Statistical Analysis Plan I5B-MC-JGDL (Version 3)

A Phase 1b (Open-Label) / Phase 2 (Randomized, Double-Blinded) Study Evaluating Gemcitabine and Docetaxel With or Without Olaratumab in the Treatment of Advanced Soft Tissue Sarcoma

NCT02659020

Approval Date: 04-Sep-2019

1. Statistical Analysis Plan for Clinical Study: I5B-MC-JGDL: A Phase 1b (Open-Label) / Phase 2 (Randomized, Double-Blind) Study Evaluating Gemcitabine and Docetaxel With or Without Olaratumab in the Treatment of Advanced Soft Tissue Sarcoma

Olaratumab (LY3012207) Soft Tissue Sarcoma

I5B-MC-JGDL is a Phase 1b/2 study in the treatment of patients with advanced or metastatic soft tissue sarcoma. The study consists of a Phase 1b part and Phase 2 part. The Phase 1b part is a single-arm, open-label, dose-escalation study to determine the recommended dose of olaratumab for the Phase 2 part. The Phase 2 part is a randomized, double-blinded, placebo-controlled study to evaluate efficacy and safety of olaratumab with gemcitabine plus docetaxel compared with placebo with gemcitabine plus docetaxel.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I5B-MC-JGDL
Phase 1b/2

Statistical Analysis Plan Version 1 approved by Lilly: 03 February 2016
Statistical Analysis Plan Version 2 approved by Lilly on 04-Oct-2017
Statistical Analysis Plan Version 3 approved by Lilly on 04-Sept-2019

Approval Date: 04-Sept-2019 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Appendix **Page**

3. Revision History

The main objective of SAP version 2.0 is to incorporate the changes specified in Protocol Amendment (c). These changes have been incorporated into the data analysis as appropriate. In addition to the substantive changes in content listed below, minor editorial changes were made throughout the document, including formatting and stylistic alterations. SAP Version 3.0 added interim analysis hazard ratio cutoffs to be reported to Lilly.

Section #	Type of Change	Description of Change
4	Content	Add the secondary objectives for olaratumab pretreated patients according to the protocol amendment (c).
5.1	Content	Add the sample size description for olaratumab pretreated groups according to the protocol amendment (c).
5.2	Content	Clarify the statistical significance level at the final analysis for the primary endpoint, considering the added interim efficacy analysis in the protocol amendment (c); Also clarify that all the analyses are intended to be performed on olaratumab naïve and olaratumab pre-treated patients separately, unless specified otherwise.
5.2.1.3	Content	Details for PRO analysis variables are updated to be consistent with other olaratumab studies.
5.7.1	Content	LMS subset is removed from the table for analysis population to align with the protocol defined objectives.
5.11.2.3	Content	Add potential pooled efficacy analyses combing olaratumab-naïve and olaratumab pre-treated cohorts according to the protocol amendment (c).
5.13	Content	Details for PRO analysis are updated to be consistent with other olaratumab studies.
5.14.2	Content	Leukopenia has been added in the list of consolidated AEs, separated from neutropenia consolidated terms according to the current version of consolidated AEs.
5.14.4	Content	New analysis for abnormal laboratory toxicity is added.

Section #	Type of Change	Description of Change
5.17	Content	Efficacy interim analysis is added according to the protocol amendment (c).
6	Content	Details on unblinding process for interim analyses have been added.

4. Study Objectives

4.1. Primary Objective

Phase 1b

The primary objective of the Phase 1b part is to determine a recommended Phase 2 dose of olaratumab that may be safely administered in combination with gemcitabine and docetaxel to patients with locally advanced or metastatic Soft Tissue Sarcoma (STS).

Phase 2

The primary objective of the Phase 2 part is to compare the overall survival (OS) in olaratumab-naïve patients with locally advanced or metastatic STS treated with olaratumab plus gemcitabine and docetaxel versus placebo plus gemcitabine and docetaxel.

4.2. Secondary Objectives

Phase 1b

The secondary objectives of Phase 1b part are the following:

- characterize the safety and toxicity profile of olaratumab in combination with gemcitabine and docetaxel
- evaluate the pharmacokinetics (PK) and immunogenicity of olaratumab in combination with gemcitabine and docetaxel
- evaluate the PK of gemcitabine and docetaxel in combination with olaratumab
- document any antitumor activity of gemcitabine and docetaxel in combination with olaratumab

Phase 2

A secondary objective of the Phase 2 part is to compare OS in olaratumab-pretreated patients with locally advanced or metastatic STS treated with olaratumab plus gemcitabine and docetaxel versus placebo plus gemcitabine and docetaxel.

