STATISTICAL ANALYSIS PLAN

A MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PLACEBO CONTROLLED STUDY OF VARLITINIB PLUS CAPECITABINE VERSUS PLACEBO PLUS CAPECITABINE IN PATINTS WITH ADVANCED OR METASTATIC BILIARY TRACT CANCER AS SECOND LINE SYSTEMIC THERAPY

Date: 06 December 2018
Version: 3.0

SAP Approvals:

[redacted] [redacted]

Date

[redacted]
Chief Medical Officer, ASLAN Pharmaceuticals

[redacted] [redacted]

Date

[redacted]
Senior Director of Biostatistics, ASLAN Pharmaceuticals

STATISTICAL ANALYSIS PLAN

A MULTICENTER, DOUBLE-BLIND, RANDOMIZED, PLACEBO CONTROLLED STUDY OF VARLITINIB PLUS CAPECITABINE VERSUS PLACEBO PLUS CAPECITABINE IN PATINTS WITH ADVANCED OR METASTATIC BILIARY TRACT CANCER AS SECOND LINE SYSTEMIC THERAPY

Date: 06 December 2018
Version: 3.0

SAP Approvals:

_____ _____

[redacted] *Date*
Chief Medical Officer, ASLAN Pharmaceuticals

_____ _____

[redacted] *Date*
Senior Director of Biostatistics, ASLAN Pharmaceuticals

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	adverse event
AKT	protein kinase b
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
AUC	area under the plasma concentration-time curve
AUC_{τ}	area under the plasma concentration-time curve during a dosage interval (τ)
BID	twice daily
BMI	body mass index
BP	blood pressure
BTC	biliary tract cancer
CI	confidence interval
CL	observed systemic plasma clearance
C_{max}	maximum concentration
C_{trough}	Trough plasma concentration (measured concentration at the end of a dosing interval at steady state [taken directly before next administration])
CRF	case report form
CRO	contract research organization
CTC	Common Toxicity Criteria
DCO	Data cut-off
DCR	disease control rate
5'-DFCR	5'-Deoxy-5-fluorocytidine
5'-DFUR	5'-Deoxy-5-fluorouridine
DoR	duration of response
DSMB	data safety monitoring board
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EDC	electronic data capture
EFR	Evaluable for Response
EGFR	epidermal growth factor receptor
FAS	Full Analysis Set
FDA	Food and Drug Administration
5-FU	5-Fluorouracil

GB	gall bladder
GCP	Good Clinical Practice
HER	human epidermal growth factor receptor
HR	hazard ratio
ICF	informed consent form
ICH	International Council on Harmonisation
ICR	Independent central review
KM	Kaplan Meier
MAD	multiple ascending dose
MedDRA	Medical Dictionary for Regulatory Activities
NTL	Non-Target Lesion
ORR	objective response rate
OS	overall survival
PFS	progression free survival
PK	pharmacokinetic
PP	Per Protocol
Rac AUC	Accumulation ratio calculated from $AUC_{\tau,ss}$ and AUC_{τ} after single dosing
Rac C_{max}	Accumulation ratio calculated from $C_{max,ss}$ and C_{max} after single dosing
ROW	rest of the world
SAE	serious adverse event
SAP	Statistical Analysis Plan
SLI	Safety lead-in
SOP	standard operating procedure
SRC	safety review committee
SUSAR	suspected unexpected serious adverse reaction
$t_{1/2}$	terminal elimination half-life
T_{max}	time of maximum concentration
TL	Target Lesion
TS	tumour size
$\%\Delta TS_{Wk12}$	percentage change from baseline in tumor size at week 12
$\%\Delta TS_{best}$	best percentage change from baseline in tumor size at any time point
US	United States

V_z	Apparent volume of distribution during terminal phase
V_{ss}	steady-state volume of distribution

1. STUDY OBJECTIVES

1.1 Primary Objectives

The primary objectives of the study are:

Safety lead-in:

- To assess the safety and tolerability of varlitinib 300 mg BID (every day), in combination with capecitabine 1000 mg/m^2 (BID for 14 days followed by a 7-day rest) as measured by incidence of AEs, and changes from baseline in safety parameters.

Part 1:

- To assess the efficacy of varlitinib in combination with capecitabine as measured by co-primary endpoints of objective response rate (ORR) and progression-free survival (PFS), both assessed by Independent Central Review (ICR).

Part 2:

- To assess the efficacy of varlitinib in combination with capecitabine as measured by OS

1.2 Secondary Objectives

Safety lead-in

1. To evaluate the pharmacokinetics (PK) of varlitinib and capecitabine (and its metabolite 5-FU) when given in combination

2. To evaluate the effect of varlitinib on QT/QTc

3. To evaluate the efficacy of varlitinib in combination with capecitabine, as measured by ORR, DoR and DCR, all based on site assessment

Part 1

1. To evaluate the efficacy of varlitinib in combination with capecitabine, as measured by OS, DoR, and DCR as assessed by ICR and ORR as assessed by site

2. To assess the safety and tolerability of varlitinib when combined with capecitabine

3. To explore exposure-response relationships for varlitinib (and any relevant circulating metabolites) for measures of efficacy, safety, and pharmacological responses.

4. To examine the effects of varlitinib, when added to capecitabine on ECG parameters including QTcF, QTcB, HR, PR, and QRS

Part 2

1. To evaluate the efficacy of varlitinib in combination with capecitabine, as measured by ORR, DoR, DCR and PFS, all based on site assessment

2. To assess the safety and tolerability of varlitinib when combined with capecitabine

3. To explore exposure-response relationships for varlitinib (and any relevant circulating metabolites) for measures of efficacy, safety, and pharmacological responses via sparse PK sampling and population PK analyses

1.3 Exploratory Objectives

Part 1:

1. To explore the role of human epidermal growth factor receptor (HER) status as a predictor of benefit to varlitinib

2. To explore possible relationships between HER family and downstream signaling protein and phospho-protein expression levels and clinical outcomes

3. To explore possible relationships between gene mutational status and clinical outcomes

Part 2:

If a relationship is found between biomarker(s) expression and clinical outcomes in Part 1 of the study, the biomarker(s) could be prospectively evaluated in Part 2 of the study.

2. Study Design

2.1 Overall Study Design

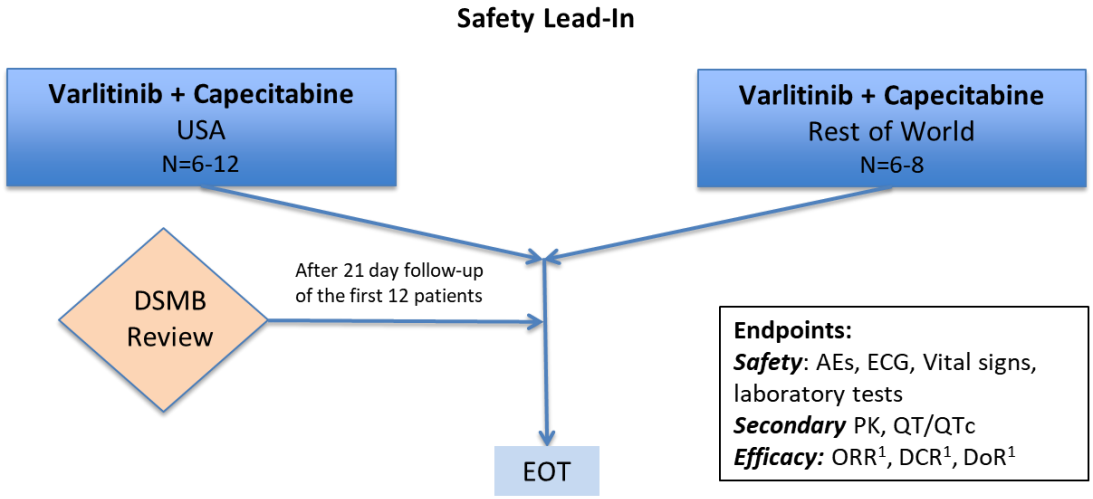
This study is designed as a multicenter, double-blind, randomized, placebo controlled, study of varlitinib plus capecitabine versus placebo plus capecitabine in subjects, with advanced or metastatic BTC as second line of systemic therapy.

Safety Lead-in:

This study will commence initially with a Safety lead-in group (Figure 1). This part of the study is a single arm, open label design to assess the safety of varlitinib (BID, every day) plus capecitabine (administered BID every day for 14 days, followed by a 7 day rest period) in a small set of subjects (12 to 20 subjects) with 12 subjects completing the PK and ECG evaluation. At least 6 subjects will be from the US. Treatment will continue until disease progression or development of toxicity or consent withdrawal or death.

A data safety monitoring board (DSMB) meeting will occur after the first 12 patients in the Safety lead-in have completed 21 days of follow-up in order to review the safety and tolerability data prior to commencement of Part 1 of the study.

Figure 1: Safety Lead-in Study Design

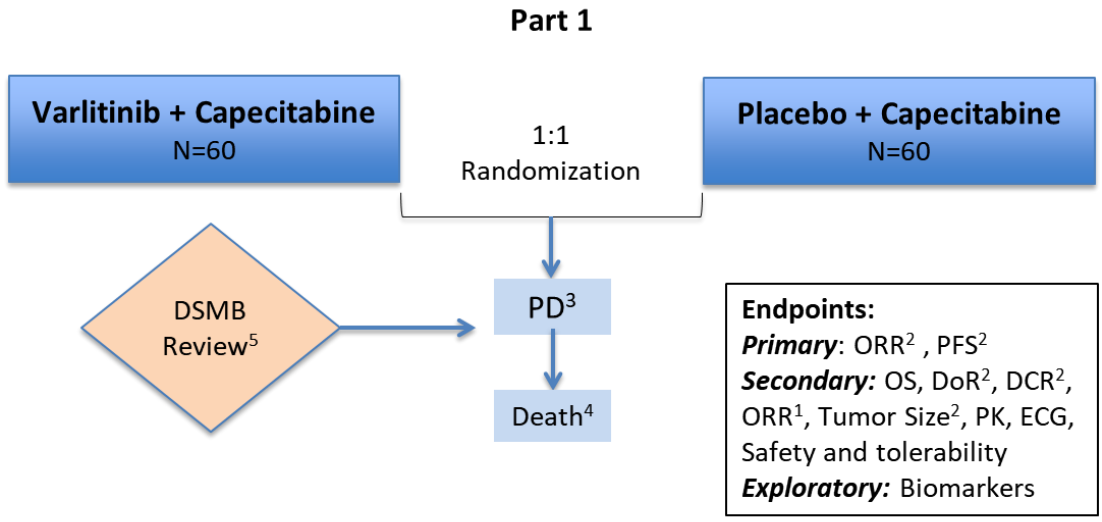


1. As assessed by site review

Part 1:

Part 1 of the study is designed as double-blind, randomized, placebo controlled study to assess the efficacy and safety of varlitinib plus capecitabine in comparison with placebo plus capecitabine in approximately 120 subjects (Figure 2). Eligible subjects will be randomized to varlitinib or placebo (BID, every day)), each in combination with capecitabine (administered BID every day for 14 days, followed by a 7 day rest period). Treatment will continue until disease progression or development of toxicity or consent withdrawal or death. A DSMB will meet at a selected time point (to be confirmed) during Part 1 for review of safety data.

Figure 2: Part 1 Study Design

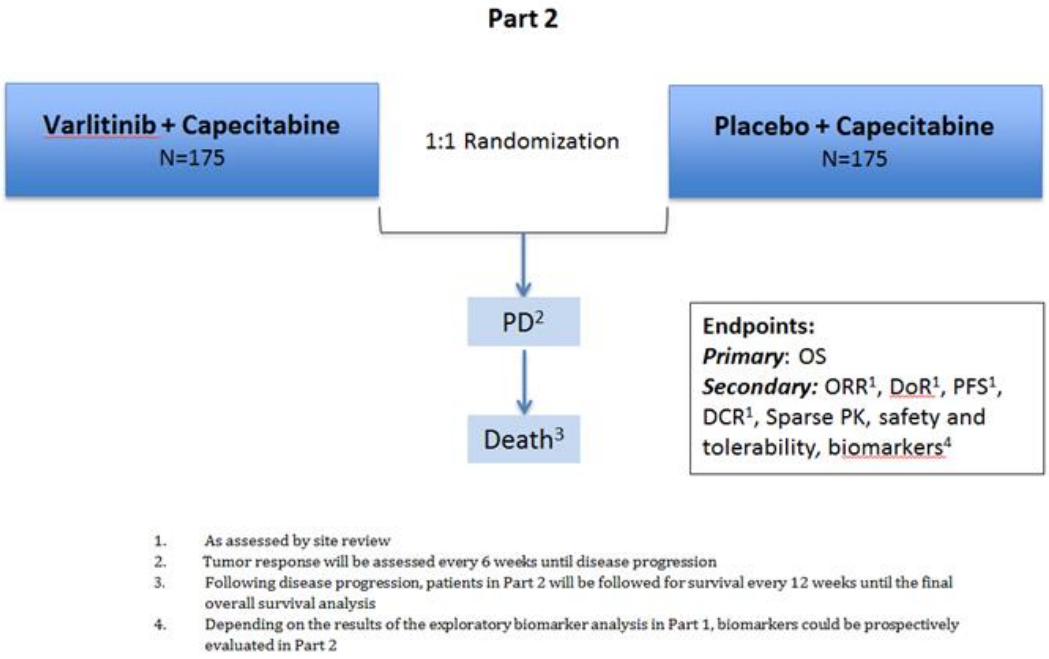


1. As assessed by site review
2. As assessed by Independent Central Review
3. Tumor response will be assessed every 6 weeks until disease progression
4. Following disease progression, patients in Part 1 will be followed for survival every 12 weeks until the DCO for the primary analysis of Part 2
5. An additional DSMB review on safety data will be performed during Part 1, as described in DSMB charter

Part 2:

Part 2 of the study is designed as double-blind, randomized, placebo controlled study to confirm the efficacy of varlitinib plus capecitabine compared to capecitabine in approximately 350 subjects (Figure 3). Part 2 will follow the same design and treatment schedule as Part 1. This part of the study will commence following review of Part 1 results and will be independently powered.

Figure 3: Part 2 Study Design



1. As assessed by site review
2. Tumor response will be assessed every 6 weeks until disease progression
3. Following disease progression, patients in Part 2 will be followed for survival every 12 weeks until the final overall survival analysis
4. Depending on the results of the exploratory biomarker analysis in Part 1, biomarkers could be prospectively evaluated in Part 2

2.2 Methods of Assigning Subjects to Treatment Groups

Safety Lead-in:

This is a non-randomized single arm design. All eligible subjects will receive varlitinib 300 mg twice daily, every day plus capecitabine 1000mg/m^2 twice daily, Day 1 to 14, followed by a 7-day rest period (21 day cycle).

