Statistical Analysis Plan: I5B-MC-JGDP (version 3)

A Phase 1b (Open-Label) / Phase 2 (Randomized, Double-Blinded) Study Evaluating Nab-Paclitaxel and Gemcitabine With or Without Olaratumab in the Treatment of First-Line Metastatic Pancreatic Cancer

NCT03086369

Approval Date: 03-Apr-2020

Statistical Analysis Plan: I5B-MC-JGDP:

Protocol I5B-MC-JGDP A Phase 1b (Open-Label) / Phase 2 (Randomized, Double-Blinded) Study Evaluating Nab-Paclitaxel and Gemcitabine With or Without Olaratumab in the Treatment of First-Line Metastatic Pancreatic Cancer

Olaratumab (LY3012207) Pancreatic Cancer

Study JGDP is an open-label, multicenter, nonrandomized Phase Ib study of olaratumab in combination with nab-paclitaxel and gemcitabine followed by a randomized, double-blinded, placebo-controlled, Phase 2 study of olaratumab plus nab-paclitaxel and gemcitabine versus placebo plus nab-paclitaxel and gemcitabine in patients with metastatic (Stage IV) unresectable pancreatic cancer, who have not received prior treatment for metastatic disease.

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol I5B-MC-JGDP
Phase 1b/2

Approval Date: 03-Apr-2020 GMT

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

1. Table of Contents

Section **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

2. Revision History

Statistical Analysis Version	Approval Date	Notes
1	26APR2017	First version.
2	19SEP2019.	1. Administrative changes 2. Clarification of AESI/IRRs 3. Removal of text related to Advisory Committee (AC). 4. Addition of text related to the iDMC
3	See approval stamp	1. Addition of algorithmic identification of IRRs (algorithmic versus investigator reported).

3. Study Objectives

3.1. Primary Objective

Objectives	Endpoints
Phase 1b: • to determine a recommended Phase 2 dose of olaratumab in combination with nab-paclitaxel and gemcitabine **Phase 2**: • to compare the efficacy of olaratumab plus nab-paclitaxel and gemcitabine with placebo plus gemcitabine and nab-paclitaxel	**Phase 1b** • DLTs • Safety (including but not limited to) TEAEs, SAEs, and clinical laboratory abnormalities **Phase 2** • OS

3.2. Secondary Objectives

Objectives	Endpoints
Phase 1b • to characterize the safety and toxicity profile of olaratumab plus nab-paclitaxel and gemcitabine • to evaluate the PK and immunogenicity of olaratumab plus nab-paclitaxel and gemcitabine • to document the antitumor activity observed with olaratumab plus nab-paclitaxel and gemcitabine	**Phase 1b** • Safety monitoring, including TEAEs, SAE, and deaths • Minimum serum/plasma concentration of olaratumab plus nab-paclitaxel and gemcitabine • OS, PFS, DoR, ORR
Phase 2 • to assess time-to-event variables • to document the antitumor activity observed with olaratumab plus nab-paclitaxel and gemcitabine	**Phase 2** • PFS, DoR • ORR
• to assess the following PROs: pain, HRQoL, and health status	• mBPI-sf, EORTC-QLQ-C30, and EQ-5D-5L
• to determine safety and tolerability of olaratumab in combination with nab-paclitaxel and gemcitabine	• TEAEs, AESIs, SAEs, clinical laboratory tests, vital signs, physical examinations, hospitalizations, and death
• to assess the PK and immunogenicity of olaratumab plus nab-paclitaxel and gemcitabine	• Minimum serum/plasma concentration of olaratumab plus nab-paclitaxel and gemcitabine

3.3. Tertiary Objectives

Objectives	Endpoints
• to characterize tumor tissue and blood biomarkers relevant to study drugs including but not limited to immune cells/immune and tumor microenvironment functioning, mechanism of action of study drugs, cancer-related pathways, and disease state • to assess the relationship between biomarkers and clinical outcomes	• Markers of immune function, tumor microenvironment, drug targets, signaling pathways, and disease status

Abbreviations: AESI = adverse event of special interest; DLT = dose-limiting toxicity; DoR = duration of response; EORTC = European Organisation for Research and Treatment of Cancer HRQoL = Health-related Quality of Life; mBPI-sf = modified brief pain index-short form; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; PK = pharmacokinetic; PRO = patient-reported outcomes; SAE = serious adverse events; TEAE = treatment-emergent adverse events.

4. Study Design

Study JGDP is Phase 1b/2 trial. The Phase 1b portion of the trial will consist of an open-label, dose-finding study and the Phase 2 portion will be a multicenter, randomized, double-blind, parallel, placebo-controlled study in patients with metastatic unresectable pancreatic cancer (Stage IV), who have not received prior treatment for metastatic disease.

4.1. Summary of Study Design

4.1.1. Phase 1b

The Phase 1b will commence prior to Phase 2. Study cycles will be 28 days in duration.

In the Phase 1b portion of the study, a 3 + 3 dose escalation design will be used to evaluate the safety and tolerability of olaratumab in combination with gemcitabine and nab-paclitaxel in patients with pancreatic cancer and determine a recommended Phase 2 dose of olaratumab.

Three patients will be treated initially at each dose level. If no DLTs occur in a cohort of 3 patients, a new cohort of 3 patients will be treated at the next higher dose level. If 1 of 3 patients at any dose level experiences a DLT, that cohort will be expanded to 6 patients.

- If no further patient experiences a DLT, the dose escalation can proceed.
    - The dose escalation can proceed if fewer than 2 out of 6 evaluable patients experience a DLT.
- If a DLT is observed in ≥2 out of a maximum of 6 patients at any given dose, dose escalation will cease and the next lower dose will be the MTD.

During the dose escalation, no more than one cohort will be open for enrollment at any given time.