The secondary objectives of the Phase 2 part are to compare olaratumab plus gemcitabine and docetaxel versus placebo plus gemcitabine and docetaxel in both olaratumab-naïve and olaratumab pre-treated groups for the following:

- 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 (HRQoL), and health status
- safety and tolerability
- evaluate the PK and immunogenicity of olaratumab

4.3. Exploratory Objectives

Phase 1b and Phase 2

- to explore biomarkers associated with clinical outcome and/or pathogenesis of STS
- to explore the exposure-response relationship of olaratumab for efficacy and/or safety

Phase 2 Only

- to evaluate change in tumor size from baseline to best overall response
- assessment of the association between clinical variables, such as histological subtypes, and clinical outcomes

5. A Priori Statistical Methods

5.1. Sample Size

Phase 1b

The primary objective of the Phase 1b part is to determine a recommended Phase 2 dose of olaratumab that may be safely administered in combination with gemcitabine (900 mg/m^2) and docetaxel (75 mg/ m^2) to patients with advanced or metastatic STS. A total of approximately 45 patients will ensure that at least 15 patients will be treated at the 15-mg/kg and 30 patients at the 20-mg/kg cohort. The number of at least 15 patients at each cohort provides sufficient data for both safety and PK analyses, and support the Phase 2 dose selection decision.

Phase 2

Based on the outcome of the Phase 1b part, the dose selected for the Phase 2 part of Study JGDL will be 15 mg/kg, 20 mg/kg, or 2 loading doses of 20 mg/kg during Cycle 1 followed by 15 mg/kg in every subsequent cycle. The dosing regimen for olaratumab will remain Days 1 and 8 of a 21-day cycle regardless of the dosing strategy adopted.

The primary objective of the Phase 2 part is to compare olaratumab plus gemcitabine and docetaxel (experimental arm) versus placebo plus gemcitabine/docetaxel (control arm) in terms of OS in patients with advanced or metastatic STS, who have not previously been treated with olaratumab (the "olaratumab-naïve" cohort).

The Phase 2 part of the study will screen approximately 200 olaratumab-naïve patients to randomize 166 patients in 1:1 randomization (83 patients in the experimental arm and 83 patients in the control arm). The primary intent-to-treat (ITT) sample size of 166 was selected assuming the final analysis of OS will occur when at least 108 OS events in randomized olaratumab-naïve patients have been observed (35% censoring).

The final total of 108 OS events (deaths) in olaratumab-naïve patients provides 80% statistical power for a two-sided log-rank test at a 0.20 significance level (assuming the true OS hazard ratio [HR] is 0.665). An OS HR of 0.665 corresponds approximately to an increased median survival from 15 months (estimated from published clinical data in various types of patients with advanced or metastatic STS) in placebo plus gemcitabine and docetaxel to 22.5 months for olaratumab plus gemcitabine and docetaxel.

A key secondary objective of the study will be to compare OS between the experimental arm and the control arm in patients with locally advanced or metastatic STS, who have previously been treated with olaratumab (the "olaratumab-pretreated" cohort). The Phase 2 part of the study will screen approximately 114 olaratumab-pretreated patients to randomize 90 olaratumab-pretreated patients in 1:1 randomization. The sample size of 90 patients in the secondary cohort of "olaratumab-pretreated" patients was selected based on both statistical and qualitative considerations.

5.2. General Considerations

This document describes the statistical analyses planned prior to the first patient visit of the Phase 1 part. 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.

There will be an interim efficacy analysis planned for the Phase 2 part of the study. All available data on patient characteristics, efficacy, and safety outcomes will be included for consideration as part of the interim efficacy analysis. The primary efficacy hypothesis will not be tested at the interim efficacy analysis; this analysis is being conducted (i) to allow the iDMC an opportunity to review safety in the context of any observed efficacy or lack of efficacy; and (ii) to provide Eli Lilly and Company with information that might be useful for future clinical development and business planning. Interim efficacy information will not be used to alter the operating characteristics or conduct of the current study, unless the iDMC makes specific recommendations to change the study because of concern for patient safety.

At the final analysis, in the event that the primary analysis of OS is statistically significant at the two-sided 0.20 alpha level, OS will also be compared to the more stringent two-sided 0.05 level. In the event that OS is statistically significant at a two-sided 0.05 alpha level, then PFS will also be formally tested at a two-sided 0.05 level. In the event that PFS is statistically significant at a two-sided 0.05 alpha level, then ORR will also be formally tested at a two-sided 0.05 level. This statistical "gate-keeping" among OS, PFS, and ORR ensures that an overall 0.05 alpha level is maintained, in the event that one or more endpoints are statistically significant at the 0.05 level. 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).

For Phase 2 part, unless specifically described otherwise, all baseline, efficacy, safety, and health outcomes analyses will be performed separately for the olaratumab-naïve and olaratumab-pretreated cohorts. The analyses to combine the two cohorts may be conducted as exploratory analyses if deemed appropriate. 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.
- **Post discontinuation Follow-Up:** begins the day after the patient and the investigator agree that the patient will no longer continue study treatment.

Short-term follow-up begins 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).

- Follow-up for progression: Patients that discontinue study treatment for reasons other than progression will be followed every 6 weeks (±7 days) until PD.
- Follow-up for survival: Patients will be followed every 2 months (±7 days) for the first 2 years, then every 6 months (±14 days) until the patient's death or overall study completion.

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 CRF.
- **Baseline measurement**
  - **Phase 1b:** the last non-missing measurement prior to first study dose.
  - **Phase 2:** 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
- **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

Definitions of efficacy analysis variables are listed below.