Each subject's eligibility will be reviewed by medical monitor during screening period, subjects will be sequentially enrolled after being screened for eligibility and providing consent.

In the case of discontinuation of chemotherapy owing to toxic effects, varlitinib or placebo therapy will be continued until disease progression, the development of unacceptable toxic effects, withdrawal of consent or death.

At screening, potential subjects will be assigned a unique screening number that starts with "S" and 4 digits (i.e. S1001) by electronic data capture system.

Parts 1 and 2

At the initial screening visit, the investigator or suitably trained delegate will:

1. Obtain signed informed consent (main study) from the potential subject before any study specific procedures are performed.

2. Potential subjects will be assigned a unique screening number that starts with "S" and 4 digits (i.e. S1001) by electronic data capture system. This screening number will be sequential from safety lead in, and will not be repeated.

3. Determine subject eligibility in accordance with the protocol. Once the subject is confirmed to be eligible, the PI or suitably trained delegate will:

4. Assign eligible subject a unique randomization number via Randomization and Trial Supply Management (RTSM) system. Randomization numbers will start with "R" followed by 4 digits (i.e., R1001) and will be assigned strictly sequentially by RTMS (starting at R1001 for Part 1 and R2001 for Part 2), as subjects are eligible for randomization. This number is the subject's unique identifier and is used to identify the subject with the assigned treatment. This randomization number will be available to the investigator in electronic data capture system.

2.2.1 Procedures for Randomization

Subjects must not be randomized unless all eligibility criteria have been met.

Subjects who satisfy all the entry criteria will be centrally assigned to study treatment by the RTSM, according to the randomization scheme generated by the CRO. Subjects will be randomized in a 1:1 ratio to either varlitinib or placebo, both on a background of capecitabine. The random scheme will be stratified by the following two covariates:

- Geographical region - US or Rest of World (ROW)
- Primary tumor location – Gall bladder (GB) or non-GB

The actual treatment given to subjects will be determined by the randomization scheme in RTSM. The randomization scheme will be produced using SAS® version 9.4, which incorporates a standard procedure for generating randomization numbers. One randomization list will be produced containing sufficient random numbers for each of the four randomization strata. A blocked randomization will be generated and all centers will use the same list in order to minimize any imbalance in the number of subjects assigned to each treatment group.

The dose and regimen of randomized therapy to be administered is as follows:

- Varlitinib 300 mg twice daily, every day
- Placebo twice daily, every day

Subjects in both arms will receive background chemotherapy with capecitabine 1000 mg/m^2 twice daily, Day 1 to 14, followed by a 7-day rest period (21 day cycle).

Subjects will be identified to the Centralized Randomization Center using site number, screening number, randomization number, gender and date of birth. Subjects will be randomized strictly sequentially within each stratum, as subjects are eligible for randomization. The RTSM will inform the investigator of the Kit ID number to be allocated to the subject at each dispensing visit via electronic data capture (EDC).

For both study parts, the primary and key secondary endpoints will be evaluated using the Full Analysis Set (FAS) based on the intention-to-treat principle.

3. ENDPOINTS

3.1 Efficacy Endpoints

3.1.1 Objective Response Rate

Objective Response Rate is a co-primary endpoint in Part 1 and a secondary endpoint in Part 2, and the Safety lead-in cohort.

ORR rate is defined as the number (%) of subjects with at least one visit response of CR or PR. Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of ORR. This will be irrespective of whether or not subjects discontinued treatment or received a subsequent therapy prior to progression.

In all study parts, the primary assessment of ORR will be based on the full analysis set (FAS). Sensitivity analysis populations are described in Section 4.

The ORR is derived programmatically from the Best overall RECIST Response (BoR) for each subject.

Best overall RECIST Response (BoR) is calculated based on the overall visit responses from each RECIST assessment. It is the best response a subject has had following randomization and prior to RECIST progression or the last evaluable assessment in the absence of RECIST progression.

Categorization of BoR will be based on the RECIST criteria using the following response categories (listed in hierarchical order): complete response (CR), partial response (PR), stable disease (SD), progressive disease (PD) and inevaluable (NE) (CR>PR>SD>PD>NE).

Target Lesions (TL)	**Non-Target Lesions (NTL)**	**New Lesions**	**Overall Response**
CR	CR	NO	CR
CR	Non-CR/Non- PD	NO	PR
CR	Not evaluated	NO	PR
PR	Non- PD or not all evaluated	NO	PR

Target Lesions (TL)	Non-Target Lesions (NTL)	New Lesions	Overall Response
SD	Non- PD or not all evaluated	NO	SD
Not all evaluated	Non- PD	NO	NE
PD	ANY	YES or NO	PD
ANY	PD	YES or NO	PD
ANY	ANY	YES	PD

For the evaluation of the co-primary endpoint ORR in Part 1 of the study, all radiological data will be assessed by an Independent Central Review (ICR), and the BoR will be determined programmatically using the data provided by the ICR. For Part 2, the Safety lead-in and for the Part 1 secondary endpoint, ORR based on site assessments, BoR will be determined programmatically using data from the site assessments

For determination of a best response of SD, the earliest of the dates contributing towards a particular overall visit assessment will be used. A best response of SD cannot be assigned unless a visit response of SD has been recorded at least 6 weeks (-5 days) after randomization (day 1), i.e. at least 38 days (to allow for the assessment window). For CR/PR, the initial overall visit assessment that showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

For subjects whose progression event is death, BoR will be calculated based on data up until the last evaluable RECIST assessment prior to death.

For subjects who die with no evaluable RECIST assessments, if the death occurred ≤ 12 weeks + 5 days after randomization, then BoR will be assigned to the progression (PD) category. For subjects who die with no evaluable RECIST assessments, if the death occurred >12 weeks +5 days after randomization then BoR will be assigned to the inevaluable (NE) category.

Progression events that have been censored due to them being more than two missed visits after the last evaluable assessment will not contribute to the BOR derivation.

A subject will be classified as a responder if the RECIST criteria for a CR or PR are satisfied at any time following randomization, prior to RECIST progression (regardless of the number of missing visits prior to the assessment of CR or PR). In accordance with RECIST 1.1, response confirmation is not required.

For each treatment group, the objective response rate (ORR) is the number of responders divided by the number of subjects in the treatment group in the FAS analysis set. This will be irrespective of whether or not subjects discontinued treatment or received a new antitumor therapy prior to disease progression.

The NE category will be assigned in the following circumstances:

- For subjects who die with no evaluable post-baseline assessments, if the death occurred > 12 weeks + 5 days after randomization (i.e., if more than 2 assessments are missed).
    - Note: Subjects who die up to the week 12 assessment (scheduled up to 12 weeks + 5 days after randomization), with no evaluable post-baseline assessment will be classified as progressors and will be treated as progressors for the waterfall plots of the percentage change from baseline in tumor size, and the statistical analysis of tumor size (see 3.1.6 and Section 7.6)
    - For subjects with no evaluable post-baseline assessments, and for whom the date of progression (including death) is either unknown or censored due to missed assessments (see Section 3.1.2).

3.1.2 Progression-Free Survival

PFS is defined as the time from randomization (or starting treatment for the Safety lead-in) until the date of objective disease progression or death (by any cause in the absence of disease progression) regardless of whether the subject withdraws from randomized therapy or receives another antitumor therapy prior to disease progression. Subjects who have not experienced disease progression or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST assessment. However, if the subject progresses or dies after 2 or more missed visits (12 weeks + 5 days maximum), the subject will be censored at the time of the latest evaluable RECIST Version 1.1 assessment.

The PFS time will always be derived based on the scan/assessment dates rather than visit dates and the following rules will be applied:

- Date of disease progression will be determined based on the **earliest** of the dates of the component that triggered the disease progression, i.e., if both the target lesions and the non-target lesions indicate disease progression but were scanned on different days, the earlier of the 2 dates would be applied.
- When censoring a subject for PFS the subject will be censored at the **latest** of the dates contributing to a particular overall visit assessment.

For Part 1, the co-primary endpoint, PFS, will be derived programmatically using data from the ICR. To support sensitivity analyses, PFS will also be derived programmatically using the site assessments of the Part 1 radiological data.

For Part 2 and the Safety lead-in, PFS will be derived programmatically using the site assessments of the radiological data.

In all study parts, the primary analysis of PFS will be assessed based on the FAS (see Section 4.2).

3.1.3 Overall Survival

OS is defined as the time from the date of randomization until death due to any cause. Any subject not known to have died at the time of the data cut-off will be censored based on the last recorded date on which the subject was known to be alive.

Note, survival calls may be made in the week following the date of Data Cut Off (DCO) for the analysis, and if subjects are confirmed to be alive or if the death date is after the DCO these subjects will be censored at the date of DCO.

3.1.4 Duration of Response

The DoR is defined as the time from the date of first documented response until the date of documented disease progression or death in the absence of disease progression. The end of response should coincide with the date of disease progression or death from any cause used for the PFS endpoint.

For Part 1, DoR will be calculated based on data from the ICR of radiological data. For Part 2 and the Safety lead-in, DoR will be calculated based on the site assessments of radiological data.

The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of PR or CR. For example, if the subject was first noted to have a PR at week 6, and the target and non-target lesions were assessed on different dates at this visit, then the later of the two assessment dates would be used as the start date of the response.

If a subject does not progress following a response, then their DoR will use the PFS censoring time.

Duration of response will not be defined for those subjects who do not have documented response.

3.1.5 Disease Control Rate

DCR rate is defined as the number (%) of subjects with at least one visit response of CR or PR, or with stable disease for a minimum of twelve weeks (−5 days) from randomization. Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of DCR. This will be irrespective of whether or not subjects discontinued treatment or received a subsequent therapy prior to progression. For all study parts, the primary assessment of DCR will be based on the FAS. For Part 1, DCR will be calculated based on data from the ICR. For Part 2 and the Safety lead-in, DCR will be calculated from the site radiological data.

3.1.6 Tumor Size

The percentage change from baseline in tumor size at Week 12 ($\%\Delta TS_{Wk12}$) is a secondary endpoint for Part 1, and will be used to present waterfall plots of the target lesion data in the Evaluable for Response (EFR) set (see Section 4.3), and also to help evaluate the exploratory objectives of investigating the possible role of HER family expression levels as predictors of clinical benefit. Tumor size is not considered to be a key secondary endpoint; for the purpose of marketing authorization, no statistical claims will be made based on tumor size.

The $\%\Delta TS_{Wk12}$ will be defined as follows:

- Baseline tumor size (TS): defined as the sum of longest diameters of target lesions at baseline
- Week 12 TS: defined as the sum of longest diameters of target lesions at week 12

$$\%\Delta TS_{Wk12} = 100 \times \frac{Week\ 12\ TS - Baseline\ TS}{Baseline\ TS}$$

An increase from baseline of 20% will be imputed for any subjects who die by any cause, or withdraw from the study due to symptomatic disease progression prior to Week 12 (+5 days). Any subjects who withdraw from the study due to radiological disease progression prior to Week 12 will be included in the analysis, and will be assigned a percentage change from baseline of 20%, or the value that was observed at the time of disease progression (whichever is larger).

The best percentage change from baseline in tumor size ($\%\Delta TS_{best}$) at any time point, will also be calculated and presented using waterfall plots.

The $\%\Delta TS_{best}$ will be defined as follows:

- Baseline tumor size (TS): defined as the sum of longest diameters of target lesions at baseline
- Best TS: defined as the smallest sum of longest diameters of target lesions as observed at any time point post-randomization up until disease progression (regardless of whether it was a scheduled assessment).

$$\%\Delta TS_{best} = 100 \times \frac{Best\ TS - Baseline\ TS}{Baseline\ TS}$$

Subjects in the EFR set without any evaluable post-randomization target lesion data will be excluded from the waterfall plot of $\%\Delta TS_{best}$, except if the subject is known to have died or withdrawn from the study due to symptomatic disease progression prior to Week 12 (+ 5 days), or if non-target lesion or new lesion data is available prior to week 12 (+5 days) which is indicative of disease progression. In such cases, the subject will have an increase from baseline of 20% imputed.

3.2 Safety and Tolerability

The safety and tolerability endpoints of the study, including the ECG parameters, are listed in the following sections. Further details of the schedule of assessments for these endpoints can be found in the Clinical Study Protocol.

Due to the non-comparative design and small sample size, the evaluation of the impact of varlitinib on ECG parameters from the Safety lead-in will be considered exploratory.

The primary electrocardiographic analysis of the protocol will be from the time-based ECG analysis performed in Part 1 at baseline, Cycle 1 Day 1, Cycle 1 Day 8, and the safety follow-up visit. Triplicate ECGs will be collected with a PK sample collected immediately following ECG acquisition at the time points outlined in the Clinical Study Protocol. ECGs will be analyzed by a Core ECG Laboratory. The planned methods of analysis for the ECG parameters, including the exposure-response modeling, are outlined in Sections 7.7.3 and 7.11.4.

3.2.1 Safety Endpoints

The assessment of safety and tolerability is the primary objective of the Safety lead-in, and a secondary objective of Parts 1 and 2. In all study parts, the safety of varlitinib 300 mg BID in combination with capecitabine, in subjects with BTC, will be assessed using the following endpoints:

- Adverse Events (AEs)
- Vital Signs (blood pressure [systolic and diastolic], heart rate, respiration rate, and body temperature)
- Laboratory assessments (clinical chemistry, hematology, coagulation and urinalysis)
- 12-lead Electrocardiogram (ECG) (P and QRS duration, HR, PR, RR, PQ, QT and QTc intervals using Bazett's and Fridericia's formulae)
- Physical examination (including a standard neurological examination and assessment of weight)

In addition, tolerability will be further explored by assessment of dose reductions, interruptions and modifications (see Section 3.2.2).

Adverse events will be coded by System Organ Class (SOC) and Preferred Term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA) for concomitant diseases and AEs and WHO Drug for medications. All AEs will be evaluated for severity according to NCI-CTCAE version 4.03

Baseline will be defined as the last assessment prior to the start of study treatment. If multiple assessments are performed at the same pre-treatment time point, the mean of the values will be used as the baseline reading.

Specifically, for the ECG assessments, on Day 1, pre-dose is defined as 2 sets of triplicate ECGs taken 45 minutes prior to dosing and approximately 5 minutes prior to dosing. The ECG values from these 2 sets of ECGs will be averaged to create the baseline value. Similarly, for all other time points, for each ECG parameter, each patient's mean will be calculated from the triplicate readings and the mean value will be used.

