I3Y-MC-JPCJ Statistical Analysis Plan Version 2

An Adaptive, Open-Label, Randomized Phase 2 Study of Abemaciclib as a Monotherapy and in Combination with Other Agents Versus Choice of Standard of Care (Gemcitabine or Capecitabine) in Patients with Previously Treated Metastatic Pancreatic Ductal Adenocarcinoma

NCT02981342

Approval Date: 12-Feb-2018

1. Statistical Analysis Plan: I3Y-MC-JPCJ
An Adaptive, Open-Label, Randomized Phase 2 Study of Abemaciclib as a Monotherapy and in Combination with Other Agents Versus Choice of Standard of Care (Gemcitabine or Capecitabine) in Patients with Previously Treated Metastatic Pancreatic Ductal Adenocarcinoma

Abemaciclib (LY2835219)

This is an adaptive, open-label, randomized Phase 2 study in patients with metastatic pancreatic ductal adenocarcinoma. This study will be conducted in 2 stages. Stage 1 will randomize patients into 4 treatment arms to assess if either abemaciclib monotherapy or abemaciclib in combination with other agents provides a level of disease control that is at least comparable to the standard of care. In Stage 2, additional patients will be randomized to treatment arms advancing to Stage 2 for further evaluation of safety and efficacy.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I3Y-MC-JPCJ
Phase 2

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly: 18 November 2016
Statistical Analysis Plan Version 2 electronically signed and approved by Lilly on date provided below.

Approval Date: 12-Feb-2018 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to the first patient visit.

Statistical Analysis Plan (SAP) Version 2 was approved prior to Stage 1 analysis.

The overall changes and rationale for the changes incorporated in Version 2 are as follows:

- Provided the details on the primary and sensitivity analyses of progression-free survival (PFS).

4. Study Objectives

4.1. Primary Objective

This study will be implemented in 2 stages:

Stage 1: To evaluate disease control rate (DCR) of the abemaciclib treatment arms versus the standard-of-care arm (gemcitabine or capecitabine)

Stage 2: To evaluate PFS of the abemaciclib treatment arms versus the standard-of-care arm (gemcitabine or capecitabine)

4.2. Secondary Objectives

Stage 1:

- To evaluate objective response rate (ORR) of the abemaciclib treatment arms versus the standard-of-care arm
- Evaluate safety and tolerability of the abemaciclib treatment arms
- Pharmacokinetics (PK) of abemaciclib and its metabolites, LY3023414, and galunisertib

Stage 2:

- To evaluate DCR of the abemaciclib treatment arms versus the standard-of-care arm
- To evaluate clinical benefit rate (CBR) of the abemaciclib treatment arms versus the standard-of-care arm
- To evaluate ORR of the abemaciclib treatment arms versus the standard-of-care arm
- To evaluate duration of response (DoR) of the abemaciclib treatment arms versus the standard-of-care arm
- To evaluate overall survival (OS) of the abemaciclib treatment arms versus the standard-of-care arm
- Evaluate the kinetics of carbohydrate antigen (CA) 19-9
- Evaluate safety and tolerability
- To evaluate pain and symptom burden of the abemaciclib treatment arms by best response group (partial response [PR], stable disease [SD], or progressive disease [PD]) versus the standard-of-care arm
- PK of abemaciclib and its metabolites, LY3023414, and galunisertib
- Exposure-response for abemaciclib, LY3023414, and galunisertib

4.3. Exploratory Objectives

Stage 1 and Stage 2:

- Assess the relationship between biomarkers and clinical outcome

The biomarker-based analyses will be described in a separate Translational Research Statistical Analysis Plan.

5. A Priori Statistical Methods

5.1. Determination of Sample Size

Prior to randomization for Stage 1, approximately 6 to 12 patients will be enrolled in the safety lead-in part of the study.

During Stage 1, approximately 25 patients will be treated per arm to provide a preliminary assessment of tumor response and assessment of safety. The null hypothesis is based on the assumption that the DCR is no greater than 50%; Table JPCJ.5.1 shows the probability of stopping at the end of Stage 1 (ranges from 11% to 72%) for DCR differences ranging from -10% (experimental arm DCR worse than standard of care) to 15% (experimental arm DCR better than standard of care). Disease control rate (DCR) is defined in Section 5.8.1.1.