In the first cohort, olaratumab will be dosed on Days 1, 8, and 15 (that is, on the same days as nab-paclitaxel and gemcitabine dosing as described in their USPI and SPC) at a starting dose of 15 mg/kg. If the 15 mg/kg dose (Cohort 1) of olaratumab is tolerated, enrollment will commence in Cohort 2 at the 20 mg/kg dose of olaratumab dosed on Days 1, 8, and 15. If the 20 mg/kg dose is tolerated approximately 9-12 additional patients will be enrolled in the cohort expansion to confirm the safety and tolerability of the dose. Planning for Phase 2 of the study will commence assuming safety and tolerability has been confirmed. See figure JGDP.1a for details.

If the Cohort 1 olaratumab dose of 15 mg/kg on Days 1, 8, and 15 is not tolerated, dosing of olaratumab at 20 mg/kg on Days 1 and 15 may be evaluated (Cohort 3). If the 20 mg/kg dose of olaratumab is tolerated, enrollment will commence with planned dose escalation to 25 mg/kg olaratumab on Days 1 and 15 in the subsequent dose cohort (Cohort 4). The Day 1 and Day 15 schedule will be evaluated to determine if decreasing the frequency of olaratumab dosing improves the tolerability of olaratumab in combination with nab-paclitaxel and gemcitabine (see Figure JGDP.1a).

After the MTD has been identified in the dose-escalation phase in cohort 1, 2, 3 or 4, approximately 9 patients will be enrolled in a cohort expansion. The purpose of the cohort expansion is to confirm the safety of the MTD in approximately 12 to 15 patients (including those in the dose escalation cohort) prior to proceeding to the Phase 2 portion of the study. Continuous evaluation of toxicity in the cohort expansion will be performed throughout enrollment in the cohort. If the rate of DLT-like events exceeds 33% in the first cycle, the findings will be discussed, and further enrollment may be interrupted. If the expansion cohort is discontinued because of toxicity, a new cohort may be initiated at a previously tested lower dose level.

After determination of the MTD in the dose escalation portion of the study and confirmation of the safety and tolerability of that dose in the cohort expansion, a review of safety data, including the number and type of DLTs, other safety information, and relevant PK, will be conducted prior to proceeding to Phase 2. The decision to proceed to Phase 2 and selection of the recommended Phase 2 dose will be made following discussions between the investigators and Lilly clinical research personnel.

While the MTD was not attained in Phase I, based on the available data, olaratumab treatment plan for phase 2 portion of the study was 20mg/kg for cycle 1 and 15mg/kg for cycle 2 onward.

4.1.2. Phase 2

The Phase 2 portion of the study is a 1:1 randomized, double-blinded, 2-arm study of olaratumab at a dose and schedule determined from Phase 1b (20mg/kg for cycle 1 and 15mg/kg for cycle 2 onward) in combination with nab-paclitaxel and gemcitabine (ExperimentalArm) versus placebo plus nab-paclitaxel-gemcitabine (Control Arm). Patients will receive over 30 minutes, an infusion of nab-paclitaxel (125 mg/m^2) followed by a gemcitabine (1000 mg/m^2) infusion over 30 minutes on Days 1, 8, and 15 of each 28-day cycle. Nab-paclitaxel and gemcitabine will be administered following olaratumab/placebo administration.

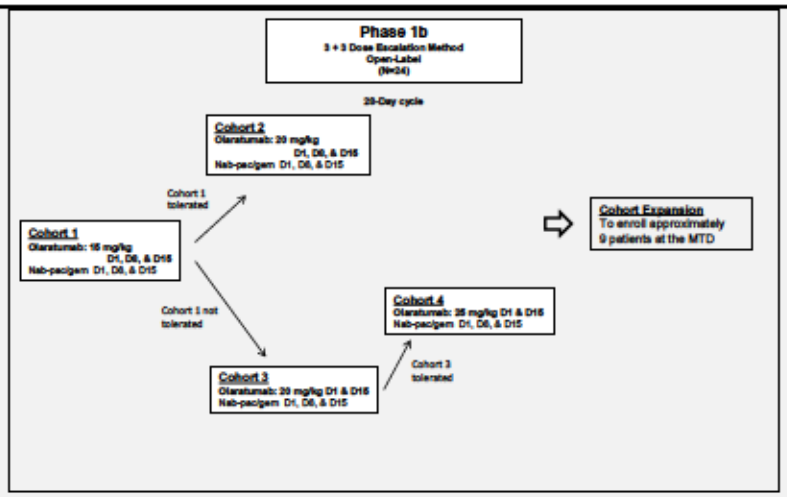
Patients will receive the study treatment until disease progression or a criterion for discontinuation is met. The primary end point is OS; secondary end points are PFS, duration of response (DoR), objective response rate (ORR), patient-reported outcomes (PROs), and safety.

Patients will be assessed for tumor response every 8 weeks. Patients without disease progression may continue to receive treatment until the development of unacceptable toxicity, death, or other discontinuation criteria are met.

Figure JGDP.b illustrates the study design of Phase 2.

After study completion, patients on study treatment who continue to experience clinical benefit and no undue risks, in the opinion of the investigator, may continue to receive study treatment (Continued Access period) until one of the criteria for discontinuation is met, as described in Protocol Section 7.8.1 (Continued Access). A continued access follow-up visit will occur 30 days (±7 days) after discontinuation.

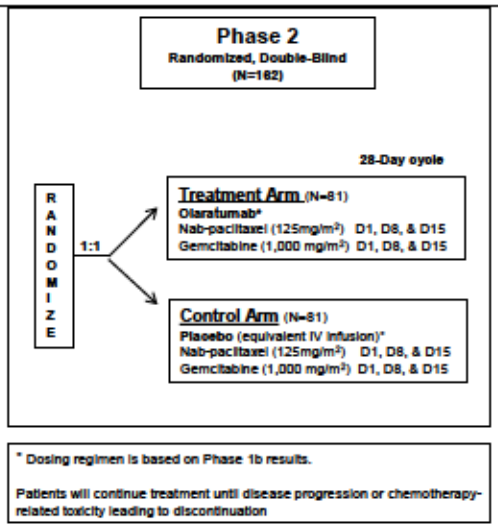
Figures JGDP.1 Illustration of study design.



Abbreviations: D = Day.

