AstraZeneca

Statistical Agreement

Study Code	D3610C00002 (Part A)
Edition Number	3
Date	29 October 2015

A Phase I/II, Multicentre, Study Comprising a Safety Run-In of AZD5363 when Combined with Paclitaxel in Patients with Advanced or Metastatic Breast Cancer; Followed by a Randomised Expansion of AZD5363 when Combined with Paclitaxel vs. Paclitaxel plus Placebo in Patients with ER-Positive Advanced or Metastatic Breast Cancer, Stratified by *PIK3CA* Mutation Status (BEECH).

Document Author Signature [redacted] Dated 29.10.15

(PRINT NAME AND ORGANISATION) [redacted]

Statistical Science Director Signature [redacted] Dated 29/10/15

on behalf of AZ

(PRINT NAME) [redacted]

(ON BEHALF OF [redacted])

Statistical Agreement

Study Code	D3610C00002 (Part A)
Edition Number	3
Date	29 October 2015

A Phase I/II, Multicentre, Study Comprising a Safety Run-In of AZD5363 when Combined with Paclitaxel in Patients with Advanced or Metastatic Breast Cancer; Followed by a Randomised Expansion of AZD5363 when Combined with Paclitaxel vs. Paclitaxel plus Placebo in Patients with ER-Positive Advanced or Metastatic Breast Cancer, Stratified by *PIK3CA* Mutation Status (BEECH).

Document Author Signature _____ Dated _____

(PRINT NAME AND ORGANISATION)

Statistical Science Director Signature _____ Dated _____

on behalf of AZ

(PRINT NAME)

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BOR	Best Objective Response
cfDNA	circulating free plasma DNA
CP&P	Clinical Pharmacology & Pharmacometrics
CRF	Case Report Form (electronic/paper)
CR	Complete Response
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CT	Computed Tomography
CTCs	Circulating Tumour Cells
CTCAE	Common Terminology Criteria for Adverse Event
CV	Coefficient of variation
DLT	Dose-Limiting Toxicity
DNA	Deoxyribonucleic acid
ECG	Electrocardiogram
GSK3β	Glycogen synthase kinase
IP	Investigational Product
ITT	Intention to Treat
LD	Longest Diameter
LLOQ	Lower Limit of Quantification
LVEF	Left ventricular ejection fraction
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
MTD	Maximum Tolerated Dose
MUGA scan	Multi Gated Acquisition scan
NA	Not Applicable
NC	Non calculable
NE	Not Evaluable

Abbreviation or special term	Explanation
NQ	Non-quantifiable
NTD	Non-Tolerated Dose
NTL	Non-Target Lesion
NYHA	New York Heart Association
OAE	Other Significant Adverse Event
ORR	Objective Response Rate
PD	Progression of Disease
PDc	Pharmacodynamic
PK	Pharmacokinetic
PK/PD	Pharmacokinetic/Pharmacodynamic
PGx	Pharmacogenetic research
PIK3CA	Phosphoinositide-3-kinase, catalytic, alpha polypeptide
PR	Partial Response
PR	ECG interval measured from the beginning of the P wave to the beginning of the QRS complex
PRAS40	Proline-rich AKT substrate of 40 kDaltons
PTEN	Phosphatase and Tensin Homolog
QRS complex	A name for the combination of three of the graphical deflections seen on a typical electrocardiogram (ECG)
QT	ECG interval measured from onset of the QRS complex to the end of the T wave
QTc	QT interval corrected for heart rate
QTcB	interval duration corrected for changes on heart rate using Bazett's formula
QTcF	interval duration corrected for changes on heart rate using Friderica’s formula
RD	Recommended Dose
RDI	Relative Dose Intensity
RECIST 1.1	Response Evaluation Criteria In Solid Tumours, version 1.1
RR	the interval between successive Rs, where R is a point corresponding to the peak of the QRS complex of the ECG wave
SAE	Serious adverse event
SD	Stable Disease
SD	Standard Deviation

Abbreviation or special term	Explanation
SRC	Safety Review Committee
TL	Target Lesion
ULN	Upper Limit of Normal
WHO	World Health Organisation

AMENDMENT HISTORY

Date	Brief description of change
31 May 2013	SA amended as per revised CSP dated 30 April 2013
11 Dec 2014	SA amended as per revised CSP dated 13 May 2014
29 Oct 2015	Section 4.2.4 amended to add the paragraph on Burden of adverse event of Diarrhoea

1. STUDY OBJECTIVES

This statistical plan covers the First part (Part A) of the study only. The second part (Part B) of the study will be covered in a separate SAP.

1.1 Primary objectives

The primary objective of the First part (Part A) of this study is

- To assess the safety and tolerability of two intermittent dosing schedules of AZD5363 when combined with weekly paclitaxel in patients with advanced or metastatic breast cancer; and to recommend by assessment of dose limiting toxicities and other safety, tolerability, pharmacokinetic (PK) and pharmacodynamic (PDc) data, a dose and schedule of AZD5363 for further study when combined with weekly paclitaxel.

1.2 Secondary objectives

The secondary objective of the First part (Part A) of this study is

- To make a p reliminary assessment of the anti-tumour activity of AZD5363 when combined with paclitaxel by assessment of objective response rate (ORR), and the percentage of patients without progressive disease, at 12 weeks.
- To assess the PK of AZD5363 when combined with paclitaxel
- To assess the PK of paclitaxel alone and when combined with AZD5363
- To assess the pharmacokinetic/pharmacodynamic (PK/PD) relationship between plasma AZD5363 exposure and plasma concentrations of biomarkers (including phospo-PRAS40, total PRAS40, pAKT and pGSK3β) anti-tumour activity (assessed by RECIST 1.1)

1.3 Exploratory objectives

The exploratory objective of the First part (Part A) of this study is

- To investigate the relationship between plasma AZD5363 exposure and plasma concentrations of exploratory biomarkers and efficacy. Biomarkers may include, but are not restricted to, somatic mutation or amplification of genes on the PI3 kinase and related pathways in cfDNA.
- To obtain a p reliminary assessment of AZD5363 treatment effect by quantitative change in circulating tumour cells (CTCs).

- To investigate the concordance of PIK3CA mutation status between per-patient analyses of blood and archival tumour tissue samples.

- To explore changes in WHO performance status in patients treated with AZD5363 in combination with weekly paclitaxel compared with weekly paclitaxel plus placebo.

- To collect optional matched pre-and post- treatment tumour biopsy samples to conduct assessment of the PD effect of therapy compared to baseline.

- To collect and store archival tumour samples and analyse surplus blood or tissue, for potential future exploratory research into factors that may influence development of cancer and/or response to AZD5363 (where response is defined broadly to include efficacy, tolerability or safety). Biomarkers may include, but are not restricted to, somatic mutation or amplification of genes on the PI3 kinase and related pathways, PTEN protein expression and AKT protein expression. This exploratory analysis will be reported separately.

- To obtain blood samples for deoxyribonucleic acid (DNA) extraction for future exploratory research into genes/genetic variation that may influence response (i.e., distribution, safety, tolerability and efficacy) to AZD5363 treatment and/or susceptibility to cancer. This exploratory analysis will be reported separately.

2. DEFINITION OF ANALYSIS SETS

The analysis of data will be based on different subsets of overall patient population according to the purpose of analysis. Five analysis populations are defined for this study: Safety analysis set, DLT analysis set, PK analysis set, Full analysis set and Pharamacodynamic analysis set as defined in the table below (Table 1).

Table 1 Analysis sets

Analysis Set	Definition
Safety analysis set	All patients who received at least 1 dose of AZD5363.
DLT analysis set	All patients who either completed the DLT evaluation period with sufficient dosing (at least 80% of specified dose of AZD5363 or paclitaxel during the DLT evaluation period), or who experienced a DLT during that period.
PK analysis set	Dosed patients for whom concentration time data/PK profile data for AZD5363 or paclitaxel has been obtained.
Full analysis set	All patients who received at least 1 dose of AZD5363.
Pharmacodynamic analysis set	All patients dosed with AZD5363 that provided biological samples for pharmacodynamic research.

3. DESCRIPTION OF VARIABLES

Full details of the definitions of all efficacy and safety endpoints are given in Section 11 of the Clinical Study Protocol.

Table 2 Summary of outcome variables and analysis sets

Outcomes variables	Analysis set
Safety and tolerability(Primary) – Based on assessment of common terminology criteria for adverse event (CTCAE) grade and type of AE, Changes in laboratory values and Changes in vital signs and Changes in ECG parameters	Safety
Maximum tolerable dose/Recommended dose (Primary) – Number of DLTs	DLT
Efficacy(Secondary) – Tumour response based on RECIST data	Full
Pharmacokinetics (Secondary) – Plasma concentration and PK parameters based on blood and urine samples	PK
Pharmacodynamics (Secondary) – Pharmacodynamic Biomarker variables[1]	Pharmacodynamic

[1] Patients who have relevant data at baseline will be included in specific analyses based on Pharmacodynamic analysis set

Analysis on all exploratory endpoints (Biomarkers and Pharmacogenetics) will be reported separately from the CSR.

3.1 Safety and Tolerability

The safety data from all patients will be assessed on an ongoing basis. Safety and tolerability will be assessed using Dose limiting Toxicities (DLTs), adverse events (AEs), physical examination, ECGs, LVEF, laboratory and vital sign data as recorded on the CRF.