Overall survival (OS) is defined for each patient as the time from the date of first study dose (Phase 1b) or randomization (Phase 2) 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 first study dose (Phase 1b part) or randomization (Phase 2 part) to the first date of radiologic disease progression (as defined by Response Evaluation Criteria In Solid Tumors, Version 1.1 [RECIST v.1.1]) or death due to any cause. Table JGDL.5.1 defines the rules of censoring to be applied to PFS. During Phase 2, sensitivity analyses of PFS will be performed using different rules for censoring (as defined by Table JGDL.5.2).

Objective response rate (ORR) is defined as the proportion of safety population (Phase 1b) or randomized population (Phase 2) 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 the proportion of safety population (Phase 1b) or randomized population (Phase 2) 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.

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. The censoring rules for DoR will be the same as the censoring rules of PFS.

Duration of Disease Control (DDC) is defined for each patient with a best response of CR, PR, or SD as the time from the date of first study dose (Phase 1b) or randomization (Phase 2) to the first date of radiologic disease progression or death due to any cause. The censoring rules for the DDC will be the same as the censoring rules of PFS.

Table JGDL.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 first study dose (Phase 1b) /randomization (Phase 2; whichever is later)
unless		
No baseline radiological tumor assessment available	Censored	Date of first study dose (Phase 1b) or randomization (Phase 2)
No adequate postbaseline radiological tumor assessment available and death reported after 2 scan intervals following first study dose (Phase 1b) or randomization (Phase 2)	Censored	Date of first study dose (Phase 1b) or randomization (Phase 2)
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 first study dose (Phase 1b) /randomization (Phase 2; whichever is later)
Tumor progression or death documented immediately after 2 or more consecutive missing scan intervals following last adequate radiological tumor assessment or first study dose (Phase 1b) /randomization (Phase 2; whichever is later)	Censored	Date of last adequate radiological assessment prior to the missing assessment or date of first study dose (Phase 1b) /randomization (Phase 2 ; whichever is later)

Abbreviations: CR = complete response; PD = progressive disease; PFS = progression-free survival; 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 JGDL.5.2. Censoring Rules for PFS Sensitivity Analysis Definitions (Phase 2)

Sensitivity Analysis (SA) 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 occurred	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.

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. If the patient is alive at the cutoff date for the analysis and a disease progression on next-line treatment has not been observed, PFS2 will be censored on the last date the patient was known to be alive. 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).

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 JGDL.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 JGDL.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). New-metastases-free survival 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 JGDL.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 JGDL.5.1, with the exception that censoring will be applied at the date of death (if no previous event or censoring).

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)

Notes:

Categories of AESIs 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 Clinical Study Report (CSR).

- **Consolidated AEs** are composite AE terms consisting of synonymous Medical Dictionary of Regulatory Activities (MedDRA) preferred terms (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 between first dose of study treatment and 30 days after the last dose of study treatment and related SAEs reported beyond 30 days after the last dose of study treatment, where the last dose stands for actual dose, that is, 0 dose is not counted as the last dose.

Exposure-related variables are listed below:

- **Dose exposures:** As reported in the electronic case report form (eCRF)
- **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:** As reported in the eCRF

5.2.1.3. Patient-Reported Outcome Analysis Variables

Phase 2 Only

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	

Time to first worsening will be calculated independently for each scale as the time from randomization to the first observation of worsening. For the symptom scales, worsening is defined as an increase of at least 10 points from baseline. For the functional scales and the global health status/QoL scales, worsening is defined as a decrease of at least 10 points from baseline. If worsening is observed after a missing value, it will be assumed that the worsening occurred at the time of the missing value. Otherwise, the patient will be considered lost to follow-up and censored at the date of last adequate assessment if no worsening has occurred. Patients with no post-baseline assessments will be censored at the date of randomization. Patients who have a score at baseline such that worsening cannot occur 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 post-baseline 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 post-baseline change scores.
- 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. **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 randomization to the first date of either a "worst pain" score increase of ≥2 points from baseline or 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 9 or more will not be included in the time to first worsening analysis.

A clinical pain response will be defined as a ≥2 points reduction from pretreatment on the mBPI-sf worst pain severity score without any increase in pain medication or a decrease in analgesic drug class of ≥1 level without increase in worst pain score. Patients with a baseline worst pain score of 0 or 1 will not be included in the analysis of this endpoint. Patients with a clinical pain response and no prior worsening will be categorized as having improved during the study. The cumulative distribution of the percentage of patients who have improved 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 post-baseline 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 post-baseline change scores.

EQ-5D-5L

The EQ-5D-5L responses may be incorporated into a cost-utility analysis. 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 specific geographies.

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.

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.

5.3. Adjustments for Covariates

Phase 2 Only

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 three randomization stratification factors listed below, given each strata having sufficient number of patients. Sensitivity analyses may also be performed that include additional stratification factor(s), or are non-stratified, or are covariate adjusted (for example, Cox models with covariates).

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

Prospectively planned sensitivity, subgroup, and multivariate analyses are described in more detailed in Section 5.11.3.

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

- liver metastases (presence at baseline versus absence at baseline)
- lung metastases (presence at baseline versus absence at baseline)
- sex (females versus males)

- age
- weight
- duration of disease since diagnosis
- grade at diagnosis (1 versus 2 versus 3)
- albumin level
- alanine aminotransferase (ALT)
- bone metastases (presence at baseline versus absence at baseline)
- prior palliative radiation therapy (0 versus ≥1)
- 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 after the year of the first study treatment.