Note: Approximately 24 patients will be enrolled into the Phase 1b. A 3 + 3 dose-escalation design will be used. It is planned that between 3 to 6 patients will be enrolled at each dose level. After determination of the MTD, a cohort expansion will further evaluate the MTD in approximatey 9 additional patients.

Figure JGDP.1a. Illustration of study design of Phase 1b.



Abbreviations: D = Day; IV = intravenous; N = number of patients.

Note: Dose and schedule of olaratumab will be determined based upon tolerability in the Phase 1b portion of the study.

Figure JGDP.1b. Illustration of study design of Phase 2.

4.2. Determination of Sample Size

Planned enrollment for each phase is as follows:

Phase 1b: Approximately 15 to 27 patients with metastatic pancreatic cancer who have not received chemotherapy for metastatic disease will be evaluated for dose-limiting toxicities (DLTs). Patients who are not evaluable for DLTs will be replaced.

Phase 2: Approximately 162 patients with metastatic pancreatic cancer who have not received chemotherapy for metastatic disease will be randomized in a 1:1 ratio to receive either olaratumab plus nab-paclitaxel and gemcitabine (Treatment Arm) or placebo plus nab-paclitaxel and gemcitabine (Control Arm).

The Phase 1b portion of the study will sequentially evaluate two alternative dosing schedules using a "3 + 3" dose-escalation design plus an expansion cohort after MTD is attained. Based upon the design, the Phase 1b portion may enroll up to approximately 24 patients to determine the dose and regimen for olaratumab in combination with nab-paclitaxel and gemcitabine recommended for use in the Phase 2 portion of the study.

The Phase 2 portion of the study will enroll approximately 162 additional patients, randomized in a 1:1 fashion between study arms. The primary analysis of OS will be performed when a minimum of 113 OS events have been observed.

An overall study-wise, two-sided alpha level of 0.20 will be applied to the primary endpoint of OS. One interim safety and efficacy analysis and one final efficacy analysis will be performed. The interim efficacy analysis will occur after 70 OS events, with OS compared between study

arms. There will be no formal hypothesis testing or alpha-spending at the interim analysis with respect to efficacy in this study. The purpose of this interim analysis is solely to inform decision-making regarding whether or not to initiate early development of a separate Phase 3 trial.

Assuming the true OS hazard ratio (HR) is 0.67, there is at least 80% statistical power to show a statistically significant difference in OS between study arms.

4.3. Method of Assignment to Treatment

The Interactive Web Response System (IWRS), which is web-based and accessible 24 hours a day will be used in both the Phase 1b and Phase 2 portions of the study. The IWRS registration consists of assigning the patient a unique study identification number. Once the patient is registered through the IWRS, he/she is considered to be enrolled in the study.

See SAP Section 6 for the unblinding plan for this study.

The following treatments will be administered to enrolled patients in this study every 4 weeks (28-day cycle).

4.3.1. Phase 1b

Investigators are required to contact Lilly clinical research personnel prior to screening a patient in Phase 1b. Once enrollment availability is confirmed, sites may proceed with screening of patient and registration into Interactive Web Response System (IWRS), which is web-based and accessible 24 hours a day. No dose escalations (that is, to the next cohort) can occur without prior discussion and agreement with the responsible Lilly CRP or clinical research scientist.

Phase 1b patients will not be eligible to participate in the Phase 2 portion of the study.

4.3.2. Phase 2

Patients who meet all criteria for enrollment will be randomly assigned via IWRS to receive either olaratumab plus nab-paclitaxel and gemcitabine (Treatment Arm) or placebo plus nab-paclitaxel and gemcitabine (Control Arm). Approximately 162 patients will be randomized in a 1:1 ratio.

Once the patient is registered through the IWRS, he/she is considered to be enrolled in the study. After randomization, patients should receive their first dose of treatment within 72 hours (3 days) whenever possible.

4.3.2.1. Stratification Factors

Randomization will be stratified into four groups, one for each combination of the following two baseline factors:

- Age group (<70 years versus ≥70 years)

- Prior adjuvant/neo-adjuvant gemcitabine use (yes versus no)

5. A Priori Statistical Methods

5.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly). Analyses will be conducted within each phase of the study (phase Ib and II). The analyses for this study will be both descriptive and inferential, except for possible exploratory analysis as deemed appropriate. Unless otherwise specified, treatment effects will be evaluated based on a two-sided significance level of 0.05. For continuous variables, summary statistics will include number of patients, mean, median, standard deviation, minimum, and maximum. Categorical endpoints will be summarized using number of patients, frequency, and percentages. Unless otherwise specified, missing data will not be imputed.

The interpretation of the study results will be the responsibility of the investigator with the Lilly Clinical Research Physician (CRP)/Clinical Research Scientist (CRS), pharmacokineticist, and statistician. The CRP or CRS and statistician will also be responsible for the appropriate conduct of an internal review for both the final study report and any study-related material to be authorized by Lilly for publication.

Exploratory analyses of the data not described below will be conducted as deemed appropriate.

5.2. 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 progression 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.
    - If month is missing, the date should be treated as completely missing.
- 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 month is missing, the date should be treated as completely missing.

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.

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 to the day.
- If month is missing, then treat as completely missing.

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.

Health Outcomes/Quality-of-Life: See SAP section 5.9.

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.