At the end of Stage 1, an additional 50 patients will be enrolled in each of the advancing treatment arms from Stage 1, giving a total of approximately 75 patients in each treatment arm (combined Stage 1 and 2). Analysis for Stage 2 will be performed when approximately 120 total PFS events have occurred for the combination of each individual abemaciclib-containing arm and the standard-of-care arm, or all planned patients have been enrolled in Stage 2, whichever comes later. Assuming a hazard ratio (HR) of 0.65 in PFS, the sample size yields approximately 76% statistical power to detect superiority of the abemaciclib-containing arm over the standard-of-care arm with the use of a 2-sided log-rank test at 0.10 significance level. If the true median PFS for the standard-of-care arm is 1.5 months, then the HR of 0.65 amounts to an approximately 0.8 month improvement in median PFS for the abemaciclib-containing arm under an additional assumption of exponential survival distribution.

Table JPCJ.5.1. Probability of Stopping at Stage 1

Null DCR	Alternative DCR	DCR Difference	Sample Size	Probability of abemaciclib containing treatment arm to stop at Stage 1 (that is, not advancing to Stage 2)
0.50	0.40	-0.10	25	0.72
0.50	0.50	0	25	0.44
0.50	0.65	0.15	25	0.11

Abbreviation: DCR = disease control rate.

5.2. General Considerations

5.2.1. Populations

The following populations will be defined for the study:

Entered population: will include all patients who sign the informed consent document.

Enrolled or 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. This population will be used for all baseline, efficacy, and health economics analyses.

Safety population: will include all randomized patients who received any quantity of study treatment, regardless of their eligibility for the study. The safety evaluation will be performed based on the first dose of study treatment a patient actually received, regardless of the arm to which he or she was randomized. The safety population will be used for all dosing/exposure, safety, and resource utilization analyses.

Safety Lead-in population: will include all patients from the safety lead-in portion of the study who received any quantity of study treatment, regardless of their eligibility for the study. This population will be used for all summaries for the safety lead-in.

Dose-limiting toxicity (DLT) evaluable population: will include all patients from the safety lead-in portion of the study who received at least 75% of planned doses of abemaciclib and galunisertib in Cycle 1. Exposure and safety summaries will be repeated for this population in addition to the safety lead-in population.

Pharmacokinetic population: will include all randomized patients who received at least 1 dose of study treatment and have baseline and at least 1 postbaseline evaluable PK sample.

5.2.2. Definitions and Conventions

The **baseline value of a safety assessment** is the last value observed prior to the first dose of abemaciclib, LY3023414, galunisertib, gemcitabine, or capecitabine.

The **baseline value of an efficacy assessment** is the last value observed prior to the date of randomization. If a patient's first assessment occurs after randomization but prior to the first dose, this assessment will be used as the baseline.

The **study day of a safety event or assessment** will be calculated as:

- the difference between the date of the event or assessment and the date of first dose plus 1 for all events or assessments occurring on or after the day of first dose. For example, if an event occurs on 08 June 2016 and the date of first dose was 06 June 2016, the study day of the event is 3.
- the difference between the date of the event or assessment and the date of first dose for all events or assessments occurring before the day of first dose. For example, if an event occurs on 05 June 2016 and the date of first dose was 06 June 2016, the study day of the event is -1.

The **study day of an efficacy event or assessment** will be calculated as:

- the difference between the date of the event or assessment and the date of randomization plus 1 for all events or assessments occurring on or after the date of randomization.

- the difference between the date of the event or assessment and the date of randomization for all events or assessments occurring before the date of randomization.

One month is defined as 365/12 days.

Unless otherwise noted, **summaries of continuous variables** will include a mean, median, standard deviation, minimum, and maximum. When appropriate, lower and upper quartiles will also be presented.

Unless otherwise noted, **summaries of categorical variables** will include the frequency and percentage (relative to the population being analyzed) of each category.

Statistical analysis of this study will be the responsibility of Lilly or its designee.

Unless otherwise stated, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05 and all confidence intervals (CIs) will be given at a 2-sided 95% CI.

No adjustments in p-values or confidence intervals will be made for multiple testing attributed to multiple arm comparisons and multiple endpoints.

For all summaries, data will be presented separately for: safety lead-in, Stage 1, and Stage 2.

