J1L-AM-JZGB Statistical Analysis Plan Version 1

A Randomized Phase 3 Study of AM0010 (pegilodecakin) in Combination with FOLFOX Compared with FOLFOX Alone as Second-line Therapy in Patients with Metastatic Pancreatic Cancer that has Progressed During or Following a First-Line Gemcitabine Containing Regimen

NCT02923921

Approval Date: 8-Apr-2019

1. Statistical Analysis Plan: J1L-AM-JZGB: A Randomized Phase 3 Study of AM0010 (pegilodecakin) in Combination with FOLFOX Compared with FOLFOX Alone as Second-line Therapy in Patients with Metastatic Pancreatic Cancer that has Progressed During or Following a First-Line Gemcitabine Containing Regimen

Pegilodecakin (LY3500518) 2nd Line Metastatic Pancreatic Cancer

This is a randomized, open-label, global Phase 3 Study of AM0010(Pegilodecakin) in combination with FOLFOX compared with FOLFOX alone as second-line therapy in patients with metastatic pancreatic cancer that has progressed during or following a first-line gemcitabine-containing regimen.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol J1L-AM-JZGB
Phase 3

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

Approval Date: 08-Apr-2019 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

3. Revision History

Statistical analysis plan (SAP) Version 1 was approved prior to Lilly study team unblinded to study results.

4. Introduction

This is a randomized, open-label, global Phase 3 study. This document describes the planned statistical analyses for protocol J1L-AM-JZGB (JZGB), A Randomized Phase 3 Study of AM0010 in Combination with FOLFOX Compared with FOLFOX Alone as Second-line Therapy in Patients with Metastatic Pancreatic Cancer that has Progressed During or Following a First-Line Gemcitabine Containing Regimen. Any deviations from the analysis plan will be described in the clinical study report. Study JZGB is also known as "SEQUIOIA" and "AM0010-301." AM0010 is also known as pegilodecakin or LY3500518.

5. Study Objectives

5.1. Primary Objective

To compare the efficacy of pegilodecakin in combination with FOLFOX versus FOLFOX alone in patients with metastatic pancreatic cancer as measured by overall survival (OS)

5.2. Secondary Objectives

- To compare the efficacy of pegilodecakin in combination with FOLFOX versus FOLFOX alone in patients with metastatic pancreatic cancer as measured by:
    - progression-free survival (PFS)
    - Objective response rate (ORR), disease control rate (DCR), Duration of Response (DoR) by RECIST v1.1
    - One-year overall survival rate.
- To compare the safety and tolerability of Pegilodecakin in combination with FOLFOX versus FOLFOX alone.

5.3. Exploratory Objectives

- To explore biomarkers that may correlate with tumor response, immune activation, and relationships to clinical efficacy outcomes.
- Study the immunogenicity of anti-drug antibody of pegilodecakin.
- To explore associations between patient-reported symptoms, functioning, and global health status/Quality of Life (QoL) using the European Organisation for Research and Treatment (EORTC) Quality of Life of cancer Patients (QLQ-C30) questionnaire as well as current health status and the EuroQol-5D (EQ-5D) Index used in the economic evaluation of health care using the 5-level EQ-5D (EQ-5D-5L) questionnaire.

6. Investigational Plan

6.1. Overall Study Design and Plan

This is a randomized, open label, global Phase 3 study of AM0010(pegilodecakin) in combination with FOLFOX compared with FOLFOX alone as second-line therapy in patients with advanced metastatic pancreatic cancer that has progressed during or following a first-line gemcitabine containing regimen.

Approximately 566 patients will be randomized. Two interim analyses are planned. The first interim analysis will be performed when at least 60 randomized patients have had the opportunity to receive 4-months of therapy from the date of randomization. Based on the first interim analysis, the Data Monitoring Committee (DMC) will review the pharmacokinetic (PK) exposure-safety and PK exposure-efficacy on the composite aggregate data for a Go/No Go decision to enroll the entire Phase 3 study.

The second interim analysis will occur when approximately 276 deaths (70% of total 393 deaths needed for final analysis) have occurred. Based on the second interim analysis, the DMC will inform the study team whether the interim boundary are met and recommend modifying, completing, or discontinuing the trial.

Patients will be randomized in a 1:1 ratio to pegilodecakin combined with FOLFOX or FOLFOX alone.

The randomization will be stratified by 2 stratification factors:

- prior single agent gemcitabine therapy versus gemcitabine/nab-paclitaxel therapy
- North America versus Europe versus Asia-Pacific(APAC)

6.2. Study Endpoints

6.2.1. Primary Endpoint

The primary endpoint of this study is OS as defined as the time from date of randomization to death due to any cause.

6.2.2. Secondary Endpoints

The secondary endpoints of the study include:

- Progression-free survival is defined as the time from date of randomization to the earlier of first documentation of definitive disease progression (if the initial progressive disease [PD] is pseudopregression, it need to be confirmed by the consecutive scan) or death due to any cause.
- Objective response rate is defined as the proportion of patients who achieve a confirmed complete response (CR) or partial response (PR) as assessed by Response Evaluation Criteria in Solid Tumors (RECIST) v1.1.

- Duration of Response is defined as the time from the date of the first documentation of objective tumor response (CR or PR) to the date of the first documentation of objective tumor progression or death due to any cause, whichever occurs first.
- One-year OS rate is defined as the survival rate as estimated using the Kaplan-Meier method at the end of the first year.
- Compare the safety and tolerability of pegilodecakin in combination with FOLFOX versus FOLFOX alone.