The ECG parameters of interest include, but are not limited to: QT including QTcF and QTcB, HR (heart or ventricular rate), PR, QRS intervals.

Creatinine clearance will be derived as:

$$eC_{Cr} = \frac{(140 - Age) \times Mass\ (in\ kilograms) \times Constant}{Serum\ Creatinine\ (in\ \mu mol/L)}$$

where constant=1.23 for males and 1.04 for females, and will be categorized as 30-<60 ml/min, 60-<90 ml/min and =>90 ml/min. Creatinine clearance will be included in summaries of the clinical chemistry data.

QTcB will be derived as:

$$QTcB = \frac{QT}{\sqrt{RR}}$$

QTcF will be derived as:

$$QTcF = \frac{QT}{\sqrt[3]{RR}}$$

For derivation of post-baseline visit values, taking account of possible visit windows, see Section 3.2.1.1.

3.2.1.1 General Considerations for Safety Assessments

Missing safety data will not be imputed. However, safety assessment values of the form of "< x" (i.e., below the lower limit of quantification) or > x (i.e., above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings.

In order to produce data summaries by visit, visit time windows are required, and will be derived using the following conventions:

- The time windows will be exhaustive so that data recorded at any time point have the potential to be summarized. Inclusion within the time window should be based on the actual date and not the intended date of the visit.
- All unscheduled visit data should have the potential to be included in the summaries.
- The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 1). If an even number of days exists between two consecutive visits then the upper limit will be taken as the midpoint value minus 1 day.

Furthermore, the following principles will apply to the presentation of safety data:

- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).
- For summaries of the ECG data, the mean value of the triplicate readings will be included in the summaries.
- Listings should display all values contributing to a time point for a subject.
- For visit based summaries:
    - If there is more than one value per subject within a time window then the closest value should be summarized, or the earlier in the event the values are equidistant from the nominal visit date. The listings should highlight the value for that subject that went into the summary table, wherever feasible.
    - For Parts 1 and 2 only, to prevent very large tables or plots being produced that contain many cells with sparse data, for each treatment group, visit data should only be summarized if the number of observations is greater than 1/3 of subjects dosed.

- For summaries at a subject level, all values should be included, regardless of whether they appear in a corresponding visit based summary, when deriving a subject level statistic such as a maximum.
- Baseline will be defined as the last non-missing measurement prior to dosing with study treatment. For laboratory data, any assessments made on day 1 will be considered pre-dose unless explicitly indicated otherwise. Where safety data are summarized over time, study day will be calculated in relation to date of first treatment (which will be termed day 1).

3.2.2 Treatment Exposure and Dose Intensity

To understand the general tolerability of varlitinib 300 mg BID, and the impact of interruptions and dose reductions, summaries of treatment exposure and dose intensity will be produced. For the purpose of this study, the term *exposure* relates to the number of days of treatment received (regardless of dose reductions), whereas *dose intensity* relates to the proportion of intended dose (300 mg BID) received, taking account of interruptions and reductions.

For all subjects, exposure to the randomized therapy (varlitinib or placebo) in Parts 1 and 2, or study therapy (varlitinib) in the Safety lead-in, will be calculated. To assess treatment exposure, the following will be calculated:

Intended Exposure to randomized therapy (days)

- Intended Exposure = last dose date – first dose date + 1

Actual Exposure to randomized therapy (days)

- Actual exposure = Intended exposure – total duration of dose interruptions

where the intended exposure is calculated as defined above, and dose interruptions include all days in which the subject receive no study medication, regardless of whether the interruption was intentional, or whether the subject forgot to take a dose.

For capecitabine intended exposure, the calculation of intended exposure and actual exposure will take into account the intermittent dosing schedule, as described below:

For intended exposure, if the last dose date is in the first 14 days of cycle X, then the intended exposure will be:

14 * (X-1) + last dose date – first dose date of cycle X +1

If the last dose date is in days 15-21 of cycle X (when there should be no dosing), then the intended exposure will be 14 * X.

For actual exposure, this will be determined as the intended exposure minus the total duration of any non-protocolled interruptions. The protocolled 7 days off in each cycle are accounted for in the derivation of intended exposure so it should not be included in the non-protocolled interruptions in this calculation.

For Parts 1 and 2, dose intensity will be calculated for subjects in both treatment arms, and will be calculated for both the randomized therapies (varlitinib vs. placebo) and also for the background chemotherapy, capecitabine.

For the open-label Safety lead-in, dose intensity will be calculated for both the study treatment, varlitinib, and for the background chemotherapy, capecitabine.

Two measures of dose intensity will be calculated:

1. Percentage of Intended Dose (PID) – defined as the percentage of the actual dose delivered relative to the intended dose, *until disease progression*.
2. Relative dose intensity (RDI) – defined as the percentage of the *actual* dose intensity delivered, relative to the *intended* dose intensity, *until the earlier of treatment discontinuation and progression* (as defined by RECIST and based on the site assessments).

The key difference between the two measures outlined above is that the former is based on the intended treatment plan of treatment until objective progression, and the latter takes account of treatment discontinuations prior to progression, and calculates the dose intensity during the actual dosing period.

In the event that the investigator does not discontinue randomized therapy at disease progression, any dosing information beyond progression will not be taken into account when calculating either measure of dose intensity.

PID and RDI will be defined as follows:

PID = 100% * d/D, where d is the actual cumulative dose delivered up to progression (or a censoring event) and D is the intended cumulative dose up to progression (or a censoring event). D is the total dose that would be delivered, if there were no modification to dose or schedule.

RDI = 100% * d/D, where d is the actual cumulative dose delivered up to the earlier of progression (or a censoring event) or the actual last day of dosing and D is the intended cumulative dose up to the earlier of progression (or a censoring event) or the actual last day of dosing.

To account for the intermittent dosing schedule for capecitabine, the intended cumulative dose at day x for capecitabine will be calculated as follows:

If day x is in the first 14 days of a cycle:

(Ceil(x/21)-1)*14*2*dose level + (x-(Ceil(x/21)-1)*21)*2*dose level

If day x is day 15 to 21 of a cycle:

(Ceil(x/21))*14*2*dose level

3.3 Pharmacokinetic Endpoints

Safety Lead-in

The PK of varlitinib (and any relevant circulating metabolites) and capecitabine and 5-FU (active metabolite of capecitabine) will be determined at selected time points as outlined in the Clinical Study Protocol.

The following PK parameters will be evaluated where possible:

Cycle 1 Day 1, Day 8 and Day 14:

- Maximum plasma concentration (C_{max})
- time to C_{max}(t_{max})
- plasma concentration before next dose (C_{trough})
- area under the plasma concentration-time curve from 0 to 12 hours ($AUC_{0-\tau}$)
- half-life ($t_{1/2}$)
- apparent clearance (Cl/F)
- apparent volume of distribution (V_z/F)
- apparent volume of distribution at the steady state (V_{ss}/F)
- where τ = 12 hours (dosing interval)

Cycle 1 Day 8

- accumulation ratio of $AUC_{0-\tau}$ ($RacAUC_{0-\tau}$) (Day 8/Day 1)
- accumulation ratio of C_{max} $RacC_{max}$ (Day 8/Day 1)

Cycle 1 Day 14

- accumulation ratio of AUC0- τ (RacAUC0- τ) (Day 14/Day 1)
- accumulation ratio of C_{max} $RacC_{max}$ (Day 14/Day 1)

In addition, unscheduled PK samples will be collected at least 7 days following dose reduction as described in the clinical study protocol, if required.

Part 1 and Part 2

PK samples will be obtained for the PK evaluation of varlitinib (and any relevant circulating metabolites). PK samples will be collected at the time points outlined in the Clinical Study Protocol.

PK sampling of all patients is planned for Part 1 of the study while sparse PK sampling is planned for Part 2.

In order to maintain the double-bind design of the study, blood samples will be collected from all subjects in Part 1 for plasma measurements of varlitinib and its potential relevant circulating metabolites. For the sparse PK sampling in Part 2, the double-blind design will also be maintained and blood samples will be collected for a selected number of subjects, without limiting the sampling to the varlitinib subjects only.

In addition, unscheduled PK samples will be collected at least 7 days following dose reduction as described in the clinical study protocol.

3.4 Biomarker Endpoints

The following IHC biomarker endpoints will be assessed, where possible:

- epidermal growth factor receptor (EGFR), pEGFR
- HER2, pHER2
- HER3, pHER3
- HER4, pHER4
- ERK, pERK
- protein kinase b (AKT) AKT, pAKT

For the HER-family biomarkers (EGFR, pEGFR, HER2, pHER2, HER3, pHER3, HER4, pHER4), the original biomarker levels (0, +, ++ or +++) will be used to create binary positive/negative categorizations, to support the exploratory analyses to assess the role of baseline HER-family expression levels as predictors of clinical benefit. For each endpoint, the binary categorizations will be defined as follows:

Positive: +, ++ or +++

Negative: 0

Biomarkers assessed via PCR/Sequencing, for example the mutational status of genes, such as *KRAS, NRAS, BRAF*, *EGFR* and other genetic factors that may affect response to therapy, will not be reported in the CSR and are not covered in this SAP.

4. Analysis Populations

4.1 Safety Population

For each part of the study, the safety population includes all subjects in that part of the study who received at least 1 dose of randomized therapy (or study medication in the Safety lead-in) and will be the primary analysis set for the assessment of safety and tolerability in the study.

For the purpose of data summaries, subjects will be included in the Safety Population according to the treatment arm and dose level initially received, regardless of any subsequent dose adjustments.

4.2 Full Analysis Set

For each study part, the full analysis set is based on the *intention-to-treat* principle and includes all randomized subjects (or treated subjects for the Safety lead-in) analyzed in accordance with the intended treatment arm, regardless of the treatment actually received.

With the exception of Tumor Size (which is not considered a key secondary endpoint; see Section 3.1.6 and 4.3) in Part 1 of the study, the primary analysis of all primary and key secondary efficacy endpoints, will be based on the full analysis set (FAS) for all study parts.

4.3 Evaluable for Response Set (Part 1 only)

The EFR set is defined as a subset of the FAS and includes subjects with measurable disease at baseline, as determined by ICR assessments of radiological data based on the RECIST criteria, version 1.1. For Part 1 only, a sensitivity analysis of ORR and DCR will be performed in the evaluable for response (EFR) set to investigate the impact of subjects without measurable disease on the primary endpoint. Since measurability is an inclusion criteria, all subjects in Parts 1 should have measurable disease at baseline. Thus, if no subjects were randomized in error, results in the EFR set will not be presented.

Since tumor size is an endpoint based on subjects' tumor measurements, it can only be defined for the subset of the FAS with measurable disease at baseline. Thus the EFR set will be the primary analysis set used for the assessment of tumor size for Part 1 (See Section 3.1.6).

4.4 Per Protocol Analysis Population

The per-protocol analysis population is defined for Part 1 only. Efficacy is not the focus of the Safety lead-in and Part 2 will be assessed on an intention to treat basis only. Thus PP analyses are not required for the Safety lead-in and Part 2 of the study.

The per-protocol analysis population will include all randomized subjects according to the treatment actually received, and excluding any subjects with major deviations. Major deviations that would lead to exclusion from the per-protocol analysis population include:

- Subjects who did not have the intended disease (failure of inclusion criteria 2)
- Subjects who did not have the intended indication, i.e. were not second line patients (failure of inclusion criteria 3)
- Subjects who did not received any randomized therapy
- Subjects with baseline radiological scan performed before the start of protocol-specified interval prior to randomization
- Subjects who did not have measurable disease at baseline as determined by ICR
- Subjects who have not received at least six doses of gemcitabine in a first line setting (i.e. failure of inclusion criteria 4, protocol version 4.0 onwards)
- Subjects with multiple (≥2) peritoneal metastases at baseline or ascites at baseline which cannot be attributed to a non-malignant cause (i.e. failure of exclusion criteria 3, protocol version 3.0 onwards), based on ICR review
- Subjects with baseline albumin <3 g/dL (failure of the albumin component of inclusion criteria 9c)

In addition to the programmatic determination of the major deviations, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification will be made blinded, prior to database lock. For example, details of disallowed concomitant medication use will be reviewed by a physician, blinded to treatment allocation, and may be deemed major.

Subjects who received the incorrect randomized therapy should be included in the PP analysis set according to the therapy actually received. Similarly, subjects who were mis-stratified at randomization will be included in the PP analyses according to the correct stratification levels, regardless of the levels declared at randomization.

Note, failure of an inclusion/exclusion criteria will not automatically be classified as a major deviation.

The per-protocol analysis population will be used to assess the sensitivity of the Part 1 ORR, DoR, DCR, PFS and OS analyses to major deviations.

4.5 PK Analysis Set

The PK population contains all subjects who have been dosed with active drug(s) and provided at least one usable post-dose PK sample. All PK data will be analyzed according to treatment received.

For the Safety lead-in, this population will comprise all data subjects who receive study treatment as per protocol (300 mg BID) and did not violate or deviate from the protocol and planned dosing regimen (300 mg) in ways that would significantly affect the PK analyses (for example skipping doses, or taking reduced doses or taking concomitant medications with the potential to cause a drug-drug interaction) during the PK sampling period. Subjects who did deviate from the planned dosing regimen, may still provide some data for inclusion in the PK set, if they have at least one usable post-dose PK sample. The population, and decisions regarding which profiles are usable, will be defined by the Study Team Physician, Pharmacokineticist and Statistician prior to any analyses being performed based on review of protocol deviations.

PK sampling of all patients is planned for Part 1 of the study while sparse PK sampling is planned for Part 2 (See Section 3.3). The number of patients who will provide samples in Part 2 for the PK analysis will be defined based on the results of Part 1 of the study. All PK data will be analyzed according to the treatment received. Deviations that have the potential to significantly affect the PK will be reviewed, and subjects, or specific PK profiles for a subject, may be excluded from the PK set if appropriate. The population, and decisions regarding which profiles are usable, will be defined by the Study Team Physician, Pharmacokineticist and Statistician prior to unblinding and any analyses being performed.

4.6 Summary of Analysis Populations

A summary of the study outcome variables, and associated analysis sets is provided in Table 1.