Stage 1 analysis will be performed approximately 16 weeks after the last planned Stage 1 patient enters treatment. Enrollment may continue while Stage 1 analysis is ongoing and any additional patients enrolled during the time that Stage 1 analysis is ongoing will not be included in Stage 1 analysis but will be included in the analysis at the end of Stage 2 where data will be pooled together from same arms in Stages 1 and 2.

5.3. Handling of Dropouts or Missing Data

With the exception of dates, missing data will not be imputed. The method of imputation for any dates that are imputed is described in the relevant section.

5.4. Patient Disposition

The number and percentage of patients entered into the study, enrolled in the study, and treated, as well as reasons for discontinuation from study treatment and reasons for discontinuation from study, will be summarized by treatment arm. A listing of patient disposition will be provided.

5.5. Patient Characteristics

5.5.1. Demographics

Patient demographics will be summarized for all enrolled patients. Patient demographics will include sex, race, ethnicity, country, age, height, weight, body mass index (BMI), and body surface area (BSA).

5.5.2. Baseline Disease Characteristics

Eastern Cooperative Oncology Group (ECOG) performance status (PS), initial pathological diagnosis, basis for initial diagnosis, histological grade, time since initial diagnosis, and CA19-9 levels will be summarized for all treated patients using descriptive statistics and listed in a data listing.

5.5.3. Historical Illnesses

Historical illnesses are clinically relevant events in the past that ended before the screening visit. Historical illnesses (using Preferred Term(s) [PTs] from the most current version of the Medical Dictionary for Regulatory Activities [MedDRA™]) will be summarized.

5.5.4. Prior Therapies

Prior radiotherapy, surgery, and systemic therapy will be summarized. Prior radiotherapy will be categorized by reason (neo-adjuvant, adjuvant, locally advanced, metastatic) for the regimen and prior surgery will be categorized by intent (curative, palliative). Prior systemic therapies will be categorized by treatment intent (curative, palliative) and setting (neo-adjuvant, adjuvant, locally advanced, metastatic). Frequency of each specific therapy will be tabulated within each type of regimen and reason for regimen. Frequency of gemcitabine-based and fluoropyrimidine-based therapies will be presented.

Most recent systemic therapy, duration of that therapy, and time since last systemic therapy will be summarized. Duration of therapy will be calculated as: date of end of therapy – date of start of therapy + 1; time since last systemic therapy will be calculated as: date of randomization – date of end of therapy + 1. If only the month and year of a treatment date or progression date is available, the day will be imputed to the 15th.

5.5.5. Poststudy Treatment Discontinuation Therapies

Therapies received following study treatment discontinuation will be summarized overall and by type of therapy by treatment arm.

5.6. Treatment Compliance

Compliance will be calculated as the ratio of total dose taken to the total assigned dose (minus any dose adjustments and doses omitted/withheld for medical or logistical reasons).

5.6.1. Abemaciclib Compliance

Treatment compliance of abemaciclib will be measured by pill counts and summarized. The total assigned dose for a patient in Arm A with no adjustments or omissions is 200 mg per dose × 2 doses per day × 28 days = 11200 mg. The total assigned dose for a patient in Arm B or Arm C with no adjustments or omissions is 150 mg per dose × 2 doses per day × 28 days = 8400 mg.

5.6.2. LY3023414 (PI3K/mTOR dual inhibitor) Compliance

Treatment compliance of LY3023414 will be measured by pill counts and summarized. The total assigned dose for a patient in Arm B with no adjustments or omissions is 200 mg per dose × 2 doses per day × 28 days = 11200 mg.

5.6.3. Galunisertib (TGF-βRI inhibitor) Compliance

Treatment compliance of galunisertib will be measured by pill counts and summarized. The total assigned dose for a patient in Arm C with no adjustments or omissions is 150 mg per dose × 2 doses per day × 14 days = 4200 mg.

5.6.4. Gemcitabine Compliance

Gemcitabine will be administered intravenously at the investigational site, under the direction of the investigator. As a result, a patient's compliance with study drug administration is ensured. Patients should attend scheduled clinic visits and must comply with study criteria under their control. Deviation(s) from the prescribed dosage regimen should be recorded on the case report form (CRF).