6.2.3. Exploratory Endpoints

The exploratory endpoints of the study include:

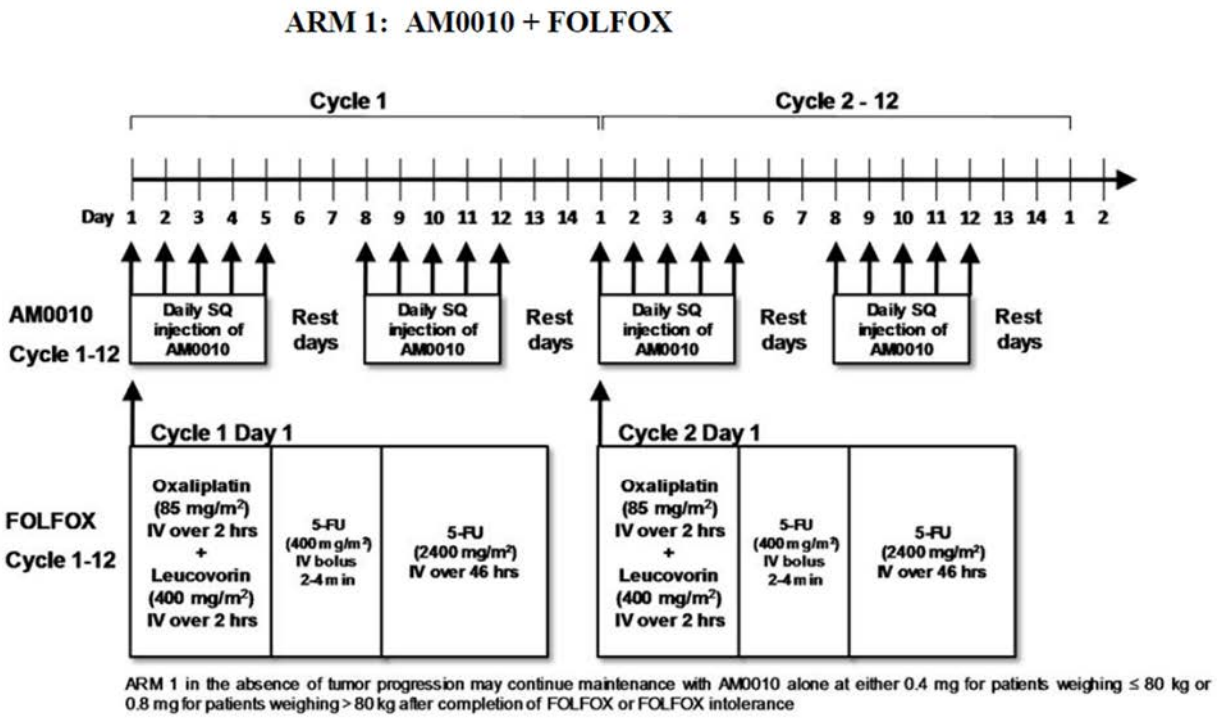
- Change from baseline for serum CA19-9.
- In addition to baseline and change from baseline in the EORTC core questionnaire, the (QLQ-C30), (EORTC QLQ-C30) and EuroQol Group 5-dimension 5-level measure of health-related quality of life (EQ-5D-5L Index score and Visual Analog Score [VAS] score). Time to deterioration in EORTC QLQ-C30 patient-reported symptoms and functioning will be explored using symptom items that correspond with the treatment-emergent adverse event (TEAE) profile (eg, fatigue, anorexia, and nausea/vomiting) as well as the subscales for functioning and Health-Related Quality of Life (HRQoL).
- Pegilodecakin PK (sparse sampling) in pancreatic cancer patients and immunogenicity .

6.3. Treatments

ARM 1: AM0010(pegilodecakin) (5 µg/kg) SQ dosed on Days 1–5 (with rest on Days 6 and 7) and Days 8-12 SQ (with rest on Days 13 and 14) plus FOLFOX (dl-LV 400 mg/m^2 and oxaliplatin 85 mg/m^2 followed by bolus 5-FU 400 mg/m^2 and a 46- to 48-hour infusion of 5-FU 2400 mg/m^2) initiated on Day 1 of a 14-day cycle, for up to 12 cycles or until disease progression by RECIST v.1.1.

While receiving FOLFOX during Cycles 1–12, pegilodecakin at 5 µg/kg will be administered as 1 of 2 fixed doses, either 0.4 mg for patients weighing ≤80 kg or 0.8 mg for patients weighing >80 kg.

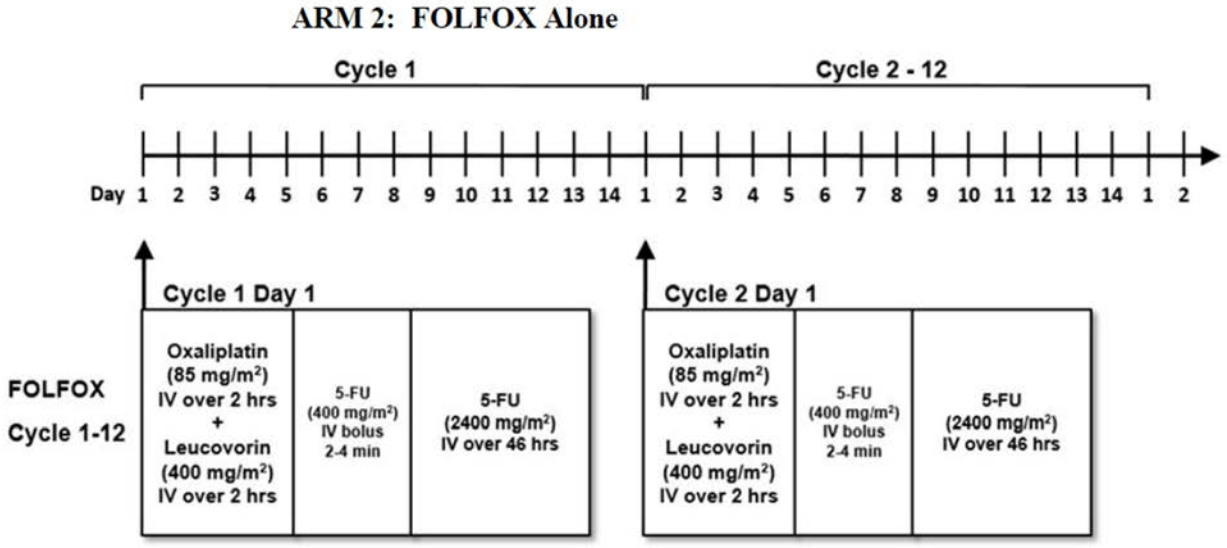
Patients in ARM 1 may continue on maintenance cycles with pegilodecakin (see Figure JZGB.6.1). After discontinuation of FOLFOX in the absence of tumor progression (ie, completion of the planned 12 cycles or unacceptable FOLFOX related toxicity), pegilodecakin at 10 µg/kg will be administered as 1 of 2 fixed doses, either 0.8 mg for patients weighing ≤80 kg or 1.6 mg for patients weighing >80 kg.