  - The date of the first study treatment, if the onset year is the same as the year of the first study treatment.
  - The date of informed consent, if the onset year is before the year of the first 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. For additional therapies, if the start date is missing then the therapy will be assumed concomitant.

Patient data listings will show partial dates without applying the above imputation rules.

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 of the month is missing, the 1st of the month will be used to replace the missing day.
- If only the month is missing, January will be used to replace the missing information.

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 1st 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 5.2.1.3 Patient-Reported Outcome Analyses 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 (any 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

The Phase 2 part 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

For the Phase 2 part, 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), randomized in Phase 2 part (that is, ITT population), treated in Phase 1 part or randomized and treated in Phase 2 part (that is, safety population)
- The primary reasons for discontinuation from study treatment and patients still. receiving treatment will be summarized by study treatment arms 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

Table JGDL.5.3 and Table JGDL.5.4 lists analysis population definitions and associated data type for analysis.

Table JGDL.5.3. Analysis Populations in Phase 1b

Population	Definition	Analysis Type / Variable	Note
Safety Population (SP)	All entered patients who received any quantity of study drug	Baseline characteristics, concomitant medication, efficacy analyses, safety analyses, e.g. dosing/exposure, AE and resource utilization.	For efficacy analysis, patients will be grouped according to the assigned dose level cohort. For safety analysis, patients will be grouped according to the actual dose level received.
DLT-evaluable Population	Patients who complete Cycle 1 or discontinue due to a DLT prior to completing Cycle 1 treatment.	DLT assessment/AE	Patients will be grouped according to the actual dose level received.

Abbreviations: AE = adverse event; DLT = Dose-Limiting Toxicity.

Table JGDL.5.4. Analysis Populations in Phase 2

Population	Definition	Analysis Type / Variable	Note
Intention-to-treat* (ITT) Population of Olaratumab-Naïve Patients	All randomized patients among have no prior olaratumab exposure	Baseline characteristics, concomitant medication, all efficacy analyses of the olaratumab-naïve cohort	Patients will be grouped according to randomized study treatment in the olaratumab-naïve cohort.
Intention-to-treat (ITT) Population of Olaratumab Pre-treated Patients	All randomized patients among have prior olaratumab exposure	Baseline characteristics, concomitant medication, all efficacy analyses of the olaratumab pre-treated cohort	Patients will be grouped according to randomized study treatment in the olaratumab pre-treated cohort
Safety Population (SP)	All patients who received any quantity of study drug	Safety, e.g. dosing/exposure, AE and resource utilization	Patients will be grouped according to actual study treatment received.

Abbreviations: AE = adverse event;
* The ITT population of olaratumab-naïve patients is the primary analysis population.

In the Phase 2 part, 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 post-baseline (either during study treatment period or 30-day post-discontinuation 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 entered (ie. signed informed consent) will appear on this listing.

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 versus ≥65), gender, ECOG performance status, country, race (White, Black, Asian, All Other), ethnicity, height (cm), weight (kg), and BSA (m2)
- potential prognostic factors in the Phase 2 part as listed in Section 5.3
- baseline disease characteristics:
  - at initial diagnosis only: disease stage
  - at study entry only: 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/medical history (no versus at least 1 diagnosis) by 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.
- For the Phase 2 potion, comparison between the CRF 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 the study treatment period or the 30-day postdiscontinuation follow-up period will be summarized by numbers and percentages by dose cohort in Phase 1b and treatment group in Phase 2, presented in decreasing frequency of the World Health Organisation drug term:

- all concomitant medications
- premedication for study drug

The proportions of patients reporting use of concomitant medications will be compared between the treatment groups in Phase 2 part. Patient listing of all concomitant therapies and premedications will be provided for both Phase 1b and Phase 2 parts.

5.10. Treatment Compliance

The number of dose omissions, reductions, and delays, the number of cycles received, and dose intensity will be summarized for all safety populations.

5.11. Efficacy Analyses

5.11.1. Primary Efficacy Analyses

Phase 1b

The following analyses will be performed to document antitumor activity of gemcitabine and docetaxel in combination with olaratumab at each dose cohort.

- Overall survival and PFS curves, the median with 95% CI for each dose cohort will be estimated using Kaplan-Meier method (Kaplan and Meier 1958)
- Objective response rate and DCR with 95% exact CI for each dose cohort will be summarized.
- Patient listings of tumor assessments (target and non-target lesion assessments and tumor response), OS, and PFS will be provided.

Phase 2

Analysis of the primary efficacy endpoint for OS will be based on the stratified log rank test, comparing OS in the in the olaratumab-naïve cohort of ITT population between assigned study treatment arms, stratified by the 3 randomization strata based on the eCRF values; that is, number of prior systemic therapies for advanced or metastatic disease (0 versus ≥1), histological tumor type (LMS versus non-LMS), and ECOG PS (0 versus 1). To avoid potential over-stratification problem, the other stratification factor used in the randomization; that is, prior pelvic radiation (yes versus no), will not be included in the final primary efficacy analysis due to the small sample size. An unstratified log-rank test will also be performed as sensitivity analysis.