5.6.5. Capecitabine Compliance

Treatment compliance of capecitabine will be measured by pill counts and summarized. The total assigned dose for a patient in Arm D with no adjustments or omissions is 1250 mg/m^2 per dose × 2 doses per day × 14 days = 35000 mg/m^2.

5.7. Concomitant Therapy

All medications will be coded to the generic preferred name according to the current World Health Organization drug dictionary. All concomitant medications will be summarized by number and percentage of patients for the safety population using the base name.

Number of patients who undergo surgery while on treatment will be summarized.

5.8. Efficacy Analyses

Unless otherwise noted, all efficacy analyses will be performed on the ITT population.

The stratification factor for the analysis of primary and secondary analyses is number of prior systemic therapies (1 versus 2). The stratification factor is captured in the IWRS and also derived from information collected on eCRFs. Unless otherwise specified, all stratified analyses will be based on the stratification factor per eCRFs. A cross tabulation of the frequency of each level of the stratification factor per IWRS and eCRF will be produced.

5.8.1. Primary Outcome and Methodology

Stage 1: Primary efficacy Endpoint for Stage 1 is DCR. Analysis of DCR will be performed approximately 16 weeks after the last planned Stage 1 patient enters treatment or has discontinued (whichever comes earlier). All tumor assessment data accumulated during this

period (beyond 16 weeks for some patients, if available), will be included in the assessment of DCR.

Any treatment arm(s) with a DCR difference ≥ 0 as compared to the standard of care (Arm D) will be selected to advance to Stage 2. Enrollment may continue while Stage 1 analysis is ongoing. Any patients enrolled during the Stage 1 analysis will not be included in Stage 1 analysis but will be included in the analysis at the end of Stage 2.

Stage 2: Primary efficacy endpoint for Stage 2 is PFS. The primary analysis of Stage 2 will be conducted when at least approximately 120 total PFS events have occurred for the combination of each individual abemaciclib-containing arm and the standard-of-care arm, or all planned patients have been enrolled in Stage 2, whichever comes later. Data from both Stages 1 and 2 will be pooled for this analysis.

5.8.1.1. Objective Response Rate and Disease Control Rate

Objective response rate (ORR), DCR, and CBR are summary measures of best overall response (BOR) as defined by RECIST v1.1. BOR is derived from time point responses. All time point responses observed while on study treatment and during the short-term follow-up period (but before the initiation of postdiscontinuation therapy) will be included in the derivation.

Each patient's BOR will be categorized as complete response (CR), PR, SD, PD, or not evaluable (NE). A BOR of CR or PR will not require confirmation.

Objective response rate (ORR) is the proportion of patients with a BOR of CR or PR. Disease control rate (DCR) is the proportion of patients with a BOR of CR, PR, or SD. In the case of SD, postbaseline tumor assessments must have met the SD criteria at least once after randomization at a minimum interval of 6 weeks to assign a best response of SD.

Patients with SD will be further classified as SD ≥6 months or SD <6 months. Stable disease ≥6 months includes all patients with a best response of SD and a PFS time of ≥6 months. Clinical benefit rate (CBR) is the proportion of patients with a BOR of CR or PR, or SD ≥6 months.

For each of these rates, point estimates and CIs (using the normal approximation to the binomial) will be calculated by treatment arm. The rates between each abemaciclib-containing treatment arm and the standard-of-care arm (Arm D) will be compared using a stratified Cochran-Mantel-Haenszel test.

5.8.1.2. Progression-Free Survival

The primary endpoint for Stage 2 of this study is PFS. PFS time is measured from the date of randomization to the date of investigator-determined objective progression as defined by Response Evaluation Criteria In Solid Tumors Version 1.1 (RECIST v1.1), or death from any cause whichever comes earlier. The detailed censoring rules are described in Table JPCJ.5.2.

Table JPCJ.5.2. Rules for Determining Date of Progression or Censor for Progression-Free Survival

Rule	Situation	Date of Progression or Censor	Outcome
1	No baseline tumor assessments	Date of Randomization	Censored
2	No post baseline assessments and no death	Date of Randomization	Censored
3	No documented progression and no death (with a post-baseline tumor assessment)	Date of last adequate tumor assessment	Censored
4	Patient lost to follow-up (or withdrew consent from study participation) before documented progression or death	Date of last adequate tumor assessment	Censored
5	Documented progression	Date of documented progression. If a tumor assessment was done on multiple days, use the earliest date for that visit.	Progressed
6	Death without documented progression	Date of death	Progressed
7	Documented progression or death after missing ≥2 consecutive post-baseline tumor assessments	Date of last adequate tumor assessment before missed assessments or date of randomization, whichever is later	Censored

Note: Progression-free survival and associated outcome is determined by the earliest of the dates above, if more than 1 situation applies.

