19%.

5.11.2. Primary Efficacy Analyses

For the primary comparison of OS, the primary endpoint, between the assigned study treatment arms, a stratified log-rank test will be performed to test the following statistical hypotheses about the OS HR for olaratumab over placebo:

H_0: OS HR $\geq$ 1.00 (Olaratumab not superior to placebo)

H_a: OS HR $<$ 1.00 (Olaratumab superior to placebo)

The stratification will be based on IWRS data that used for randomization (see Section 5.3 for more details). An unstratified log-rank test will also be performed as sensitivity analysis.

The following analyses of OS will also be performed:

- Summary of OS events (number and percentage), censoring rate, and reasons for censoring
- Restricted mean difference in OS between the treatment groups and its 95% CI, with the area under the Kaplan-Meier survival curve calculated up to the minimum across treatment arms of the maximum observed (that is, event or censored) time
- Kaplan-Meier survival curve (Kaplan and Meier, 1958) by treatment group will be provided.
- The Kaplan-Meier method will be used to estimate parameters (medians, quartiles, and percentages), difference of percentage and associated 95% CI and p-values for time-to-event analyses on each treatment group at 3, 6, 9, 12, and 24 months. Patients who did not have the event at the corresponding time point will be considered right-censored observations.
- HR for treatment effect will be estimated using Cox proportional hazards (PH) model stratified identically to the primary log-rank test with assigned treatment as the only covariate, reported with 2-tailed 95% CIs and Wald's test p-value. This Cox PH model will be referred to as the primary Cox PH model henceforth.

5.11.3. Secondary Efficacy Analyses

5.11.3.1. Supportive Analyses of Primary Efficacy Endpoint

The following supportive analyses of OS will be performed:

- HR for treatment effect will be estimated using an unstratified Cox PH model.
- As a sensitivity analysis, the primary OS analysis will be repeated using stratification based on the eCRF values.
- HR for treatment effect will be estimated using a multivariate Cox PH model, stratified by the randomization factors, with covariates selected among the additional factors listed in Section 5.3 using a stepwise selection method. Factors will be analyzed as continuous variables, except for those factors specifically identified with categories in Section 5.3. The stepwise selection will use an entry p-value <0.05 and exit p-value ≥ 0.10. The "assigned treatment arm" variable will not be used within the stepwise procedure but rather added to the final model. The OS HR for treatment effect and corresponding 95% CI will be estimated from the final model. Any covariate listed in Section 5.3 may be removed from this planned analysis if the number of patients representing 1 level of that variable is insufficient or data collected on that variable are insufficiently complete.

- As a sensitivity analysis, the primary OS analysis will be repeated for the Per-Protocol population.
- To evaluate whether new anticancer treatment affects OS, a sensitivity analysis will be conducted censoring OS at a start date of new anticancer therapy.
- To evaluate whether the number of cycles of doxorubicin affects overall survival, a comparative OS sensitivity analysis will be conducted that excludes patients completing less than 4 cycles of doxorubicin (then repeated excluding those completing less than 5 cycles, then 6 cycles).
- Efficacy analyses (OS, PFS, and ORR) will be analyzed for a subset of patients who have received at least one dose of either olaratumab or placebo monotherapy after discontinuation of study treatment with doxorubicin. Baseline characteristics and drug exposure will be summarized as supportive analyses for this subgroup of patients receiving monotherapy.

5.11.3.2. Key Secondary Efficacy Analyses

For PFS, the same analyses used for the analyses of the primary endpoint OS will be performed. In addition, as sensitivity analyses, the primary PFS analysis will be repeated using different PFS censoring rule as defined in Table JGDJ.5.2, to evaluate whether and to what extent the conclusion of the PFS analysis under the primary definition would be affected under the different censoring rules.

This comparison of PFS using the same method as that for the primary analysis of PFS will be considered inferential only in case of significant results for OS analysis (that is, as a gatekept analysis so as not to inflate the overall type I error rate).

The main PFS analysis will also be repeated for a subset of patients started monotherapy.

If PFS analysis is significant, then testing on ORR will be conducted. Objective tumor response (CR+PR) rate (ORR) will be reported along with exact confidence bounds (CI: 95%) and compared using the Cochran-Mantel-Haenszel test adjusting for the stratification variables. A sensitivity sensitivity analysis on ORR will be performed to consider only confirmed response (using 4-week interval for confirmation).

5.11.3.3. Analyses of Other Secondary Efficacy Variables

All time-to-event variables (including those defined in Sections 5.2.1.1 and 5.2.1.3) will be analyzed using stratified log-rank tests (analogous to the primary analysis) for comparisons between study arms, stratified Cox models (for between-arm statistics including the treatment HR), and Kaplan-Meier method (for within-arm statistics).

Disease control (CR+PR+SD) rate will be reported along with exact confidence bounds (CI: 95%) and compared using the Cochran-Mantel-Haenszel test adjusting for the stratification variables.

Patient listings of tumor assessments (target and non-target lesion assessments and tumor response), OS, and PFS will be provided.