Table 1 Summary of Outcome Variables and Analysis Populations

Outcome Variable	Analysis Population (Sensitivity Analysis Population)
Efficacy Data (Primary and key secondary endpoints)	
ORR, DCR,	FAS (EFR and PP Part 1 only)
PFS, OS, DoR	FAS (PP Part 1 only)

Part 1 ORR based on site data	FAS
Part 1 Tumor Size[1]	EFR
Demography	FAS
Pharmacokinetics	
PK	PK
Safety	
Exposure	Safety
Adverse Events	Safety
Laboratory measurements	Safety
Vital Signs	Safety
ECG	Safety
Physical Examination	Safety

[1] Tumor size, by definition, is only defined in patients with measurable disease. For marketing authorization, no statistical claims will be based on tumor size (See Section 3.1.6).

5. Protocol Deviations

5.1 Major Deviations

For Part 1, major protocol deviations are defined as those that would lead to exclusion from the per-protocol set (see Section 4):

- Subjects who did not have the intended disease (failure of inclusion criteria 2)
- Subjects who did not have the intended indication, i.e. were not second line patients (failure of inclusion criteria 3)
- Subjects who did not received any randomized therapy.
- Subjects with baseline radiological scan performed before the start of protocol-specified interval prior to randomization.
- Subjects who did not have measurable disease at baseline as determined by ICR
- Subjects who have not received at least six doses of gemcitabine in a first line setting (i.e. failure of inclusion criteria 4, protocol version 4.0 onwards)
- Subjects with multiple (≥2) peritoneal metastases at baseline or ascites at baseline which cannot be attributed to a non-malignant cause (i.e. failure of exclusion criteria 3, protocol version 4.0 onwards), based on ICR review
- Subjects with baseline albumin <3 g/dL (failure of the albumin component of inclusion criteria 9c)

For the Safety lead-in and Part 2, the following will be classified as major deviations:

- Subjects who did not have the intended disease (BTC) or indication (2nd line patients).
- Subjects who did not received any randomized therapy (Part 2) or Study treatment (Safety lead-in).

- Subjects with baseline radiological scan performed before the start of the protocol-specified interval prior to randomization (Part 2) or start of treatment (Safety lead-in).

Additionally, the following will be considered major deviations for Part 2 only (but not the SLI):

- Subjects who have not received at least six doses of gemcitabine in a first line setting (i.e. failure of inclusion criteria 4, protocol version 4.0 onwards)
- Subjects with multiple (≥2) peritoneal metastases at baseline or ascites at baseline which cannot be attributed to a non-malignant cause (i.e. failure of exclusion criteria 3, protocol version 4.0 onwards), based on ICR review
- Subjects with baseline albumin <3 g/dL (failure of the albumin component of inclusion criteria 9c, protocol version 4.0 onwards)

For each stage of the study, these protocol deviations will be listed. For Parts 1 and 2, major deviations will be summarized by treatment group.

In addition to the programmatic determination of the major deviations, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification will be made blinded to treatment allocation (Parts 1 and 2) and prior to database lock. For example, details of disallowed concomitant medication use will be reviewed by a physician and may be deemed major.

Note, failure of an inclusion/exclusion criteria will not automatically be classified as a major deviation.

5.2 Important Deviations

In addition to the programmatic determination of the major deviations as described in Section 5.1, the following *important deviations* will be programmatically derived and listed and summarized:

- Failure of any inclusion/exclusion criteria (with exception of any inclusion/exclusion criteria that are already listed as major deviations)
- Prior anti-cancer treatments which stopped less than 4 weeks before randomization (Parts 1 and 2)/starting study medication (Safety lead-in).
- Subjects taking concurrent therapies for BTC.
- Subjects taking continuous treatment with proton pump inhibitors during the study period
- Subjects taking strong CYP3A4 inhibitors
- For the Safety lead-in only: <85% compliance during the first 14 days

Errors in dispensing and incorrect stratifications will be summarized and listed as *important deviations*. These include:

- Subjects who receive no study treatment whatsoever for a period of more than one week due to errors in dispensing of medication. Note, this is not due to tolerability issues where subjects may stop taking drug

- Subjects who at some point receive the incorrect treatment (e.g. varlitinib rather than placebo or vice-versa)
- Subjects who were inadvertently dispensed the wrong dose of capecitabine

The summary will include all subjects with a dispensing error but will also include information on how many of those subjects received at least one dose of the incorrect treatment (varlitinib rather than placebo or vice-versa) at any time. Subjects who receive the incorrect treatment at any time will be included in the Safety Set as described in Section 4.1.

In addition, other non-programmable important deviations will be collected on the CRF, listed and summarized including, but not limited to:

- For the safety-lead-in only, subjects who were replaced due to <85% varlitinib compliance during first 14 days
- Subjects of child-bearing potential not practicing acceptable methods of birth control from the time they enter the screening period until 3 months after the last cycle of treatment.
- Subjects breast feeding at any point between the time they enter the screening period until 3 months after the last cycle of treatment.

Compliance will be assessed by Study Team review of protocol procedures with specific attention paid to missed dosing due to reasons other than AEs. (E,g, subjects receiving <75% of intended doses) and will not be assessed programmatically. Any compliance issues considered likely to impact on the interpretation of safety or efficacy will be added to the deviation CRF page.

Finally, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification will be made prior to database lock. For example, details of disallowed concomitant medication use will be reviewed by a physician and if deemed important, may also be included in the listing of important deviations.

All subjects who failed any inclusion/exclusion criteria will be listed along with details of the failed criteria. This information will also be summarized in terms of the number (%) of subject failing any of the inclusion/exclusion criteria and will be based on the FAS. Note, failure of an inclusion/exclusion criteria will not automatically be classified as a major deviation.

Other deviations may occur during the trial, which are considered minor and not believed to have any significant impact on the interpretation of the study results. All of these deviations will be recorded by the study monitors but will not be listed or summarized as part of the CSR.

Examples of minor deviations include:

- Visits outside of protocol visit windows
- Incomplete assessments
- Use of concomitant medication related to safety

This list is not exhaustive and other minor deviations may occur.

6. General Statistical Considerations

All statistical analyses will be performed by or under the direct auspices of ASLAN Pharmaceuticals, using SAS®, Version 9.2 or higher.

Unless specified otherwise, data from the Safety lead-in, Part 1 and Part 2 will be presented separately in the Tables, Figures and Listings. Subjects whose data contributed to the Part 1 analysis will not be included in the Part 2 analysis.

The database for the Safety lead-in will be locked at the time of the Part 1 analyses. However, interim data from the Safety lead-in will be made available to the DSMB prior to the final data cut-off (DCO), in order to enable the DSMB to review the emerging safety and tolerability data and to advise whether the study can proceed into Part 1. Details of the data to be provided to the DSMB will be provided in the DSMB charter.

The primary analysis of Part 1 will be the later of 3 months after Last Subject In (LSI) or when 70% of the subjects (84 subjects) have experienced a PFS event. At this point, there will be a formal database lock, and the primary analysis of all primary and secondary endpoints will be performed.

The data cut-off for Part 2 will be when 247 OS events have occurred. All secondary endpoints, including ORR, PFS and DoR, will be reported at the time of the primary analysis of OS.

6.1 Sample Size Estimation

6.1.1 Safety Lead-in

The Safety lead-in is not formally powered to assess any statistical hypotheses. In this part, 12 to 20 eligible subjects (with at least 6 from the US) will be enrolled to receive varlitinib plus capecitabine.

Subjects in this group will be replaced if varlitinib compliance is < 85% in the first 14 days.

6.1.2 Part 1

Part 1 is designed with co-primary endpoints, ORR and PFS. In order to maintain an overall, one-sided 10% type I error rate for Part 1, a **Hochberg** procedure will be used, such that the trial will be deemed to have met its primary objective if either endpoint is significant at the one-sided 5% level, or if both endpoints are significant at the one-sided 10% significance level.

One hundred and twenty subjects provides approximately 80% power to detect a true difference of 17% in response rate, based on a one-sided 5% significance level and assuming 10% response rate for the placebo group and a 27% response rate for varlitinib. To ensure the adequate data are available to evaluate the effects of varlitinib on both co-primary endpoints, the data cut-off for the primary analysis will be the later of 3 months after LSI, and when 70% of subjects (84 subjects) have experienced a PFS event. Based on a minimum of 84 PFS events at the time of the primary analysis,

the trial would have a minimum of 80% power to detect a true HR of 0.58 for PFS, based on a one-sided 5% significance level.

Furthermore, if the primary objective of the study is met, the following approach for type I error control will be applied in order to support marketing approval:

- The overall type I error rate for marketing approval will be controlled at the two-sided 5% level, using a Hochberg procedure
- Using this approach, statistical significance may be claimed if *either* ORR or PFS is significant at the two-sided 2.5% significance level or if *both* ORR and PFS are significant at the two-sided 5% significance level.

One hundred and twenty subjects provides approximately 80% power to detect a true difference of 20% in response rate, based on a two-sided 5% significance level and assuming 10% response rate for the placebo group and a 30% response rate for varlitinib.

A total of 84 PFS events provides 80% power to detect a true HR of 0.54, based on a two-sided 5% significance level. Assuming an 8 month non-linear recruitment period, a 2.67 month median PFS for capecitabine alone, and a true HR of 0.54, it is estimated that the data cut-off for the primary analysis of Part 1 will occur approximately 12 months after the first subject is randomized into the study (FSI).

6.1.3 Part 2

Part 2 is designed with a primary endpoint of OS. Approximately 350 subjects will be randomized into Part 2 to obtain 247 death events (70% maturity). If the true OS HR for the comparison of varlitinib+capecitabine versus placebo+capecitabine is 0.7 (likely to correspond to a 43% prolongation of OS), the study has 80% power to demonstrate a statistically significant difference for OS, assuming a two-sided 5% significance level.

Assuming non-linear recruitment of 350 subjects over a twenty-four month period, and median OS times of 8.6 months and 6 months for the varlitinib and placebo arms respectively, it is estimated 247 OS events will occur approximately 31 months after FSI for Part 2.

All secondary endpoints will be analyzed at the time of the primary analysis of OS.

6.1.4 Sample Size Re-estimation

There are no planned sample size re-estimations.

6.2 Adjustments for Covariates

It is planned that all statistical analyses will be stratified by geographical region (US/Non-US) and primary tumor location (GB/non-GB), in accordance with the stratification factors declared at randomization.

6.3 Handling of Dropouts or Missing Data

In general, missing data will remain missing and will not be included in analyses. Exceptions are described below.

- Missing Baseline Data

If a baseline (Day 1) value is not available and a screening value is available for the same parameter, then the last screening value will be used as baseline. This value will also be used for calculations of changes from baseline.

- Missing % Change from Baseline in Sum of Longest Diameters of Target Lesions (TLs) at Week 12

An increase from baseline of 20% will be imputed for any subjects who die by any cause, or withdraw from the study due to symptomatic disease progression prior to Week 12+ 5 days. Any subjects who withdraw from the study due to radiological disease progression prior to Week 12 will be included in the analysis, and will be assigned a percentage change from baseline of 20%, or the value that was observed at the time of disease progression (whichever is larger).

- Missing Best % Change from Baseline in Sum of Longest Diameters of TLs at any time

For the waterfall plot of the best % change from baseline, subjects with measurable disease at baseline (see Section 4.3), for whom no post-baseline TL data are available, will have an increase from baseline of 20% imputed in the following circumstances:

  - The subject has died prior to week 12 (+ 5 days)
  - Progression has been detected based on NTLs or new lesions prior to week 12 (+ 5 days)
  - Subject has withdrawn due to symptomatic progression prior to week 12 (+ 5 days)

6.4 Monitoring and Interim Analyses

6.4.1 Monitoring

The study will be monitored to ensure that the study is conducted and documented properly according to the protocol, good clinical practice (GCP), and all applicable regulatory requirements.

6.4.2 Interim Analysis

Data from the open-label Safety lead-in will be reviewed on an ongoing basis, and a Data Safety Monitoring Board (DSMB) meeting will occur once the first 12 subjects have completed one cycle of treatment (with a minimum 85% compliance over the first 14 days). The purpose of the initial DSMB meeting is to review the safety and tolerability data prior to commencement of Part 1 of the study.

There are no interim analyses planned during Part 1 or 2. The primary analysis of Part 1 will be the later of 3 months after Last Subject In (LSI) or when 70% of the subjects (84 subjects) have experienced a PFS event. At this time, a database lock will be performed and the primary analysis of all primary and secondary endpoints will be performed. The data cut-off for Part 2 will be following the 247th OS event. Whilst Part 2 is ongoing, further efficacy and safety data may continue to be accrued for Part 1 subjects. For example, further OS events occurring after the primary analysis of Part 1 will continue to be collected and databased.

The database for both parts of the study will be locked at the data cut-off for the Part 2 analyses.

Data for any subjects who continue to receive treatment beyond this point will be reported as a further addendum to the CSR.

6.5 Multiple Testing Strategy

6.5.1 Part 1

Part 1 is designed with co-primary endpoints, ORR and PFS. In order to maintain an overall, one-sided 10% type I error rate for Part 1, a **Hochberg** procedure will be used, such that the trial will be deemed to have met its primary objective if either endpoint is significant at the one-sided 5% level, or if both endpoints are significant at the one-sided 10% significance level. If the trial meets its primary objective as described above, all secondary efficacy endpoints will also be tested at the one-sided 10% significance level to better understand the clinical benefit offered by varlitinib treatment.

Furthermore, if the primary objective of the study is met, as described above, statistical significance at the two-sided 5% level will be tested to support marketing approval. For this assessment , the following approach to type I error control will be applied:

- The overall type I error rate for marketing approval will be controlled at the two-sided 5% level, using a Hochberg procedure
- Using this approach, statistical significance may be claimed if *either* ORR or PFS is significant at the two-sided 2.5% significance level or if *both* ORR and PFS are significant at the two-sided 5% significance level.

If statistical significance is demonstrated using the approach described above, the secondary endpoints OS and DCR will also be assessed at the two-sided 5% significance level, in order to further describe the nature of the benefits offered by varlitinib treatment, to support marketing authorisation.

For the purpose of supporting marketing approval, all hypothesis testing at significance levels higher than 5% will be considered exploratory.

6.5.2 Part 2

In order to describe the nature of the benefits of varlitinib treatment, OS, ORR, PFS and DCR will all be tested at a 2-sided significance level of 5%.

However, in order to strongly control the type I error at the two-sided 5% level, the following hierarchical testing order is defined for the Part 2 efficacy endpoints: OS – ORR – PFS – DCR.