Sensitivity analyses will be undertaken for calculation of the primary endpoint in order to evaluate the robustness of the analysis. The following sensitivity analyses will be performed for PFS:

Progression-Free Survival Sensitivity Analysis 1 (censoring for receiving subsequent systemic anticancer therapy): if a patient is initiated on another anticancer therapy prior to objective progression, including any postdiscontinuation treatment systemic therapy, radiotherapy, or surgical intervention, PFS will be censored at the date of the last complete objective progression-free disease assessment before initiation of the new therapy.

Progression-Free Survival Sensitivity Analysis 2 (nonobjective progression as a PFS event): if a patient is discontinued from study treatment due to investigator determined non-objective progression (for example, symptomatic deterioration), then the patient's PFS time will be calculated using the date of non-objective progression as the progression date.

Since radiologic imaging is required every 8 weeks (+/- 3 days) from the first dose of study therapy, 2 scan intervals will be considered as 112 days.

The planned primary analysis of PFS at the end of Stage 2 will be conducted when at least approximately 120 total PFS events have occurred for the combination of each individual abemaciclib-containing arm and the standard-of-care arm, or all planned patients have been

enrolled in Stage 2, whichever comes later. Data from both Stages 1 and 2 will be pooled for this analysis.

The Kaplan-Meier (KM) method (Kaplan and Meier 1958) will be used to estimate the PFS curves as well as PFS rates at 3, 6, and 9 months for each treatment arm. Comparison between abemaciclib-containing treatment arms and the standard-of-care arm (Arm D) will be done using the log-rank test stratified by randomization strata. An unstratified log-rank test may also be performed. The stratified Cox proportional hazard model (Cox 1972) will be used to estimate the HR and its corresponding 95% CI.

5.8.2. Secondary Efficacy Analyses

5.8.2.1. Duration of Response

The DoR time is defined only for responders (patients with a BOR of CR or PR). It is measured from the date of first evidence of CR or PR to the date of objective progression or the date of death due to any cause, whichever is earlier. It is calculated as date of progression or death – date of first response evaluation of CR or PR + 1. The DoR will be censored according to the same rules as PFS.

A KM analysis of DoR will be performed to estimate the DoR curve for each arm. Point estimates and CIs for DoR quartiles and DoR rates will be calculated every 3 months for the first 12 months.

5.8.2.2. Overall Survival

Overall survival (OS) is defined as the time from randomization until death from any cause. 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. The first OS analysis will be at the time of the PFS analysis. The final analysis of OS will occur 12 months after the last patient enters treatment.

The KM method (Kaplan and Meier 1958) will be used to estimate the OS curves as well as OS rates at 6 months and 12 months for each treatment arm. Comparison between abemaciclib-containing treatment arms and the standard-of-care arm (Arm D) will be done using the log-rank test stratified by randomization strata. An unstratified log-rank test may also be performed. The stratified Cox proportional hazard model (Cox 1972) will be used to estimate the HR and its corresponding 95% CI.

5.8.2.3. Other Efficacy Analyses

Depending on the number of patients who undergo surgery while on treatment, additional analysis may be performed to assess efficacy endpoints listed above in patients with and without surgery while on treatment.

In patients where method of tumor assessment while on study was switched from computerized tomography (CT) scan with contrast to CT scan without contrast due to hypersensitivity or other conditions, additional analysis of ORR, DCR, DoR, and PFS may be performed where tumor assessment data after the switch to CT scan without contrast will not be considered in the analysis.

5.9. Health Outcomes/Quality-of-Life Analyses

Patient-reported outcomes are measured through paper versions of the following:

- mBPI-sf (modified Brief Pain Inventory, Short Form)
- EORTC QLQ-C30 (The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30)

For each instrument, the compliance rate overall, overall by treatment arm, and by treatment arm at protocol-specified time points will be calculated as the number of completed assessments divided by the number of expected assessments (that is, patients still on study). Compliance rates, reasons for noncompliance, and data collected for each instrument will be summarized by treatment arm.