5.11.4. Subgroup Analyses

OS and PFS HR for treatment effect (with 95% CIs) will be estimated using an unstratified Cox PH model for each of the following subgroups (defined based on eCRF data):

- number of prior systemic therapies for advanced/metastatic disease (0 versus ≥1)
- prior systemic treatment in the neo-adjuvant or adjuvant setting (yes vs. no)
- histological tumor type (LMS vs. liposarcoma vs. undifferentiated pleomorphic sarcoma vs. other STS types)
- LMS primary site (uterine vs. non-uterine)
- ECOG performance status (0 vs. 1)
- region (North America vs. Europe vs. ROW)
- disease stage at randomization (metastatic disease versus only locally advanced disease)
- liver lesions (presence at baseline vs. absence at baseline)
- lung lesions (presence at baseline vs. absence at baseline)
- sex (females vs. males)
- age (< 65 years vs. ≥ 65 years)
- weight (above and below median)
- duration of disease since diagnosis (above and below median)
- grade of STS at diagnosis (1/low vs. 2/intermediate vs. 3/high)
- albumin level (above and below 35 g/dL)
- ALT (above and below median)
- bone lesions (presence at baseline vs. absence at baseline)
- prior radiation therapy (none vs. any)
- duration of most recent prior systemic therapy (above and below median)
- hemoglobin (above and below median)
- platelets (above and below 350 μliters)
- leukocytes (above and below 10,000 μliters)
- PDGFRα status (positive and negative [additional details in Section 5.16])

If a level of a factor consists of fewer than 5% of randomized patients, analysis within that level will be omitted. Additional subgroup analyses may be performed as deemed appropriate. 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. Appropriate interpretation is important since, even if all patient subgroups benefit to exactly the same extent in truth, 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) will be followed to provide background information on the extent of variability that might be expected by chance alone.

Additional exploratory subgroup analyses will consider specifically those histologic subtypes (listed immediately below) that are rare and/or possibly less responsive to chemotherapy with doxorubicin. Patient listings with efficacy outcomes will be generated for each of these subtypes. Summary analyses such as Kaplan-Meier or Cox modeling may be performed, either for the overall combination or certain combinations of these subtypes, depending on the number of patients with each of these subtypes.

- alveolar soft-part sarcoma
- synovial sarcoma
- clear cell sarcoma of soft tissue
- malignant solitary fibrous tumour
- PEComa NOS, malignant
- plexiform fibrohistiocytic tumour
- giant cell tumour of soft tissue
- extraskeletal myxoid chondrosarcoma
- angiosarcoma
- dedifferentiated liposarcoma

Selected analyses as listed below will be performed for subgroups gender (male vs. female), age group (<65 years vs. ≥65 years), and race (White vs. Black vs. Asian vs. All Other). ().

- overview of treatment-emergent AESIs
- summary of patient demographics and baseline characteristics
- summary of TEAEs by system organ class (SOC) and PT
- summary of TEAEs by worst CTCAE grade and PT
- summary of TEAEs by PT and decreasing frequency
- summary of TEAEs by consolidated category and PT

Additional subgroup analyses will be performed on the following variables if minimuium of 40 subjects was achieved:

- age: <65 years, ≥65 years and <75 years, ≥75 years and <85 years, and ≥ 85 years

Patients demographic and baseline characteristics, drug exposure, key efficacy (OS, PFS, ORR), safety, and PRO variables will also be summarized for patients with no prior systemic therapies for advanced/metastatic disease.

5.11.5. Analyses by Baseline Lung Lesions

Section 5.11.4 identified the presence or absence of lung lesions at baseline as one factor to be considered for efficacy subgroup analyses. Retrospective data analyses of the phase 2 study I5B-MC-JGDG (conducted after initiating enrollment of the phase 3 study I5B-MC-JGDJ, but prior to unblinding of aggregate phase 3 efficacy data) showed strong evidence of a statistical interaction for overall survival, between study treatment arm and the presence or absence of baseline lung lesions. A much stronger OS treatment effect was observed in patients with

baseline lung lesions than in patients without baseline lung lesions. These results suggest the hypothesis that olaratumab, when used as part of a treatment for patients with advanced or metastatic disease, may be generally more effective in patients with lung lesions than in patients without lung lesions; and therefore in this section we pre-specify analyses for the phase 3 JGDJ study to look for confirmation of this statistical interaction.

Statistical interaction should initially be evaluated by constructing a Cox model with main effects for (i) the assigned study treatment arm and (ii) radiologic evidence of baseline lung lesions, plus an interactive term for (iii) treatment arm by baseline lung lesions. Other terms for prognostic factors may be included as additional cofactors in the model as deemed appropriate depending on the strength of prognostic effect. Statistical interaction will be judged primarily based on the statistical significance of the interaction term in the model.

If the interaction is statistically significant, then separate Cox models should be constructed for patients with baseline lung lesions and for those without (and repeated for other time to event efficacy variables including PFS, nMFS, and time to new metastases). Also in the event of a statistically significant interaction, further exploratory analyses should be conducted to evaluate whether the interactive effect is consistent across histologic subtypes and across other key subgroups (e.g. men versus women, ECOG performance status 0 versus 1, etc.). These further exploratory analyses will be important for evaluating the clinical relevance of any observed statistical interaction.

Additional exploratory analyses will be conducted considering tissue and biomarker data with respect to lung lesions. For example, available tissue from lung lesion biopsies may be analyzed for PDGFR alpha expression. Depending on available data, these results may be analyzed for association with efficacy outcomes.

5.12. Post-Study Drug Discontinuation Therapy

The numbers and percentages of patients reporting post-study therapies will be provided overall, by type of therapy (surgery, radiotherapy, or systemic therapy), and by regimen for all systemic anticancer regimens used. Patients will also be analyzed by post-study systemic anticancer regimens with respect to whether the treatment was the first post-study regimen, second post-study regimen, etc.

5.13. Patient-Reported Outcome Analyses

For each instrument, percentage compliance will be calculated as the number of completed assessments divided by the number of expected assessments (that is, patients still on study). Percentage compliance will be summarized by treatment group and overall. Similarly, the reasons for non-compliance will also be summarized descriptively.

Time-to-event variables will be analyzed using stratified log-rank tests (analogous to the primary analysis) for comparisons between study arms, stratified Cox models (for between-arm statistics including the treatment HR), and Kaplan-Meier method (for within-arm statistics).

Data will also be summarized descriptively for all of the variables identified Section 5.2.1.3, including shift tables.