The following analyses of OS will also be performed on ITT population of the olaratumab-naïve cohort:

- 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 12, 18 and 24 months. Patients who did not have the event at the corresponding time point will be considered right-censored observations.

- Hazard ratio 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.

5.11.2. Secondary Efficacy Analyses

5.11.2.1. Supportive Analyses of Primary Efficacy Endpoint

The following supportive analyses of OS will be performed:

- Hazard ratio for treatment effect will be estimated using an unstratified Cox PH model.
- Hazard ratio for treatment effect will be estimated using a multivariate Cox PH model, stratified by the aforementioned 3 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 censoring OS at the time of starting post-discontinuation anticancer treatment.

5.11.2.2. Key Secondary Efficacy Analyses

A secondary objective of the Phase 2 part is to compare OS in olaratumab-pretreated patients with locally advanced or metastatic STS treated with olaratumab plus gemcitabine and docetaxel versus placebo plus gemcitabine and docetaxel. The same analyses used for the analyses of the primary analysis for oalratumab-naïve cohort will be performed.

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 JGDL.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 gatekeepered analysis so as not to inflate the overall type I error rate).

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 CIs (CI: 95%) and compared using the

Cochran-Mantel-Haenszel test adjusting for the aforementioned three randomization factors. A sensitivity analysis on ORR will be performed to consider only confirmed response.

5.11.2.3. Analyses of Other Secondary Efficacy and Patient-Reported Outcome 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 CIs (CI: 95%) and compared using the Cochran-Mantel-Haenszel test adjusting for the aforementioned three stratification factors.

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

In the event that efficacy is observed to be very similar between the olaratumab-naïve and olaratumab-pretreated cohorts, it may be reasonable to conclude a uniform efficacy across cohorts; in that case, additional efficacy analyses may be performed pooling these 2 populations, in order to obtain pooled estimates of efficacy parameters.

5.11.3. 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 based on CRF data collection:

- number of prior systemic therapies for advanced/metastatic disease (0 versus ≥1)
- histological tumor type (LMS versus non-LMS)
- Eastern Cooperative Oncology Group performance status (0 versus 1)
- prior pelvic radiation (yes versus no)
- liver metastases (presence at baseline versus absence at baseline)
- sex (females versus males)
- age (<65 years versus ≥65 years)
- weight (above and below median)
- duration of disease since diagnosis (above and below median)
- grade at diagnosis (1 versus 2 versus 3)
- albumin level (above and below median albumin level)
- alanine aminotransferase (above and below median)

- bone metastases (presence at baseline versus absence at baseline)
- prior palliative radiation therapy (0 versus ≥1)
- duration of most recent prior systemic therapy (above and below median)
- hemoglobin (above and below median)
- platelets (above and below 350 x 10^9/L)
- leukocytes (above and below 10 x 10^9/L)

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 Gemcitabine and Docetaxel. 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
- clear cell sarcoma of soft tissue
- malignant solitary fibrous tumor
- perivascular epithelioid cell tumor (PEComa) NOS, malignant
- plexiform fibrohistiocytic tumour
- giant cell tumour of soft tissue
- extraskeletal myxoid chondrosarcoma
- dedifferentiated liposarcoma
- myxoid/round cell liposarcoma

- Synovial Sarcoma
- Malignant Peripheral Nerve Sheath Stumor (MPNST)
- Rhabdomyosarcoma
- Leiomyosarcoma (especially uterine LMS)
- Undifferentiated Pleomorphic Sarcoma (UPS)
- Pleomorphic Liposarcoma

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 in 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 each of the 15 scales, the median time to first worsening will be analyzed using Kaplan-Meier method and compared between study treatment arms using the stratified log-rank test (analogous to the primary analysis). For a given patient, if worsening is not observed in any assessment after randomization, time to first worsening will be censored at the date of the last assessment provided by the patient, or censored at randomization date in the event the patient has no post-randomization assessments.

We anticipate the possibility that the time to first worsening analyses described above may be affected by transient effects or lack adequate statistical power, due to noise that may be

introduced because of toxicity on both study arms during approximately the first 3 months of chemotherapy. For this reason, we will also analyze the time to first worsening (post 3 months), defined for each patient as the duration starting from 3 months after randomization until the first reported worsening from baseline thereafter. For a given patient, if worsening is not observed in any assessment dated more than 3 months after randomization, time to first worsening (post 3 months) will be censored at the date of the last assessment more than 3 months after randomization, or otherwise censored at exactly 3 months after randomization if there are no further assessments.

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 based on safety population and separately for olaratumab-naïve and olaratumab pre-treated cohorts. 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 are defined below.