5.3. Populations for Analyses

Table JGDM.5.1. Analysis Populations

Population	Definition	Analysis Type / Variable
All Entered Patients	All patients who signed Inform Consent.	
Phase I		
DLT-evaluable Population	The DLT-evaluable population will include all enrolled patients who complete Cycle 1 or discontinue due to a DLT prior to completing Cycle 1 treatment.	
Phase II		
Intention-to-Treat (ITT) population	will include all randomized patients. The ITT analysis of efficacy data will consider allocation of patients to treatment groups as randomized and not by actual treatment received.	Baseline characteristics, efficacy, health economics.
Per Protocol (PP)	will include all randomized patients who receive at least 1 dose of study treatment and do not have any major protocol violations that could potentially affect the efficacy conclusions of the study. This population will be defined in detail in the SAP prior to database lock.	
mBPI-sf	The mBPI-sf population will include all patients who completed the baseline assessment (2 assessments within 7 days of Cycle 1 Day 1) followed by at least 1 mBPI-sf "worst pain" assessment after 1 cycle of study drug (Cycle 2 Day 1 or later).	mBPI
EORTC QLQ-C30	QLQ-C30 population will include all patients who completed the baseline assessment (Cycle 1 Day 1) followed by at least 1 QLQ-C30 assessment after 1 dose of study drug (Cycle 2 Day 1 or later).	
EQ-5D-5L	The EQ-5D-5L population will include all patients who completed the baseline assessment (Cycle 1 Day 1).	
Phase I and II		
Safety Population	All enrolled patients who receive any quantity of study treatment, regardless of their eligibility for the study, will be included in the safety analysis. Safety evaluation will be performed based on the actual initial therapy a patient has received, regardless of any other cohort or treatment arm to which he or she was assigned.	Baseline characteristics, concomitant medication, efficacy analyses, Safety, e.g. dosing/exposure, AE and resource utilization.
Biomarker Evaluable Population (BE)	will include the subset of patients from the ITT population from whom a valid assay result has been obtained. Samples for biomarker research will be collected at times specified in Protocol Appendix 4.	

Abbreviations: AE = adverse event; BE = Biomarker Evaluable; IP = Investigator Product; ITT = Intent-to-Treat; PP= Per protocol.

5.4. Patient Disposition

A detailed description of patient disposition will be provided, including a summary of the number and percentage of patients entered into the study, enrolled in the study, and treated as well as number and percentage of patients completing the study, or discontinuing by phase, site, assigned treatment and overall. A patient is considered completed when overall survival event has been reported (death). A listing of patient discontinuation will also be presented.

5.5. Demographic and Baseline Characteristics

Demographic data are collected and reported to demonstrate that the study population represents the target patient population considered for regulatory approval.

A summary of baseline patient and disease characteristics, historical diagnoses, preexisting conditions, and prior therapies will be reported using descriptive statistics.

Summary tables will include:

- patient demographics: age (years) and age group, gender, race (White, Black, Asian, All Other), ethnicity, height (cm), weight (kg), and BSA (m^2), ECOG performance status
- baseline disease characteristics:
  - at study entry only: current disease stage, grade, 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
- Medical history by MedDRA PT, presented in decreasing frequency

  Note: Subjects reporting more than 1 condition/diagnosis within a PT will be counted only once for that PT.

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

Other patient characteristics will be summarized as deemed appropriate.

5.6. Treatment Compliance

The number of cycles received, dose omissions, dose reductions, dose delays, and dose intensity will be summarized for all treated patients by randomized treatment. If any patients received treatment that they were not randomized to, an additional analysis by treatment received will be conducted.

A by-patient listing of treatment compliance data will be provided.

5.7. Concomitant Therapy

The following concomitant medications used during study treatment period or up to the 30-day postdiscontinuation follow-up period will be summarized by numbers and percentages by dose cohort (Phase Ib) and treatment arm (Phases Ib and II), presented in decreasing frequency of the World Health Organisation drug term by treatment:

- Summary of prior medications
- Premedication for study drug
- Concomitant medications
- Analgesic medications

Patient listings of all concomitant therapies, premedications and analgesics will be provided.

5.8. Efficacy Analyses

Tumor assessments will be performed according to Response Evaluation Criteria In Solid Tumors Version 1.1 (RECIST 1.1) (Eisenhauer et al., 2009) for each patient at the times shown in the Schedule of Activities (Protocol Section 2). Digital images are to be sent to a third-party organization for storage.

For patients with solid tumors, computed tomography (CT) scans, including spiral CT, are the preferred methods of measurement (CT scan thickness recommended to be ≤5 mm); however, magnetic resonance imaging (MRI) is also acceptable in certain situations and if there is a concern about radiation exposure associated with CT. Intravenous and oral contrast is required unless medically contraindicated.

The CT portion of a positron emission tomography (PET)-CT scan may be used as a method of response assessment if the site can document that the CT is of identical diagnostic quality to a diagnostic CT (with intravenous and oral contrast). A PET scan alone or as part of a PET-CT may be performed for additional analyses but cannot be used to assess response according to RECIST v.1.1 (Eisenhauer et al. 2009).

The method of tumor assessment used at baseline must be used consistently throughout the study. Radiological scan of the thorax, abdomen, and pelvis is required.

5.8.1. Primary Outcome and Methodology

5.8.1.1. Overall Survival

Overall survival (OS) is defined as the time from the date of randomization to the date of death from any cause. For each patient, prior to each data analysis, a reasonable effort will be made to obtain the most up-to-date status of the patient (date of death or last date known to be alive). If the patient is alive or lost to follow-up at the time of data analysis, OS data will be censored on the last date the patient is known to be alive. For any patient who has withdrawn consent for

further follow-up of survival data, OS will be censored at the last date for which the patient consented to be followed for the study.

Overall survival will be compared between treatment arms using a stratified log-rank test, stratified by the randomization strata. The corresponding hazard ratio between treatment arms will be estimated using a stratified Cox regression model (Cox 1972), stratified by the randomization strata. Overall survival curves, the median and survival rates at various time points with 95% CI, for each treatment arm will be estimated using the Kaplan-Meier method (Kaplan and Meier 1958). An overall study-wise, two-sided alpha level of 0.20 will be applied.

Sensitivity analyses will be conducted for overall survival where events will only include disease related deaths. Patients without the reported event will be censored at the last date known to be alive or the date of death (if death is not disease related)(OS1). In addition, a third sensitivity analysis will be carried out where all deaths are included and patients are censored when new anticancer treatment is started (OS2). Similar analyses methods used for the primary endpoint will be conducted.

5.8.2. Secondary Efficacy Analyses

5.8.2.1. Progression-Free Survival

Progression-free survival (PFS) is defined as the time from randomization until the first radiographic documentation of progression as defined by RECIST (Version 1.1) or death from any cause in the absence of progressive disease (PD). Patients who die without a reported prior progression will be considered to have progressed on the day of their death. Patients who did not progress or are lost to follow-up will be censored at the day of their last radiographic tumor assessment. If no baseline or postbaseline radiologic assessment is available, the patient will be censored at the date of randomization. If death or PD occurs after 2 or more consecutive missing radiographic visits, censoring will occur at the date of the last radiographic visit prior to the missed visits. The use of a new anticancer therapy prior to the occurrence of PD will result in censoring at the date of last radiographic assessment prior to initiation of new therapy. Additional details can be found in Table JGDM.5.1.