QLQ-C30

Percentages of patients categorized as "improved" will be summarized and compared between study arms. Maximum improvement and worsening scores will be analyzed as continuous variables and compared between study arms using analysis of covariance (with both parametric and non-parametric p-values reported).

For the Physical functioning (PF2) as well as Emotional functioning (EF) scales, the change from baseline will be further analyzed by using mixed-effect repeated measures models based on restricted maximum likelihood estimation. Fixed effects in the model will include study treatment arm, baseline score, cycle number, and interaction term for study treatment arm and cycle number. The variance-covariance matrix that results in the minimum Akaike Information Criterion (Akaike 1973) from among unstructured, variance components, auto-regressive, and compound symmetric will be incorporated in the model. To implement the variance structure, patients will be included in the model as a random effect. The magnitude of the main effects and interactions will be evaluated and discussed. Based on the model-based means (LSMeans) from the repeated measures model, treatment group contrasts will be tested for each cycle number separately. Treatment group contrasts also will be tested for the treatment group marginal means.

Floor and ceiling effects will be summarized for each of the 15 scales. 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 worsening is less likely.

Modified Brief Pain Inventory – short form (mBPI-sf)

Percentages of patients categorized as "improved" will be summarized and compared between study arms. Similarly, percentage of patients achieving a clinical pain response will be summarized and compared between study arms. The cumulative distribution of the percentage of pain responders by treatment arm as a function of time is to be presented graphically. Maximum improvement and worsening scores will be analyzed as continuous variables and compared between study arms using analysis of covariance (with both parametric and non-parametric p-values reported).

Individual pain items on the mBPI-sf (that is, worst, least, average, and current pain) will be described using descriptive statistics by treatment arm and cycle. A mixed effects repeated measures model will be applied to compare between treatment arms, which may be adjusted for other covariates. Similar analyses will also be conducted for the mean of 7 pain interference with function items.

EQ-5D-5L

The EQ-5D-5L responses for each item will be summarized by frequency and corresponding percentages by treatment arm and cycle. Descriptive statistics (mean, standard deviation, median, minimum and maximum) for the index and VAS will be calculated and presented by treatment arm and cycle. Additionally, the change from baseline will also be presented. The index score

between treatment arms will be compared using mixed models. The model will include baseline score as a covariate and an unstructured covariance matrix will be utilized. A similar analysis will be performed on the VAS scores.

Of interest is a significant time-by-group interaction for each of the items, addressing whether treatment group profiles are different over time (from randomization through the last assessment following discontinuation).

5.14. Safety Evaluation

5.14.1. Exposure

Exposure to study drug will be analyzed for all patients treated with any non-zero amount of study drug. Analyses will be summarized for the treated subset of the ITT population and the treated subset of the LMS population. A summary of study drug exposure will include number of infusions, duration of therapy, cumulative dose level, weekly dose intensity, and relative dose intensity. The exposure formulas of olaratumab and doxorubicin are defined below.

Olaratumab:

- Duration of therapy (in weeks) = ([Date of last cycle Day 1 - date of first dose of olaratumab] + 21) ÷ 7
- Cumulative dose (mg) = Sum of all doses
- Calculated dose level administered (mg/kg) = Actual total dose of olaratumab (mg) / Closest body weight prior to that administration (kg)
- Cumulative dose level (mg/kg) = sum of all calculated dose level
- Dose intensity (mg/kg/week) = (cumulative dose level) ÷ (duration of treatment)
- Planned dose intensity (mg/kg/week) = planned dose per infusion (mg/kg) / infusion cycle
- Relative dose intensity (%) (based on planned dose) = (dose intensity/planned dose intensity)*100

All analyses of olaratumab exposure will be performed for both overall treatment period as well as for monotherapy after the completion of combination treatment with doxorubicin.

Doxorubicin:

- Duration of therapy (in weeks) = ([date of last dose of doxorubicin - date of first dose of doxorubicin) + 21) ÷ 7
- Cumulative dose (mg) = sum of all doses
- Calculated dose level administered (mg/ m^2) = actual total dose of doxorubicin (mg)/closest BSA prior to that treatment (m^2)
- BSA (m^2) = [weight (kg) ^0.425 * height (cm) ^0.725] * 0.007184
- Cumulative dose level(mg/ m^2) = sum of all calculated dose levels
- Dose intensity (mg/ m^2/weeks) = (cumulative dose level) ÷ (duration of treatment)
- Planned dose intensity (mg/ m^2/weeks) = planned dose per infusion (mg/ m^2) / infusion cycle

- Relative dose intensity (%) (based on planned dose) = (dose intensity/planned dose intensity)*100.

Dexrazoxane:

- Duration of therapy (in weeks) = ([Date of last dose of dexrazoxane – date of first dose of dexrazoxane] + 21) ÷ 7
- Calculated dose level administered (mg/m^2) = actual total dose of dexrazoxane (mg)/ closest BSA (mg^2) prior to that treatment
- Cumulative dose (mg) = sum of all doses
- Cumulative dose level (mg/m^2) = sum of all calculated dose levels
- Dose intensity (mg/m^2/weeks) = (cumulative dose level) ÷ (duration of treatment)
- Planned dose intensity (mg/m^2/weeks) = planned dose per infusion (mg/m^2) ÷ infusion cycle
- Relative dose intensity (%) (based on planned dose) = (dose intensity/planned dose intensity)*100.

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

5.14.2. Adverse Events

An adverse event (AE) is defined as any untoward medical occurrence in a patient administered a pharmaceutical product, without regard to the possibility of a causal relationship.

Treatment-emergent adverse event (TEAEs) are events that first occurred or worsened in severity after baseline and up to 30-day short-term follow-up visit.The MedDRA PT derived from the verbatim term will be used and severity is measured using the grade defined by the National Cancer Institute - Common TerminologyCriteria for Adverse Events (NCI-CTCAE).