3.1.1 Dose Limiting Toxicities (DLT)

An AE or laboratory abnormality considered to be related to paclitaxel, that commences at any time during the DLT evaluation period (Cycle 1, comprises a 28-day evaluation period) and meets either of the following criteria

- Requires a dose reduction or omission to <60% of the intended Cycle 1 total paclitaxel dose (90 mg/m2 x 3).
- Any delay to the administration of weekly paclitaxel chemotherapy on D1 of Cycle 2 by ≥7 days as a consequence of paclitaxel induced toxicity

A DLT is defined as an AE or laboratory abnormality considered to be related to AZD5363 that commences at any time during the DLT evaluation period (Cycle 1) which includes:

- Haematological toxicity:
    - CTCAE grade 4 haematological toxicity of any duration
    - Febrile neutropenia (CTCAE grade ≥3 with temperature ≥38.5°C which is unresponsive to antipyretics)
    - CTCAE grade ≥3 neutropenia requiring hospitalisation
    - CTCAE grade ≥3 thrombocytopenia associated with non-traumatic bleeding (but not applicable to patients on therapeutic anticoagulation)

- Clinical chemistry toxicity:
    - CTCAE ≥ Grade 3 hyperglycaemia (glucose >13.9 mmol/L) for more than 1 week, despite optimal intervention which is not attributable to another co-morbidity
    - Grade 4 hyperglycaemia (glucose >27.8 mmol/L)
    - AST or ALT >10x ULN and AZD5363 is considered the most likely cause
    - AST or ALT >8x ULN, when combined with doubling of bilirubin from baseline, and AZD5363 is considered the most likely cause

- Cardiovascular toxicity:
    - QTc (Fridericia's or Bazett's correction) interval > 500 msec or QTc increase >60 msec from baseline on two ECGs at least 30 minutes apart, that cannot be attributed to another cause
    - Symptomatic congestive cardiac failure (New York Heart Association [NYHA] class III/IV) and a drop in left ventricular ejection fraction (LVEF) which cannot be attributed to another cause. Note: Concomitant use of vasotonic

drugs, adrenergic blockers, negative inotropic agents and anti-hypertensives should be taken into consideration when assigning causality.
  - A decrease in LVEF of ≥20% to a level below the institution's lower limit of normal range
  - CTCAE grade ≥3 hypotension which cannot be attributed to another cause

- Other toxicities:

  - Clinically significant rash that despite optimal treatment remains CTCAE grade ≥3 for 5 days or longer and that cannot be attributed to another cause
  - CTCAE grade ≥3 nausea, vomiting or diarrhoea, despite optimal anti-emetic or anti-diarrhoeal therapy, and which cannot be attributed to another cause
  - Any other CTCAE grade ≥3 toxicity which, in the opinion of the investigator, is clinically significant and related to AZD5363
  - Any delay to the administration of weekly paclitaxel chemotherapy on D1 of Cycle 2 by ≥7 days as a consequence of AZD5363 induced toxicity
  - Any other toxicity that is greater than that at baseline, is clinically significant and/or unacceptable, does not respond to supportive care and results in a disruption of dosing schedule of more than 14 days

- Any event, including significant dose reductions or omissions, judged to be a DLT by the SRC

A DLT excludes:

- Alopecia of any grade
- Isolated laboratory changes of any grade without clinical sequelae or clinical significance
- An immune allergic reaction <CTCAE Grade 4, thought to be related to either AZD5363, paclitaxel or the combination of the two agents which is manageable and not life-threatening.

Intra-patient dose escalation is not permitted. Dose reductions or delays are allowed during DLT evaluation period as per CSP (section 5.5.4). If during the DLT evaluation period, dose reductions or a delay is beyond allowed limits as described in CSP, the patient will be considered either as not evaluable or as having attained a DLT.

3.1.2 Adverse Event

Please refer to section 6.4.1 to 6.4.4 of Clinical Study Protocol (CSP).

3.1.3 Duration of exposure

Total duration of exposure to AZD5363 for each schedule will be calculated as

Date (last day of week of last dose of AZD5363) – Date (first dose of AZD5363) + 1

For paclitaxel:

For patients who discontinued paclitaxel therapy during the 1st or 2nd week of a dosing cycle

Date (last day of week of last dose of paclitaxel) – Date (first dose of paclitaxel) + 1

Otherwise

Date (last day of cycle of last dose of paclitaxel) – Date (first dose of paclitaxel) + 1

3.1.4 Dose intensity

Dose intensity of AZD5363 and paclitaxel will be addressed by considering relative dose intensity (RDI) defined as follows:

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 plus the protocol-defined post-dose rest period.

If there is no progression and a patient had no evaluable assessments then to calculate RDI for AZD5363, the patient will be censored at day 2 (i.e. the first day of AZD5363 dosing) and for calculating RDI for paclitaxel, the patient will be censored at day 1.

RDI for AZD5363 and paclitaxel will be calculated for the entire intended treatment period.

3.1.5 Laboratory data

Blood and urine samples for determination of clinical chemistry, haematology, and urinalysis will be taken at the times indicated in the Study Plan at the time points described in section 6.4.5 of the Clinical Study Protocol.

The baseline value of each laboratory variable will be derived, as described in the Clinical Study Protocol, values captured 3 days before baseline visit will be considered as baseline value. If multiple records are present between -3 days and baseline visit then last value obtained prior to the first dose of medication will be considered as baseline. Alternatively, if two visits are equally eligible to assess patient status at baseline (e.g., two assessment at screening or baseline both prior to first dose with no w ashout or other intervention in the screening period), the average can be taken as a baseline value.

The other post-baseline assessment non missing value lab parameter closest to the scheduled visit date, will be considered as visit value. The visit will be missing if no assessment was reported within the specified visit window around the scheduled date. If two assessments are equidistant from a scheduled visit, then the earlier of the two will be used.

Designation of visits for Lab data assessment are given in table below intermittent dosing schedule (2/5) (Table 3a)

Table 3a Day Ranges for LAB assessments

Cycle No.	Visit day	Target day	Day Range
	Screening	-1	-28 – -1
1	Day 1 (Baseline)	1	1
1	Day 2	2	2
1	Day 3	3	3 - 6
1	Day 9	9	7 – 12
1	Day 16	16	13 – 19
1	Day 23	23	20 – 26
2	Day 2	30	27 – 33
2	Day 8	36	34 – 39
2	Day 15	43	40 – 46
2	Day 22	50	47 – 54
3	Day 2	58	55 – 72
4	Day 2	86	73 – 100
Y	Day 2	X	X-13 – X+14

Where Y = 5,6, …. And X=(Y-1)*28+2

Designation of visits for Lab data assessment are given in table below for intermittent dosing schedule (4/3) (Table 3b)

Table 3b Day Ranges for LAB assessments

Cycle No.	Visit day	Target day	Day Range
	Screening	-1	-28 – -1
1	Day 1 (Baseline)	1	-3 – 1
1	Day 2	2	2 – 3
1	Day 5	5	4 – 7
1	Day 9	9	8 – 12
1	Day 16	16	13 – 19
1	Day 23	23	20 – 26
2	Day 2	30	27 – 33
2	Day 8	36	34 – 39

Table 3b Day Ranges for LAB assessments

Cycle No.	Visit day	Target day	Day Range
2	Day 15	43	40 – 46
2	Day 22	50	47 – 54
3	Day 2	58	55 – 72
4	Day 2	86	73 – 100
Y	Day 2	X	X-13 – X+14

Where y = 5,6, …. And X=(Y-1)*28+2

For Serum/Plasma Glucose, Insulin and Insulin c-peptide measurement a visit window rules are defined in table 3c and 3d for respective dosing cohorts.

Table 3c Day Ranges for Serum/Plasma Glucose, Insulin and Insulin c-peptide assessments (intermittent dosing schedule 2/5)

Cycle No.	Visit day	Target day	Day Range
	Screening	-1	-28 – -1
1	Day 1 (Baseline)	1	1
1	Day 3	3	2 – 6
1	Day 9	9	7 – 12
1	Day 16	16	13 – 19
1	Day 23	23	20 – 26
2	Day 2	30	27 – 44
3	Day 2	58	45 – 72
Y	Day 2	X	X-13 – X+14

Table 3d Day Ranges Serum/Plasma Glucose, Insulin and Insulin c-peptide assessments (intermittent dosing schedule 4/3)

Cycle No.	Visit day	Target day	Day Range
	Screening	-1	-28 – -1
1	Day 1 (Baseline)	1	-3 – 1
1	Day 2	2	2 – 3
1	Day 5	5	4 – 6

Table 3d Day Ranges Serum/Plasma Glucose, Insulin and Insulin c-peptide assessments (intermittent dosing schedule 4/3)

Cycle No.	Visit day	Target day	Day Range
1	Day 8	8	7 – 11
1	Day 15	15	12 – 18
1	Day 22	22	19 – 26
2	Day 2	30	27 – 44
3	Day 2	58	45 – 72
Y	Day 2	X	X-13 – X+14

3.1.6 Visit windows for other safety assessments

For Physical examination, weight, ECG, LVEF, Lipids, Glycosylated haemoglobin, and additional glucose assessments following conventions should apply to define visit response.

- The time windows should be exhaustive so that data recorded at any time point has the potential to be summarised. 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.
- 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).
- Listings should display all values contributing to a time point for a patient.
- For visit based summaries:
  - If there is more than one value per patient within a time window then the closest value should be summarised, or the earlier in the event the values are equidistant from the nominal visit date.
  - To prevent very large tables or plots being produced that contain many cells with meaningless data, for each treatment group visit data should only be summarised if the number of observations is greater than the minimum of 20 or > 1/3 of patients dosed.

- For summaries at a patient level, all values should be included, regardless of whether they appear in a corresponding visit based summary, when deriving a patient level statistic such as a maximum.

Unless otherwise stated, baseline will be defined as the last non-missing measurement prior to dosing with study treatment (AZD5363). For laboratory data and vital signs data, any assessments made on day 1 will be considered pre-dose. Where safety data are summarized over time, study day will be calculated in relation to date of first treatment (AZD5363).

Missing safety data will generally 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.

3.1.7 ECG

A 12-lead ECG will be performed on all patients as described in protocol section 6.4.9. The following parameters will be recorded for each ECG: date and time of ECG, heart rate (beats/min), PR, R-R, QRS, QT (ms), QTc (ms), sinus rhythm (yes/no), and overall evaluation (normal/abnormal).