Progression-free survival will be compared between treatment arms using a stratified log-rank test, stratified by the randomization strata. The corresponding HR between treatment arms will be estimated using a stratified Cox regression model (Cox 1972), stratified by randomization strata. Progression-free survival curves, median PFS, and PFS rates at various time points with 95% CI for each treatment arm will be estimated using the Kaplan-Meier method (Kaplan and Meier 1958).

Table JGDM.5.2. Censoring Rule of PFS Primary Analysis

Situation	Event / Censor[d]	Date of Event or Censor[e]
Tumor progression[a] 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.
unless		
No baseline radiological tumor assessment available	Censored	Date of first study dose
No adequate[b] postbaseline radiological tumor assessment available and death reported after 2 scan intervals[c] following first study dose	Censored	Date of first study dose
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? (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 (whichever is later)	Censored	Date of last adequate radiological assessment prior to the missing assessment(s) or date of first study dose

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 SAP 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.

Sensitivity analyses for PFS will also be conducted including:

1) PFS1: In addition to tumor progression or death as events as listed in the primary PFS endpoint, discontinuation due to symptomatic PD will be considered an event. Rules of missingness will apply as indicated in the primary PFS.
2) PFS2: Progression or death events are included regardless of tumor scan missingness specified in primary PFS endpoint.
3) PFS3: Progression or death events regardless of missingness or new anticancer treatment.

5.8.2.2. Objective Response Rate

Objective response rate is defined as the number of patients who achieve a best overall response of complete response (CR) or partial response (PR) from randomization until PD/recurrence divided by the total number of patients randomized to the corresponding treatment arm (ITT population). Patients who do not have a tumor response assessment for any reason will be considered non-responders and will be included in the denominator when calculating the response rate. The objective response rate (ORR), with 95% CI, will be summarized for each

treatment arm and compared between treatment arms using the Cochran-Mantel-Haenszel test adjusted for the randomization strata.

5.8.2.3. Duration of Response

Duration of response (DOR) is defined as the time from the date that measurement criteria for CR or PR (whichever is first recorded) are first met until the first date that disease is recurrent or objective progression is observed, per RECIST 1.1 criteria, or the date of death from any cause in the absence of objectively determined disease progression or recurrence.

5.9. Health Outcomes/Quality-of-Life Analyses

Patient-reported outcomes are measured through the following:

- mBPI-sf
- EORTC QLQ-C30
- EQ-5D-5L

For each instrument (mBPI-sf, EORTC QLQ-C30, and EQ-5D-5L), percentage compliance will be calculated as the number of completed assessments divided by the number of expected assessments (that is, patients still on study and alive at expected assessment timepoint). For this purposes, an "assessment" will refer to one patient's responses for one questionnaire at a single time-point. The phrase "assessed" means that a response was recorded for at least 1 of the questionare items. Percentage compliance and reasons for non-compliance will be summarized overall as well as by treatment arm and timepoint. Data will be separately summarized by treatment and time point using descriptive statistics.

mBPI-sf

The main efficacy measure for the pain endpoint will be the time to first worsening of the mBPI-sf "worst pain" score. Time to first worsening in pain will be described using the method of Kaplan and Meier and a comparison will be made between the treatment arms using a log-rank test. In a Cox regression model, covariates included in the model will be baseline mBPI-sf score as well as stratification factors, age, prior adjuvant/neo-adjuvant gemcitabine use and weight loss within 3 months of enrollment. Patients who do not have a worsening event will have time to worsening censored at the time of the last mBPI-sf assessment. "Worsening" will be defined as either a "worst pain" increase of $\geq$2 points postbaseline or an analgesic drug class increase of $\geq$1 level (Farrar et al., 2001; Rowbotham, 2001). However, other approaches to defining clinically meaningful worsening in pain (such as a more distributional based estimate of 1/2 SD or 1 standard error of the mean (SEM) might be considered.

A clinical pain response will be defined as a 30% reduction from pretreatment on the mBPI-sf 24-hour worst pain severity score without any increase in pain medication. Patients with a baseline worst pain score of 0 (best possible score) will not be included in the analysis of this endpoint. A sustained response will be defined as a clinical pain response maintained for additional cycles without any increase in pain medication. The cumulative distribution of the percentage of pain responders by treatment arm as a function of time from enrollment is to be presented graphically.

For the mBPI-sf, 2 baseline assessments will be collected, the first one within 7 days of the first cycle (not on Cycle 1 Day 1) and the second one at Day 1 of the first cycle prior to study treatment administration. The baseline mBPI-sf score will be calculated as the average of the 2 baseline assessments. If a trend is evident, possible ad hoc analyses may be conducted.

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, such as stratification factors, age, prior adjuvant/neo-adjuvant gemcitabine use as well as weight loss within 3 months prior to study enrollment. Similar analyses will also be conducted for the mean of 7 pain interference with function items.

- Missing Data:
    - Composite scores ("pain severity" and "pain interference"): If $\geq$ 50% of the values are missing, then the composite score should be coded as 'missing'. If < 50% of the values are missing, then the average of the remaining scored items should be used per MD Anderson BPI scoring manual Chapter 2 (MD Anderson, 2009).
    - The "pain severity" composite score will require that at least 2 of the items are present and that an average of those two items can suffice to estimate average pain severity.
    - For the 'pain interference' composite score, at least 4 of the 7 items must be present to be averaged for a patient's 'average pain interference' score.

The analysis will include all cycles for which at least 25% of patients in each arm have an mBPI-sf assessment. Additional sensitivity analyses may be conducted including analyses to investigate the floor and ceiling effects of the mBPI-sf measure.