Study drug-related AEs are AEs that were considered to be at least possibly related to study drug by an investigator. Missing relationship is considered related to all study drugs.

Serious adverse events (SAEs) are any AEs that result in one of the following outcomes:

- death
- initial or prolonged inpatient hospitalization
- a life-threatening experience (that is, immediate risk of dying)
- persistent or significant disability/incapacity
- congenital anomaly/birth defect
- considered important by the investigator for any other reason

AEs of special interest (AESIs)

AESIs are events which have been identified as safety signals during preclinical or early clinical trials or based on class effects of similar drugs. These events will be monitored prospectively in

the clinical developmental program. Each event is defined by a careful assessment and grouping of individual related MedDRA PTs. The list of PTs for AESIs is in Appendix 2.

AESI for olaratumab
- IRRs

AESI for doxorubicin
- Cardiac arrhythmias
- Cardiac dysfunction

AESI for the combination of olaratumab and doxorubicin
- IRRs
- Cardiac arrhythmias
- Cardiac dysfunction

Consolidated AEs include Abdominal Pain, Anemia, Fatigue, Hyperbilirubinaemia, Hypertension, Hypoalbuminaemia, Hypokalaemia, Hypomagnesaemia, Hyponatraemia, Hypoproteinemia, Intestinal Obstruction, Leukocytosis, Leukopenia, Lymphopenia, Mucositis, Neuropathy, Neutropenia, Musculoskeletal pain, Rash, and Thrombocytopenia. Each consolidated AE contains PTs identified as clinically identical or synonymous. The list of PTs for consolidated AEs is in the Appendix 3.

The most current version of MedDRA at time of analysis will be used when reporting AEs by MedDRA terms. 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 1 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.

5.14.2.1. Overall Summary of Adverse Events

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, or CTCAE Grade 3 or 4 or 5 TEAE
- patients with AEs that led to death (all, up to 30 days after last dose of study drug), or discontinuation of study drug regimen
- patients with SAEs that led to discontinuation of study drug regimen

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. Comparison between the treatment groups will be performed using Fisher's exact test.

5.14.2.2. Treatment-Emergent Adverse Events (TEAEs)

An overview of TEAEs will be provided to summarize the number and percentage of patients with any:
- TEAE
- treatment-emergent SAE

- CTCAE Grade ≥3 TEAE
- >5% by experimental arm
- TEAE leading to death (on treatment and within 30 days of last dose of study drug)
- TEAE leading to discontinuation of olaratumab, chemotherapy, or any study drug
- TEAE leading to dose modification of any study drug, olaratumab, or chemotherapy
- TEAE leading to hospitalizations
- TEAE leading to transfusions
- TEAE by cycle

The numbers and percentages will be calculated based on overall (regardless of causality), possibly related to olaratumab, chemotherapy, or any study drug for the overview of TEAEs.

In addition, the following TEAE summaries will be provided (regardless of causality, and study-drug related):

- summary of TEAEs by SOC and PT
- summary of TEAEs by high level group term (HLGT) and high level term HLT
- summary of TEAEs by worst CTCAE grade and PT
- summary of TEAEs by PT and decreasing frequency
- summary of TEAEs by AESI and consolidated category and PT

A patient listing of all AEs will be provided.

5.14.3. Deaths, SAEs, and Other Significant AEs

Deaths

The following death reports will be provided:

- summary of deaths (all deaths and deaths within 30 days of last dose of study drug) and their primary cause (study disease progression, AE, other)
- listing of treatment-emergent adverse events leading to death

SAEs

The following SAE summaries will be provided:

- summary of treatment-emergent SAE by SOC and PT
- summary of study drug-related treatment-emergent SAE by SOC and PT
- summary of AESI and consolidated treatment-emergent SAEs
- summary of study drug-related consolidated treatment-emergent SAEs

A listing of SAEs will be produced.

AE of special interest (AESIs)

The following AESI analyses will be provided:

- overview of treatment-emergent AESI (regardless of causality and study drug-related)
- summary of treatment-emergent AESI by AESI group and PT (regardless of causality and study drug-related)

- listing of treatment-emergent AEs of IRR
- listing of treatment-emergent AESIs (cardiac arrhythmia and cardiac dysfunction)

Consolidated AEs

Any AE summary table that has the need for consolidation will include a summary of the corresponding consolidated AEs in the table. The associated synonymous PTs will also be presented under each consolidated AE. Events will be ordered by decreasing frequency of PTs and consolidated AEs will be presented alphabetically. The following analyses include consolidated AEs:

- summary of TEAEs by worst CTCAE grade and PT
- summary of TEAEs by PT and decreasing frequency
- summary of TEAEs by consolidated category and PT
- summary of consolidated treatment-emergent SAEs
- summary of study drug-related consolidated treatment-emergent SAEs
- summary of TEAE on Core Safety Information criteria

Other significant adverse events

The following analysis will be provided:

- summary of TEAEs that led to discontinuation of any study drug, olaratumab, or chemotherapy by SOC and PT
- summary of TEAEs that led to dose modification of any study drug, olaratumab, or chemotherapy by SOC and PT
- listing of TEAEs leading to discontinuation of study drug
- listing of TEAEs leading to study drug dose modifications

5.14.4. Clinical Laboratory Evaluation

The severity of laboratory results will be classified according to CTCAE Version 4.0. The shifts in CTCAE toxicity grading from baseline to worst grade postbaseline (first dose up to 30 days after the last dose of study treatment) will be produced.

A patient listing of all laboratory data will be provided with a flag for values outside of the laboratory normal range as well as investigator site, patient identifier, age, gender, race, weight and visit.