Abbreviations: F-FU = 5-fluorouracil; IV = intravenous; SQ = subcutaneous.

Figure JZGB.6.1. ARM 1: pegilodecakin +FOLFOX.

ARM 2: Patients receive only FOLFOX which is initiated on Day 1 of a 14-day cycle, for up to 12 cycles or until disease progression by RECIST v.1.1 (see Figure JZGB.6.2).



Abbreviations: F-FU = 5-fluorouracil; IV = intravenous.

Figure JZGB.6.2. ARM 2: FOLFOX alone.

Crossover of patients from the FOLFOX treatment arm into the FOLFOX plus pegilodecakin treatment arm is not permitted. Supportive care per the institution's normal standard of care, including concomitant medications, can be provided at the Investigator's discretion.

6.4. Treatment Duration

Each cycle is 14 days, with Days 6, 7, 13, and 14 being rest days for pegilodecakin. FOLFOX will be administered on Day 1 of each cycle. Patients randomized to the investigational combination treatment (pegilodecakin plus FOLFOX) ARM 1, after completion of up to 12 cycles of combination chemotherapy or if experiencing chemotherapy intolerance (as defined as Grade 3 or 4 non-hematologic toxicity that has not resolved to baseline in 28 days or Grade 4 hematologic toxicity that has not resolved to baseline in 28 days) may continue to receive maintenance pegilodecakin if the patient continues to receive clinical benefit (CR, PR, or stable disease [SD]). Patients will be treated until tumor progression by RECIST v.1.1. For ARM 2 (FOLFOX only) patients may receive FOLFOX for up to 12 cycles or 24 weeks or until tumor progression by RECISIT v.1.1 or until patients experience unacceptable toxicity, require palliative radiotherapy, withdraw consent, or their physician feels it is no longer in their best interest to continue on study treatments whichever occurs earlier.

To mitigate the chance of detecting false-progression (ie, pseudoprogression) early in the course of treatment on ARM 1 with pegilodecakin in combination with FOLFOX, patients whose scans show radiographic progression in the absence of clinical deterioration including worsening Eastern Cooperative Oncology Group (ECOG) Performance Status (PS) as assessed by the Investigator may remain on study treatments and should have an additional scan 4 weeks (±3 days) later as outlined in protocol Section 8.1.1. If this subsequent scan shows disease progression, the patient will be discontinued from study treatments.

7. General Statistical Considerations

7.1. Population:

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

The intent-to-treat (ITT) population includes all randomized patients per interactive voice response system (IVRS).

The safety population includes all randomized patients who received any amount of pegilodecakin or FOLFOX.

Per-Protocol Population: All patients who are in the safety population and have no major protocol deviations.

The major protocol deviations which will result in exclusion from the Per-Protocol population are listed below. This is not an exhausted list and may be amended over the course of the trial.

- patients with disease other than metastatic pancreatic adenocarcinoma
- patients with ECOG >1 at baseline
- patients who received more than 1 prior regimen in metastatic setting
- patients who do not have measurable disease at baseline
- patients who received the incorrect randomized treatment assignment (for any duration)

Anti-drug antibody (ADA) evaluable population includes all patients with at least 1 baseline and postbaseline ADA assessment.

Response evaluable population includes patients who have at least 1 baseline tumor scan and 1 adequate postbaseline tumor scan.

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

All analyses will be performed by treatment arm. Unless otherwise noted, all analyses on the ITT population will be performed by assigned treatment arm and all analyses on the safety population will be performed by actual treatment received.

7.2. Definitions and Conventions

The **date of randomization** is the date the patient was randomly assigned to pegilodecakin + FOLFOX arm or FOLFOX arm using the IVRS.

The **date of first dose** is the date of the first dose of pegilodecakin or any component of FOLFOX whichever is earliest.

The **baseline value assessment** for safety **in general** is the last value observed on or prior to the first dose of pegilodecakin or FOLFOX. For efficacy, the baseline value assessment is the last

value on or prior to the randomization date or the first value on or prior to the first dose of pegilodecakin or FOLFOX whichever occurs earlier.

The change from baseline will be calculated by subtracting the baseline values from the individual post-treatment values. If either the baseline or post-treatment value is missing, the change from baseline is set to missing.

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 2014 and the date of first dose was 06 June 2014, 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 2014 and the date of first dose was 06 June 2014, 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 = 30.4375 days

One year = 365.25 days

Unless otherwise noted, summaries of continuous variables will include a mean, median, standard deviation, minimum, and maximum.

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

Visit window approaches will not be used for this study. Visit based summaries will include scheduled assessments only. All scheduled and unscheduled post-baseline assessments will be used for derivation of minimum, maximum, and worst-case post-baseline values. All available data will be listed.

P-values will be rounded to 4 decimal places. P-values that round to 0.0000 will be presented as '<0.0001' and p-values that round to 1.000 will be presented as '>0.9999.' Confidence intervals (CIs) will be presented as 2-sided 95% CIs unless specified differently in specific analysis.

All data listings will present subjects in the ITT population, unless otherwise noted. For data listings, imputed dates will not be included and only the actual reported values will be displayed. As applicable, durations will be calculated and presented using imputed dates. When this occurs, a footnote will be added to the data listing indicating this.