QLQ-C30

Assessments should 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 (one global health status/QoL scale, five functional scales, and nine symptom scales/items). A linear transformation is used to obtain scales ranging from 0 to 100 where:

- A high score for any of the 5 functional scales represents a high or healthy level of functioning
- A high score for the global health status/QoL represents a high level of QoL
- A high score for any of the 9 symptom scales/item represents a high level of symptoms/problems

Table JGDM.5.3. 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	

Scale	Raw Score: Mean of items	Score
Financial difficulties (FI)	Q28	

The analysis described below will include all cycles for which at least 25% of patients in each arm have an assessment. Change from baseline for each scale score will be calculated by subtracting baseline value from the value of the scale at each time point for each of the 15 scales and the information will be summarized descriptively by treatment arm.

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.

In addition, for each of the 15 scale scores, time to deterioration (TTD) with respect to baseline will be analyzed using Cox proportional hazards model with assigned treatment and baseline scores as covariates as well stratification factors age, prior adjuvant/neo-adjuvant gemcitabine use as well as weight loss within 3 months prior to study enrollment will be included. Kaplan-Meier graphs by treatment arm will be produced. Sensitivity analyses will be performed to explore the effects of prognostic factors. Furthermore, the choice of covariates may differ for each of the 15 scales and such analyses will be exploratory.

Deterioration 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 deterioration can be calculated as the time from randomization to the first observation of deterioration. If deterioration is observed after a missing value, it will be assumed that the deterioration occurred at the time of the missing value. In an additional sensitivity analysis, the patient will be considered lost to follow-up and censored at the date of last assessment. Patients with no post-baseline assessment will be censored at the date of randomization. Patients who have the worst possible score at baseline will not be included in this analysis.

Floor and ceiling effects will be summarized for each of the 15 scales. The ceiling effect is defined as the percentage of patients who have a baseline score of <10 (assuming that a meaningful change is 10 points) on the symptom scales, and > 90 on the functional scales and

global health status/QoL scale. The floor effect is defined as the percentage of patients who have a baseline score of > 90 points on the symptom scales, and < 10 points on the functional scales and global health status/QoL scale. The presence of a significant ceiling effect suggests that not much improvement is possible for that scale. Likewise the presence of a floor effect suggests that deterioration is less likely.

EQ-5D-5L

The EQ-5D-5L responses may be incorporated into a cost-utility analyses but will not be included in the clinical study report. 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.
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. An appropriate covariance structure will be assessed and utilized. A similar analysis will be performed on the VAS scores. covariates in this analyses will includ stratification variables, age and prior adjuvant/neo-adjuvant gemcitabine use as well as weight loss within 3 months of enrollment.
The analysis will include all cycles for which at least 25% of patients in each arm have an assessment. Additional ad hoc analyses may be conducted including evaluation floor and ceiling effects.

5.10. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Pharmacokinetic analyses will be conducted on patients who have received at least 1 dose of the study drug and have had sufficient postdose samples collected to allow estimation of PK parameters.

In the Phase 1b part, PK parameter estimates will be computed by standard noncompartmental methods of analysis for olaratumab. The maximum concentration (C_{max}), time of C_{max} (t_{max}), area under the concentration-time curve (AUC), half-life ($t_{1/2}$), steady-state volume of distribution (V_{ss}), clearance (CL), and other relevant parameters that can be calculated from the data will be reported from these noncompartmental analyses.

In the Phase 2 part, PK parameters for olaratumab (CL, exposure, V_{ss}, and $t_{1/2}$) and inter-individual PK variability will be computed using nonlinear mixed effect modeling implemented in NONMEM. The version of any software used for the analysis will be documented and the program will meet the Lilly requirements of software validation.

If warranted by the data, PK/pharmacodynamic analyses using OS, PFS, and/or other appropriate clinical endpoints will also be conducted to characterize the exposure-response relationship in this study.

5.11. Healthcare Resource Utilization

Hospitalizations, transfusions, and concomitant medications during study treatment (see Section 5.7 above) will be summarized descriptively by treatment arm.

Utilization data will be summarized for each category by treatment arms. The following resource utilizations will be described:

- Analgesics (on study treatment and during short term follow up)
- Transfusions (on study treatment and during short term follow up)
- Growth factors (on-study treatment and during short term follow-up)
- Surgery (on study treatment and during short term follow up)
- Hospitalizations (on study treatment and during short term follow up)
- Post discontinuation radiotherapy and systemic therapy.

For categorical variables, frequency and the corresponding proportions will be calculated and tests for differences in proportion between groups will be performed using a chi-squared test. Continuous variables will be described by the mean, median, standard deviation, minimum and maximum. A t-test will be used to compare mean utilization.

5.12. Safety Analyses

All patients who receive at least 1 dose of any study therapy will be evaluated for safety and toxicity.

The MedDRA Version 19.1 or later will be used when reporting AEs by MedDRA terms. The MedDRA Lower Level Term will be used in the treatment-emergent computation. Treatment-emergent adverse events will be summarized by System Organ Class (SOC) and by decreasing frequency of Preferred Term within SOC.

5.12.1. Phase 2

Safety analyses will include summaries of the following:

- AEs, including maximum severity and relationship to study drug
- SAEs, including relationship to study drug
- AEs leading to dose adjustments
- Adverse events of Special Interest (AESI)
- Infusion-related reactions (IRR)
- discontinuations from study treatment due to AEs or death
- treatment-emergent abnormal changes in laboratory values
- treatment-emergent abnormal changes in vital signs and ECGs

5.12.2. Extent of Exposure

Study drug exposure will be summarized based on Safety Population. The summary will include duration of treatment, number of infusions, the number of cycles received per patient, cumulative dose, cumulative dose level, weekly dose intensity, relative dose intensity, by actual treatment received. Details of study drug administration will be included in patient listings.

Dose modifications, delays and discontinuations will also be summarized.

The exposure formulas are defined below.