5.14.5. Hospitalizations and Transfusions

The frequency and percentage of patients with any hospitalizations experienced during the study treatment period or 30-day postdiscontinuation follow-up period will be summarized by treatment group. Hospitalization incidence rates will be compared between the treatment groups using Fisher's exact test. In addition, total number of days in hospital and admissions will be summarized and compared using the Wilcoxon rank sum test.
Note. Discharge date will be imputed with last contact date for hospitalizations that are still ongoing at time of analysis.

The frequency and percentage of patients with any blood transfusions received during the study treatment period or 30-day postdiscontinuation follow-up period will be summarized by treatment group. Transfusions will be further characterized by transfused blood product (for example, packed red blood cells, platelets, fresh frozen plasma, or whole blood). The proportions of patients having blood transfusions will be compared between the treatment groups using Fisher's exact test.

Details of hospitalizations and transfusions will be included in patient listings.

5.14.6. Vital Signs, Physical Findings, and Other Observations Related to Safety

A summary of ECOG performance status at each scheduled time point will be provided. Actual value and change from baseline for vital sign measurements will be summarized at each assessment time point using summary statistics. Electrocardiogram (ECG) measurements will be summarized at each assessment time point using summary statistics. Listings of ECOG performance status, vital signs, and ECG data will be provided.

5.15. Pharmacokinetics and Immunogenicity

Serum concentrations of olaratumab prior to infusion (minimum concentration) and at 1 hour post-end of olaratumab infusion (approximately maximum concentration) will be summarized using descriptive statistics. Additional analysis utilizing the population PK approach may also be conducted if deemed appropriate. Relationships between olaratumab exposure and measures of efficacy and safety will be explored. A separate analysis plan will be provided for further PK analysis.

For immunogenicity, the number and percentage of patients with positive olaratumab antibody response will be summarized. Additional efficacy or safety analyses may be performed in the subgroup of patients with positive olaratumab antibody response. The antibody response and any alteration in olaratumab PK may also be explored, as well as any relationship with experiencing an infusion reaction. Further exploratory analyses may be performed as appropriate.

5.16. Translational Research

Planned translational research analysis includes the effect of platelet derived growth factor receptor α (PDGFRα) status on OS and PFS. PDGFRα status will be determined by using the Cell Signaling Technology rabbit monoclonal antibody (clone D13C6) proven to be specific for PDGFRα with no cross-reactivity for PDGFRβ. This PDGFR α protein expression (pretreatment) immunohistochemistry is assessed at Lilly Clinical Diagnostics Laboratory in tumor cells, and was provided as a dichotomous variable with "positive" and "negative" expression, where a "positive" result shows at least 10% of the tumor (rounded to the nearest decile) demonstrating at least weak but specific membranous staining (1+ on a 0, 1+, 2+, 3+ scale of staining intensity). "Negative" corresponds to staining that does not meet these criteria. All other translational research analysis will be considered exploratory.

5.17. Interim Analysis

An IDMC will be established prior to first patient visit. The IDMC will review unblinded interim analyses of safety and efficacy data. These unblinded interim analyses will be performed by independent statistician. Only the IDMC and regulatory authorities are authorized to evaluate unblinded interim efficacy and safety analyses; Lilly (sponsor) will not be unblinded until the final analyses. Information that may unblind the study during the interim analyses will not be reported to study sites or study team until the study has been unblinded.

All interim analyses will include complete assessments of safety-related data; one interim analysis will include efficacy analysis at 60% of target OS events. Patient enrollment will continue during the conduct of these analyses. The interim analyses will be performed as follows:

- An unblinded safety data review after approximately 40 evaluable patients in each arm. This review will be performed by IDMC.
- After the first safety data review, regular safety reviews will be conducted by the IDMC approximately every 6 months.
- An efficacy analysis for the interim OS analysis will be conducted by the IDMC after 194 OS events (60% of the final OS events) have been observed in the ITT population.
- At the interim efficacy analysis, a safety data analysis will also be performed.

The IDMC meeting will occur within 2 months of the interim data cutoff date. The details on the process flow/communication plan among study team, IDMC, and senior management are provided in the IDMC charter.

5.17.1. Interim Efficacy Analysis

One interim efficacy analysis will be performed on/after 194 OS events (60% of the final OS events) observed in the ITT population. This analysis will be performed in order to provide the IDMC and regulatory authorities an opportunity to review interim safety and efficacy data together. There will be no formal statistical hypotheses tested at the interim for primary and secondary efficacy outcomes. There will be no alpha-spending associated with the interim analysis; study follow-up and data collection will continue as planned until the final analysis regardless of the nature of the interim efficacy results. Only the IDMC and regulatory authorities will be allowed access to unblinded interim data ; Lilly (sponsor) will not be unblinded to the interim analyses.

The following core analysis will be performed in addition to safety interim analysis described in Section 5.17.2:

Demographic and Other Baseline Characteristics

- patient demographics
- potential baseline prognostic factors
- baseline disease characteristics
- prior cancer therapies
- historical illness

Efficacy Analyses

- Kaplan-Meier survival curves and median survival and 95% CI for OS and PFS
- HR estimated using unstratified Cox PH model (because there will be many small cells for the primary Cox PH model at the time of the interim analysis) for OS and PFS
- PFS sensitivity analysis with different censoring rules according to Table JGDJ.5.2
- summary of ORR and DCR

5.17.2. Safety Interim Analysis

The following analyses will be performed for safety interim analysis, with analysis population as specified in Table JGDJ.5.4.