3.1.8 Calculation or derivation of safety variables

ECG Changes
QTc will be calculated using both Bazett's (QTcB) and Fridericia's (QTcF) formulae as follows:

- $QTcB = \frac{QT}{\sqrt{\frac{RR}{1000}}}$
- $QTcF = \frac{QT}{\sqrt[3]{\frac{RR}{1000}}}$

where QT and RR is in msec.

Creatinine clearance will be calculated using the Cockcroft and Gault formula below, from creatinine:

For creatinine values in mol/L

Men: [(140 – age[1]) x weight[2] (kg) x 1.23] / creatinine (µmol/L)

Women: [(140 – age[1]) x weight[2] (kg) x 1.04] / creatinine (µmol/L)

[1] Age at baseline
[2] Last known weight

Corrected calcium will be calculated using the formula below, from total calcium and albumin:

Corr. Calcium = Total Calcium (mmol/L) + ([40 – Albumin (g/L)] x 0.02)

Baseline for laboratory variables will be defined as last non missing measurement prior to the first dose of Medication and change from baseline will be calculated for each scheduled visit using the formula:

Change = [(Post dose measurement - baseline measurement) / Baseline measurement]

And percentage change from baseline using,

Percentage change = [(Post dose measurement - baseline measurement) / Baseline measurement] *100

3.2 Calculation or derivation of pharmacokinetics variables:

Pharmacokinetic analysis of the plasma concentration data for AZD5363 and paclitaxel will be performed by CP&P, AstraZeneca. The plasma concentrations, determined using the sparse PK sampling scheme data will be analyzed using population non-linear mixed effects approach (Beal and Sheiner 1988-1998). The population PK analysis of data from this study will from part of the overall population analysis for the AZD5363 program and is described in a population PK analysis plan associated with this program.

The actual sampling times will be used in the parameter calculations. Details of sample collections are provided in CSP section 6.6.

In the calculation of plasma concentration summary statistics, plasma concentrations below the lower limit of quantification (LLOQ) will be set to zero at pre-dose. Summary statistics will be presented according to the following rules:

- If, at a given time point, 50% or less of the plasma concentrations are nonquantifiable (NQ) then geometric mean (gmean), coefficient of variation (CV), gmean ± standard deviation (SD), arithmetic mean, SD and median will be calculated by substituting the LLOQ for values which are NQ. The minimum at that time point will be reported as NQ.
- If more than 50%, but not all, of the concentrations are NQ then gmean, CV, gmean ± SD, arithmetic mean and SD will be reported as not calculable (NC). The minimum and median at that time point will be reported as NQ.
- If all the concentrations are NQ then gmean, arithmetic mean, median, minimum and maximum will be reported as NQ and CV, gmean ± SD and SD as NC.
- If data are available for less than 3 patients, no summary statistics other than minimum, maximum and N will be presented.
- If the calculation of the gmean - SD results in a value less than the LLOQ, NQ will be displayed.

The lower limits of quantification (LLOQ) for AZD5363 and paclitaxel in plasma are 1.0ng/mL and 10ng/mL respectively. Full details of the analytical methods used will be described in separate bioanalytical reports.

3.3 Derivation of tumour response

For all patients, the RECIST tumour response data will be used to determine each patient's visit response according to RECIST version 1.1. Patients must present with at least one lesion (measurable and/or nonmeasurable) that can be accurately assessed at baseline by (CT) or magnetic resonance imaging (MRI) and is suitable for repeated assessment.

Baseline radiological tumour assessments are to be performed no more than 28 days before the start of study treatment. If more than one baseline scan is recorded, then measurements from the one that is closest to start of study treatment will be used. Follow-up assessments will be performed every 12 weeks (±7 days) after start of treatment until objective disease progression as defined by RECIST 1.1.

At each visit, an overall visit response will be determined programmatically - using the information from target lesions (TL), non-target lesions (NTL) and new lesions. RECIST outcomes will be calculated by AZ programming or designee.

3.3.1 Target lesions (TLs)

Measurable disease is defined as having at least one measurable lesion which is ≥10 mm in the longest diameter (except lymph nodes which must have short axis ≥15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI). Previously irradiated lesions will not be considered measurable. A patient can have a maximum of 5 measurable lesions representative of all involved organs (maximum of 2 lesions per organ) recorded at baseline and these are referred to as target lesions.

At each visit, sites will enter the longest diameter (LD) in mm for each TL. If tumour is too large to measure then the maximum measurable size has been recorded. If tumour is too small to measure then the minimum measurable size (5 mm or another value the radiologist indicates is reliable) will be recorded. If a TL is too small or too large to be measured, that will be indicated in the patient data listing.

3.3.1.1 Sum of LD

For each visit, the sum of LDs will be calculated as the sum of the longest diameters recorded for each TL. If the LD is missing then it will be handled as per section 3.3.1.2.

CT, MRI and clinical examination are the only methods of assessment that can be used within a trial, with CT and MRI being the preferred methods and clinical examination only used in special cases. If a change in method of assessment occurs between CT and MRI this will be considered acceptable and no adjustment within the programming is needed. [redacted]

[redacted]

3.3.1.2 Missing TL data

All TL data missing at a visit

If all target lesion measurements are missing then the target lesion visit response is Not Evaluable (NE). The overall visit response will also be NE, unless there is a progression of non-target lesions or new lesions, in which case the response will be progressive disease (PD).

Some TL data missing at a visit (scaling-up rule)

- [redacted]
- [redacted]
- [redacted]
- [redacted]

3.3.1.3 Percent change from baseline in sum of LD

Percent change from baseline in sum of LD value will be derived as follows:

The percent change from baseline in sum of LD

$$= \frac{\text{Visit sum of LD} - \text{Baseline sum of LD}}{\text{Baseline sum of LD}} \times 100$$

The baseline sum of LD is calculated as the sum of longest diameters from the assessment closest to start of study treatment.

The percentage change from baseline in sum of LD should be rounded to 1 decimal place. For example 19.95% should be rounded to 20.0% but 19.94% should be rounded to 19.9%.

3.3.1.4 Percent change from minimum in sum of LD

Percent change from minimum in sum of LD value will be derived as follows:

The percent change from minimum in sum of LD

$$= \frac{\text{Visit sum of LD} - \text{Minimum sum of LD}}{\text{Minimum sum of LD}} \text{ x } 100$$

The minimum sum of LD is calculated as the minimum of sum of LD from any previous visit including baseline.

The percentage change from minimum in sum of LD should be rounded to 1 decimal place. For example 19.95% should be rounded to 20.0% but 19.94% should be rounded to 19.9%.

3.3.1.5 TL visit response

TL visit response is calculated based on assessment of TLs and will be determined for each visit as CR, PR, PD, SD or NA.

Table 4 gives definition of TL visit responses.

Table 1 Definition of TL visit responses

Visit Responses	Description
Complete Response (CR)	Disappearance of all target lesions. Any pathological lymph nodes selected as target lesions must have a reduction in short axis to <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of target lesions and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Not applicable (NA)	No target lesions are recorded at baseline

TL Visit responses subsequent to CR

A CR response can only be followed by CR, PD or NE. If a CR has occurred then the following rules at the subsequent visits must be applied:

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

3.3.2 Irradiated lesions/lesion intervention

Previously irradiated lesions (i.e. lesion irradiated prior to entry into the study) should be recorded as NTLs and should not form part of the TL assessment.

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation / palliative surgery / embolisation), should be handled in the following way and once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumours:

- [redacted]
- [redacted]
- [redacted]

[redacted]

3.3.3 Non-target lesions

The non-target lesion response will be based on the Investigators overall assessment of NTLs. Investigators will characterize NTL response as PD, CR, Non-CR/Non-PD, NE or NA.

3.3.4 New lesions

New lesions will be identified via a Yes/No tick box. If new lesion response (yes/no) is missing and the new lesion details are blank it should be treated as NE.

3.3.5 Overall visit response

A complete list of overall visit responses based on different TL responses, NTL responses and yes/no of new lesions is given in Table 5. Overall visit response will be determined for each visit.

Table 5 Overall visit response

Target	Non-Target	New Lesions	Overall Visit Response
CR	CR (or NA)	No	CR
CR	Non-CR/Non-PD or NE	No (or NE)	PR
PR	Non-PD or NE	No (or NE)	PR
SD	Non-PD or NE	No (or NE)	SD
PD	Any	Any	PD
Any	PD	Any	PD
Any	Any	Yes	PD
NE	Non-PD	No	NE
NA	CR	No	CR
NA	Non-CR/Non-PD	No	SD
NA	NE	No (or NE)	NE
NA	Non-PD	NE	SD

Note: For patients with non-measurable disease at baseline, categorisation of objective tumour response assessment will be based on the RECIST 1.1 criteria of response: complete response (CR), PD (progression of disease) and Non CR/Non PD.

3.3.6 Best objective response (BOR)

BOR will be calculated based on the overall visit responses obtained up until RECIST progression is documented. In the absence of RECIST progression, BOR is determined using visit responses up until the last evaluable overall visit response.

For patients who die in the absence of progression and have no evaluable overall visit responses prior to death, set the BOR to PD.

A patient's overall best objective response will be determined:

- CR: Overall visit response of CR confirmed at least 4 weeks later by another overall visit response of CR
- PR: Overall visit response of PR confirmed at least 4 weeks later by an overall visit response of at least PR (i.e. CR/PR) with no intervening PD
- SD≥ 12 weeks: Stable disease recorded at least 84 days after treatment (ks from start of treatment). If patients had an overall visit response of SD on day 28 but then progressed on day 84 this will not be considered as SD ≥ 12 weeks.
- PD: Progression, or death in the absence of CR/PR or SD
- NE: No evidence of CR/PR or SD or PD or death
- For patients who progress and subsequently have a response, then the best objective response is only derived from assessments
  up to and including the time of the progression (i.e. it will not include the response after the patient has progressed).