Exploratory analysis may be performed to investigate associations between patient-reported data (mBPI-sf and EORTC QLQ-C30) and additional clinical, efficacy, and /or utilization measures as appropriate.

5.9.1.1. Pain Assessment

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 treatment arms by cycle with respect to change from baseline for each item. The model will include baseline score as a covariate and an unstructured covariance matrix will be utilized. The analysis will include all cycles for which at least 25% of patients in each arm have an mBPI-sf assessment.

Corresponding analyses will also be conducted for the mean of 7 pain interference with function items. If a patient does not complete Questions 5a through 5g on the mBPI-sf, the mean score for the 7 pain interference items will be calculated based on those answered questions when at least 4 out of 7 questions were completed (that is, ≥50% of the questions were answered). Pain analysis will be based on all enrolled patients with at least 1 baseline and 1 postbaseline score when change from baseline is analyzed.

Time to deterioration in a total score/domain score/item score will be summarized descriptively using the Kaplan-Meier method (Kaplan and Meier 1958), and a comparison between treatment arms will be made using the Cox proportional hazards model (Cox 1972).

Additionally, for BPI “worst pain”, structural equation modeling (SEM) and extended pattern mixture modeling (ePMM) may be applied to describe the association of patient-reported outcomes (PRO) and efficacy (PFS and OS) outcomes, and will be further detailed in the global patient outcomes and real world evidence (GPORWE) SAP.

5.9.1.2. Health-Related Quality of Life

EORTC QLQ-C30 instrument data will be scored as described by Aaronson and colleagues (Aaronson et al. 1993), yielding scores for a global health status scale, 5 functional scales, and 9 symptom scales. Descriptive statistics for each EORTC QLQ-C30 scale will be calculated. A mixed effects, repeated measures model will be applied to compare treatment arms by cycle with respect to change from baseline for each scale. The model will include baseline score as a

covariate and an unstructured covariance matrix will be utilized. The analysis will include all cycles for which at least 25% of patients in each arm have an assessment.

Time to deterioration in a total score/domain score/item score will be summarized descriptively using the Kaplan-Meier method (Kaplan and Meier 1958), and a comparison between treatment arms will be made using the Cox proportional hazards model (Cox 1972).

Additionally, for EORTC QLQ-30 "worst score" (for each scale), structural equation modeling (SEM) and extended pattern mixture modeling (ePMM) may be applied to describe the association of PRO and efficacy (PFS and OS) outcomes, and will be further detailed in the GPORWE SAP.

5.10. Utilization

Utilization data will be summarized overall by category and by treatment arm by category. The following categories will be described:

- analgesics (on study treatment and during short-term follow-up)
- antidiarrheal therapy (on study treatment and during short-term follow-up)
- transfusions (on study treatment and during short-term follow-up)
- hospitalizations (on study treatment and during short-term follow-up)
- surgery (while on-treatment and post-treatment), radiotherapy, and systemic therapy.

Utilization in additional drug classes may be explored (for example, anti-diabetic agents, stomatitis treatments).

For categorical variables, frequency and the corresponding proportions will be calculated. Continuous variables (days of hospitalization) will be described by the mean, median, and standard deviation.

Additional summaries for utilization by best overall response may be generated.

5.11. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Pharmacokinetic and pharmacodynamic analyses will be performed according to a separate PK analysis plan.

5.12. Safety Analyses

5.12.1. Extent of Exposure

Drug exposure, dose intensity, and drug adjustment (dose omissions and reductions) for abemaciclib, LY3023414, galunisertib, gemcitabine, and capecitabine will be summarized for all treated patients per treatment arm. Drug exposure will include summaries of cycles received per patient, duration on therapy, and cumulative dose. Dose intensity will be calculated as the actual cumulative amount of drug taken divided by the duration of treatment. Relative dose intensity

will be calculated as the actual amount of drug taken divided by the amount of drug prescribed times 100% (that is, expressed as a percentage). The summary of dose adjustments and omissions will include the reason for adjustment or omission.