Demographic and Other Baseline Characteristics

- patient demographics

Patient Disposition

- patient disposition
- reasons for treatment discontinuation as well as patients continuing on the study

Exposure

- summary statistics for exposure-related variables
- dose intensity of study drugs
- reasons for dose adjustments and dose delays

Adverse Events

- overview of AEs
- TEAEs summarized by PT*
- CTCAE v 4.0 Grade 3 or 4 or 5 AEs*
- SAEs summarized by PT*
- AESIs by PT*
- reasons for deaths
- AEs leading to study treatment discontinuations summarized by PT
- AEs leading to study treatment dose modification summarized by PT
- listing of SAEs
- listing of preexisting conditions and AEs

* Repeat for events deemed by the investigator to be possibly related to study medication.

Additional analyses may be conducted, at the request of the IDMC.

5.18. 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 AEs, provided as a dataset which will be converted to an XML file. Both SAEs and “Other” AEs are summarized by treatment group and by MedDRA Preferred Term.
- An AE is considered “Serious” whether or not it is a TEAE.

- An AE is considered in the "Other" category if it is both a TEAE and is not serious.
- For each SAE and "Other" AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event (if certain subjects cannot be at risk for some reason, for example, gender-specific AEs, then the number will be adjusted to only include the patients at risk)
    - the number of participants who experienced each event term
    - the number of events experienced.
- For each SAE, for each term and treatment group, the following are also provided for the EudraCT results submission:
    - The number of occurrences (events) causally related to treatment
    - The total number of deaths
    - The number of deaths causally related to treatment
- Consistent with www.ClinicalTrials.gov requirements, a threshold for frequency of "Other" AEs can be implemented rather than presenting all "Other" AEs. For example, "Other" AEs that occur in fewer than 5% of patients in any treatment group may not be included if a 5% threshold is chosen. The frequency threshold must be less than or equal to the allowed maximum of 5%.
- A participant flow will be created that will describe:
    - Number of participants per treatment arm. Screen failures do not need to be included. Number of participants who did not complete the study per treatment arm. This analysis will be based on study discontinuation, not treatment discontinuation.
    - Reasons participants did not complete the study.

5.19. Development Safety Update Report

The following reports are needed for the Development Safety Update Report (DSUR):

- Exposure information
- Listing of subjects who died during the DSUR period
- Discontinuations due to AEs during the DSUR Period

6. Unblinding Plan

This unblinding plan refers to the process to be followed for the final OS analyses.

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. This blinding will be maintained until the primary data lock.

Data sets will be created for the purpose of aggregate data review in which treatment assignment and related data, such as study drug administration dates and amounts, are 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.

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 discontinued from study therapy or excluded from any safety or efficacy analysis.

In order to maintain the scientific integrity of this double-blind trial and the prospectively planned alpha-controlled analyses, access to study data will be strictly controlled. Treatment assignment will be scrambled in the reporting database until the database lock for final OS analysis.

7. References

Akaike, H. Information theory and an extension of the maximum likelihood principle, in Petrov, B.N.; Csáki, F., *2nd International Symposium on Information Theory*, Akadémiai Kiadó, Budapest. 1973, pp. 267–281.

Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35(11):1095-1108.

Fayers P, Aaronson NK, Bjordal K, Curran D, Broenvold M; on behalf of the EORTC Quality of Life Study Group. EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels (Belgium): EORTC Quality of Life Group; 2001.

Farrar JT et al. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001 Nov;94(2):149-58.Fleming TR. Interpretation of subgroup analyses in clinical trials. *Drug Information Journal*. 1995;29:1681S-1687S.

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457-481.

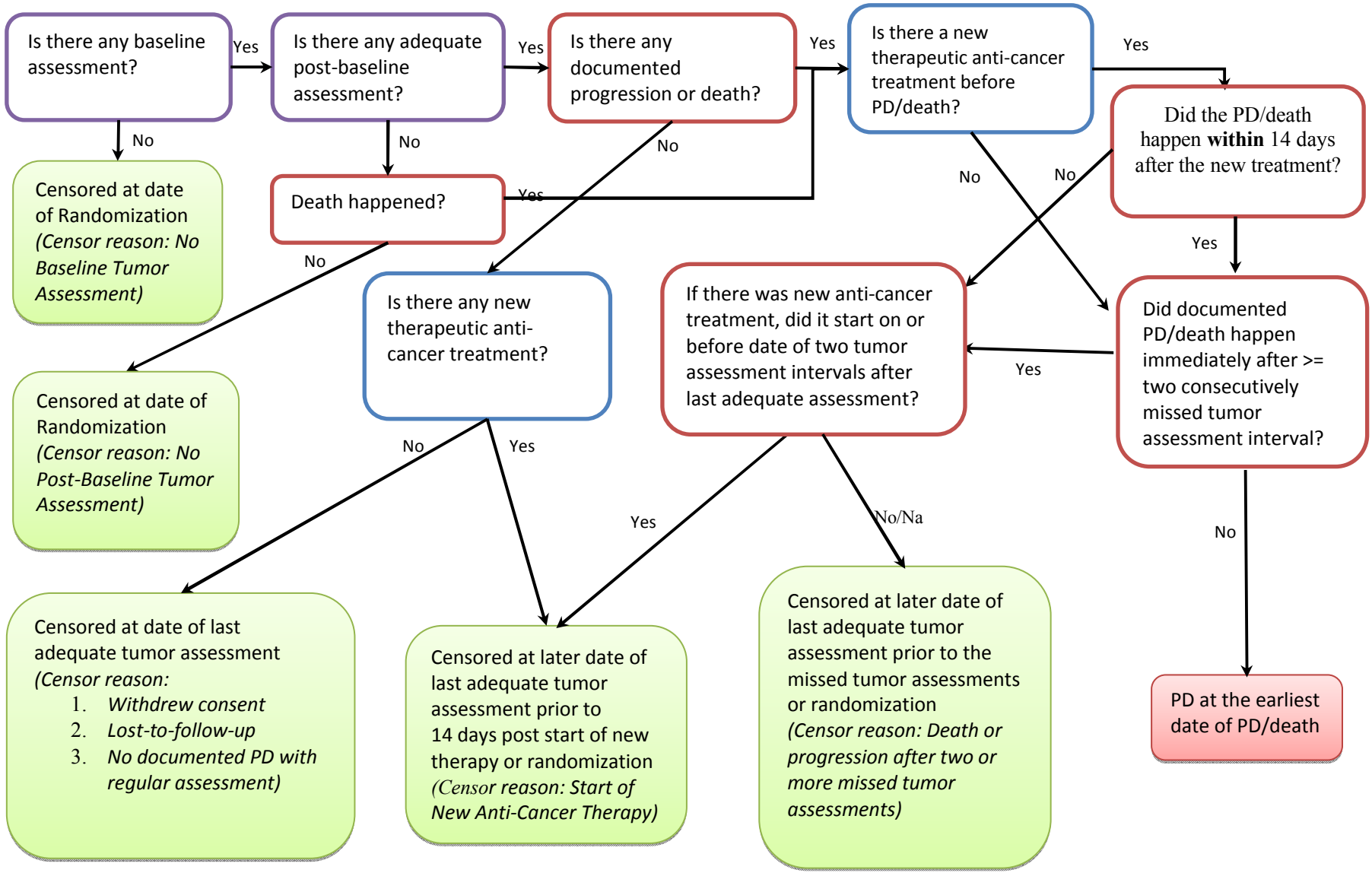
Maurer W, Bretz F. Multiple testing in group sequential trials using graphical approaches. *Stat Biopharmaceut Res*. 2013;5(4):311-320.