3.3.7 Objective response rate (ORR) at week 12

ORR at week 12 is defined as the percentage of patients who have a week 12 visit response of CR or PR prior to any evidence of progression (as defined by RECIST 1.1) and regardless of whether the responses are confirmed or not.

A visit response of CR is defined when all TL and NTL lesions present at baseline have disappeared (with the exception of lymph nodes which must be <10mm to be considered non-pathological) and no new lesions have developed since baseline. A visit response of PR is defined when the sum of diameters of the TLs has decreased by 30% or more compared to baseline (with no evidence of progression) and the NTLs are at least stable with no evidence of new lesions.

The denominator for all response rate calculations will be based on Full analysis set.

3.3.8 Proportion of patients without progressive disease at 12 weeks

The proportion of patients without progressive disease at 12 weeks is defined as the percentage of patients with a 12 week visit response of CR, PR or SD (as defined by RECIST 1.1) with no evidence of previous progression. The denominator will be the number of patients in Full analysis set.

3.3.9 Changes in Tumour size

Percentage change in tumour size will also be determined for patients with measurable disease at baseline and is derived at each visit as the percentage change from baseline in the sum of the diameters of TLs.

Percentage change = [(Post dose TL sum-baseline TL sum)/Baseline TL sum]*100

3.4 Derivation of pharmacodynamic variables

Secondary biomarkers, phospo-PRAS40, total PRAS40, pAKT, pGSK3β from platelet rich plasma samples will be reported in the Clinical Study Report (CSR) other exploratory biomarker research will be reported separately and will not be part of the CSR.

3.5 Derivation of other exploratory Biomarker research variables

Results from the other exploratory biomarker and pharmacogenetic research will be reported separately from the CSR.

4. ANALYSIS METHODS

4.1 General principles

All summary data will be presented by initial dose levels (by cohort) and dosing schedule (2/5 and 4/3). All captured data will be listed and summarized as appropriate.

There is no formal statistical analysis required for Part A of this study. The below mentioned general principles will be followed throughout the study:

- All summaries (safety, tolerability, PK, PD and efficacy) will be presented by the initial dose level and schedule, unless otherwise specified.

- Descriptive statistics will include number of non-missing patients (n), mean, standard deviation (SD), median, minimum, and maximum values for continuous variables, and for categorical variables the frequencies and percentages of patients will be presented.

- For continuous data, mean, SD and median will be rounded to 1 additional decimal place compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.

- For categorical data, percentages will be rounded to 1 decimal place.

- In addition, for PK data geometric mean and coefficient of variance (CV) will also be displayed as appropriate. CV will be rounded to 2 additional decimal places compared to the original data.

4.2 Description of Analysis methods

4.2.1 Demographic, Baseline Characteristics and Disposition

Demographics, baseline characteristics, disposition, inclusion in analysis populations, previous anti-cancer treatments, disease characteristics at baseline, chemotherapy and concomitant medications will be summarised by initial dose level (i.e. by cohort) & schedules and listed for all patients. Important deviations from protocol and medical history data will be listed for all patients.

4.2.2 Exposure

Exposure to investigational product, i.e., total amount of AZD5363 received will be listed and summarized. For intermittent dosing average weekly dose of AZD5363 received will be summarized by treatment cohorts.

Total exposure to investigation product and total time on study will be summarised by mean, standard deviation, minimum, maximum, median and number of observations by initial dose level.

The following summaries related to study treatment will be produced by initial AZD5363 dose level:

- Total exposure to AZD5363
- Number (%) of patients summarised by total number of paclitaxel cycles
- Relative dose intensity of AZD5363 and paclitaxel
- Number of days on AZD5363 study treatment (calculated two ways; by including and then excluding dose interruptions)
- Number of patients with at least one dose interruptions, dose reductions will be presented separately for the initial period of evaluability defined as 28 days or end of cycle 1 dosing and for any time following this initial period of the study.
- Duration of initial AZD5363 dose

In addition, a plot of AZD5363 dose over time will be presented, showing one line for each patient, with different line types for different dose levels and symbols to indicate when doses got interruptions, reduction and discontinued.

Reasons for discontinuation of investigational product AZD5363 and paclitaxel will be listed including the study day of treatment discontinuation.

4.2.3 Safety

Safety and tolerability will be assessed in terms of AEs, laboratory data, vital signs, ECG, LVEF/MUGA scans/echocardiograms assessments. These safety measures will be listed and summarized descriptively in the safety analysis set. Data from all cycles of initial treatment will be combined in the presentation of safety data. Safety analysis set will be used to present all safety data.

The following sections describe the planned safety summaries. However, additional safety tables (not specified in this statistical agreement) may need to be produced to aid interpretation of the safety data.

4.2.4 Adverse Events

Data from all cycles of initial treatment will be combined in the presentation of safety data. All adverse events will be listed individually for each patient and dose group (dose and schedule) and summarised according to the system organ class (SOC) and preferred term (PT) assigned to the event using the MedDRA. AEs will be graded according to the CTCAE version 4.0 for AEs. The CTCAE grade will be assigned by the investigator.

Any AEs occurring after the first dose of study treatment and within 28 days of the last dose of study treatment will be included in the AE summaries. AEs occurring before the first dose of study treatment or more than 28 days after the last dose of study treatment will not be included in AE summaries but will be included and presented in the patient listings.

AEs will be categorized into one or more of the following categories depending on the type of events reported:

- All AEs
- All AEs (by episode)
- All AEs with causality related to study medication (AZD5363/paclitaxel/both)
- All serious adverse events (SAEs)
- All SAEs with causality related to study medication (AZD5363/paclitaxel/both)
- AE leading to study medication (AZD5363/paclitaxel/both) interrupted
- AEs leading to discontinuation of study medication (AZD5363/paclitaxel/both)

- AEs leading to discontinuation of study medication (AZD5363/paclitaxel/both), causally related to study medication (AZD5363/paclitaxel/both).
- AEs with an outcome of death
- AEs with an outcome of death, causally related to study medication (AZD5363/paclitaxel/both)
- AEs with CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, causally related to study medication (AZD5363/paclitaxel/both)
- Other significant AEs

An overall summary of the number and percentage of patients in each category will be presented, as well as an overall summary of the number of events in each category.

The number and percentage of patients reporting adverse events in each category above will be summarised by MedDRA system organ class and preferred term, by dose group.

AEs will be assigned a CTCAE version 4.0 grade and summaries of the number and percentage of patients will be provided by worst CTCAE grade, preferred term and dose group.

All adverse event data will be listed for all patients. Investigator terms with their MedDRA preferred terms will be listed for each patient. Listings will include the last dose for the patient along with the number of days since last dose. In addition, serious adverse events, other significant adverse events, adverse events that led to withdrawal, adverse events with causality AZD5363 and adverse events of special interest will be listed. Adverse events that occur after the last AZD5363 dose and within 28 days of the last AZD5363 dose will be attributed to the visit in which the last dose was given.

A summary of deaths (all death, AE outcome = death) will be provided by number and percentage of patients by treatment group.

Burden of adverse event of Diarrhoea

The burden of AE of Diarrhoea will be determined for each patient in the safety population and summarized. [redacted]

[redacted]

Dose Limiting Toxicities (DLTs)

All observed DLTs will be listed individually by patient and summarised by initial dose level. In addition, to facilitate the discussion of DLTs, all CTCAEv4 grade ≥3 AEs, grade ≥3 laboratory data and QTc values >500msec during the initial 29 day evaluation period will be listed.

AEs of special interest

Grouped summary tables of certain MedDRA preferred terms will be produced to depict AEs of special interest. An example grouping of AEs will be rash, diarrhoea etc and a combined total of all. Other groupings based on preferred terms will be provided by the medical team, and a listing of the preferred terms in each grouping will be provided.

The following summary tables and plots will be provided for each of these groups of AEs of special interest:

- Patients who had at least 1 AE of special interest in particular time intervals (example, 1-7 days, 8-14 days etc) will be summarized by first onset of AE by highest CTCAE grade by preferred term only.

- The duration of AEs of special interest will be summarized (N, Mean, SD, Median, Min, Max and total treatment days) for each dose cohort and for each preferred term.

- Events leading to discontinuation of AZD5363 will be summarized. This summary will be repeated for events leading to dose modification.

4.2.4.1 Cardiac measurements

12-Lead ECG

Overall evaluation of ECG will be summarised by cycle using frequency counts and percentage of patients for each treatment group as normal or abnormal, and the relevance of the abnormality will be summarised by “clinically significant” or “not clinically significant”.

In addition, a shift table comparing baseline (normal, abnormal - not clinically significant, abnormal - clinically significant) to last observation on treatment will also be presented.

Continuous ECG measurements will be summarised over time in terms of absolute values and change from baseline at each scheduled measurement by dose group. When there are multiple observations in a visit, average will be taken.

The baseline value for each parameter is defined as the last available value before the first dose of study treatment in cycle 0.

The number and percentage of patients with increases of QTcB and QTcF of >30msec and >60msec from baseline, and QTcB and QTcF values of >450msec, >480msec and >500msec will be produced.

MUGA scans/echocardiograms

MUGA scans/echocardiograms will be listed.

4.2.4.2 Vital signs

Vital signs (blood pressure and pulse) will be summarised using descriptive statistics over time in terms of absolute values and changes from baseline at each scheduled time point by dose group using descriptive statistics.

4.2.4.3 Laboratory assessments

All laboratory data will be summarised at each scheduled time point and by dose group using descriptive statistics. Change from baseline on continuous data will be summarised using descriptive statistics at each scheduled time point by dose group. For categorical data, shift from baseline will be summarised using frequency and proportion at each scheduled time point by dose group.