For abemaciclib, LY3023414, galunisertib, and capecitabine, dose intensity will be expressed in mg/day whereas dose intensity for gemcitabine will be expressed in mg/m^2/day. The assigned cumulative dose for abemaciclib while on study is 2 doses per day × 200 mg per dose × number of days on treatment for Arms A; 2 doses per day × 150 mg per dose × number of days on treatment for Arms B and C. For Arm B, the assigned cumulative dose for LY3023414 while on study is 2 doses per day × 200 mg per dose × number of days on treatment. For Arm C, the assigned cumulative dose for galunisertib while on study is 2 doses per day × 150 mg per dose × number of days on treatment. For Arm D, the assigned cumulative dose for gemcitabine while on study is 1 dose per day × 1000 mg/m^2 × number of days of infusion and the assigned cumulative dose for capecitabine is 2 doses per day × 1250 mg/m^2 × number of days on treatment.

5.12.2. Adverse Events

Adverse event (AE) terms and severity grades will be assigned by the investigator using Common Terminology Criteria for Adverse Events (CTCAE) Version 4. In addition, AE verbatim text will also be mapped by the sponsor or designee to corresponding terminology within MedDRA. Adverse events will be reported using the following reporting process:

- In CTCAE Version 4, each CTCAE term is a MedDRA lower level term (LLT), except in the case where the CTCAE term is a MedDRA System Organ Class (SOC) followed by 'Other – specify'.
- The CTCAE Version 4 term reported by the investigator will be mapped to the MedDRA PT and SOC of the corresponding MedDRA LLT, unless the reported CTCAE term is 'Other – specify'.
- If the reported CTCAE term is 'Other – specify' the MedDRA LLT, PT, and SOC mapped from the verbatim AE term will be used.
- All listings and summaries will use the PT resulting from this process.

Preexisting conditions are defined as AEs that begin prior to the first dose of study drug.

A treatment-emergent adverse event (TEAE) is defined as any AE that begins between the day of first dose and 30 days after treatment discontinuation (or up to any time if serious and related to study treatment), or any preexisting condition that increases in CTCAE grade between the day of first dose and 30 days after treatment discontinuation (or up to any time if serious and related to study treatment).

Comparisons of preexisting conditions to on-treatment events at the LLT level will be used in the treatment-emergent computation.

A serious adverse event (SAE) is any AE during this study that results 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 significant by the investigator for any other reason.

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered SAEs when, based on appropriate medical judgment, they may jeopardize the patient and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

The following summaries and listings will be produced:

- Overview of AEs
- Summary of TEAEs by PT (any grade and Grade ≥ 3)
- Summary of DLTs (safety lead-in portion)
- Summary of TEAEs by SOC and PT (any grade and Grade ≥ 3)
- Summary of TEAEs by SOC and PT and maximum grade (1-5)
- Summary of treatment-emergent SAEs by SOC and PT (any grade and Grade ≥ 3)
- Summary of AEs as reason for study treatment discontinuation by SOC and PT
- Summary of TEAEs leading to dose omissions, reductions, hospitalizations
- Listing of SAEs

The TEAE and SAE summaries will be produced for all TEAEs/SAEs and repeated for TEAEs/SAEs related to study treatment, where relationship of the AE to the study treatment will be assessed by the investigator (yes or no).

5.12.3. Deaths, Other Serious Adverse Events

A summary of all deaths, including reasons for deaths, will be provided. All deaths, deaths on therapy, deaths within 30 days of discontinuation of study therapy, deaths on therapy or within 30 days of discontinuation of study therapy, and deaths after 30 days of discontinuation of study therapy will be summarized by reason for death. For deaths due to AE, the preferred term will be provided. In addition to the tabular summary, a by-patient listing of all deaths on study not attributed to study disease by the investigator will be provided.

5.12.4. Clinical Laboratory Evaluation

All relevant hematology and chemistry laboratory values will be graded according to CTCAE Version 4. These calculated grades will be summarized by cycle and maximum postbaseline grade over the entire study for each treatment arm. Treatment-emergent changes will be summarized by the maximum postbaseline grade, and a shift table of baseline grade by maximum postbaseline grade will be produced.

5.12.5. Vital Signs and Other Physical Findings

Temperature, blood pressure, pulse rate, respiration rate, weight, and ECOG PS will be summarized by cycle.