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 dose − date of first dose] + 21) ÷ 7
- Cumulative dose (mg) = Sum of all doses
- Cumulative dose level (mg/kg) = Sum of (dose administered at each infusion [mg] ÷ Last available weight [kg] prior to that infusion)
- Weekly dose intensity (mg/kg/week) = (Cumulative dose level[mg/kg]) ÷ (Duration of Treatment [week])
- Planned weekly dose intensity (mg/kg/week) = planned dose per infusion (mg/kg) / weeks per infusion cycle = 2 x 15 mg/kg / 3 weeks = 10 mg/kg/week
- Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 100

Gemcitabine treatment:

- Duration of treatment (weeks) = ([Date of last dose − date of first dose] + 21) ÷ 7
- Cumulative dose (mg) = Sum of all doses
- Cumulative dose level (mg/m2) = Sum of (dose administered at each infusion [mg] ÷ Last available body surface area [BSA] [m2] prior to that infusion)
- BSA (m2) = 0.007184 * weight (kg) ^0.425 * height (cm) ^0.725
- Weekly dose intensity (mg/kg/week) = (Cumulative dose level[mg/ m2]) ÷ (Duration of treatment [week])
- Planned weekly dose intensity (mg/m2/week) = 2 x 900 mg/m2/ 3 weeks = 600 mg/m2/week
- Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 100

Docetaxel treatment:

- Duration of treatment (weeks) = ([Date of last dose − date of first dose] + 21) ÷ 7
- Cumulative dose (mg) = Sum of all doses
- Cumulative dose level (mg/m2) = Sum of (dose administered at each infusion [mg] ÷ Last available body surface area [BSA] [m2] prior to that infusion)

- BSA (m2) = 0.007184 * weight (kg) ^0.425 * height (cm) ^0.725
- Weekly dose intensity (mg/kg/week) = (Cumulative dose level[mg/ m2]) ÷ (Duration of treatment [week])
- Planned weekly dose intensity (mg/m2/week) = 75 mg/m2/ 3 weeks = 25 mg/m2/week
- Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 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.

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 Terminology Criteria 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

Adverse events of special interest (AESIs)

Adverse events of special interest 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.

Adverse events of special interest for olaratumab

- infusion-related reactions (IRR)

Consolidated AEs include Abdominal Pain, Anemia, Fatigue, Hyperbilirubinaemia, Hypertension, Hypoalbuminaemia, Hypokalaemia, Hypomagnesaemia, Hyponatraemia, Hypoproteinemia, Intestinal Obstruction, Leukocytosis, Leukopenia, Lymphopenia, Mucositis, Musculoskeletal pain, Neuropathy, neutropenia, 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 the consolidated AEs at the time of analysis will be used.

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 system organ class (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, 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 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:

- Treatment-emergent AEs
- treatment-emergent SAEs
- Common Terminology Criteria for Adverse Events Grade ≥3 TEAEs
- >5% by experimental arm
- Treatment-emergent AEs leading to death (on treatment and within 30 days of last dose of study drug)
- Treatment-emergent AEs leading to discontinuation of olaratumab, chemotherapy, or any study drug

- Treatment-emergent AEs leading to dose modification of any study drug, olaratumab, or chemotherapy
- Treatment-emergent AE 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 worst CTCAE grade and PT
- summary of TEAEs by PT and decreasing frequency
- summary of TEAEs by consolidated category and PT

A patient listing of all AEs will be provided.

In the Phase 1b part, DLTs will be summarized by cohort in the patients who are evaluable for DLT assessments and listed by patient.

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

A listing of SAEs will be produced.

Adverse events of special interest (AESIs)

The following AESI analyses will be provided:

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

- listing of treatment-emergent AESIs

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 TEAEs 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

Selected safety analyses as listed below may be performed as needed for subgroups gender (male versus female), age group (<65 years versus ≥65 years), race (White versus non-White), histology (LMS versus non-LMS), and geographic region (North America versus /Europe/Rest of the world [ROW]).

- 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
- summary of treatment-emergent AESIs by AESI group and PT

Additional age subgroup analyses may be performed on the following variables:

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

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. The summary of abnormal laboratory toxicity shift from baseline to worst post-baseline CTCAE Grade will also be provided.

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, for randomized and treated patients, the total number of days in hospital and admissions will be summarized. **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 experienced 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.

The immunogenicity analyses will be conducted on all immunogenicity evaluable patients (evaluable population) within the defined safety population. The frequency and percentage (incidence) of evaluable patients with positive, negative, or missing ADA (Anti-drug Antibody) to olaratumab at baseline, and with positive, negative, or inconclusive ADA at post-baseline will be summarized by dose cohort or treatment group. Patients who are TE (treatment-emergent)-ADA positive (persistent positive or transient positive), TE-ADA persistent positive, and TE-ADA transient positive will also be summarized. Positive neutralizing ADA, negative neutralizing ADA and inconclusive neutralizing ADA will also be reported for patients with TE-ADA positive. The detailed immunogenicity definitions and terms are included in the Program SAP for LY3012207 Version 2, Section 5.5.

The following analyses will also be performed based on evaluable or safety population:

- Listing of TEAEs for evaluable patients with at least one sample of ADA positive antibody to olaratumab (evaluable population)
- Listing of antibody to olaratumab and drug concentration data for evaluable patients who have at least one ADA positive sample (evaluable population)
- Listing of anti-olaratumab antibody for patients with IRR (safety population)
- Listing of antibody to olaratumab and drug concentration data (safety population)
- Listing of olaratumab-treated patients with either at least one sample of ADA positive antibody to olaratumab or IRR or both (safety population)
- Summary of ADA antibody to olaratumab and IRR for olaratumab- treated patients (safety population)

Additional efficacy or safety analyses may be performed for patients or the subgroup(s) 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

Translational research analyses will be performed according to a separate analysis plan.