For all laboratory variables, which are included in the CTCAE version 4.0, the CTCAE grade will be calculated and summarised using frequency counts and percentages in the form of shifts from baseline to maximum grade post baseline.

For urinalysis, shift table comparing baseline to maximum value by treatment will be presented (i.e. using number of patients with results of negative, trace or positive).

4.2.4.4 Physical examination

All individual physical examination data will be summarised and listed.

4.2.4.5 Other safety data

Other safety data for serum/plasma glucose measurement, weight, Lipids profile, Glycosylated haemoglobin, Glucose, insulin and insulin c-peptide will be listed individually by patient and summarised by dose group.

4.2.5 Pharmacokinetics

Plasma concentrations of AZD5363 and paclitaxel will be summarised by nominal sample time points (Pre-dose, 2 hr, 4 hr, 6hr & 8 hr). Plasma concentrations will be summarised by

dose level and schedule. Plasma concentrations at each time point will be summarised according to actual dose received by the following summary statistics:

- The geometric mean (gmean, calculated as exp [μ], where μ is the mean of the data on a logarithmic scale)
- Coefficient of variation (CV, calculated as 100 √ [exp(s^2)-1], where s is the standard deviation of the data on a log scale)
- Gmean ± standard deviation (calculated as exp[μ±s])
- Arithmetic mean calculated using untransformed data
- Standard Deviation calculated using untransformed data
- Minimum
- Maximum
- Median
- Number of observations

The plasma concentration data for AZD5363 and paclitaxel will also be displayed graphically. Displays will include plasma concentration patient profiles (on the linear and log-scale) versus time and gmean concentration (+/-standard deviation) versus time, stratified by dose.

Results from population PK analysis will be reported separately to the Clinical Study Report (CSR). Data from this study may also form part of a pooled analysis with other AZD5363 studies.

4.2.6 Tumour response

The following will be summarised by dose and schedule, using best objective response categories: CR, PR, SD, PD and NE.

- Best objective response during the study including the breakdown of SD into unconfirmed CR and PRs, week 12 visit response, ORR at week 12, proportion of patients without progressive disease at week 12 (derived as number of CR, PR and SD at week 12)
- Changes in tumour size at 12 weeks for patients with measurable disease would be presented by waterfall plots and summaries - if there is not sufficient data for waterfall plots individual patient profile plots may be considered

4.2.7 Biomarker analysis

All biomarker analysis will be performed on the Pharmacodynamic analysis set patients for whom data is available. The following biomarkers will be reported in CSR. Other exploratory biomarkers will be reported separately and will not be part of the CSR.

- phospo-PRAS40
- Total PRAS40
- pAKT
- pGSK3β

For each biomarker, the absolute value and the change from baseline will be summarised at each time point, by dose and schedule.

Standard summary statistics n, mean, standard deviation, median, minimum and maximum will be presented.

Box-plots and line plots will be produced for each biomarker by dose cohort, schedule and time. Line plots over time will be produced for each biomarker, for all patients along with an extra line reflecting median biomarker levels, for each cohort separately.

5. CHANGES OF ANALYSIS FROM PROTOCOL

Not applicable.

6. REFERENCES

Beal and Sheiner 1988-1998
Beal, SL, and Sheiner, LB (1988-1998). NONMEM Users Guides. Part I-VIII.NONMEM Project Group C255, University of California at San Francisco.

AstraZeneca

Statistical Analysis Plan

Study Code	D3610C00002 (B)
Edition Number	3.0
Date	03/02/2017

A Phase I/II, Multicentre, Study Comprising a Safety Run-In of AZD5363 when Combined with Paclitaxel in Patients with Advanced or Metastatic Breast Cancer; Followed by a Randomised Expansion of AZD5363 when Combined with Paclitaxel vs. Paclitaxel plus Placebo in Patients with ER-Positive Advanced or Metastatic Breast Cancer, Stratified by *PIK3CA* Mutation Status (BEECH).

A Phase I/II, Multicentre, Study Comprising a Safety Run-In of AZD5363 when Combined with Paclitaxel in Patients with Advanced or Metastatic Breast Cancer; Followed by a Randomised Expansion of AZD5363 when Combined with Paclitaxel vs. Paclitaxel plus Placebo in Patients with ER-Positive Advanced or Metastatic Breast Cancer, Stratified by *PIK3CA* Mutation Status (BEECH).

Study Statistician [redacted] 03/02/17
Date

A Phase I/II, Multicentre, Study Comprising a Safety Run-In of AZD5363 when Combined with Paclitaxel in Patients with Advanced or Metastatic Breast Cancer; Followed by a Randomised Expansion of AZD5363 when Combined with Paclitaxel vs. Paclitaxel plus Placebo in Patients with ER-Positive Advanced or Metastatic Breast Cancer, Stratified by *PIK3CA* Mutation Status (BEECH).

Global Product Statistician [redacted] 6th Feb 2017.
Date

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
AZ	AstraZeneca
BOR	Best Objective Response
BP	Blood Pressure
CP&P	Clinical Pharmacology & Pharmacometrics
CR	Complete Response
CTC	Common Terminology Criteria
CTCs	Circulating Tumour Cells
CTCAE	Common Terminology Criteria for Adverse Event
CV	Coefficient of variation
DBP	Diastolic Blood Pressure
DoR	Duration of Response
ECG	Electrocardiogram
ER	Estrogen Receptor
GSK3β	Glycogen synthase kinase
HDL	High Density Lipoprotein
HR	Hazard Ratio
IP	Investigational Product
ITT	Intention to Treat
IVRS	Interactive Voice Response System
LLOQ	Lower Limit of Quantification
LVEF	Left ventricular ejection fraction
NE	Not Evaluable
NQ	Non-quantifiable
NTL	Non-Target Lesion
OAE	Other Significant Adverse Event
ORR	Objective Response Rate
PD	Progression of Disease
PK/PDc	Pharmacokinetic/Pharmacodynamic
PFS	Progression Free Survival

Abbreviation or special term	Explanation
PIK3CA	Phosphoinositide-3-kinase, catalytic, alpha polypeptide
PR	Partial Response
PR	ECG intervals measured from the beginning of the P wave to beginning of the QRS complex.
PRAS40	Proline-rich AKT substrate of 40 kDaltons
PTEN	Phosphatase and Tensin Homolog
QoL	Quality of Life
QT	ECG interval measured from onset of the QRS complex to the end of the T wave
QTc	QT interval corrected for heart rate
QTcB	interval duration corrected for changes on heart rate using Bazett's formula
QTcF	interval duration corrected for changes on heart rate using Friderica's formula
RECIST 1.1	Response Evaluation Criteria In Solid Tumours, version 1.1
R-R	the interval between successive Rs, where R is a point corresponding to the peak of the QRS complex of the ECG wave
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SD	Stable Disease
SD	Standard Deviation
ULN	Upper Limit of Normal

1. STUDY DETAILS

This statistical analysis plan (SAP) covers the Phase II (Part B) of the study only. The Phase I safety run-in part (Part A) of the study is not covered in this SAP.

1.1 Study objectives

1.1.1 Primary objective

The primary objective of the second part (Part B) of this study is:

To assess the relative efficacy of AZD5363 when combined with weekly paclitaxel compared with weekly paclitaxel plus placebo by assessment of progression free survival (PFS) in the overall advanced or metastatic ER+ve breast cancer population and in the *PIK3CA* mutation-positive sub-population.

1.1.2 Secondary objective(s)

The secondary objectives of the second part (Part B) of this study are:

- To assess the relative efficacy of AZD5363 when combined with weekly paclitaxel compared with weekly paclitaxel plus placebo by assessment of objective response rate (ORR) at 12 weeks, best objective response (BOR), durable response rate (DRR) and duration of response (DoR).
- To assess the relative anti-tumour activity of AZD5363 when combined with weekly paclitaxel vs. weekly paclitaxel plus placebo by comparison of change in tumour size at 12 weeks (target lesion assessment using Response Evaluation Criteria In Solid Tumours [RECIST 1.1]).
- To compare overall survival (OS) in patients treated with AZD5363 in combination with weekly paclitaxel compared with weekly paclitaxel plus placebo by assessment of time to death.
- To further assess the safety and tolerability of AZD5363 when combined with weekly paclitaxel compared with weekly paclitaxel plus placebo.
- To investigate the effect on patients' quality of life (QoL) of AZD5363 when combined with weekly paclitaxel compared with weekly paclitaxel alone by change from baseline, utilising a patient-completed QoL questionnaire.
- To assess the pharmacokinetics (PK) of AZD5363 when combined with paclitaxel.
- To assess the PK of paclitaxel alone and when combined with AZD5363.
- To assess the Pharmacokinetic /Pharmacodynamic (PK/PDc) relationship between plasma AZD5363 exposure and plasma concentrations of biomarkers (including

phospo-PRAS40, total PRAS40, pAKT and pGSK3β) anti-tumour activity (assessed by RECIST 1.1).

- To assess the toxicity burden associated with diarrhoea in relation to the number of episodes, duration of episodes and variations in intensity within episodes of the event's occurrence.

1.1.3 Exploratory objectives

- To investigate the relationship between plasma AZD5363 exposure and plasma concentration of exploratory biomarkers and efficacy. Biomarkers may include, but are not restricted to, somatic mutation or amplification of genes on the PI3 Kinase and related pathways in circulating tumour plasma DNA (ctDNA).

- To obtain a preliminary assessment of AZD5363 treatment effect by quantitative change in circulating tumour cells (CTCs).

- To investigate the concordance of *PIK3CA* mutation status between per-patient analyses of blood and archival tumour tissue samples.

- To explore changes in WHO performance status in patients treated with AZD5363 in combination with weekly paclitaxel compared with weekly paclitaxel plus placebo.