5.12.6. Electrocardiograms and cardiac assessments

Local electrocardiograms (ECGs) are to be performed according to protocol-specified time points for all treatment arms, whereas echocardiograms and chest CT/MRI scans are to be performed for safety lead-in and treatment arm C only. A listing of ECG, echocardiogram, and chest CT/MRI scan findings at baseline and postbaseline visits that are considered to be a medical history condition or an AE will be provided. For echocardiogram and chest CT/MRI scans, number of patients in normal, abnormal (of clinical significance) categories will be summarized.

5.12.7. CA19-9 levels

Summary statistics of change from baseline in CA19-9 levels will be presented.

5.13. Subgroup Analyses

Subgroup analyses of DCR, PFS, ORR, and OS may be performed for each of the potential prognostic subgroup variables listed below. Subgroup analysis will be performed at the end of Stage 2 and will include pooled data from the same arms of Stages 1 and 2.

- Baseline stratification factors (number of prior systemic therapies 1 versus 2)
- Sex (male; female)
- Age (<65 years; ≥65 years, <70 years; ≥70 years)
- Baseline ECOG PS (0; 1)
- BMI (<median; ≥median)
- Ethnicity (White; East Asian; others)
- Liver metastases (Yes; No)
- Region (North America; Asia; Europe; other)
- Stage at diagnosis (Stage IV, other)
- Time since diagnosis (<median; ≥median)
- Time since last systemic therapy (<median; ≥median)

- Previous lines of metastatic therapy (0; 1; 2)
- Previous fluoropyrimidne-based therapy (yes; no)
- Previous gemcitabine-based therapy (yes; no)
- Previous irinotecan (yes; no)
- Previous platinum (yes; no)
- Previous radiotherapy (yes; no)
- Previous surgery (yes; no)
- Albumin (<40; ≥40 g/L)
- CA19-9 (<40; ≥40 U/mL)

If a level of a factor consists of insufficient number of patients, analysis within that level will be omitted.

Analyses will be done within subgroup and, separately, across subgroups with a test of interactions of subgroups with treatment performed. Estimated HRs and CIs for the within-subgroup analyses will be presented as a Forest plot along with p-values for tests of interactions between subgroup variables and treatment.

Other subgroup analyses may be performed as deemed appropriate. If any safety analyses identify important imbalances between arms, subgroup analyses of these endpoints may be performed.

5.14. Protocol Violations

Significant protocol violations that potentially compromise the data integrity and patients' safety will be summarized by treatment arm for all randomized patients. These violations will include deviations that can be identified programmatically and those that can only be identified by the clinical research associate during monitoring. Significant protocol deviations are described in another document within the study Trial Master File.

5.15. Interim Analyses and Data Monitoring

First safety analysis will be performed after the completion of safety lead-in portion (6 to 12 patients) of the study for abemaciclib plus galunisertib arm. From thereon, periodic safety reviews will be performed with the first one to occur after 10 patients have been randomized to each treatment arm in Stage 1 and completed 1 cycle (28 days) or have discontinued treatment.

The Stage 1 analysis of efficacy and safety will be conducted under the guidance of an internal Assessment Committee (AC) after the last planned Stage 1 patient has enrolled and completed at least 16 weeks of treatment or have discontinued (whichever comes earlier). All data accumulated during this period (beyond 16 weeks for some patients, if available), will be included in the analysis. The purpose of this analysis is to evaluate safety and efficacy (DCR) to select which arms will continue to Stage 2. The AC will include the study statistician, a Global Patient Safety physician, and a Medical Director.

5.16. Annual Report Analyses

Annual report analyses, including Developmental Safety Update Report and Investigator's Brochure analyses, are described in the LY2835219 Program SAP.

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 AEs, provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AEs are summarized: by treatment arm, by MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
  - the number of participants at risk of an event
  - 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 clinical study report, manuscripts, and so forth.

6. References

Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JCJM, Kaasa S, Klee MC, Osoba D, Razavi D, Rofe PB, Schraub S, Sneeuw KCA, Sullivan M, Takeda F. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst.* 1993;85(5):365-376.

Cox DR. Regression models and life-tables. *J Royal Stat Soc Ser B*. 1972;74(2):187-220.

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

Leo Document ID = 471eda0f-fb22-4871-a46d-30b48d755bd4

Approver: PPD
Approval Date & Time: 12-Feb-2018 18:54:47 GMT
Signature meaning: Approved