7.3. Sample Size

The sample size is calculated in order to compare the OS between patients randomized to receive pegilodecakin in combination with FOLFOX versus FOLFOX alone. Approximately 566 patients will be randomized to the 2 arms in a 1:1 ratio to observe at least 393 death events. Assuming a median OS time for the FOLFOX arm of 5.9 months, 393 events are needed to detect a 35% increase with 85% power using a log-rank test (2-sided) with an overall type 1 error of 0.05, which corresponds to a median survival of 8 months in the pegilodecakin in combination with FOLFOX arm (hazard ratio [HR] = 0.7375).

7.4. Planned Interim Analyses/Final Analysis

Two interim analyses are planned for this study. The first interim analysis will be performed when approximately 60 randomized patients have had the opportunity to complete 4 months of therapy from the date of randomization. The nominal significance will be calculated based on O'Brien and Fleming alpha spending function in EAST v6.4. The second interim analysis is planned after approximately 276 deaths (70% of the 393 events) have been observed. The stopping boundaries will be derived based on the actual observed deaths at the first and second interim analysis. If the second interim analysis is performed approximately at 276 deaths, the study could be stopped by the DMC for efficacy if the p-value is <0.015. The nominal significance level for the final look at OS after 393 events would then be 0.045.

Table JZGB.7.1. Schedule of Analyses

	Interim Analysis 1	Interim Analysis 2	Final Analysis
Conditions	Approximately 60 randomized patients who have had the opportunity to receive 4-months of therapy	Approximately 276 OS events	393 OS events
Alpha level	0.0001[a]	0.015	0.045
Upper HR Boundaries	2.145	1.342	1.224
Lower HR Boundaries	0.466	0.745	0.817

Abbreviations: HR = hazard ratio; OS = overall survival.

[a] A very small alpha is planned to be spent at the first interim analysis. Hazard Ratio boundaries are calculated based on the alpha spent. Hazard ratio (HR) boundaries are calculated based on the alpha spent. The O'Brien-Fleming (OBF) spending function is used to allocate the alpha at analysis 2 and final analysis with total alpha level of 0.0499. The actual alpha at each interim analysis will be recalculated based on the number of events observed at each analysis. The HR boundaries are for reference only and they are not for decision making.

7.5. Randomization, Stratification

Patients will be randomized in a 1:1 ratio to pegilodecakin combined with FOLFOX or FOLFOX. The randomization will be stratified by 2 stratification factors:

- prior single agent gemcitabine therapy versus gemcitabine/nab-paclitaxel therapy
- North America versus Europe versus APAC

Strata from IVRS (randomized strata) will be used in all efficacy analysis.

7.6. Methods for Handling Missing Data

In general, missing data will be treated as missing and no data imputation will be applied, unless otherwise specified. Additional data handling rules including those for imputation of partially missing dates are provided in a separate document pertaining to programming specifications.

8. Patient Disposition

8.1. Disposition

Patient disposition will be summarized using the All Enrolled population, randomized or not randomized, along with the reason for not being randomized.

The number or patients in the study populations, ITT, Safety, Per Protocol (PP) and reasons for exclusion from the PP population (no baseline scan, no post-baseline scan, or has a major protocol deviation) will be presented by randomized arm.

The number and percentage of patients who continued on pegilodecakin maintenance cycles will be summarized and the number and percentage of patients in survival follow-up will be summarized as well.

A summary table of randomized stratifications (geographic region and prior therapy) by treatment arm will be provided for the ITT population using data provided by IVRS. This data will be compared to data provide via the electronic case report form (eCRF) for consistency. If any inconsistencies exist between the IVRS and eCRF data, a separate summary for the stratification factors will be provided using eCRF data.

A table summarizing study discontinuation, along with reason for discontinuation will be provided. Primary reason of study discontinuation may include 1 of following: withdrawal of consent; death; discontinuation of the study at the request of ARMO, a regulatory agency, or an institutional review board (IRB)/independent ethics committee (IEC); and other. This analysis will be performed only on the ITT population.

A summary table of patients treated, treatment ongoing, treatment discontinued, along with reasons of treatment discontinuation for each study drug: pegilodecakin and components of FOLFOX (oxaliplatin, Leucovorin, and 5-Fluorouracil) will be provided. This table will be performed only on the Safety population. In addition, the number of patients who received maintenance therapy will also be summarized.

All disposition data will also be provided via data listings.

8.2. Protocol Deviations

Identification of important protocol deviations may be derived from data or may be from observation from clinical monitoring. All patients with important protocol deviations or major protocol deviations, will be summarized by randomized arm. In addition, the listing of patients excluded from PP population along with the corresponding reason will be provided.

8.3. Duration of Follow-Up

The duration of follow-up is defined as the time from randomization to the last known alive date.

9. Demographics and Baseline Characteristics

9.1. Demographics

The demographic characteristics will be summarized by randomized arm and overall for ITT, Safety populations. The demographic characteristics consist of age, sex, race, ethnicity, height, weight, body surface area (BSA), body mass index (BMI), and ECOG PS.

9.2. Baseline Disease Characteristics

Cancer history data will be summarized and listed, including metastatic pancreatic cancer at baseline (Yes or No), time from initial diagnosis to randomization (months), stage at initial diagnosis, locally advanced or metastatic stage cancer at initial diagnosis (Yes or No), prior radiation therapy (Yes or No), prior Whipple procedure (Yes or No), prior biliary stent performed (Yes or No), location of the primary lesion(s) in the pancreas, histology of pancreatic cancer (Adenocarcinoma or Other), number and location of metastatic sites at screening will be summarized by randomized arm and overall.

9.3. Prior Systemic Cancer Therapy

The number of prior systemic cancer therapy regimens patient received, prior systemic cancer therapy (by regimen), intent of prior cancer therapy (Adjuvant/Neo-Adjuvant or Advanced/Metastatic), duration of prior cancer therapy (months), the best response to the prior systemic cancer therapy, time from progressive disease of first line therapy to randomization (days), reason for discontinuation, instrument used to measure the PD will be summarized by randomized arm.