- To collect optional matched pre-and post- treatment tumour biopsy samples to conduct assessment of the PDc effect of therapy compared to baseline.

- To collect and store archival tumour samples and analyse surplus blood or tissue, for potential future exploratory research into factors that may influence development of cancer and/or response to AZD5363 (where response is defined broadly to include efficacy, tolerability or safety). Biomarkers may include, but are not restricted to, somatic mutation or amplification of genes on the PI3 kinase and related pathways, PTEN protein expression and AKT protein expression. This exploratory analysis will be reported separately.

- To obtain blood samples for DNA extraction for future exploratory research into genes/genetic variation that may influence response (ie, distribution, safety, tolerability and efficacy) to AZD5363 treatment and/or susceptibility to cancer. This exploratory analysis will be reported separately.

1.2 Study designs

A double-blind, stratified and randomised evaluation of AZD5363 when combined with weekly paclitaxel vs. weekly paclitaxel plus placebo. Approximately 100 female patients, with ER+ve advanced or metastatic breast cancer will be enrolled in this phase, of which

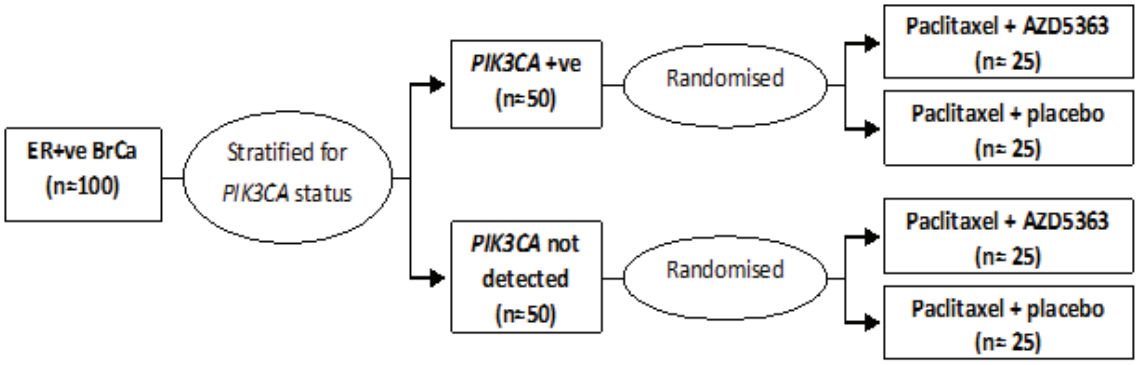
approximately 50 must have a tumour detectable as carrying a *PIK3CA* mutation. The dose and schedule of AZD5363 will be selected as an outcome from the safety run-in, Part A.

All enrolled patients will be stratified to either the *PIK3CA* mutation 'detected' or 'not detected' group by the outcome of two tests (tissue and blood (ctDNA)). If a patient is detected as having the *PIK3CA* mutation in both or either of the two tests, they will be assigned to the mutation detected strata. If the mutation is not detected in both of the two tests, (or should one test be not detected and the other is either unknown or not completed for any reason) the patient will be assigned to the not detected strata. This will be referred to as the protocol definition of *PIK3CA* mutational status. Patients within each stratum will be randomised to 28-day cycles comprising either of:

- Paclitaxel 90 mg/m^2 IV once weekly plus AZD5363 capsules taken orally, twice daily.
- Paclitaxel 90 mg/m^2 IV once weekly plus AZD5363-matching placebo capsules taken orally, twice daily.

The overall study plan for PART B is shown in Figure 1.

Figure 1 Study design



1.3 Number of subjects

For this part of study (PART B), the sample size calculation was based on the primary endpoint of progression free survival (PFS). A sample size is determined for each *PIK3CA* mutation status strata and irrespective of *PIK3CA* status (i.e. an overall sample size to compare treatment groups).

[redacted]

[redacted]

2. ANALYSIS SETS

2.1 Definition of analysis sets

The main analysis sets are defined below for this study:

2.1.1 Intention to treat (ITT)

The intention to treat analysis set (ITT) will include all randomised patients. The ITT analysis set will be the primary analysis set for the analysing the efficacy parameters. Patients in the ITT analysis set will be analysed according to the treatment they were randomised to receive by the interactive voice response system (IVRS) regardless of the treatment they actually received.

2.1.2 Safety analysis set

The safety analysis set will include all patients who received at least 1 dose of AZD5363/matching placebo. Any patient who only receives one or more doses of paclitaxel and does not commence dosing with AZD5363/matching placebo will not be included in the assessment of safety data.

The safety analysis set will be the primary set used for the safety analyses. Patients in the safety analysis set will be analysed by actual treatment initially received. If patients were dose reduced during the study, all data will be summarised/analysed based on the initial dose of study treatment received. However, some listings such as adverse event (AE) listings may give the actual dose the patient received at the time of the AE (e.g. dose interruption etc).

2.1.3 Pharmacokinetic (PK) analysis set

The AZD5363 PK analysis set will comprise all patients who received at least one dose of AZD5363 and have who have reportable plasma concentration data for AZD5363 and have no important adverse events that would impact PK data (e.g. vomiting).

The paclitaxel PK analysis set will comprise all patients who received at least one dose of paclitaxel and have who have reportable plasma concentration data for paclitaxel and have no important adverse events that would impact PK data (e.g. vomiting).

2.1.4 Pharmacodynamic (PD) analysis set

The PD analysis set will comprise all patients who received at least one dose of AZD5363 and/or paclitaxel and have who have reportable PD data with no important protocol deviations that may impact PD.

A summary of outcome variables and analysis sets is defined in Table 1.

Table 1 Summary of outcome variables and analysis sets

Outcomes variables	Analysis set
Efficacy (Primary)	
Progression free survival (PFS)	ITT
Efficacy (Secondary):	
Objective Response Rate (ORR)	ITT
Best Objective Response (BOR), Duration of Response (DoR), Durable Response Rate (DRR)	
Change in tumour size	
Overall Survival	
Quality of Life (QoL)	
Pharmacokinetics (Secondary)	
Plasma Concentration of AZD5363	PK
Plasma Concentration of paclitaxel	
Pharmacodynamics (Secondary)	PD
Biomarker: phospo-PRAS40, total PRAS40, pAKT and pGSK3β	
Safety and tolerability (Secondary):	Safety
AEs including specific safety data (i.e. Skin reactions, diarrhoea score)	
Laboratory data (haematology, clinical chemistry, urinalysis); Vital signs, ECG, Physical examination, cardiac examination	
Biomarkers (Exploratory)	PD
Circulating tumour cells (CTCs).	

2.2 Violations and deviations

Protocol deviations are recorded directly on the eCRF, the important protocol deviations will be identified prior to data base lock by the study team. All important protocol deviations will be listed summarised on the basis of randomised treatment group.

None of the deviations will lead to any patients being excluded from any of the analysis sets described in Section 2.1 (with the exception of the PK analysis sets, if the deviation/AE (e.g. vomiting) is considered to impact upon PK). If the deviations are serious enough to have the potential to impact the primary analysis, sensitivity analyses may be performed.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST visit responses

For all subjects, the RECIST tumour response data will be used to determine each subject's visit response according to RECIST version 1.1. It will also be used to determine if and when a subject has progressed in accordance with RECIST and also their best objective response.

Baseline radiological tumour assessments are to be performed no more than 28 days before the start of randomised treatment. Tumour assessments are then performed every 12 weeks following the start of treatment until objective disease progression (using RECIST v1.1).

At each visit, an overall visit response will be determined programmatically - using the information from TLs, NTLs and new lesions. RECIST outcomes will be calculated using a computer program.

Patients with at least one lesion, not previously irradiated and not chosen for an optional fresh biopsy during the study screening period, that can be accurately measured at baseline as ≥10mm in longest diameter (except lymph nodes which must have short axis ≥15mm) by computed tomography (CT) or magnetic resonance imaging (MRI) and is suitable for accurate repeat assessment will be included in this study.

Measurable disease:

A lesion, not previously irradiated, that can be accurately measured at baseline as ≥10 mm in the longest diameter (except lymph nodes which must have short axis ≥15 mm) with CT or MRI and which is suitable for accurate repeated measurements.

Non-measurable disease:

- All other lesions, including small lesions (longest diameter <10 mm or pathological lymph nodes with ≥10 to <15 mm short axis at baseline).
- Previously irradiated lesions

3.1.1 Target lesions (TLs)

A subject can have a maximum of 5 measurable lesions recorded at baseline and these are referred to as the target lesions. If more than one baseline scan is recorded then measurements from the one that is closest to and prior to randomisation will be used to define the baseline sum of TLs.

Baseline assessments should be performed no more than 28 days before the start of study treatment, and ideally should be performed as close as possible to the start of study treatment.

Note: For patients who do not have measurable disease at entry (i.e. no TLs), evaluation of overall visit responses will be based on the overall non-target lesion (NTL) assessment and the absence/presence of new lesions. If a patient does not have measurable disease at baseline then the TL visit response will be not applicable (NA).

Table 2 TL visit responses

Visit Responses	Description
Complete Response (CR)	Disappearance of all target lesions. Any pathological lymph nodes selected as target lesions must have a reduction in short axis to <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of target lesions and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Not Evaluable (NE)	All target lesion measurements are missing or had a lesion intervention at this visit, or >1/3 target lesion measurements are missing and sum of diameters of non-missing target lesions does not qualify for PD.
Not applicable (NA)	No target lesions are recorded at baseline

Rounding of TL data:

For calculation of PD and PR for TLs percentage changes from baseline and previous minimum should be rounded to 1 decimal place before assigning a target lesion response. For example 19.95% should be rounded to 20.0% but 19.94% should be rounded to 19.9%.