6.6 Examination of Subgroups

For ORR (Part 1), PFS (Part 1) and OS (Part 2), sub-group analyses will be performed to investigate the consistency of the treatment effect across expected prognostic factors, and possible predictive factors. The sub-groups to be explored are the stratification factors region (US/Non-US) and primary tumor location (GB/Non-GB), plus the following factors:

- Gender (Male vs. Female)
- Race (Asian/Caucasian/Other)
- Baseline ECOG Status
- Extent of Disease (Locally Advanced/Metastatic)
- Age (<60/>=60)
- Baseline HER2 expression (positive defined as +++ or above by IHC, otherwise negative)
- Any HER positivity (positive defined as + or higher by IHC for HER1, HER2, HER3 or HER4, otherwise negative)

These sub-group analyses will be performed using the primary analysis population, and considered supportive of the primary analyses, thus no adjustment to the significance level will be made.

Sub-group analyses will not routinely be performed for the sensitivity analysis populations, unless the main sensitivity analysis results do not concur with the primary analysis, and further investigation is required.

6.7 Stratification Factors

As described in Section 2.2, for Parts 1 and 2, the randomization scheme will be stratified by region (US/Non-US) and primary tumor location (GB/Non-GB). It is planned that all formal statistical analyses will be stratified by the same two factors, using the values declared at randomization.

In the event that any individual stratum is very small, and contains limited information (i.e. events or responders), certain stratified analysis techniques may result in the entire stratum's data being "omitted" from the fitted model. Furthermore, even when the primary analysis can be fitted, if the information within the stratum is limited, the treatment by stratum interaction tests are unlikely to be meaningful, and the planned sub-group analyses may also become problematic.

If there are fewer than 8 patients in any of the 4 strata, the number of events (PFS and OS) and responders (ORR and DCR) within the affected stratum will be reviewed, prior to study unblinding. If at the blind review stage there is deemed to be insufficient information within the stratum to make the planned stratified analysis meaningful, geographical region will be "dropped" as a stratification factor from the statistical analyses, and the primary analyses would remain stratified by primary tumor location (GB/non-GB). The reason for removing geographical region rather than primary tumor

location as a stratification factor, is that it is anticipated that the percentage of GB:non-GB patients will be closer to 50:50, therefore geographical region is the stratification factor most likely to be driving the small stratum.

In this scenario, geographical location (US/non-US) would be added to the list of factors to be explored as part of the planned sub-group analyses. As noted above, any decision to exclude geographical region as a stratification factor in the analysis would be made prior to study unblinding, and would be fully documented in the study file.

7. Methods of Analysis

7.1 Objective Response Rate

All RECIST-based data will be listed, and the BoR and ORR will be summarized by treatment group. For those subjects with a BOR of PD, the summary will include sub-categories to distinguish between PD as defined by the RECIST criteria, and early deaths that have been assigned to the PD category (see Section 3.1.1). A listing of early deaths will also be provided.

As described in Section 3.1.1, and in accordance with RECIST v1.1, response confirmation is not required in order for a subject to be classified as a responder for the assessment of ORR. For Part 1 only, descriptive statistics will also be produced for BoR and ORR, based on confirmed responses only using the FAS set. For this summary table, to be classified as a responder, the subject would need to have at least two assessments where the visit response is PR or CR, without any intermediate assessments indicative of progression.

There will be no formal analyses of efficacy data from the Safety lead-in. For Parts 1 and 2, the primary analysis will test for the superiority of varlitinib-containing treatment arm relative to the placebo-containing arm.

Objective Response Rate will be analyzed using a stratified exact binomial test (Mehta et al, Emerson) which extends Fisher's Exact test to more than one strata. The analysis will be stratified by region (US/Non-US) and primary tumor location (Gall Bladder/Non-GB). The 2-sided p-value will be obtained by doubling the 1-sided p-values.

The conditional maximum likelihood estimate of the odds ratio will be presented together with its exact 2-sided 95% confidence limit, which excludes one if and only if the stratified exact binomial test is significant at a 2-sided 5% level. For Part 1 only, corresponding significance levels and confidence intervals will be calculated using tests performed at a 1-sided 10% level. The one-sided P-value will also be presented.

To test for heterogeneity across strata, the exact test of Zelen will be performed by testing whether the treatment effect differs between the 4 strata included in the primary analysis.

For Part 1 only, subgroup analyses will be conducted comparing ORR between treatments in the subgroups of the FAS set defined by the stratification factors Region (US/Non-US) and primary tumor location (GB/Non-GB), plus the following factors:

- Gender (Male vs. Female)
- Race (Asian/Caucasian/Other)
- Baseline ECOG Status
- Extent of Disease (Locally Advanced/Metastatic)
- Age (<60/>=60)
- Baseline HER2 expression (positive defined as +++ or above by IHC, otherwise negative)
- Any HER positivity (positive defined as + or higher by IHC for HER1, HER2, HER3 or HER4, otherwise negative)

Other baseline variables may also be included if there is clinical justification or an imbalance is observed across the treatment arms.

For Part 1, for each sub-group, the odds ratio and corresponding 80% confidence interval will be presented on a forest plot including the odds ratio and 80% CI from the overall population. Additionally, if the primary analysis of ORR is significant at the two-sided 5% level in the overall population, the forest plot will be replicated using 95% confidence intervals.

No adjustment to the significance level for testing will be made since all these analyses will be considered supportive of the primary analysis of ORR.

Subjects whose data contributed to the Part 1 analysis will not contribute to the Part 2 analysis.

For Part 1, the primary analysis of ORR based on ICR will be performed using the FAS set, and two sensitivity analyses will be performed using the EFR and PP sets. Note that the subgroup analyses will not be repeated for the sensitivity analyses. ORR based on site assessments is a secondary endpoint and will be analyzed using the FAS set. No sub-group analyses are planned for this secondary endpoint.

For Part 2, the primary analysis of ORR will be performed using the FAS set, based on the site assessments of radiological data. No subgroup analyses are planned for ORR in Part 2.

7.2 Progression-Free Survival

In Parts 1 and 2, the primary analysis will test for the superiority of the varlitinib-containing treatment arm relative to the placebo-containing arm, as assessed by PFS.

In Part 1, the data cut-off for the primary analysis of PFS will be the later of 12 weeks after LSI and 84th documented PFS event. The primary analysis of PFS will be based on data from the ICR. An analysis of PFS based on site assessments will be used to assess the sensitivity of the primary analysis to informative censoring. See Section 7.2.1.2.

In Part 2, PFS will be analyzed at the time of the primary analysis of OS. The Part 2 PFS analysis will be based on the site assessments of radiological data.

In both study parts, PFS will be analyzed using a stratified log-rank test using the Breslow method to handle ties (Breslow, 1974). The stratification factors are

geographical region (US/Non-US) and primary tumor location (GB/Non-GB). The results will be presented in terms of the HR (with a HR <1 favoring the varlitinib-containing arm) and associated two-sided 95% confidence interval (CI) and two-sided p-value. For Part 1 only, the 80% confidence interval and one-sided p-value will also be presented. The HR and its CI can be estimated from the log-rank as follows (Berry et al 1991, Collett, 2003, Sellke and Siegmund 1983):

$$HR = \exp(U/V)$$

$$95\%\ CI\ for\ HR = (\exp\{U/V - 1.96/\sqrt{V}\}, \exp\{U/V + 1.96/\sqrt{V}\})$$

Where $U = \sum_k U_k = \sum_k \sum_i (d_{1ki}, -e_{1ki})$ is the stratified log-rank test statistic obtained from the SAS LIFETEST procedure, $\sqrt{V} = \sqrt{\sum_k V_k}$, is its standard deviation, k denotes the stratum and d_{1ki} and e_{1ki} are the observed and expected events in group 1, stratum k.

Kaplan-Meier (KM) plots of PFS, with tick marks to identify censored observations, will be presented by treatment group. Summaries of the number and percentage of subjects experiencing a PFS event, and the type of event (RECIST 1.1 progression or death) will be provided along with median PFS for each treatment.

The assumption of proportionality will be assessed.

For Part 1 only, subgroup analyses will be conducted comparing PFS between treatments in the subgroups of the full analysis set defined by the stratification factors Region (US/Non-US) and primary tumor location (GB/Non-GB), plus the following factors:

- Gender (Male vs. Female)
- Race (Asian/Caucasian/Other)
- Baseline ECOG Status
- Extent of Disease (Locally Advanced/Metastatic)
- Age (<60/>=60)
- Baseline HER2 expression (positive defined as +++ or above by IHC, otherwise negative)
- Any HER positivity (positive defined as + or higher by IHC for HER1, HER2, HER3 or HER4, otherwise negative)

Other baseline variables may also be assessed if there is clinical justification. For each subgroup, the HRs (varlitinib: placebo) and associated CIs will be calculated from a log rank test within each individual subgroup. These will be presented on a forest plot including the HR and 80% CI from the overall population. Additionally, if the primary analysis of PFS is significant at the two-sided 5% level in the FAS, the forest plot will be replicated using 95% confidence intervals.

The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected and possible prognostic and predictive factors. No adjustment to the significance level for testing will be made since all these analyses will be considered supportive of the primary analysis of PFS.

For both Part 1 and Part 2, Cox proportional hazards modeling will be employed to assess the effect of covariates on the HR estimate. Before embarking on more detailed modeling, an initial model will be constructed, containing treatment and the two stratification factors alone, to ensure any output from the Cox modeling is likely to be consistent with the results of the stratified log-rank test.

The presence of quantitative interactions will be assessed formally by means of an overall global interaction test. This will be performed in the full analysis set by comparing the fit of a Cox proportional hazards model including treatment, all covariates, and all covariate-by treatment interaction terms, with one that excludes the interaction terms and will be assessed at the 2-sided 10% significance level. If the fit of the model is not significantly improved then it will be concluded that overall the treatment effect is consistent across the subgroups.

If the global interaction test is found to be statistically significant, an attempt to determine the cause and type of interaction will be made. Stepwise backwards selection will be performed on the saturated model, whereby (using a 10% 2-sided level throughout) the least significant interaction terms are removed one-by-one and any newly significant interactions re-included until a final model is reached where all included interactions are significant and all excluded interactions are non-significant. Throughout this process all main effects will be included in the model regardless of whether the corresponding interaction term is still present. This approach will identify the factors that independently alter the treatment effect and prevent identification of multiple correlated interactions.

Any quantitative interactions identified using this procedure will then be tested to rule out any qualitative interaction using the approach of Gail and Simon 1985.

For both parts, the primary analysis of PFS will be based on the programmatically derived PFS using all scans regardless of whether they were scheduled or not.

The primary analysis will be performed on the FAS.

Subjects whose data contributed to the Part 1 analysis will not also contribute to the Part 2 analysis.

7.2.1 Sensitivity Analyses for PFS – Part 1 only

The following sensitivity analyses will be performed to investigate the impact of potential sources of bias in the primary assessment of PFS.

7.2.1.1 Assessment of Major Deviations

To assess the impact of major protocol deviations on the primary analysis of PFS, a sensitivity analysis will be performed using the PP analysis set. PFS based on the PP set will be analyzed using a stratified log-rank rest, using the methods described for the primary analysis of PFS in Section 7.2. Note, sub-group analyses will not be performed for the sensitivity analysis.

7.2.1.2 Assessment of Informative Censoring

For Part 1, the primary analysis of PFS is based on an ICR of radiological data. Informative censoring can occur if site believe that a subject has progressed and perform no further radiological assessments in that subject, but the ICR disagree, thus the subject becomes censored at the last available radiological assessment for the ICR assessment of PFS.

To investigate the sensitivity of the primary analysis to this source of bias, a sensitivity analysis will be performed using PFS based on the site assessment of the radiological data. PFS based on site assessments will be analyzed using a stratified log-rank rest, using the methods described for the primary analysis of PFS in Section 7.2. Note, the sub-group analyses will not be performed for the sensitivity analysis.

7.2.1.3 Evaluation-Time Bias Assessment

Bias can be introduced if one treatment arm is assessed more frequently than the other.

For both Parts 1 and Part 2, sensitivity analyses will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled timepoints. The midpoint between the time of progression and the previous evaluable RECIST assessment will be analyzed using a stratified log-rank test, as described in Section 7.2. This approach has been shown to be robust to even highly asymmetric assessment schedules (Sun and Chen 2010). Note that the sub-group analyses will not be performed for the sensitivity analysis.

7.2.1.4 Attrition-Bias Assessment

Attrition bias will be assessed by repeating the primary PFS analysis except that the actual PFS event times, rather than the censored times, of subjects who progressed or died in the absence of progression immediately following two, or more, non-evaluable tumor assessments will be included. In addition, subjects who take subsequent therapy prior to progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy. This analysis will be performed using the methods described in Section 7.2, supported by a Kaplan-Meier plot of the time to censoring where the censoring indicator of the primary PFS analysis is reversed. Note that the sub-group analyses will not be performed for the sensitivity analysis.

7.3 Overall Survival

For Parts 1 and 2, the analysis of OS will test for the superiority of varlitinib-containing treatment arm relative to the placebo-containing arm.

For Part 1, OS will be analyzed at the time of the primary analysis of ORR and PFS. For Part 2, the primary analysis of OS will be performed after the 247th documented OS event.

OS will be analyzed using the methods described in Section 7.2 for the primary analysis of PFS. For Part 2 only, sub-group analyses, as described in Section 7.2, will also be performed.

Subjects whose data contributed to the Part 1 analysis will not contribute to the Part 2 analysis.

7.4 Duration of Response

The DoR will not be formally analyzed. DoR will be listed and presented using Kaplan-Meier curves. Descriptive statistics, based on the KM estimates, will be presented by treatment group. For Part 1, DoR will be based on data from the ICR. For Part 2, DoR will be based on site assessments.

7.5 Disease Control Rate

Disease Control Rate will be analyzed using a stratified exact binomial test (Mehta et al, Emerson) which extends Fisher's Exact test to more than one strata. The analysis will be stratified by region (US/Non-US) and primary tumor location (Gall Bladder/Non-GB). The 2-sided p-value will be obtained by doubling the 1-sided p-values.

The conditional maximum likelihood estimate of the odds ratio will be presented together with its exact 2-sided 95% confidence limit, which excludes one if and only if the stratified exact binomial test is significant a 2-sided 5% level. For Part 1 only, corresponding significance levels and confidence intervals will be calculated using tests performed at a 1-sided 10% level. The one-sided p-value will also be presented.

To test for heterogeneity across strata, the exact test of Zelen will be performed testing whether the treatment effect differs between the 4 strata included in the primary analysis.

The primary analysis of DCR will be based on the FAS set. For Part 1 and 2, a sensitivity analysis will also be performed using the EFR.