9.4. Medical History

Medical history terms will be coded using the latest version of the Medical Dictionary for Regulatory Activities[MedDRA].

A summary for medical history will be presented including the count and percent of patients with any medical history record and also by body system and Preferred Term (PT).

10. Treatments and Medications

10.1. Concomitant Medications

Concomitant medications will be coded using the most recent version of World Health Organization (WHO) Drug Dictionary. Concomitant medications are defined as

- The medication was started on or after the date of first dose of study treatment and before treatment discontinuation date + 30day;
- The medication was started prior to first dose study treatment but was ongoing at the time of the first dose of study treatment.

Concomitant medications will be summarized separately and presented by drug class and PT by randomized arm on safety population. In addition use of selected concomitant medication, such as steroid, use will be summarized separately. Study Treatments

All study treatment summaries will be based on the Safety population.

10.1.1. Treatment Exposure

The following parameters will be summarized by treatment arm:

Overall Treatment:

- duration of treatment (months):
  - Pegilodecakin: (last dose date – first dose date+1)/ 30.4375
  - Any component of FOLFOX: (last dose date – first dose date + 14)/30.4375
- number of cycles administered (patient with total number of cycles administered, 1 cycle, 2 cycles, 3 cycles, etc.)
- total number of cycles administered

The following parameters will be summarized (descriptive statistics) by study drug component within each treatment arm:

- total cumulative dose (mg for pegilodecakin and mg/m^2 for each FOLFOX drug component)
- total dose per cycle (dose intensity)
- relative dose intensity (%), summarized by descriptive statistics and by categories of [60%, 80%), [80%, 100%), [100%, 120%) and ≥120%, etc.

In ARM 1(Pegilodecakin + FOLFOX), pegilodecakin's will be treated up to 12 cycles with 4 mg/cycle (2weeks) for patients with weight ≤80 kg and 8 mg/cycle (2weeks) for patients with weight >80 kg. After the 12 cycle, patient receiving benefit on ARM 1 can continue on maintenance cycle with double of the dose 8 mg/2 weeks for patients with weight ≤80 kg and 16 mg/2 weeks for patients with weight >80 kg. Therefore the exposure will be summarized in the following categories:

- for FOLFOX, exposure will only be summarized up to 12 cycles.
- for Pegilodecakin:
    - Cumulative dose, total dose per cycle and relative dose intensity will be summarized up to 12 cycle prior to maintenance period and repeated within the maintenance period.
    - The duration of treatment, and number of cycle administered will be summarized across all treatment durations and repeated for the maintenance cycle.

The key parameters used to calculate dosing data are shown in Table JZGB.10.1:

Table JZGB.10.1. Patient Exposure Prior to Maintenance Period

		FOLFOX		
Treatment	pegilodecakin	**Oxaliplatin**	**Leucovorin**	**5-FU**
Planned Dose intensity	4 mg/cycle (14 days) (≤ 80kg) 8 mg/cycle (14 days) (> 80kg)	85 mg/m^2 per cycle(14 days)	400 mg/m^2 per cycle (14 days)	2800 mg/m^2 per cycle (14 days)
Cumulative dose	sum of the doses mg for pegilodecakin, or mg/m^2 for FOLFOX administered to a patient during the treatment period			
Actual Dose intensity	Cumulative dose/ duration of treatment (mg/2week)	Cumulative dose/total number of cycle received (mg/m^2/cycle)		
Relative dose intensity (%)	Actual Dose intensity/planned dose intensity			

10.1.2. Treatment Modifications

Treatment modifications included any dose delays, reductions, and interruptions will be summarized. The following parameters will be summarized by study drug, as applicable.

10.1.3. Dose Delays

Number of dose delays per patient, length of delay, and reason for delay will be summarized for each drug component for each treatment arm.

10.1.4. Dose Interruptions

Only drug components in FOLFOX can be interrupted. Number of patients with at least 1 dose interruption, reason for interruption, and number of interruptions per patient will be summarized for FOLFOX arm.

10.1.4.1. Dose Reductions

Each drug component may be reduced. Number of patients with at least 1 dose reduction, reason for reductions, and number of dose reductions will be summarized for each treatment arm for

each component of treatment. Each FOLFOX dose reduction level is defined in protocol Table 12.

11. Efficacy Analysis

All efficacy analyses will be based on the ITT population.

11.1. Primary Efficacy Analysis

The primary endpoint is OS. Overall survival is defined as the time from date of randomization to the date of death (due to any cause). For patients whose last known status is alive at the data cutoff date for the analysis, time will be censored as the last contact date prior to the data cutoff date.

The null (H_o) and alternative (H_a) hypotheses for testing overall survival are:

H_o: HR pegilodecakin $_{+ \text{ FOLFOX/FOLFOX alone}} = 1$

H_a: HR pegilodecakin $_{+ \text{ FOLFOX/FOLFOX alone}} \neq 1$ (superiority)

The hypotheses will be tested using a 2-sided stratified log-rank test. The stratification factors from IVRS will be used. Kaplan-Meier (K-M)estimates for the median time along with 95% CIs for each randomized arm; the hazard ratio with 95% CI will be provided as well.

Survival rates at fixed time points (3 months, 9 months, 12 months, 15 months, etc.) will be estimated using K-M estimate for each treatment group. These rates will be compared based on a normal approximation for the difference between the treatment arms. Associated 2-sided 95% CI will be calculated.

Hazard ratios for overall survival and their 95% CIs will be provided using Cox proportional-hazard model(Cox 1972) with treatment, stratification factors (from IVRS).