Missing TL data:

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

[redacted]

[redacted] × [redacted] = [redacted]

[redacted]

Lymph nodes

[redacted]

TL Visit responses subsequent to CR

A CR response can only be followed by CR, PD or NE. If a CR has occurred then the following rules at the subsequent visits must be applied:

- [redacted]

TL too big to measure:

If a target lesion becomes too big to measure this should be indicated in the database and a size ('x') above which it cannot be accurately measured should be recorded. If using a value of x in the calculation of target lesion response would not give an overall visit response of PD, then this will be flagged and reviewed by the study team blinded to treatment assignment. It is expected that a visit response of PD will remain in the vast majority of cases.

TL too small to measure:

If a target lesion becomes too small to measure a value of 5mm will be entered into the database and used in TL calculations, unless the radiologist has indicated and entered a smaller value that can be reliably measured. If a target lesion response of PD results then this will be reviewed by the study team blinded to treatment assignment.

Irradiated lesions/ lesion intervention/ curative surgery:

Previously irradiated lesions (i.e. lesion irradiated prior to entry into the study) should be recorded as NTLs and should not form part of the TL assessment.

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation / palliative surgery / embolisation/ curative surgery), should be handled in the following way and once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumours:

- [redacted]

[redacted]

Lesions that split in two:

If a TL splits in two, then the longest diameters (LDs) of the split lesions should be summed and reported as the LD for the lesion that split.

Lesions that merge:

If two target lesions merge, then the LD of the merged lesion should be recorded for one of the TL sizes and the other TL size should be recorded as 0 cm.

Change in method of assessment of target lesions:

CT and MRI are the only methods of assessment that can be used within the trial, with CT being the preferred method. If a change in method of assessment occurs between CT and MRI this will be considered acceptable and no adjustment within the programming is needed.

[redacted]

Non-Target Lesions (NTLs) and new lesions

Non-target lesion response will be derived based on the Investigator's overall assessment of NTLs as follows:

Progressive disease:	Unequivocal progression of existing NTLs, which may be due to an important progression in one lesion only or in several lesions.
Complete response:	Disappearance of all NTLs present at baseline with all lymph nodes non-pathological in size (<10mm short axis).
Non-CR/Non-PD:	Persistence of one or more NTLs with no evidence of progression.
Not evaluable:	Only relevant when one or some of the NTLs have not been assessed and in the Investigator's opinion they are not able to provide an evaluable overall NTL assessment.
Not applicable:	Only relevant if there are no NTLs at baseline.

New lesions will be identified via a Yes/No tick box. The absence and presence of new lesions at each visit should be listed alongside the TL and NTL visit responses.

A new lesion indicates progression so the overall visit response will be PD irrespective of the TL and NTL response.

If the question 'Any new lesions since baseline' has not been answered with Yes or No and the new lesion details are blank this is not evidence that no new lesions are present and should be treated as NE in the derivation of overall visit response.

3.1.2 Overall visit response

Table 4 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 4 Overall visit responses

TARGET	NON-TARGET	NEW LESIONS	OVERALL VISIT RESPONSE
CR	CR (or NA)	No (or NE)	CR
CR	Non-CR/Non-PD or NE	No (or NE)	PR
PR	Non-PD or NE	No (or NE)	PR
SD	Non-PD or NE	No (or NE)	SD
PD	Any	Any	PD

Table 4 Overall visit responses

TARGET	NON-TARGET	NEW LESIONS	OVERALL VISIT RESPONSE
Any	PD	Any	PD
Any	Any	Yes	PD
NE	Non-PD or NE	No (or NE)	NE
NA	CR	No (or NE)	CR
NA	Non-CR/Non-PD	No	SD
NA	NE	No (or NE)	NE
NA	Non-PD	NE	SD

3.2 Outcome variables

3.2.1 Primary outcome variable: Progression free survival (PFS)

PFS is defined as the time from the date of randomisation until the date of objective disease progression (as per RECIST 1.1) or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy prior to progression. Patients who have not progressed 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. [redacted]

PFS will be calculated using a programming algorithm using the data collected on objective tumour assessment.

The PFS time will always be derived based on scan/assessment dates not visit dates. RECIST assessments/scans contributing towards a particular visit may be performed on different dates. The following rules will be applied:

- Date of progression will be determined based on the **earliest** of the dates of the component that triggered the progression
- When censoring a patient for PFS the patient will be censored at the **latest** of the dates contributing to a particular overall visit assessment

3.2.2 Secondary outcome variables

Change in tumour size at 12 weeks.

Change in tumour size at 12 weeks is based on RECIST target lesion measurements taken at baseline and at week 12 for patients with measurable disease only. Tumour size is the sum of the longest diameters of the target lesions. Target lesions are measurable tumour lesions. Baseline for RECIST is defined to be the last evaluable assessment prior to starting treatment. More details on target lesions and measurements can be found in Appendix F of the protocol.

Subjects who progress before week 12 should have had a tumour assessment performed at the time of progression prior to treatment discontinuation and unblinding. The tumour size from their latest progression assessment will be used instead of the week 12 assessment for these subjects.

Patients who discontinue study treatment for reasons other than objective disease progression should have tumour assessments scans performed as scheduled in the protocol and the tumour size from the week 12 assessment will be used in this analysis.

Handling Missing TLs measures

Target lesion imputation:

[redacted]

Apply a window around the week 12 visit:

Whenever tumour size data for the week 12 visit (Note: or visit at which progression was documented if before week 12) is available then this should be used in the analysis. A windowing rule will be applied and will follow the protocol allowed visit window; therefore any RECIST scan performed within ± 1 week of the protocol scheduled visit will be used for that visit.

Assess level of missing data and change TSA analysis if appropriate:

If after target lesion imputation and applying a window around the week 12 visit there remains reasonable amount of missing tumour size measurement data, a non-parametric method will be considered for the TSA analysis, assigning patients who have died with the worst rank and ranking other imputed data. What constitutes a reasonable amount of data will be judged on a case-by-case basis but > 10% missing data would be cause for consideration.

Following imputation, the blinded data will be assessed for normality and if appropriate, a decision to use a non-parametric method as the primary TSA analysis may still be made after an assessment of the combined actual and imputed data

If, after applying the above considerations to the missing data, there is still missing tumour size measurement data (tsm data) at week 12, the recommendation would be to follow the imputation process outlined below for each individual patient where data is missing (applied prior to blind review).

[redacted]

All imputed data should be derived within the reporting dataset with a corresponding flag against the imputed value to show that this value has been programmatically derived

Best percentage change in tumour size.

The best percentage change in tumour size from baseline will be reported, i.e. the maximum reduction from baseline or the minimum increase from baseline in the absence of a reduction from baseline based on all post baseline assessments.

If following the scaling up rules, best percentage change cannot be calculated due to missing data, and a patient has no post baseline assessments, then the following imputation rules should be applied:

a) [redacted]

b) [redacted]

c) [redacted]

All imputed data should be derived within the reporting dataset with a corresponding flag against the imputed value to show that this value has been programmatically derived.

3.2.2.1 Objective Response Rate (ORR)

ORR is defined as the percentage of patients with at least one investigator-assessed visit response of CR or PR and will based on a subset of all randomised patients with measurable disease at baseline per the site investigator. The response does not need to be confirmed. Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of ORR. Patients who discontinue randomised treatment without progression, receive a subsequent anti-cancer therapy and then respond will not be included as responders in the ORR. The denominator for ORR will be the Efficacy analysis set (ITT population), excluding any patients who do not have measurable disease at baseline as per the RECIST 1.1 criteria.

Section 3.1 gives details about the derivation of objective tumour response (CR and PR).

3.2.2.2 Duration of Response (DoR)

Duration of response will be defined as the time from the date of first documented response until date of documented progression or death in the absence of disease progression, the end of response should coincide with the date of progression or death from any cause used for the PFS endpoint. If a subject does not progress following a response, then their duration of response will use the PFS censoring time.

3.2.2.3 Durable Response Rate (DRR)

Durable response rate (DRR) is defined as the proportion of patients with complete or partial response lasting continuously for at least 24 weeks (with a window of +/- 1 week) from time of onset of response.

The denominator for DRR will be the Efficacy analysis set (ITT population), excluding any patients who do not have measurable disease at baseline as per the RECIST 1.1 criteria.

3.2.2.4 Overall survival

Overall survival is defined as the time from the date of randomisation until the date of death due to any cause, regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy (i.e date of death or censoring – date of randomisation + 1). Any patient not known to have died or who are lost to follow-up or withdraw consent at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive.

If there are no survival or death data available at the data cut-off date, then the last date the patient is known to be alive will be calculated from the last assessment date of all modules

(except the visit module) on the database and patient will treated as censored patients at last recorded date.

3.2.2.5 Response rate at 12 weeks and best objective tumour response

Response rate at 12 weeks (with a window of +/- 1 week) is defined as the percentage of patients who have a week 12 visit response of CR or PR (as defined by RECIST 1.1), regardless of whether the responses are confirmed or not.

Best objective tumour response will be based on RECIST measurements taken throughout the whole study and will classify patients as complete and partial responses, stable disease, progressive disease and not evaluable. It is the best response a patient has had following first dose but prior to starting any subsequent cancer therapy and up to and including RECIST progression or the last evaluable assessment in the absence of RECIST progression.

BoR will be determined programmatically based on RECIST using all site investigator data up until the first progression event. The denominator will be consistent with that used in the ORR analysis.

For patients whose progression event is death, BoR will be calculated based upon all evaluable RECIST assessments prior to death.

3.2.2.6 Proportion of patients without progressive disease at 12 weeks

The proportion of patients without progressive disease at 12 weeks is defined as the percentage of patients with a 12 week visit response (with a window of +/- 1 week) of CR, PR or SD (as defined by RECIST 1.1) with no evidence of previous progression. If the week 12 response is missing or NE but the next evaluable response is SD or better then the patient will be defined as having NPD at 12 weeks.

The denominator for patients without progressive disease at 12 weeks will be all randomised patients.