7.6 Tumor Size

For all study parts, tumor size will be presented graphically using waterfall plots for each treatment arm, presenting each subject's Week 12 percentage change in tumor size as a separate bar with the bars ordered from the largest increase to the largest decrease. Reference lines at the +20% and –30% change in tumor size levels will be added to the plots, which correspond with the definitions of disease progression and PR, respectively. In these waterfall plots, the subjects whose Week 12 percentage change in tumor size is based on an imputation will be clearly identified. For Parts 1 and 2, the two treatment arms will be plotted separately, but presented side-by-side on a single page.

Waterfall plots of the best percentage change from baseline will also be presented, as described above for the Week 12 percentage change from baseline.

The percentage change from baseline in tumor size at Week 12 ($\%\Delta TS_{Wk12}$) is a secondary endpoint for Part 1. In line with the exploratory objectives of the trial, the

statistical analysis of tumor size in Part 1 will be also be used to help evaluate the role of HER family expression levels as predictors of benefit to varlitinib. If a relationship is found between biomarker(s) expression and clinical outcomes in Part 1 of the study, the biomarker(s) could be prospectively evaluated in Part 2 of the study.

These analyses will be exploratory, and may be driven by the data observed, and specifically the prevalence of the HER family sub-groups of interest. However, the analysis principles to be followed are outlined below.

1. ASLAN will create and document a list of the key biomarkers of interest, based upon the biological plausibility of predictivity for clinical outcomes. The list will be limited to those biomarkers, or combinations of biomarkers for which a scientific hypothesis exists. These will represent the biomarkers that will be used to assess the listed exploratory objectives, providing that there is sufficient prevalence of that biomarker in the trial population.
2. For each biomarker of interest, each subject's baseline biomarker status will be independently evaluated blinded to study treatment and clinical outcome
3. For each biomarker sub-group, or combination of biomarkers identified on the prospectively defined list (step 1), if at least 15 subjects in the EFR population are determined to be expressors, then the role of that biomarker as a predictor of clinical outcome (tumor size) will be evaluated as described below.

For each biomarker of interest (or combination of biomarkers), the absolute values at baseline and Week 12, along with the percentage change from baseline target lesion size at Week 12 will be summarized using descriptive statistics and presented by treatment arm and baseline biomarker status.

The number and percentage of subjects in each treatment arm whose Week 12 data is imputed will also be presented (split by biomarker sub-group).

The effect of varlitinib on percentage change in tumor size will be estimated from an analysis of covariance (ANCOVA) model fitting the Week 12 percentage change from baseline as the response variable, and treatment, geographical location (US/Non-US), primary tumor location (GB/Non-GB), baseline biomarker status (positive or negative), the interaction treatment x baseline biomarker status, baseline tumor size, and the time from the baseline scan to randomization as covariates.

The significance of the interaction term will be evaluated at the one-sided 10% level, to determine whether there is evidence of a treatment biomarker interaction. Regardless of the significance of the interaction term, the number of subjects, unadjusted mean and adjusted least squares (LS) means for each treatment arm will be presented for each level of biomarker status factor (positive or negative), together with the difference in adjusted LS means between treatment arms for each level of the biomarker status factor, and corresponding p-value and 80% and 95% confidence intervals for the difference.

If the interaction term is not significant at the one-sided 10% level, the number of subjects, unadjusted mean and adjusted least squares (LS) means for each treatment arm will also be presented together with the difference in adjusted LS means between treatment arms and corresponding p-value and 80% and 95% confidence intervals for the difference.

If, at blind review, it is judged the data do not adequately follow a normal distribution assumption, then a log-transformation may be used. Specifically, in place of percent change in Week 12 value the ANCOVA model will use $\log_e$ (Week 12/baseline). In this situation the ratio of the geometric LS means corresponding confidence interval and p-value will be presented.

As a sensitivity analysis a nonparametric analysis will be performed. The p-value from the ANCOVA model will be presented together with the Hodges Lehmann estimate of the median difference and corresponding 80% confidence interval which will be derived. The median percentage change and range will be presented for each treatment arm, together with the number of subjects and percentage of subjects in each treatment arm whose Week 12 data is imputed in the nonparametric analysis.

For all biomarkers of interest (step 1), waterfall plots color-coded by baseline biomarker status will be produced.

7.7 Assessment of Safety and Tolerability

All safety data will be assessed for the safety population.

7.7.1 Assessment of Adverse Events

All AE data will be listed along with information regarding initial study dose, dose at onset, onset time, duration, severity, and relationship to study treatment.

Treatment emergent AEs (TEAEs) are defined as AEs with an onset date on or after the start of treatment. The following summaries will be produced, by treatment group and overall for all TEAEs:

- An overview table of the incidence of TEAEs, grade 3+ TEAEs, SAEs, TEAEs leading to treatment discontinuation and TEAEs leading to death, by treatment arm and overall. For each summary category, the results will be shown overall (regardless of causality), and for the incidence of causally related TEAEs. For example, the overall incidence of TEAEs will be presented, as well as the incidence of TEAEs related to randomized therapy, the incidence of TEAEs related to capecitabine and the incidence TEAEs related to both randomized therapy and capecitabine.
- Summary of TEAEs by system organ class (SOC) and preferred term: Both the number and percentage of subjects in each category (subject-level summary) and the number of episodes (episode-level summary). This summary table will be repeated for TEAEs attributed as *causally related to randomized therapy* (varlitinib or placebo)
- Summary of TEAEs occurring in at least 10% of subjects, sorted in descending order of frequency (i.e. most frequent event shown first). The order of frequency will be determined by the most frequent preferred term across both arms, regardless of CTC grade. For each event, the results will be presented for all CTC grades, and also split by grade 1-2, and grade 3+.
- Summary of TEAEs attributed as *causally related to randomized therapy* occurring in at least 10% of subjects, sorted in descending order of frequency (i.e. most frequent event shown first). The order of frequency will be determined by the most

frequent preferred term across both arms. For each event, the results will be presented for all CTC grades, and also split by grade 1-2, and grade 3+.

- Summary of CTC grade 3 and above TEAEs sorted in descending order of frequency (i.e. most frequent event shown first). The order of frequency will be determined by the most frequent preferred term across both arms.
- A summary or CTC grade 3 and above TEAEs attributed as *causally related to randomized* therapy sorted in descending order of frequency (i.e. most frequent event shown first). The order of frequency will be determined by the most frequent preferred term across both arms.
- Summary of SAEs by preferred term
- Summary of SAEs attributed as *causally related to randomized* therapy, sorted by preferred term

Additionally, the following will be listed:

- AEs with outcome of death along with the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment) and investigator's assessment of severity and relationship to study drug
- All SAEs along with the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment), date of resolution (if AE is resolved), investigator's assessment of severity and relationship to study drug
- AEs leading to discontinuation of study medication, listed along with the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment) and investigator's assessment of severity and relationship to study drug

If an AE is reported more than once during the study period the greatest severity and the worst-case attribution will be presented in summary tables. Any AEs commencing >28 days after discontinuation of study treatment will not be included in the tabulations of AE data.

7.7.2 Clinical Laboratory Data

All clinical laboratory data (clinical chemistry, hematology and urinalysis data), vital signs and ECG data will be listed. In addition, all data measured on a continuous scale (except respiration rate and pulse) will be displayed graphically as described below:

For the Safety lead-in:

- Subject profile plots of laboratory data over time, including reference lines at the LLN and ULN. Any dose reductions will be indicated on the plots using a change in line-type.
- Subject profile plots of the change from baseline over time, including a horizontal reference line at zero. Any dose reductions will be indicated on the plots using a change in line-type.

For Parts 1 and 2:

- Box plots over time, by treatment, including reference lines at the LLN and ULN
- Box plots of the change from baseline over time, by treatment, including a horizontal reference line at zero

To enable assessment of the potential for drug-induced liver injury, for all study parts the following outputs will be produced:

- A scatter plot of maximum on-treatment alanine aminotransferase (ALT) versus maximum on-treatment total bilirubin (x-axis), both expressed as multiples of the upper limit of normal (ULN), including reference lines at 3×ULN for ALT, and 2×ULN for total bilirubin. For each subject, the maximum ALT and maximum total bilirubin may occur at different visits
- A scatter plot of maximum on-treatment aspartate aminotransferase (AST) versus maximum on-treatment total bilirubin (x-axis), both expressed as multiples of the upper limit of normal (ULN), including reference lines at 3×ULN for AST, and 2×ULN for total bilirubin. For each subject, the maximum AST and maximum total bilirubin may occur at different visits.

In the event that the scatter plots of ALT or AST against bilirubin identify any potential Hy's law cases (subjects with either both ALT > 3 x ULN and total bilirubin > 2 x ULN, or both AST > 3 x ULN and total bilirubin > 2 x ULN), profile plots over time will be produced for these subject's liver function tests (ALT, AST, ALP and total bilirubin), expressed in multiples of the ULN, showing all four FLT parameters on the same plot.

7.7.3 ECG Data

For all study parts, ECG data will be summarized as described below:

- For each parameter, absolute values and the change from baseline at each scheduled assessment will be summarized by treatment group using descriptive statistics including number of observations, mean, standard deviation, minimum, median and maximum values.

For both the Safety lead-in and Part 1, the following will be summarized:

- The proportion of patients obtaining treatment-emergent absolute QTcF values > 450 ms and ≤ 480 ms; > 480 ms and ≤ 500 ms; and > 500 msec
- The proportion of patients obtaining a QTcF increase from baseline values ≥30 and <60 ms; and ≥60 ms
- The proportion of patients obtaining a QRS change from baseline > 25% resulting in QRS > 120 ms
- The proportion of patients obtaining a PR interval change from baseline >25% reaching a value >220 ms
- The proportion of patients obtaining a > 25% decrease from baseline in heart rate, resulting in a heart rate < 50 beats per minute (bpm) or a >25% increase from baseline in heart rate resulting in a heart rate > 100 bpm

For Part 1, the above categorical data summaries will be presented by treatment arm. All outliers will be summarized for each treatment group on the basis of incidence rates. A subject will be counted only once for a particular outlier event if the subject experiences more than one episode of that event.

T-wave morphology will also be analyzed and will be listed and summarized using descriptive statistics, focusing on change from baseline, i.e., treatment emergent changes.

7.7.4 Physical Examination

Physical examination data will be listed.

7.7.5 Exposure, Dose Intensity and General Tolerability

To further assess the tolerability of varlitinib 300 mg BID in combination with capecitabine, the dose intensity and treatment exposure of both the randomized therapy and capecitabine, as defined in Section 3.2.2, will be listed and summarized by treatment arm. Exposure summaries will be produced overall (across all cycles) and dose intensity will be presented by cycle for the first 3 cycles, and overall.

For Part 1, swimmer plots of the time on treatment will be presented for all subjects in the FAS. The bars on these plots will be ordered as follows:

CR subjects at the top, PR subjects next, SD next, PD next and NE subjects at the bottom. Within each category they will be sorted from longest treatment duration to shortest treatment duration. Subjects whose time on treatment is censored will be indicated.

This plot will be repeated to account for dose modifications and interruptions.

For Part 2, swimmer plots of time on treatment will be produced for responding subjects only.

7.8 Disposition of Subjects

Subject disposition, including but not limited to, the date of informed consent, randomization details (Parts 1 and 2 only), date of first dose (Safety lead-in only) and reasons for discontinuation from randomized therapy (Parts 1 and 2) or study treatment (Safety lead-in), will be listed and summarized by treatment group.

7.9 Demographics and Baseline Characteristics

Demographic and baseline characteristic data will be presented based on the FAS.

Baseline demographic data, including, but not limited to:

- age
- gender
- weight
- height
- race
- geographical location
- primary tumor location (intra-hepatic bile duct, extra-hepatic bile duct, hilar, gall bladder, ampulla of vater or other)
- ethnicity
- baseline HER1, HER2, HER3 and HER4 status

will be listed and summarized using appropriate descriptive statistics.

7.10 Medical History and Prior and Concomitant Medications

Relevant medical history and prior treatment for BTC will be listed. For Parts 1 and 2, prior therapies will also be summarized by treatment group, type of therapy (local/, surgery, or systemic therapy), reason for therapy (adjuvant, neoadjuvant or metastatic therapy) and best response.

All medications received following the start of treatment (including those that were ongoing prior to first dose) will be listed.

7.11 Pharmacokinetic Data

Safety Lead-in part:

Pharmacokinetic parameters in plasma will be derived using standard non-compartmental methods with WinNonlin® Professional Version 7.3 or higher (Pharsight Corp., Mountain View, California, US) or SAS® Version 9.2 or higher (SAS Institute, Inc., Cary, North Carolina, US). Based on previous population PK analyses, a one compartment with delayed first-order absorption and first-order elimination, and auto-inhibition on parameter CL/F will be selected as the base model. Actual elapsed time from dosing will be used for final plasma PK parameter calculations.

Part 1 and Part 2:

The bioanalytical analysis of the plasma concentration data for varlitinib will be performed at CMIC Pharmaceuticals Japan. Pharmacokinetic analyses will be conducted according to CMIC Standard Operating Procedures (SOPs) for PK analyses unless otherwise specified.

In Parts 1 and 2 of the study, Population PK analyses to describe the PK in the 2nd line BTC population, and PK/PD analyses to explore the potential of varlitinib to prolong the QTc interval, will be performed by nonlinear mixed effects model (NLME) using NONMEM version 7.3 (ICON plc, Ellicott City, MD, USA).

R version 3.3.3 or later (R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/) or SAS® Version 9.2 or higher (SAS Institute, Inc., Cary, North Carolina, United States) will be used for data preparation, visualization and inspection of the raw data and/or post-processing of NONMEM outputs. PsN 4.4.8 (Lindbom et al) may be used as aids in model assessment and for some automated modeling and simulation procedures.

Actual dates and time of samples and doses will be used for final analyses.

7.11.1 Concentration Values

All concentration values below the lower limit of quantification (LLQ) and samples with no reportable value occurring prior to dosing will be replaced by “0”. For tabulation, graphical representation and calculation of summary statistics, all samples <LLQ or with no reportable value observed after drug administration will be treated as missing.

7.11.2 Calculation of the Actual Sampling Times

For all sampling times, the actual sampling times will be calculated as the difference between the sample collection actual clock time and the actual clock time of the oral administration of the tablet.

Therefore:
Actual clock time of sample collection = actual clock time of dosing + difference

The actual sampling times expressed in hours and rounded off to 3 decimal digits will be used to calculate the PK parameters, except for pre-dose samples, which will always be reported as zero (0.000), regardless of the time difference.

7.11.3 Pharmacokinetic Analysis

For the Safety lead-in summary statistics for varlitinib will be tabulated for the PK parameters detailed in Section 3.3 by study day. The mean, standard deviation (std), coefficient of variation (%CV), geometric mean, %CV geometric coefficient of variation, median, minimum, and maximum will be presented for PK parameters detailed in Section 3.3. Median, minimum, and maximum will be presented for t_{max}.