11.1.1. Subgroup Analyses

Subgroup analysis for the primary endpoint (OS) will be performed and the HR and 95% CI will also be provided for each treatment arm in the forest plot. Subgroups are defined as follows:

- prior therapy
    - prior single agent gemcitabine therapy versus
    - Gemcitabine/nab-paclitaxel therapy
- geographic region
    - North America
    - Europe
    - Asia-Pacific
- age group at baseline
    - <65 years
    - ≥65 years

- sex
    - female
    - male
- race (White, Asian, other)
    - White
    - Asian
    - other
- baseline level of CA19-9
    - normal
    - >normal - ≤ (59 * Upper limit Normal [ULN])
    - >59 * ULN
- Number of metastatic sites
    - 1 site
    - 2 sites
    - ≥3 sites
- Location of primary lesions in pancreas
    - head
    - other
- liver metastases
    - yes
    - no
- time from initial diagnosis (<6 months, ≥6 months)
    - <6 months
    - ≥6 months
- Eastern Cooperative Oncology Group
    - 0
    - 1
- weight
    - ≤80kg
    - >80 kg

If a subgroup category has less than 10 patients per treatment arm, HR will not be computed/displayed. A forest plot will be provided based on the HRs for the subgroup analyses.

11.2. Secondary Efficacy Endpoints

The secondary efficacy endpoints are PFS and ORR, DCR, DoR and 1-year OS rate.

11.2.1. Progression-Free Survival

Progression-free survival is defined as the time from randomization to the date of the first documented tumor progression as determined by the investigator (per RECIST v1.1 criteria) or death due to any cause - whichever occurs first.

The recorded date of radiographic progression is the date of the tumor assessment visit at which progression is declared. Only adequate tumor assessments (ATAs) will be considered in the determination of radiographic progression and censoring date. Adequate tumor assessments is defined as assessment at a time point evaluation of CR, PR, SD (non-CR, non-PD), or progression.

The censoring rules are:

- Patients who do not progress or die as of the data cutoff date will be censored on the date of their last ATA on or prior to the data cutoff date.
- Patients who do not have any baseline tumor assessment will be censored on the date they are randomized with a duration of 1 day.
- Patients who do not have any on-study tumor assessment and do not die will be censored on the date they are randomized with a duration of 1 day.
- Patients who received any subsequent anti-cancer therapy will be censored on the date of last ATA on or prior to initiation of the subsequent anti-cancer therapy.
- Patients who are lost to follow for more than 1 scheduled visit (greater than (7*8*2+2*3)=118 days from the previous ATA date) prior to progression disease or death will be censored on the date of last ATA prior to the missing scheduled visits.

In the event of suspected pseudoprogression

- If a consecutive sequence of pseudoprogression is confirmed with a confirmatory scan, the date of progression used to determine PFS will be the date of the first consecutive pseudoprogression, not the date of the confirmatory scan.
- If no additional scan is available after the initial suspected pseudoprogression, the initial PD data will be considered as PD date.
- If the pseudoprogression is not confirmed for the following scan, the patient will continue to be followed up for future scans. Multiple pseudoprogression's can occur for a patient during the study until disease progression is confirmed or the patient lost to follow up.

The K-M estimates for the median, the first and third quartiles with 95% CI (if estimable), along with the HR (including 95% CI), will be provided. The figure of K-M curve will also be

presented for each randomized arm and the difference in the curves will be tested using the stratified log-rank test, with the same strata as in OS. In addition forest plot will be provided for the PFS based on the HRs for the subgroup analyses.

11.2.2. Objective Response Rate

Objective response rate is defined as the proportion of patients with a confirmed CR or confirmed PR relative to the total analysis population. Responses observed at each time point while on study treatment (but before the time of disease progression[defined in 10.2.1] or the initiation of post discontinuation therapy) will be included in the derivation.

Objective response rate will be summarized along with the exact 95% CI by randomized arm. A stratified Cochran-Mantel-Haenszel (CMH) test will be used to compare ORR between 2 randomized arms by the same stratification factors as OS. The odds ratio with 95% CI will be provided.

11.3. Other Efficacy Analysis

11.3.1. Disease Control Rate

Disease control rate, defined as the proportion of patients who achieve confirmed CR, or confirmed PR, or SD assessed by investigator according to RECIST 1.1. Disease control rate will be analyzed in the same manner as ORR.

11.3.2. Duration of Response, Time to Response

Analyses of duration of response and time to response will only include patients with confirmed CR or PR based on investigator's assessments.

The duration of response is defined as the time from the date of the first documentation of objective tumor response (CR or PR) as determined by the investigator using RECIST v1.1, that is subsequently confirmed to the date of the first documentation of objective tumor progression or to death due to any cause, whichever occurs first.

Duration of response will be censored on the date of the last adequate tumor assessment on study for patients who do not have tumor progression and who do not die due to any cause while on study. Duration of response will be censored according to the same rules as PFS.

Duration of response will be summarized descriptively using the K-M method for each randomized arm, including the median, the first and third quartiles along with 95% CI (if estimable).

Time to response is defined as the time from date of randomization to the date of the first documentation of objective tumor response that is subsequently confirmed. Descriptive statistics will be summarized for time to response.

11.3.3. Best Overall Response

Best overall response (BOR) is defined as the best response as determined by the investigator, recorded between the date of randomization and the date of objectively documented progression per RECIST v1.1 or the date of subsequent therapy, whichever occurs first. For patients without documented progression or subsequent anti-cancer therapy, all available response assessments will contribute to the BOR determination. Complete or PRs may be claimed only if the criteria for each are met at a subsequent time point as specified in the protocol (approximately 4 weeks later).

For patients who continue pegilodecakin in combination with FOLFOX treatment beyond suspected pseudoprogression, if the pseudoprogression is not confirmed in the following scan, the additional responses occurred after the pseudoprogression should be considered in deriving BOR.

Confirmed responses are those that persist on repeat imaging study 4 weeks after initial documentation of response. Designation of best response of SD requires the criteria to be met at least 8 weeks (±3 days) after randomization. Designation of best response of SD requires the criteria to be met at least 8 weeks (±3 days) after randomization.

Best overall response will be summarized along with the exact 95% CI by randomized arm.

11.3.4. Subsequent Anti-Cancer Therapy

The number of patients starting a new anti-cancer therapy will also be summarized by treatment arm. The subsequent anti-cancer therapies will be coded by WHO Drug Dictionary and summarized by regimen.