Osoba D et al. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. 1998 Jan;16(1):139-44.

Rowbotham, Michael C. What is a ‘clinically meaningful’ reduction in pain? *Pain* . 94(2):131-132, November 2001.

8. Appendices

Appendix 1. Flow Chart of PFS Censoring Rules



Abbreviation: PD = progressive disease.

Appendix 2. List of Preferred Terms for AESIs

AESI	Preferred Term
Cardiac Dysfunction	Acute left ventricular failure
	Acute pulmonary oedema
	Acute right ventricular failure
	Cardiac asthma
	Cardiac failure
	Cardiac failure acute
	Cardiac failure chronic
	Cardiac failure congestive
	Cardiac failure high output
	Cardiogenic shock
	Cardiopulmonary failure
	Cardiorenal syndrome
	Chronic left ventricular failure
	Chronic right ventricular failure
	Cor pulmonale
	Cor pulmonale acute
	Cor pulmonale chronic
	Ejection fraction decreased
	Hepatic congestion
	Hepatojugular reflux
	Left ventricular failure
	Low cardiac output syndrome
	Neonatal cardiac failure

AESI	Preferred Term
	Obstructive shock
	Pulmonary oedema
	Pulmonary oedema neonatal
	Right ventricular failure
	Ventricular failure
	Artificial heart implant
	Atrial natriuretic peptide abnormal
	Atrial natriuretic peptide increased
	Brain natriuretic peptide abnormal
	Brain natriuretic peptide increased
	Cardiac cirrhosis
	Cardiac index decreased
	Cardiac output decreased
	Cardiac resynchronisation therapy
	Cardiac ventriculogram abnormal
	Cardiac ventriculogram left abnormal
	Cardiac ventriculogram right abnormal
	Cardiomegaly
	Cardio-respiratory distress
	Cardiothoracic ratio increased
	Central venous pressure increased
	Diastolic dysfunction
	Dilatation ventricular
	Dyspnoea paroxysmal nocturnal
	Heart transplant
	Hepatic vein dilatation

AESI	Preferred Term
	Jugular vein distension
	Left ventricular dysfunction
	Myocardial depression
	Nocturnal dyspnoea
	N-terminal prohormone brain natriuretic peptide abnormal
	N-terminal prohormone brain natriuretic peptide increased
	Oedema
	Oedema due to cardiac disease
	Oedema neonatal
	Oedema peripheral
	Orthopnoea
	Peripheral oedema neonatal
	Pulmonary congestion
	Right ventricular dysfunction
	Scan myocardial perfusion abnormal
	Stroke volume decreased
	Systolic dysfunction
	Venous pressure increased
	Venous pressure jugular abnormal
	Venous pressure jugular increased
	Ventricular assist device insertion
	Ventricular dysfunction
	Ventricular dyssynchrony
Cardiac Arrhythmias	Chronotropic incompetence
	Electrocardiogram repolarisation abnormality
	Electrocardiogram RR interval prolonged

AESI	Preferred Term
	Electrocardiogram U-wave abnormality
	Sudden cardiac death
	Bradycardia
	Cardiac arrest
	Cardiac death
	Cardiac telemetry abnormal
	Cardio-respiratory arrest
	Electrocardiogram abnormal
	Electrocardiogram ambulatory abnormal
	Electrocardiogram change
	Heart rate abnormal
	Heart rate decreased
	Heart rate increased
	Loss of consciousness
	Palpitations
	Rebound tachycardia
	Sudden death
	Syncope
	Tachycardia
	Tachycardia paroxysmal
	Bradyarrhythmia
	Ventricular asystole
	Accessory cardiac pathway
	Adams-Stokes syndrome
	Agonal rhythm
	Atrial conduction time prolongation

AESI	Preferred Term
	Atrioventricular block
	Atrioventricular block complete
	Atrioventricular block first degree
	Atrioventricular block second degree
	Atrioventricular conduction time shortened
	Atrioventricular dissociation
	Bifascicular block
	Brugada syndrome
	Bundle branch block
	Bundle branch block bilateral
	Bundle branch block left
	Bundle branch block right
	Conduction disorder
	Defect conduction intraventricular
	Electrocardiogram delta waves abnormal
	Electrocardiogram PQ interval prolonged
	Electrocardiogram PQ interval shortened
	Electrocardiogram PR prolongation
	Electrocardiogram PR shortened
	Electrocardiogram QRS complex prolonged
	Electrocardiogram QT prolonged
	Electrocardiogram repolarisation abnormality
	Lenegre's disease
	Long QT syndrome
	Sinoatrial block
	Trifascicular block