Profile plots will be provided on both linear and log scale, showing the subject data.

For Part 1 and 2

Using appropriate PK software depending on whether comprehensive PK sampling or sparse PK sampling is obtained in practice, the PK data will be used to derive PK parameters such as, but not restricted to, C_{max}, AUC, and $t_{1/2}$ for varlitinib.

The final PK analyses will be the responsibility of the CRO. Pharmacokinetic analyses will be conducted according to CRO Standard Operating Procedures (SOPs) for PK analyses unless otherwise specified. The actual PK sampling times will be used in the PK calculations.

The varlitinib concentration-time profiles, along with the derived PK variables, will be listed for each patient per dose and dosing day and summarized appropriately. Population PK models may be used to derive the PK parameters and will aim to characterize variability in the population by investigating the influence of covariates such as weight, age, smoking status and/or concomitant medications. In addition, if the data are suitable, potential relationships between plasma varlitinib and efficacy or safety endpoints will be investigated using a graphical approach and/or appropriate PK/PD modeling techniques.

7.11.4 PK/PD Analyses with special consideration of effects on QT

7.11.4.1 General Principles

The final ECG assessments (including interval measurements) will be made by the Core ECG laboratory. The Part 1 electrocardiographic Core ECG Laboratory assessments constitute the study's primary assessment of ECG/QTc. Any analysis of ECG data from the Safety lead-in is considered exploratory.

Due to the non-randomized, open label design of the Safety lead-in, data from the Safety lead-in will be presented separately from data from Part 1 of the study. Furthermore, local ECG/QT outputs (not the central ECG report) will be made available to the DSMB for consideration when determining whether the study should progress to Part 1.

7.11.4.2 Methods of Analysis

Safety Lead-in

Data from the Safety lead-in will be used to explore possible relationships between the varlitinib plasma concentrations and changes from baseline in QTcF via exposure-response modeling of the QTcF and the concentrations of varlitinib, capecitabine, and their major metabolites.

To support the interpretation of these analyses, the following graphical displays will be produced:

- Time-matched scatter plots of varlitinib plasma concentrations vs. change from baseline in QTcF.
- Time-matched scatter plots of capecitabine plasma concentrations vs. change from baseline in QTcF.
- Time-matched scatter plots of plasma concentrations for concentrations 5-FU (active metabolite of capecitabine) vs. change from baseline in QTcF.
- In case that major circulating metabolites of capecitabine (5'-DFCR, 5'-DFUR and FBAL) will be determined, time-matched scatter plots of plasma concentrations for major metabolites (>10% AUCparent) vs. change from baseline in QTcF will be also conducted.

Part 1

For Part 1, a multi-step approach will be applied to evaluate the effect of capecitabine alone and varlitinib in combination with capecitabine on the ECG parameters at different time points post-dose. The comparisons will be done between ECG values at each time point post-dose and baseline and between treatment arms with and without varlitinib for each time point individually. The evaluation of the cardiodynamic effects of capecitabine and varlitinib may be explored using exposure-response modeling if there appears to be a meaningful effect of varlitinib when added to capecitabine on the ECG parameters including QTcF.

The statistical stepwise approach to be used to separate the cardiodynamic effect of varlitinib from normal circadian fluctuation in the ECG values and from the effect of capecitabine is described below:

1. ANOVA mixed effects comparison of the absolute values of ECG parameters at each time point vs baseline using treatment as a fixed effect in order to evaluate the contribution of circadian fluctuation and to determine how treatment differences contribute to the changes. The ECG parameter values will be log-transformed and the resulting output will include the geometric mean ratio for test vs reference with the corresponding confidence interval and p-value. Interaction term, time point x treatment, will be also evaluated if the data allow.

2. ANOVA comparison of the absolute values of ECG parameters as well as change from baseline (ΔECG) between treatments for each post-dose time point assessed separately in order to separate the varlitinib effect from that of capecitabine. The ECG parameter values will be log-transformed and the resulting output will include the geometric mean ratio for test vs reference with the corresponding confidence interval and p-value. ΔECG will be tested as untransformed and log-transformed parameters to avoid bias from discarding 0 ΔECG values from the comparisons.
3. Exposure-Response Analysis: Linear mixed effect model between ΔQTcF and time matched concentrations of varlitinib and other potential metabolites if measured and as appropriate. Other ECG variable may be similarly analyzed if appropriate and other models use, if appropriate.
   3.1. ΔQTcF will be calculated as the difference between each time point value and the average baseline value for each arm.
   3.2. ΔΔQTcF will be calculated for each patient in the combination arm individually as the difference between the ΔQTcF value for the arm with combination varlitinib/capecitabine treatment and the average value of all patients in the capecitabine alone arm matching time point. Other ECG parameters may also be used in the analysis including QTcB, heart rate, etc to provide supporting information.
   3.3. Base linear mixed effect model will include patients as a random effect
   3.4. The total dose of capecitabine may be included as a fixed effect in the model to account for differences in total doses received by subjects with different body surface area.
4. Linear mixed effect model between ΔQTcF parameters and time matched concentrations of capecitabine and 5-FU (metabolite of capecitabine) and other potential metabolites of varlitinib and capecitabine, if measured.
   4.1. ΔQTcF will be calculated for all patients in both arms as the difference between each time point value and the average baseline value.
   4.2. Base linear mixed effect model will include patients as a random effect and treatment (capecitabine alone or in combination with varlitinib) as a fixed effect; the total capecitabine dose may also be used as a fixed effect.
5. Different covariates may be included in the complex models in addition to the base model as fixed effects such as time point in order to ensure the model feasibility and to evaluate covariates influential to ECG parameters.

The results of the fitted models will be presented using statistical parameters such as the geometric mean ratios and corresponding confidence intervals (for ANOVA) and estimates of the slope, and associated confidence interval for the linear mixed effect models.

In addition, graphic exploration of the exposure-ECG relationship may be performed:

- Residual plots to predict results
- Double Y plots with time on the X axis and time matched values of PK concentrations for varlitinib and capecitabine on the Y1 axis, ΔECG on the Y2 axis to determine the possibility of hysteresis between drug concentration and cardiodynamic effects
- Regression/scatter plots with time matched concentrations of varlitinib and capecitabine on the X axis and ΔECG on the Y axis

- Regression/scatter plots with varlitinib or capecitabine non-compartmental PK parameters of C_{max}, C_{trough}, AUC_{tau} and others on the X axis and maximum ΔECG on the Y axis
- Forest plots for different categories in ECG changes vs mean (SD) non-compartmental PK parameters for varlitinib or capecitabine on the X axis

7.11.5 Biomarker Analysis

For Parts 1 and 2, all biomarkers assessed via IHC will be listed and summarized by treatment arm, using appropriate descriptive statistics. If a relationship is found between biomarker(s) expression and clinical outcomes in Part 1 of the study (exploratory endpoint), the biomarker(s) could be prospectively evaluated in Part 2 of the study.

Biomarkers assessed via PCR/Sequencing, for example the mutational status of genes, such as *KRAS, NRAS, BRAF*, *EGFR* and other genetic factors that may affect response to therapy, will not be reported in the CSR and are not covered in this SAP.

8. Changes in the Conduct of the Study or Planned Analysis

Following completion of the safety lead-in, Protocol V4.0 incorporated changes to the study inclusion/exclusion criteria. These changes were made based on the advice from the study investigators, to ensure that the trial population reflected a true 2nd line BTC population, and excluded primary refractory patients and other patients deemed unlikely to benefit from any further treatment. Following these changes to the study population, the major deviation definition (Section 5.1) and consequently the per protocol analysis set definition (Section 4.4) in SAP v3.0 were also updated to reflect these changes.

The original protocol and SAP included stratified randomization and analyses, based on the primary tumor location (GB or non-GB) and geographical location (US or non-US). When strata are small and contain limited or no “events” (PFS or OS) or responders (ORR and DCR), certain statistical techniques exclude all data within the affected stratum from the fitted model. To ensure that the primary analyses are meaningful and that data from all patients in the FAS contribute to the statistical analyses, Section 6.2 of SAP v3.0 has been amended to describe the conditions under which a stratification variable, e.g. geographical location, may not be included as a stratification factor in the statistical analyses. As outlined in Section 6.2, any decisions to remove a stratification factor from the statistical analyses will be reached and documented prior to unblinding.

PFS and ORR have always been, and remain, co-primary endpoints in the study. In the previous protocol and SAP, a Hochberg procedure was proposed for testing at the one-sided 10% level, whereas hierarchical testing was proposed to support marketing approval at the two-sided 5% significance level. In SAP v3.0, a Hochberg procedure is also proposed to maintain an overall two-sided 5% type I error rate, as required for marketing approval. This change is a result of the data from the ongoing ASLAN001-007 trial in first line BTC, in which durable disease stabilization has been reported.

9. REFERENCES

Berry G, Kitchin RM, Mock PA. A comparison of two simple hazard ratio estimators based on the logrank test. Statistics in Medicine 1991; 10:749-755

Breslow, NE. Covariance Analysis of Censored Survival Data. Biometrics 1974; 30:89–99.

Collett D. Modelling Survival Data in Medical Research. 2nd ed. Boca Raton (FL): Chapman & Hall/CRC; 2003.

Eisenhauer EA et al. New response evaluation criteria in solid tumors: Revised RECIST guideline (version 1.1). European Journal of Cancer 2009:45:228-247

Emerson JD. Combining estimates of the odds ratio: the state of the art. Statistical Methods in Medical Research 1994;3:157-178

Gail M, Simon R. Testing for qualitative interactions between treatment effects and subject subsets. Biometrics 1985;41:361-72.

Hochberg, Y. A Sharper Bonferroni procedure for multiple tests of significance. Biometrika 1988;75: 800-802.

Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN) - A Perl module for NONMEM related programming. Comput Methods Programs Biomed 2004; 75:85±94. https://doi.org/10.1016/j. cmpb.2003.11.003.

Mehta CR, Patel NR, Gray R. Computing an exact confidence interval for the common odds ratio in several contingency tables. Journal of the American Statistical Association 1985;80:969-973

Sellke, T. and Siegmund, D. Sequential analysis of the proportional hazards model. Biometrika 70: 315-326, 1983

Zelen M. The analysis of several 2x2 contingency tables. Biometrika 1971;58:129-137

10. LIST OF TABLES, FIGURES AND LISTINGS

10.1 Tables

Unless otherwise stated, tables will be all repeated for safety run-in, part 1 and part 2.

Table Number	Table Title
14.1.1	Subject Disposition and Withdrawals – All Subjects
14.1.2	Analysis Populations – All Subjects
14.1.3.1	Major Protocol Deviations – Full Analysis Set (Parts 1 and 2)
14.1.3.2	Important Protocol Deviations – Full Analysis Set (Parts 1 and 2)
14.1.3.3	Stratification Levels – Full Analysis Set (Parts 1 and 2)
14.1.4	Demographic Characteristics – Full Analysis Set
14.1.5	Prior First Line Therapies for BTC – Full Analysis Set (Parts 1 and 2)
14.1.6	Treatments for Early Stage Disease including Surgical and Adjuvant Treatments – Full Analysis Set (Parts 1 and 2)
14.2.1.1	Summary of Objective Response Rate – FAS Population
14.2.1.2	Primary Statistical Analysis of Objective Response Rate – FAS Population (Parts 1 and 2)
14.2.1.3	Summary of Objective Response Rate by Subgroup – FAS Population (Parts 1 and 2)
14.2.1.4	Statistical Analysis of Objective Response Rate by Subgroup – FAS Population (Parts 1 and 2)
14.2.1.6	Summary of Objective Response Rate to Support Per Protocol Sensitivity Analysis – PP Population (Part 1)
14.2.1.7	Per Protocol Sensitivity Analysis of Objective Response Rate – PP Population (Part 1)
14.2.1.8	Summary of Objective Response Rate to Support EFR Sensitivity Analysis – EFR Population (Part 1)
14.2.1.9	EFR Sensitivity Analysis of Objective Response Rate – EFR Population (Parts 1)

Table Number	Table Title
14.2.1.10	Summary of Objective Response Rate based on Site Assessments – FAS Population (Part 1)
14.2.1.11	Statistical Analysis of Objective Response Rate based on Site Assessments – FAS Population (Part 1)
14.2.1.12	Summary of Confirmed Objective Response Rate – FAS Population (Part 1)
14.2.2.1	Summary Statistics for Progression – FAS Population
14.2.2.2	Progression Status at Data Cut-Off – FAS Population
14.2.2.3	Primary Analysis of Progression-Free Survival – FAS Population (Parts 1 and 2)
14.2.2.5	Subgroup Analysis of Progression-Free Survival – FAS Population (Part 1)
14.2.2.7	Summary Statistics for Progression to Support Per-Protocol Sensitivity Analysis – PP Population (Part 1)
14.2.2.8	Progression Status at Data Cut-Off to Support Per-Protocol Sensitivity Analysis– PP Population (Part 1)
14.2.2.9	Per Protocol Sensitivity Analysis of Progression-Free Survival – PP Population (Part 1)
14.2.2.11	Summary Statistics for Progression to Support Sensitivity Analysis Based on Site Assessment – FAS Population (Part 1)
14.2.2.12	Progression Status at Data Cut-Off to Support Sensitivity Based on Site Assessment – FAS Population (Part 1)
14.2.2.13	Sensitivity Analysis of Progression-Free Survival Based on Site Assessment – FAS Population (Part 1)
14.2.2.14	Sensitivity Analysis of Progression-Free Survival – Evaluation-Time Bias Assessment – FAS Population (Part 1)
14.2.2.15	Sensitivity Analysis of Progression-Free Survival – Attrition Bias Assessment – FAS Population (Part 1)
14.2.3.1	Summary Statistics for Overall Survival – FAS Population
14.2.3.2	Survival Status at Data Cut-Off – FAS Population
14.2.3.3	Statistical Analysis of Overall Survival – FAS Population (Parts 1 and 2)