5.12.2.1. Olaratumab

- Duration of treatment (weeks) = ([Date of last dose − date of first dose] + 21) ÷ 7
- Cumulative dose (mg) = Sum of all doses (in mg)
- Cumulative dose (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 (mg/kg))÷(Duration of Treatment (weeks))
- Planned weekly dose intensity (mg/kg/week)
  - Cycle dose = 15, 20 or 25 mg/kg/three weeks (dependent on cohort)
- Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 100

5.12.2.2. Nab-paclitaxel

- 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/m^2) = Sum of (dose administered at each infusion [mg] ÷ Last available BMI [m^2] prior to that infusion)
- Weekly dose intensity (mg/m^2/week)=(Cumulative dose level)÷(Duration of Treatment)
- Planned weekly dose intensity (mg/m^2/week)
  - Cycle dose = 125 mg/m^2/three weeks
- Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 100

5.12.2.3. Gemcitabine

- 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/m^2) = Sum of (dose administered at each infusion [mg] ÷ Last available BMI prior to that infusion)
- Weekly dose intensity (mg/m^2/week)=(Cumulative dose level) ÷ (Duration of Treatment)
- Planned weekly dose intensity (mg/m^2/week)
- Cycle dose = 1000 mg/m^2/three weeks
- Relative dose intensity (%) = (Weekly dose intensity) ÷ (Planned weekly dose intensity) x 100

5.12.3. Adverse Events

Adverse events will be coded using the MedDRA dictionary. Severity grades will be assigned by investigators using the NCI-CTCAE Version 4.0.

Adverse event-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)**
  - Infusion-related reactions (IRRs) [olaratumab and nab-paclitaxel]

    The infusion-related reaction analysis uses a comprehensive approach to identify and analyze IRRs (immediate hypersensitivity reactions and delayed hypersensitivity reactions), using two distinct tools:

    [1] Standardized MedDRA Queries (SMQs)

    The following SMQs or preferred terms (PT) will be used:

    - Analphylactic reaction SMQ (narrow and algorithmic)
    - Hypersensitivity SMQ (narrow terms)
    - Angioedema SMQ (narrow terms)

    Hypersensitivity reactions will be classified as immediate or non-immediate reactions, depending on temporal relationship to the timing of olaratumab administration. Immediate reactions were treatment-emergent or worsening events that occurred on the day of the drug administration. Non-immediate reactions (or delayed hypersensitivity reactions) are events that occur after the day of drug administration but prior to the next drug administration.

    [2] Investigator Check Box

    Lilly ensures that investigators have the ability to identify any AE as a potential IRR via a check box on the eCRF. This approach will enable the identification of potential IRRs that may present with common symptoms,

such as chills, back pain or abdominal pain, that may not have been represented in the above SMQs.

AESI for nab-paclitaxel and gemcitabine:

- Pneumonitis 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 clinical study report (CSR).

- **Consolidated AEs** are composite AE terms consisting of synonymous Medical Dictionary for Regulatory Activities (MedDRA) preferred terms (PTs) to allow meaningful interpretation of the AE data. Consolidated AE categories and PTs will be maintained at compound 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 days after the last dose of study treatment, and study drug-related SAEs reported beyond 30 days after the last dose of study treatment, where last dose stands for actual dose, that is, 0 dose is not counted as last dose.

Exposure-related variables are listed below:

- **Dose exposures**: As reported in the 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

Study drug-related AEs are AEs that were considered to be at least possibly related to study drug by an investigator. See section 5.4 pertaining to missing data.

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 descending frequency of PT across treatment arms; when summarized by system organ class (SOC) and PT, AEs will be presented in descending frequency of PT within SOC across treatment arms. If more than one AE is recorded for a patient within any SOC or PT term, the patient will only be counted once on the most severe grade and the closest relationship to treatment.

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

- patients with at least one treatment-emergent adverse event (TEAE), serious adverse event (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

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.

Treatment-emergent adverse events (TEAEs) will be summarized/listed as follows:

- treatment-emergent AE by system organ class (SOC) and PT
- study drug related TEAEs
- summary of TEAEs by worst CTCAE grade
- consolidated TEAEs by consolidated category and PT
- study drug related consolidated TEAE
- summary of TEAEs by PT and descending frequency
- listing of TEAE leading to death (on treatment and within 30 days of last dose of study drug)
- listing of TEAE leading to discontinuation of olaratumab, chemotherapy, or any study drug
- listing of TEAEs leading to dose modification of any study drug, olaratumab, or chemotherapy

A patient listing of all AEs will be provided.

The following treatment-emergent 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 consolidated treatment-emergent SAEs

- summary of study drug-related consolidated treatment-emergent SAEs

A listing of SAEs will be produced.

The following death reports will be provided:

- summary of deaths (all deaths and deaths within 30 days of last dose of study drug)

The following AE of special interest (AESIs) summary 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

5.12.4. 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 postbaseline (first dose up to 30 days after the last dose of study treatment) grade will be produced.

A patient listing of all laboratory data will be provided with a flag for values outside of the laboratory normal range.

5.12.5. Vital Signs and Other Physical Findings

Actual value and change from baseline for vital sign measurements will be summarized at each assessment time point using summary statistics. Listings of vital signs data will be provided.

5.12.6. Electrocardiograms

Change of categorical values (abnormal/normal) from baseline for electrocardiograms (ECG) will be summarized at each assessment time point. Listings of ECG data will be provided.

5.13. Subgroup Analyses

Subgroup analyses will be conducted within stratification (age and prior adjuvant/neo-adjuvant gemcitabine use). See SAP section 4.3.2.1 for stratification details. Subgroup analyses of weight loss over the prior three months of greater than 10 percent (yes, no) will also be explored. Additional efficacy or safety analyses may be performed in subgroups of patients.

5.14. Important Protocol Deviations

Protocol deviations will be identified in Protocol Deviation Plan. Important protocol deviations (e.g. not meeting inclusion/exclusion criteria, noncompliance with protocol procedures, dosing errors, use of prohibited medication, continuing after meeting withdrawal criteria) will be listed by treatment group and by category of deviations.

5.15. Interim Analyses, Data Monitoring

5.15.1. Introduction

In order to minimize the operational and statistical bias that result from performing an interim analysis, the interim analyses for this study will be conducted under the auspices of an internal independent Data Monitoring Committee (iDMC). The purpose of the iDMC is to advise the Lilly study team regarding the continuing safety of study participants and the continuing validity and scientific merit of the trial.