AESI	Preferred Term
	Ventricular dyssynchrony
	Wolff-Parkinson-White syndrome
	Nodal arrhythmia
	Nodal rhythm
	Sick sinus syndrome
	Sinus arrest
	Sinus arrhythmia
	Sinus bradycardia
	Wandering pacemaker
	Arrhythmia
	Heart alternation
	Heart rate irregular
	Pacemaker generated arrhythmia
	Pacemaker syndrome
	Paroxysmal arrhythmia
	Pulseless electrical activity
	Reperfusion arrhythmia
	Withdrawal arrhythmia
	Arrhythmia supraventricular
	Atrial fibrillation
	Atrial flutter
	Atrial parasystole
	Atrial tachycardia
	Junctional ectopic tachycardia
	Sinus tachycardia
	Supraventricular extrasystoles

AESI	Preferred Term
	Supraventricular tachyarrhythmia
	Supraventricular tachycardia
	ECG P wave inverted
	Electrocardiogram P wave abnormal
	Retrograde p-waves
	Anomalous atrioventricular excitation
	Cardiac flutter
	Extrasystoles
	Tachyarrhythmia
	Accelerated idioventricular rhythm
	Cardiac fibrillation
	Parasystole
	Rhythm idioventricular
	Torsade de pointes
	Ventricular arrhythmia
	Ventricular extrasystoles
	Ventricular fibrillation
	Ventricular flutter
	Ventricular parasystole
	Ventricular pre-excitation
	Ventricular tachyarrhythmia
	Ventricular tachycardia
Infusion-related Reactions (48 core PTs)	Allergic oedema
	Anaphylactic reaction
	Anaphylactic shock
	Anaphylactoid reaction

AESI	Preferred Term
	Anaphylactoid shock
	Angioedema
	Ciculatory collapse
	Circumoral oedema
	Conjunctival oedema
	Corneal oedema
	Cytokine release syndrome
	Distributive shock
	Drug hypersensitivity
	Epiglottic oedema
	Eye oedema
	Eye swelling
	Eyelid oedema
	Face oedema
	First use syndrome
	Gingival oedema
	Gingival swelling
	Gleich's syndrome
	Hypersensitivity
	Idiopathic urticaria
	Infusion related reaction
	Kounis syndrome
	Laryngeal oedema
	Laryngotracheal oedema
	Limbal swelling
	Lip oedema

AESI	Preferred Term
	Lip swelling
	Oculorespiratory syndrome
	Oedema mouth
	Oropharyngeal swelling
	Palatal oedema
	Periobital oedema
	Pharyngeal oedema
	Scleral oedema
	Shock
	Swelling face
	Swollen tongue
	Tongue oedema
	Tracheal oedema
	Type 1 hypersensitivity
	Urticaria
	Urticaria cholinergic
	Urticaria chronic
	Urticaria papular
Infusion-related Reactions (9 additional PTs)	Abdominal pain
	Abdominal pain upper
	Abdominal pain lower
	Back pain
	Chills
	Dyspnoea
	Flushing
	Hypotension

	Pyrexia

Appendix 3. List of Preferred Terms for Consolidated AEs

Consolidated AE	Preferred Term
ABDOMINAL PAIN	Abdominal pain
	Abdominal pain lower
	Abdominal pain upper
ANAEMIA	Anaemia
	Haemoglobin decreased
	Red blood cell count decrease
FATIGUE	Asthenia
	Fatigue
HYPERBILIRUBINAEMIA	Blood bilirubin increased
	Hyperbilirubinaemia
HYPERTENSION	Hypertension
	Blood pressure increased
HYPOALBUMINAEMIA	Blood albumin decreased
	Hypoalbuminaemia
HYPOKALAEMIA	Blood potassium decreased
	Hypokalaemia
HYPOMAGNESAEMIA	Blood magnesium decreased
	Hypomagnesaemia
	Magnesium deficiency
HYPONATRAEMIA	Blood sodium decreased
	Hyponatraemia
HYPOPROTEINEMIA	Hypoproteinemia
	Protein total decreased

Consolidated AE	Preferred Term
INTESTINAL OBSTRUCTION	Gastrointestinal obstruction
	Intestinal obstruction
	Small intestinal obstruction
LEUKOCYTOSIS	Leukocytosis
	White blood cell count increased
LEUKOPENIA	Leukopenia
	White blood cell count decreased
LYMPHOPENIA	Lymphocyte count decreased
	Lymphopenia
MUCOSITIS	Aphthous stomatitis
	Mucosal inflammation
	Oropharyngeal pain
	Stomatitis
MUSCULOSKELETAL PAIN	Arthralgia
	Back Pain
	Bone Pain
	Flank Pain
	Groin Pain
	Muscle Spasms
	Musculoskeletal Chest Pain
	Musculoskeletal Pain
	Myalgia
	Neck Pain
	Pain In Extremity
NEUROPATHY	Hypoaesthesia
	Neuropathy peripheral

	Paraesthesia
	Peripheral sensory neuropathy
NEUTROPENIA	Neutropenia
	Neutrophil count decreased
RASH	Dermatitis
	Dermatitis acneiform
	Dermatitis allergic
	Dermatitis bullous
	Rash
	Rash follicular
	Rash generalised
Consolidated AE	**Preferred Term**
	Rash macular
	Rash papular
	Rash pruritic
	Rash pustular
THROMBOCYTOPENIA	Platelet count decreased
	Thrombocytopenia