5.17. Interim Analysis

Phase 1b

Upon completion of the dose escalations, an interim safety review will be conducted prior to proceeding to Phase 2 including safety, PK, and pharmacodynamics. All relevant data including data beyond Cycle 1 will be reviewed to determine the recommended Phase 2 dose of olaratumab. The decision to proceed to Phase 2 will be made following discussions between the investigators and Lilly clinical research personnel.

Phase 2

An independent data monitoring committee (iDMC) will be established prior to first patient visit of Phase 2 part to conduct safety and efficacy interim reviews. The membership, roles, and responsibilities of the iDMC are defined in the iDMC Charter (that is, a separate iDMC charter document).

There will be no prespecified rules for stopping or modifying the trial due to efficacy or safety. The iDMC members will review unblinded interim safety data to determine whether there are sufficient safety concerns to justify modifying the study or the termination of study treatment and/or enrollment.

Only the iDMC is authorized to evaluate unblinded safety analyses. Study sites will receive information about interim results ONLY if they need to know for the safety of their patients.

The first iDMC meeting to review interim safety data will occur when approximately 60 patients (approximately 30 patients from each arm) have received at least 2 cycles of treatment or discontinued all study treatment due to any reasons prior to Day 8 in Cycle 2. PK data may be provided to the iDMC upon request.

Subsequent iDMC meetings to review accumulating safety data will occur approximately every 6 months thereafter until approximately 1 year after completing enrollment. Enrollment and treatment will continue during the iDMC safety assessments. In the event a safety signal is detected, additional meetings may occur as needed. Details as to the process and communication plan will be provided in the iDMC Charter.

5.17.1. Interim Efficacy Analysis

One interim efficacy analysis, performed by the iDMC, is planned for the Phase 2 part of the study. All available phase 2 data on patient characteristics, efficacy, and safety outcomes will be included for consideration as part of the interim efficacy analysis. The primary efficacy hypothesis will not be tested at the interim efficacy analysis; this analysis is being conducted for two reasons: (i) to allow the iDMC an opportunity to consider patient safety in the context of any observed efficacy or lack of efficacy; and (ii) to provide Eli Lilly and Company with information that might be useful for future clinical development and business planning. Interim efficacy information will not be used to alter the operating characteristics or conduct of the current study, unless the iDMC makes specific recommendations to change the study because of concern for patient safety.

Per protocol, the interim efficacy analysis will occur after observing approximately 40 OS events among the olaratumab-pretreated cohort of the ITT population. Due to the larger size of the olaratumab- naïve cohort, it is expected that there will be at least another 40 OS events in the olaratumab-naïve cohort at the time of this analysis.

Interim efficacy analyses will be conducted using the intent-to-treat principle, such that all randomized patients will be included in efficacy analyses. For OS and PFS, Kaplan Meier and Cox regression analyses will be performed, including by-treatment-arm medians, HRs, and associated 95% confidence intervals for both the olaratumab-naïve cohort and the olaratumab

pre-treated cohort (analyzed separately). A summary of best tumor response will also be provided.

Immediately following this interim efficacy analysis, The DMC should communicate to Lilly (as specified in the DMC Charter) whether or not the overall survival hazard ratio (OS HR) in all randomized olaratumab-naïve patients met either of the following criteria:

- OS HR $\leq$ 0.55
- OS HR $\geq$1.25

Lilly may use such information for future business development planning, but the information is not intended to be used to alter the current ongoing trial.

5.17.2. Safety Interim Analysis

The following analyses will be performed for each safety interim analysis in Phase 1b and Phase 2, with analysis population as specified in Table JGDL.5.3 and Table JGDL.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
- Treatment-emergent AEs summarized by worst CTCAE grade, SOC and PT*
- Common Terminology Criteria for Adverse Events v 4.0 Grade 3 and 4 AEs*
- Serious AEs summarized by SOC and PT*
- Adverse events of special interests and Consolidated AEs by PT*
- Reasons for deaths
- Treatment-emergent AEs leading to study treatment discontinuations summarized by PT
- Treatment-emergent AEs leading to study treatment dose modification summarized by PT

- listing of SAEs
- listing of preexisting conditions and AEs
- Laboratory Toxicity Shift from Baseline to Worst Grade On-study Based on CTC Grade
- Listing of Dose Limiting Toxicities (Phase 1b)

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

Additional analysis including subgroup analysis may be conducted as needed.

5.18. Clinical Trial Registry Analyses

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

6. Unblinding Plan

This unblinding plan refers to the process to be followed for the final OS analyses in the Phase 2 part.

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 final 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.

The interim analyses will be performed by the independent statistical analysis center. iDMC will review unblinded interim safety data to determine whether there are sufficient safety concerns to justify modifying the study or the termination of study treatment and/or enrollment, and also conduct the unblinded initial efficacy review. iDMC chair will notify the sponsor Senior Management Designee(SMD) the recommendation based on the review. If there is an evidence of interim efficacy, a minimum number of sponsor internal personnel will conduct further review of the data as results of this interim efficacy analysis may provide information to help inform the initiation of new studies.

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.