One interim analysis is planned for the Phase 2 portion of this trial. An interim analysis that includes both safety and efficacy will be conducted. The analysis will occur after at least 70 OS events have been observed among Phase 2 patients. The interim analysis will not be used for purposes of formally testing any efficacy hypotheses, but descriptive efficacy results will be considered in order to decide whether or not a separate Phase 3 trial evaluating olaratumab in combination with nab-paclitaxel and gemcitabine should be initiated.

The interim efficacy analysis 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. 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 patients met either of the following criteria:

- OS HR ≤ 0.55
- OS HR ≥ 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.

Only the iDMC is authorized to evaluate unblinded interim efficacy and safety analyses. Study sites will receive information about interim results only if they need to know for the safety of their patients. Additional details including the flow of information will be included in a separate document.

5.16. Annual Report Analyses

The following reports are needed as requested for annual reporting purposes

5.16.1. Clinical Investigator Brochure

- Summary of SAE (patients on therapy).
- Summary of deaths reported (patients on therapy).
- Summary of Patient disposition.
- Summary of primary reason for treatment discontinuation.

- Listing and summary of treatment-emergent adverse events –by CTCAE category and term.

5.16.2. Development Safety Update Report

- Cumulative Subject Exposure by Age Group and Sex
- Cumulative Subject Exposure by Racial Group
- Estimated Cumulative Subject Exposure
- Subject Exposure by Gender
- Listing of Patients Who Discontinued Due to Adverse Event
- Listing of Deaths

5.17. Clinical Trial Registry Analyses

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

Analyses provided for the CTR requirements include the following:

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

- An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (TEAE).
- An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

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. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, 255-265.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, *34*, 187-220.

Eisenhauer, E., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., ... & Rubinstein, L. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European journal of cancer*, *45*(2), 228-247.

EuroQol (n.d.). Valuation of EQ-5D. Retrieved from http://www.euroqol.org/about-eq-5d/valuation-of-eq-5d/eq-5d-5l-value-sets.html.

Farrar, J. T., Young, J. P., LaMoreaux, L., Werth, J. L., & Poole, R. M. (2001). Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*, *94*(2), 149-158.

Fayers, P. M., Aaronson, N. K., Bjordal, K., Grønvold, M., Curran, D., & Bottomley, A. (2001). EORTC QLQ-C30 scoring manual. Retrieved from http://lillynet.global.lilly.com/sites/RealWorldAnalytics/Oncology%20PRO%20Guidance/Forms/AllItems.aspx

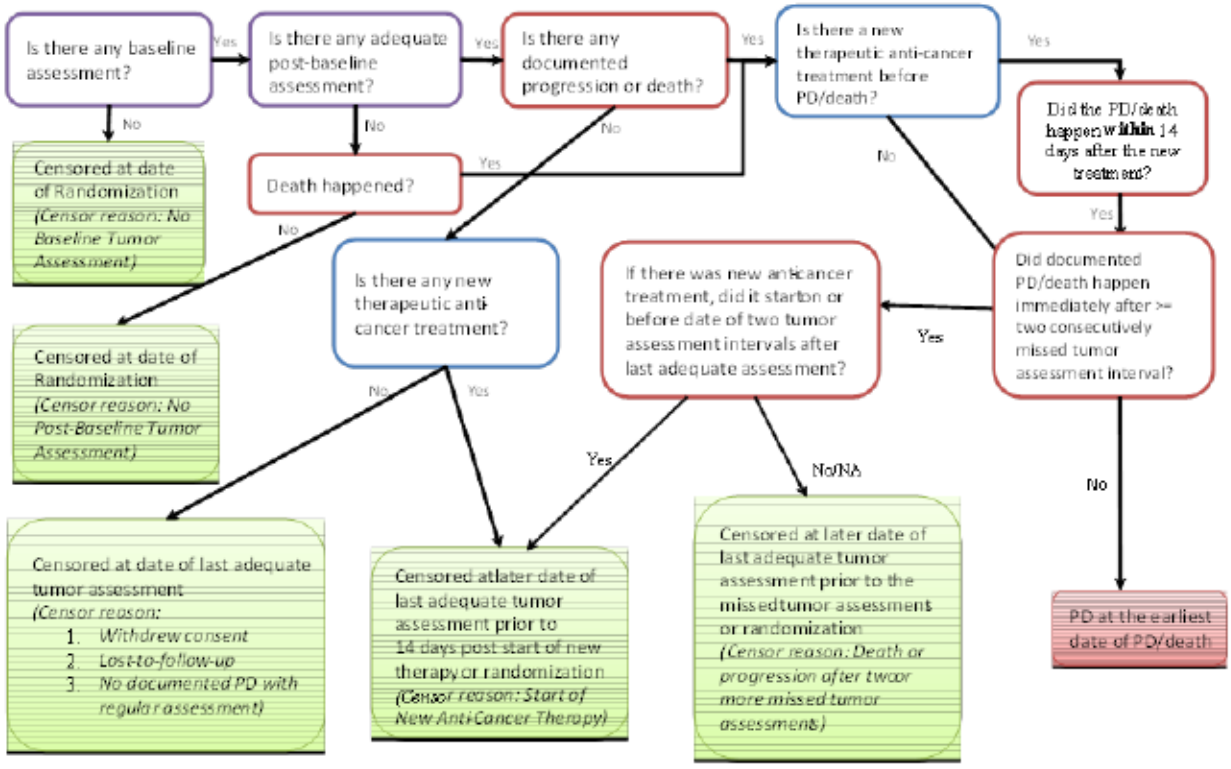
Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53*(282), 457-481.

MD Anderson (2009). The Brief Pain Inventory. Retrieved from https://www.mdanderson.org/content/dam/mdanderson/documents/Departments-and-Divisions/Symptom-Research/BPI_UserGuide.pdf

Rowbotham, M. C. (2001). What is a 'clinically meaningful'reduction in pain?. *Pain*, *94*(2), 131-132.

8. Appendices

Appendix 1. Flow Chart of PFS Censoring Rules



Abbreviation: PD = progressive disease