11.3.5. Sensitivity Efficacy Analyses

The additional sensitivity analyses for OS includes:

OS sensitivity analysis 1: An overall survival analysis censoring patients at the time of initiation of subsequent anti-cancer therapy.

OS sensitivity analysis 2: An overall survival analysis defining OS time as the time from the date of study enrollment to the date of death due to study disease. Survival time will be censored on the date the patient was last known to be alive for patients who have no reported event. For patients that have died due to reasons which are not disease related, survival time will be censored at the date of death.

OS sensitivity analysis 3: OS will be analyzed after adjusting for selected prognostic factors. Potential prognostic factors include the stratification factors (per CRF) and other factors as outlined in Section 11.1.1. The HR for treatment effect will be estimated using a multivariate Cox proportional hazard model constructed by selecting variables among all the potential variables, using stepwise selection method with entry p-value 0.05 and exit p-value 0.01. The treatment factor will be kept out of the model throughout the covariate selection process and only added into the final model.

The additional sensitivity analysis for PFS includes:

PFS sensitivity analysis 1: A PFS analysis that considers clinical progressions as events.

PFS sensitivity analysis 2: A PFS analysis that does not censor for the start of additional anticancer therapy

PFS sensitivity analysis 3: A PFS analysis that considers pseudoprogression as events. For patients with multiple progression events, the PFS is defined as from the randomization to the first progression event.

11.3.6. Exploratory Tumor Marker Analyses

A separate analysis plan for tumor marker analyses will be provided.

11.4. Testing Hierarchy for Efficacy Endpoints

In order to preserve an overall 2-sided type-I error rate of 5%, a hierarchical testing approach (Glimm et al. 2010) will be applied to key secondary endpoints following analysis of the primary endpoint of OS. The hierarchical ordering of key secondary endpoints is as follows:

1. Progression-Free Survival
2. Objective Response Rate

If primary analysis of OS is statistically significant either in interim or final, then stratified log-rank test for the PFS will be tested with a 2-sided type one error of 5%. If OS is not statistically significant. Progression-free survival and ORR will not be further tested. If PFS analysis is statistically significant, then ORR will be further tested with stratified CMH test with a 2-sided type 1 error of 5%. If any endpoint is not statistically significant, the following endpoints will not be further tested.

11.5. Restricted Mean Analysis

The common method for describing benefit on the time scale is to calculate the difference in median event time between the 2 treatment arms. An alternative method for describing benefit on the time scale is to estimate the average difference between the K-M curves. This corresponds to calculating the difference in the average time to event for the 2 treatment arms (Irwin 1949; Karrison 1997; Meier et al. 2004). Similar to the HR, this method uses all of the available information across the K-M curves, but has the additional advantage of assessing benefit on the time scale.

To estimate an improvement in OS with pegilodecakin +FOLFOX, we will follow the method of Irwin (1949) detailed in Karrison (1997) and Meier and colleagues (2004) for estimating the 'difference in average OS,' which we will refer to more formally as the restricted mean difference in OS.

The area under each K-M curve will be calculated using numerical integration (trapezium rule) per Karrison and implemented in SAS using PROC LIFETEST. The difference between treatment arms and a CI for the difference will be formed.

Since the K-M curve may be ill-determined beyond a certain range, or even undefined (if the longest observation is censored), for evaluation and comparison of means, the area under each K-M curve will be calculated between time 0 and restriction time T, which is why this is referred to as a "restricted mean." Following the suggestion of Karrison, the restriction time T will be chosen as largest time point t such that the standard error (SE) of the survival estimate at time t in each treatment group is no more than 0.075. For this purpose, we will use the simple, albeit conservative, formula proposed by Peto et al. (1977) for calculating the SE of S(t) as

$$SE\big(S(t)\big) = S(t)\sqrt{1 - S(t)/n(t)}$$

, where n(t) is the number of patients still at risk at time t.

12. Safety Analysis

Unless otherwise stated, safety summaries will be produced by actual treatment patient received for the safety population.

12.1. Adverse Events

Adverse event terms will be coded using the latest version of MedDRA (version 19.0 or later). Adverse events (AEs) will be graded according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v4.03.

Summaries of AEs will be generated for those events that are considered treatment emergent, where a TEAE is defined as an AE that begins or worsens in grade after the date of first study drug administration through 30 days after all treatment discontinuation. Any related serious adverse event (SAE) which occurs more than 30 days after all treatment discontinuation will also be considered as TEAE.

All AE data will be presented in listings; separate listings for serious AEs, AEs leading to any study treatment discontinuation (either FOLFOX or Pegilodecakin), and AEs leading to death will be generated. Adverse events by subgroup (age, gender, and race) will also be summarized.

12.1.1. Incidence of Adverse Events

The following TEAE/SAE listings and summaries will be produced:

- overview of TEAEs
- summary of TEAEs by PT (all grade and grade ≥3)
- summary of TEAEs by SOC and PT (all grade and grade ≥3)
- summary of TEAEs by PT and maximum grade (1-5)
- list of SAEs
- summary of SAEs by SOC and PT (all grade and grade ≥ 3).

12.1.2. Relationship of Adverse Events to Study Drug

A TEAE will be considered as study drug related if the investigator considered the event as related to any study drug component. If the relationship to any study drug component is missing, this AE will be considered as study drug related AE.

The 4 summaries in Section 12.1.1 will be repeated for TEAEs/SAEs related to study treatment.

12.1.3. TEAEs Leading to Treatment Discontinuation

A summary table for TEAEs leading to any/each study drug component discontinuation will be presented by SOC and PT.

12.1.4. Death

A summary of death with the primary death reason and death period (all death, death during the treatment and within 30 days after treatment discontinuation, death >30 days of treatment discontinuation) will be presented based on ITT population.