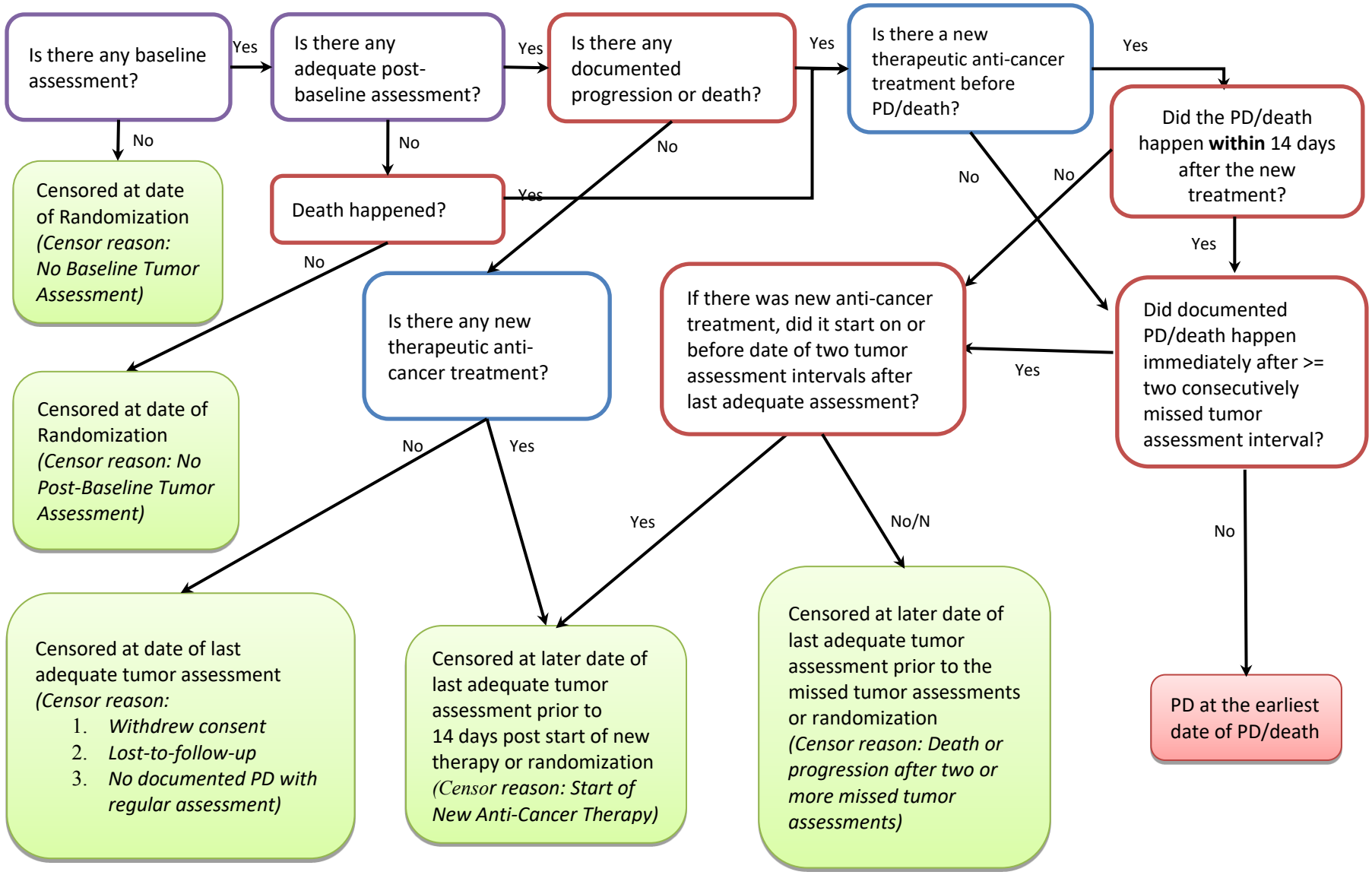
7. References

Fleming TR. Interpretation of subgroup analyses in clinical trials. *Drug Inf J*. 1995;29:1681S-1687S.

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

8. Appendices

Appendix 1. Flow Chart of PFS Censoring Rules



Abbreviation: Na = not applicable; PD = progressive disease.

Appendix 2. List of Preferred Terms for AESIs

AESI	Preferred Term
Infusion-related Reactions	Allergic oedema
	Anaphylactic reaction
	Anaphylactic shock
	Anaphylactoid reaction
	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
	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
	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
LYMPHOPENIA	Lymphocyte count decreased
	Lymphopenia
MUCOSITIS	Aphthous stomatitis
	Mucosal inflammation
	Oropharyngeal pain
	Stomatitis
NEUROPATHY	Hypoaesthesia
	Neuropathy peripheral
	Paraesthesia
	Peripheral sensory neuropathy
NEUTROPENIA	Neutropenia
	Neutrophil count decreased
LEUKOPENIA	Leukopenia
	White blood cell 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
THROMBOCYTOPENIA	Platelet count decreased
	Thrombocytopenia
MUSCULOSKELETAL PAIN	Arthralgia
	Back Pain
	Bone Pain
	Flank Pain
	Groin Pain
	Muscle Spasms
	Musculoskeletal Chest Pain
	Musculoskeletal Pain
	Myalgia
	Neck Pain
	Pain In Extremity