Table Number	Table Title
14.2.3.5	Subgroup Analysis of Overall Survival – FAS Population (Part 2)
14.2.4.1	Summary Statistics for Duration of Response – FAS Population
14.2.4.3	Summary Statistics for Duration of Response – PP Population (Part 1 only)
14.2.5.1	Summary of Disease Control Rate – FAS Population
14.2.5.2	Statistical Analysis of Disease Control Rate – FAS Population (Parts 1 and 2)
14.2.5.3	Summary of Disease Control Rate to Support Per Protocol Sensitivity Analysis – PP Population (Part 1)
14.2.5.4	Per Protocol Sensitivity Analysis of Disease Control Rate – PP Population (Part 1)
14.2.5.5	Summary of Disease Control Rate to Support EFR Sensitivity Analysis – EFR Population (Part 1)
14.2.5.6	EFR Sensitivity Analysis of Disease Control Rate – EFR Population (Part 1)
14.2.6.1	Summary of Tumor Size at Baseline and Week 12 – EFR Population
14.2.6.3	Analysis of Covariance of Percentage Change from Baseline in Tumor Size at Week 12 – EFR Population
14.2.6.4	Non-Parametric Sensitivity Analysis of Percentage Change from Baseline in Tumor Size at Week 12 – EFR Population
14.2.6.6	Summary of Tumor Size at Baseline and Week 12 by Subgroup – EFR Population
14.2.6.8	Analysis of Covariance of Percentage Change from Baseline in Tumor Size at Week 12 by Subgroup – EFR Population
14.2.6.9	Non-Parametric Sensitivity Analysis of Percentage Change from Baseline in Tumor Size at Week 12 by Subgroup – EFR Population
14.2.7.1	Summary of Plasma Concentrations - Varlitinib – PK Population
14.2.7.2	Summary of Plasma Concentrations - Capecitabine – PK Population (Parts 1 and 2)

Table Number	Table Title
14.2.7.3	Summary of Plasma Pharmacokinetic Parameters – Varlitinib - PK Population
14.2.7.4	Summary of Plasma Pharmacokinetic Parameters – Capecitabine - PK Population (Parts 1 and 2)
14.2.8.1	Summary of IHC Biomarkers – FAS Population (Parts 1 and 2)
14.2.8.2	Shift Table of IHC Biomarkers from Baseline to Week 6 – FAS Population (Parts 1 and 2)
14.3.1.1	Summary Statistics for Exposure – Safety Population
14.3.1.2	Summary Statistics for Dose Intensity – Safety Population
14.3.2.1	Overview of TEAEs – Safety Population
14.3.2.2	Incidence of TEAEs by System Organ Class and Preferred Term – Subject and Episode Level – Safety Population
14.3.2.3	Incidence of Causally Related TEAEs by System Organ Class and Preferred Term – Subject and Episode Level – Safety Population
14.3.2.4	Incidence of most frequent (>10%) TEAEs by Preferred Term – Safety Population
14.3.2.5	Incidence of most frequent (>10%) Causally Related TEAEs by Preferred Term – Safety Population
14.3.2.6	Incidence of TEAEs CTCAE Grade 3 or Higher by System Organ Class and Preferred Term – Safety Population
14.3.2.7	Incidence of Causally Related TEAEs CTCAE Grade 3 or Higher by System Organ Class and Preferred Term – Safety Population
14.3.2.8	Incidence of serious TEAEs by System Organ Class and Preferred Term – Safety Population
14.3.2.9	Incidence of Serious Causally Related TEAEs by System Organ Class and Preferred Term – Safety Population
14.3.3	Summary of Deaths – Safety Population
14.3.4.16	Categorical Summary of Haematology Parameters – Safety Population (Parts 1 and 2)
14.3.4.17	Categorical Summary of Clinical Chemistry Parameters – Safety Population (Parts 1 and 2)

Table Number	Table Title
14.3.4.18	Categorical Summary of Urinalysis Parameters – Safety Population (Parts 1 and 2)
14.3.6.1	Absolute Values and Change from baseline in ECG Data – Safety Population
14.3.6.2	Categorical ECG Data Summaries – Safety Population
14.3.6.3	ECG Overall Evaluation by Visit – Safety Population
14.3.6.5	Comparison of ECG parameters at all visits to baseline by treatment – Safety Population
14.3.6.6	Comparison of ECG parameters at all visits between treatments – Safety Population
14.3.6.7	Comparison of ECG parameters at all visits between treatments – PK Population

10.2 Listings

Unless otherwise stated, listings will be all repeated for safety run-in, part 1 and part 2.

Listing Number	Listing Title
16.2.1	Discontinuation Information – Safety Population
16.2.2.1	Inclusion/Exclusion Criteria Eligibility – All Subjects
16.2.2.2	Randomization – All Subjects (Parts 1 and 2)
16.2.2.3	Screening Failure Data – All Subjects
16.2.3.1	Major Protocol Deviations – Full Analysis Set
16.2.3.2	Important Deviations – Full Analysis Set
16.2.3.3	Analysis Populations – All Subjects
16.2.3.4	Stratification Information - FAS
16.2.4.1	Demographic Characteristics – Safety Population
16.2.4.2	Medical History – Safety Population

Listing Number	Listing Title
16.2.4.3	Baseline Disease Characteristics – Safety Population
16.2.4.4	Prior and Concomitant Medications – Safety Population
16.2.4.5	Prior Anti-Cancer Therapies – Safety Population
16.2.4.6	Local Anti-Cancer Therapies – Safety Population
16.2.4.7	Substance Use – Safety Population
16.2.4.8	Primary Tumour Diagnosis – Safety Population
16.2.4.9	Tissue Collection – Safety Population
16.2.4.10	Biopsy Collection – Safety Population
16.2.4.11	Therapeutic or Diagnostic Procedures – Safety Population
16.2.4.12	Meal Record – Safety Population
16.2.4.13	Emergency Unblinding – Safety Population
16.2.5.1	Study Drug Administration – Safety Population
16.2.5.2	Capecitabine Administration – Safety Population
16.2.5.3	Study Drug Exposure and Dose Intensity – Safety Population
16.2.5.4	Capecitabine Exposure and Dose Intensity – Safety Population
16.2.5.5	Study Drug Dispensing– Safety Population
16.2.5.6	Study Drug Compliance (Varlitinib / Placebo) – Safety Population
16.2.5.7	Study Drug Compliance (Capecitabine) – Safety Population
16.2.6.1.1	Varlitinib Plasma Concentrations – Safety Population
16.2.6.1.2	Capecitabine Plasma Concentrations – Safety Population
16.2.6.1.3	Varlitinib Plasma Pharmacokinetic Parameters – PK Population
16.2.6.1.4	Capecitabine Plasma Pharmacokinetic Parameters – PK Population
16.2.6.2.1	Target Lesion Assessments (ICR) – FAS Population
16.2.6.2.2	Non-Target Lesion Assessments (ICR) – FAS Population

Listing Number	Listing Title
16.2.6.2.3	Best Overall Response and Objective Response Rate (ICR) - FAS Population
16.2.6.2.4	Target Lesion Assessments (Site Assessments) – FAS Population
16.2.6.2.5	Non-Target Lesion Assessments (Site Assessments) – FAS Population
16.2.6.2.6	Best Overall Response and Objective Response Rate (Site Assessments) - FAS Population
16.2.6.2.7	Early Deaths Considered as Progression Events – FAS Population
16.2.6.2.8	Progression-Free Survival – FAS Population
16.2.6.2.9	Overall Survival – Safety Population
16.2.6.2.10	Duration of Response – All Responding Subjects
16.2.6.2.11	IHC Biomarker Data – Safety Population
16.2.7.1	Adverse Events – Safety Population
16.2.7.2	Adverse Events Leading to Discontinuation of Study Medication – Safety Population
16.2.7.3	Serious Adverse Events – Safety Population
16.2.7.4	Adverse Events Leading to Death – Safety Population
16.2.7.5	Deaths – Safety Population
16.2.8.1	Haematology Parameters – Safety Population
16.2.8.2	Clinical Chemistry Parameters – Safety Population
16.2.8.3	Urinalysis Parameters – Safety Population
16.2.8.4	Coagulation Parameters – Safety Population
16.2.8.5	Serology Parameters – Safety Population
16.2.8.6	Pregnancy Test Results – Safety Population
16.2.8.7	CA19 and CEA Results - Safety Population
16.2.8.8	Urine Microscopy Parameters – Safety Population

Listing Number	Listing Title
16.2.9.1	ECG Assessment– Safety Population
16.2.9.2	Patients with ECG Elevations – Safety Population
16.2.9.3	Vital Signs – Safety Population
16.2.9.3	Physical Examinations – Safety Population
16.2.9.4	ECOG Assessment – Safety Population

10.3 Figures

Unless otherwise stated, figures will be all repeated for safety run-in, part 1 and part 2.

Figure Number	Figure Title
14.2.1.5	Forest Plot of Objective Response Rate by Subgroup – FAS Population (Parts 1 and 2)
14.2.2.4	Kaplan-Meier Plot of Progression-Free Survival – FAS Population (Parts 1 and 2)
14.2.2.6	Forest Plot of Progression-Free Survival by Subgroup – FAS Population (Parts 1 and 2)
14.2.2.10	Kaplan-Meier Plot of Progression-Free Survival – PP Population (Part 1)
14.2.2.17	Kaplan-Meier Plot of Progression-Free Survival – Attrition-Bias Assessment – FAS Population (Parts 1 and 2)
14.2.3	Kaplan-Meier Plot of Overall Survival – FAS Population (Parts 1 and 2)
14.2.3.6	Forest Plot of Overall Survival by Subgroup – FAS Population (Parts 1 and 2)
14.2.4.2	Kaplan-Meier Plot of Duration of Response – FAS Population (Parts 1 and 2)
14.2.4.4	Kaplan-Meier Plot of Duration of Response – PP Population (Part 1)
14.2.6.2	Waterfall Plot of the Percentage Change From Baseline in Tumor Size at Week 12 – EFR Population

Figure Number	Figure Title
14.2.6.5	Waterfall Plot of the Best Percentage Change from Baseline in Tumor Size – EFR Population
14.2.6.7	Waterfall Plot of Percentage Change from Baseline in Tumor Size at Week 12 by Subgroup – EFR Population
14.2.6.10	Waterfall Plot of Best Percentage Change from Baseline in Tumor Size by Subgroup – EFR Population
14.2.7.5	Profile Plots of Plasma Concentrations on a Linear Scale – Varlitinib - PK Population
14.2.7.6	Profile Plots of Plasma Concentrations on a Log Scale – Varlitinib - PK Population
14.2.7.7	Mean Profile Plots of Plasma Concentrations on a Linear Scale - Capecitabine – PK Population
14.2.7.8	Mean Profile Plots of Plasma Concentrations on a Log Scale - Capecitabine – PK Population
14.2.8.9	Scatter Plots of Plasma Concentration Against Change from Baseline in QTcF on a Linear Scale – Varlitinib – PK Population
14.2.8.10	Scatter Plots of Plasma Concentration Against Change from Baseline in QTcF on a Linear Scale – Capecitabine – PK Population
14.2.8.11	Scatter Plots of Plasma Concentration Against Change from Baseline in QTcF on a Linear Scale – Varlitinib Metabolites – PK Population
14.2.8.12	Scatter Plots of Plasma Concentration Against Change from Baseline in QTcF on a Linear Scale – Capecitabine Metabolites – PK Population
14.3.1.3	Swimmer Plot of Time on Treatment by Response Category – FAS Population (Part 1)
14.3.1.4	Swimmer Plot of Time on Treatment by Response Category, Accounting for Dose Modifications and Interruptions – FAS Population (Part 1)
14.3.1.5	Swimmer Plot of Time on Treatment by Response Category – Responding Subjects - FAS Population (Part 2)
14.3.1.6	Swimmer Plot of Time on Treatment by Response Category, Accounting for Dose Modifications and Interruptions – Responding Subjects - FAS Population (Part 2)

Figure Number	Figure Title
14.3.4.1	Profile Plots of Haematology Parameters – Safety Population (Safety Lead-in)
14.3.4.2	Profile Plots of Clinical Chemistry Parameters – Safety Population (Safety Lead-in)
14.3.4.3	Profile Plots of Urinalysis Parameters – Safety Population (Safety Lead-in)
14.3.4.4	Profile Plots of Change from Baseline in Haematology Parameters – Safety Population (Safety Lead-in)
14.3.4.5	Profile Plots of Change from Baseline in Clinical Chemistry Parameters – Safety Population (Safety Lead-in)
14.3.4.6	Profile Plots of Change from Baseline in Urinalysis Parameters – Safety Population (Safety Lead-in)
14.3.4.7	Box Plots of Haematology Parameters – Safety Population (Parts 1 and 2)
14.3.4.8	Box Plots of Clinical Chemistry Parameters – Safety Population (Parts 1 and 2)
14.3.4.9	Box Plots of Urinalysis Parameters – Safety Population (Parts 1 and 2)
14.3.4.10	Box Plots of Change From Baseline in Haematology Parameters – Safety Population (Parts 1 and 2)
14.3.4.11	Box Plots of Change from Baseline in Clinical Chemistry Parameters – Safety Population (Parts 1 and 2)
14.3.4.12	Box Plots of Change from Baseline in Urinalysis Parameters – Safety Population (Parts 1 and 2)
14.3.4.13	Scatter Plots of Maximum ALT vs Maximum Total Bilirubin – Safety Population
14.3.4.14	Scatter Plots of Maximum AST vs Maximum Total Bilirubin – Safety Population
14.2.4.15.x	Hys Law Plot for Subject XXXXX – Safety Population
14.3.5.1	Profile Plots of Vital Signs Parameters – Safety Population (Safety Lead-in)
14.3.5.2	Profile Plots of Change from Baseline in Vital Signs Parameters – Safety Population (Safety Lead-in)

Figure Number	Figure Title
14.3.5.3	Box Plots of Vital Signs Parameters – Safety Population (Parts 1 and 2)
14.3.5.4	Box Plots of Change From Baseline in Vital Signs Parameters – Safety Population (Parts 1 and 2)
14.3.6.5	Profile Plots of ECG Parameters – Safety Population (Safety Lead-in)
14.3.6.6	Profile Plots of Change from Baseline in ECG Parameters – Safety Population (Safety Lead-in)
14.3.6.7	Box Plots of ECG Parameters – Safety Population (Parts 1 and 2)
14.3.6.8	Box Plots of Change From Baseline in ECG Parameters – Safety Population (Parts 1 and 2)
14.3.6.9	Scatter Plots of Time Matched Plasma Concentrations Against Change from Baseline in QTcF on a Linear Scale – Varlitinib; PK Population
14.3.6.10	Scatter Plots of Plasma Concentrations Against Change from Baseline in QTcF on a Linear Scale - Capecitabine; PK Population
14.3.6.11	Scatter Plots of Plasma Concentrations Against Change from Baseline in QTcF on a Linear Scale – Varlitinib Metabolites; PK Population
14.3.6.12	Scatter Plots of Plasma Concentrations Against Change from Baseline in QTcF on a Linear Scale – Capecitabine Metabolites; PK Population