12.1.5. Consolidated Adverse Event

Given the high level of granularity of the MedDRA dictionary, clinically identical or synonymous PTs reported under different terms in the database, in addition to being reported separately, will also be consolidated in a separate summary. The list of consolidated AE categories will be reported in the clinical study report.

12.1.6. Immune-Mediated Adverse Event

Potential immune mediated AEs, (eg, colitis, pneumonitis, endocrine, and hepatic events) will be summarized.

12.1.7. Adverse Events of Special Interest

Adverse events of special interest (AESIs) are AEs thought to be potentially associated with the study drug or the disease under study. Therefore, aggregate or single AE terms including additional medical interventions for identifying AEs possibly associated with the study drug were developed.

Adverse events of special interest are listed in the Investigator's Brochure. The MedDRA PTs that are grouped under each of the AESI terms will be provided.

12.2. Clinical Laboratory Evaluations

Hematology and chemistry laboratory values will be graded according to CTCAE v4.03. These calculated grades will be summarized by cycle and maximum post-baseline grade over the entire study. Shift tables from baseline to worst post-baseline CTCAE grade will be generated.

A listing of abnormal laboratory events will be presented.

12.3. Vital Sign Measurements

Blood pressure, heart rate, temperature, weight, BSA, along with changes from baseline will be summarized by scheduled visits.

12.4. Electrocardiogram

A twelve-lead electrocardiogram (ECG) including ventricular rate, PR, QRS, QT, and QTc intervals will be summarized. All numeric results along with changes from baseline will be summarized for all scheduled visits.

The QTc parameter will be summarized for the following categories:

- summarized by visit: QTc change from baseline 30 to <60 msec and ≥60 msec.

The patients with abnormal ECG will be listed.

12.5. ECOG Performance Status

Eastern Cooperative Oncology Group PS status will be summarized by visit.

All ECOG data will also be provided via a data listing.

12.6. Tumor Marker - CA 19-9

Patients with above normal CA19-9 at baseline will be included for the summary and percent change from baseline by visit using descriptive statistics for each treatment arm. The best change of CA19-9 from baseline is defined as the percentage change at the nadir from the baseline value, which will be summarized descriptively by treatment arm.

For patients with both baseline abnormal CA 19-9 and post-baseline assessments, the percent of patients with a maximum drop in CA19-9 of at least 20%, 50%, and return to normal will be calculated for each treatment arm.

12.7. Pregnancy

A listing of patients reporting a pregnancy during the treatment will be presented.

12.8. Utilization

Utilization data will be summarized by category across arms. The following categories will be described:

- transfusions (on study treatment and within 30 days after the treatment discontinuation)
- growth factor
- hospitalizations (on study treatment and within 30 days after the treatment discontinuation)
- post discontinuation systemic therapy

For categorical variables, frequency and the corresponding proportions will be summarized by treatment arm.

13. Pharmacokinetics

Pegilodecakin observed concentrations will be summarized by descriptive statistics.

Additional analysis utilizing the population PK approach may also be conducted if deemed appropriate. Relationships between pegilodecakin exposure and measures of efficacy and safety may be explored.

A separate analysis plan for pegilodecakin PK and exposure/response will be provided.

14. Immunogenicity

The treatment-emergent anti-drug antibody (TE ADA) evaluable population will be defined within the Safety Population as those patients with both a baseline ADA test result and a postbaseline ADA test result. As a proportion of the TE ADA evaluable population, the number and frequency of the following will be tabulated by treatment group: TE ADA+ patients, and TE ADA+ patients with detected higher-tier assay results. Included in the same tabulation will be, as a proportion of the TE ADA evaluable population, the number and frequency of patients with ADA present at baseline, as well as patients with detected higher-tier assay results at baseline.

A summary will be provided of the number and percentage of patients who receive pegilodecakin reporting specific TEAEs (overall and by PT) by patient TE ADA status (TE ADA+, TE ADA-, TE ADA Inconclusive, Patient not evaluable for TE ADA). The PT will be ordered by decreasing frequency in the TE ADA+ status group. The specific TEAE of interest are events in any one of the MedDRA Anaphylactic reactions Standardized MedDRA Query (SMQ), Hypersensitivity SMQ, Angioedema SMQ, Injection site reactions High Level Term (HLT), or Infusion site reactions HLT.

Additional immunogenicity analyses, to be presented in a standalone immunogenicity report in context with other studies of pegilodecakin, will be described in a separate SAP.

14.1. Interim Analysis

A planned first interim analysis will occur once at least 60 randomized patients who have the opportunity to receive 4 months of therapy from the date of randomization. The interim analysis is not designed to stop the study early for outstanding efficacy. Based on review of the PK exposure-safety and PK exposure-efficacy on the aggregated data of composite efficacy endpoints (PFS, ORR, and OS), the DMC will recommend continuing the Phase 3 trial, amending the study protocol, or stopping the study. The study will continue to enroll while this analysis is being completed.

A second interim analysis is planned when approximately 276 deaths (70%) of 393 deaths are reached in this event-driven trial. Details of the interim analysis results will be reviewed by the DMC. The DMC will inform the sponsor whether the interim efficacy boundary was met and provide the Sponsor with a recommendation to complete the trial as planned, modify, or discontinue the trial.

15. References

Cox DR. Regression models and life-tables. *J Royal Stat Soc Ser B.* 1972;34(2):187-220.

Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. *Stat Med.* 2010;29(2):219-228.

Irwin JO. The standard error of an estimate of expectational life. *J Hygiene*. 1949;47;188-189.

Karrison TG. Use of Irwin's restricted mean as an index for comparing survival in different treatment groups-- interpretation and power considerations. *Control Clin Trials.* 1997;18(2):151-167.

Meier P, Karrison T, Chappell R, Xie H. The Price of Kaplan-Meier. *J Amer Stat Assoc*. 2004;99(467):890-896.

Leo Document ID = 1f124bc1-036e-4fe6-ab27-907d3eb67d37

Approver: PPD
Approval Date & Time: 08-Apr-2019 00:38:08 GMT
Signature meaning: Approved