3.3 Health Related Quality of Life (HRQoL)

The EORTC QLQ-C30 + BR-23 questionnaires will be used to identify and analyse global health status/ quality of life and disease-related symptoms according to EORTC scoring manual.

EORTC QLQ-C30 is a 30-item generic instrument designed specifically to assess quality of life (QoL). The structure of questionnaire is defined as:

5 functional scales (physical, role, cognitive, emotional, and social)

3 symptom scales (fatigue, pain, and nausea/vomiting)

A global health scale (Global health status/QoL)

5 single items assessing common physical symptoms of cancer (dyspnoea, loss of appetite, insomnia, constipation, and diarrhoea)

A single item assessing the financial effect of cancer.

The raw scores for each of the 30 items of the QLQC-30 will be converted into scores for each of the 15 domains on a scale from 0 to 100. The scoring algorithm for the QLQC-30 (from the EORTC scoring manual for QLQC-30) will be used to generate the domain scores.

In case of Functioning and Global QoL scales, higher scores indicate better function whereas for symptomatic scales higher scores indicate worse symptoms.

The BR-23 breast cancer–specific questionnaire includes 23 questions to assess function (body image, sexuality, and future perspective) and symptoms (Systemic therapy side effects, Breast symptoms, Arm symptoms, Upset by hair loss) specific to breast cancer.

Seven domain scores will be derived on a 0 to 100 scale as per the EORTC manual.

The compliance rate and evaluability rate will be calculated at baseline and each subsequent assessment for each of the HRQoL and disease related symptoms scores.

The compliance rate and evaluability rate will be calculated as –

Compliance Rate = (The number of evaluable forms/ the number of expected forms) x 100.

Evaluability Rate = (The number of evaluable forms / the number of received forms) x 100.

Only patients who have a baseline score will have their scores included in any analysis.

The pre-dose of day-1 of Cycle 1 is defined as baseline for both EORTC + BR23 questionnaires.

- HRQoL measurement prior to dosing with study treatment (AZD5363/placebo and Paclitaxel) will be considered as baseline values. For each of the HRQoL and disease-related symptoms scores, a change in score from baseline of at least 10 points can be considered as a clinically relevant or a minimally important difference. A visit response will be calculated for each subscale as detailed in Table 5.

Table 5 Visit Response

Score	Change from baseline	Visit Response
Symptom scales / items	≤-10	Improved
	>-10, <10	No change
	≥10, or missing questionnaire due to "Subject too sick, other than disease under investigation" or "Subject too affected by symptoms of disease under investigation"	Worsened
	Missing for reasons other than those above	Not evaluable
Functional scales, Global health status	≥10	Improved
	>-10, <10	No change
	≤-10, or missing questionnaire due to "Subject too sick, other than disease under investigation" or "Subject too affected by symptoms of disease under investigation"	Worsened
	Missing for reasons other than those above	Not evaluable

For each patient, a best overall QoL response will be calculated for each of the HRQoL and disease related symptoms as detailed in Table 6.

Table 6 Best QoL Response

Score	Definition	Best QoL Response
Any	Two visit responses of 'improved' a minimum of 21 days apart without an intervening visit response of 'worsened'	Improved
	Does not qualify for overall score response of "improved". Two visit responses of either "no change", or "improved" and "no change" a minimum of 21 days apart without an Intervening visit response of "worsened".	No change

Table 6 Best QoL Response

Score	Definition	Best QoL Response
	Does not qualify for overall score response of "improved" or "no change". Either (1) A visit response of "worsened" without a response of "improved" or "no change" within 21 days; or (2) Died within 12 weeks of last evaluable PRO assessment.	Worsened
	Dose not qualify for one of the above	Not evaluable

Time-to-deterioration will only be reported and analysed, should initial review of the summary outputs for the QoL data indicate that it would be meaningful.

If conducted, time to deterioration for each of the HRQoL scale and disease related symptoms score will be calculated. For each endpoint separately, patients should be considered non-evaluable if any of the following apply

- The subject does not have any evaluable baseline data
- The subject does not have any evaluable post-baseline data
- The subject's baseline score is so close to the maximum/minimum subscale score that it would be impossible for them to experience a deterioration (e.g. maximum score – baseline score < deterioration threshold)

A deterioration time is defined as the first visit response of worsened without a visit response of no change or improved within the following 21 days. If a subject has no QoL data after the first visit which deterioration has occurred, the subject will still be considered to have deteriorated. Subjects who die or are unable to complete the questionnaires will also be considered to have deteriorated if it occurs within 12 weeks of last evaluable HRQoL assessment.

Time to deterioration is calculated as the time from randomisation to date of the first assessment at which deterioration started. Subjects who do not deteriorate will be censored at the time of their last evaluable subscale score for the specific subscale being analysed. If subject does not have an evaluable post baseline score for the subscale, the subject will be censored for time to deterioration at day 1. Any subject who cannot possibly worsen compared to baseline will be censored for time to worsening at day 1.

Handling missing HRQoL data

For each score, if less than 50% of the individual items are missing, the score will be divided by the number of non-missing items and multiplied by the total number of items comprising the score. If more than 50% of the items are missing, that scale score will be treated as

missing. Missing single items are treated as missing. The reason for any missing questionnaire will be identified and recorded.

If there are cases in which more than one questionnaire has been completed on the same day, and provided different answers, the questionnaire with the worst overall score will be used.

3.4 Safety Variables

Safety and tolerability will be assessed using reports of adverse events (including SAEs), deaths, vital signs (including blood pressure [BP], and pulse), cardiac monitoring (including 12-lead ECGs and LVEF (by MUGA or cardiac ECHO) , physical examination, lipids (Triglycerides, HDL, LDL and Cholesterol), laboratory findings (including Haematology, Clinical Chemistry and Urinalysis). These parameters will be collected for all patients. Adverse events will be collected throughout the study, from time of signature of informed consent until 28 days after the last dose of AZD5363/placebo. The detailed definitions and calculation of the safety and tolerability variables are given in section 6.4 and 11.2 of CSP. Appropriate summaries of these data will be presented as described in Section 4.2.4.

3.4.1 Adverse events

All AEs will be coded using latest available version of the Medical Dictionary for Regulatory Activities (MedDRA). AE will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

3.4.2 Other Significant events (OAE)

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as a SAE and/or a Discontinuation of Investigational Product due to Adverse Event. Based on the expert's judgement, significant adverse events of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered OAEs and reported as such in the CSR. A similar review of laboratory, vital signs and ECG data will be performed for identification of OAEs.

Examples of these are marked haematological and other laboratory abnormalities, and certain events that lead to intervention (other than those already classified as serious), dose reduction or significant additional treatment.

3.4.3 Laboratory variables

Blood and urine samples for determination of haematology, clinical chemistry, and urinalysis will be collected as outlined in CSP section 6.4.5. Laboratory values will be graded according to CTCAE version 4.0. Neutrophils and glucose values with a grade >= 4 will be used to determine tolerability for the go-no go criteria.

The baseline value of each laboratory variable will be derived, as described in the Clinical Study Protocol; values captured up to 3 days before baseline visit will be considered suitable for a baseline value. If multiple records are present between -3 days and baseline visit (cycle 1 day 1) then the last value obtained prior to the first dose of study medication will be

considered as baseline. Alternatively, if two visits are equally eligible to assess patient status at baseline (e.g., screening and baseline assessments both prior to first dose with no washout or other intervention in the screening period), the average can be taken as a baseline value.

The post-baseline non-missing values of laboratory parameter closest to the scheduled visit date will be considered as the visit value. If two assessments are equidistant from a scheduled visit, then the earlier of the two will be used. The visit will be missing if no assessment was reported within the specified visit window around the scheduled date.

Designation of visits for Lab data assessment are given in table below (Table 7):

Table 7 Day Ranges for lab assessments

Cycle No.	Visit day	Target day	Day Range
	Screening	-1	-28 – -1
1	Day 1 (Baseline)	1	-3 – 1 (prior to dose)
1	Day 2	2	2 – 3
1	Day 5	5	4 – 7
1	Day 9	9	8 – 12
1	Day 16	16	13 – 19
1	Day 23	23	20 – 26
2	Day 2	30	27 – 33
2	Day 8	36	34 – 39
2	Day 15	43	40 – 46
2	Day 22	50	47 – 54
3	Day 2	58	55 – 72
4	Day 2	86	73 – 100
Y	Day 1	X	X-13 – X+14

Where Y = 5, 6 …….. and X = (Y-1)*28+1

Table 8 Day Ranges for Serum/Plasma glucose assessments

Cycle No.	Visit day	Target day	Day Range
	Screening	-1	-28 – -1
1	Day 1 (Baseline)	1	-3 – 1 (prior to dose)
1	Day 2	2	2 – 3
1	Day 5	5	4 – 7
1	Day 9	9	8 – 12

Table 8 Day Ranges for Serum/Plasma glucose assessments

Cycle No.	Visit day	Target day	Day Range
1	Day 16	16	13 – 19
1	Day 23	23	20 – 26
2	Day 2	30	27 – 44
3	Day 2	58	45 – 71
Y	Day 2	X	X-13 – X+14

Where Y = 5, 6 …….. and X = (Y-1)*28+1

3.4.4 ECG

A 12-lead ECG will be performed as per intermittent dosing described in protocol section 6.4.9. The following parameters will be recorded for each ECG: date and time of ECG, heart rate (beats/min), PR, R-R, QRS, QT (ms), QTc (ms), sinus rhythm (yes/no), and overall evaluation (normal/abnormal).

3.4.5 Calculation or derivation of safety variables

ECG Changes

QTc will be calculated using both Bazett's (QTcB) and Fridericia's (QTcF) formulae as follows:

- $$QTcB = \frac{QT}{\sqrt{\frac{RR}{1000}}}$$
- $$QTcF = \frac{QT}{\sqrt[3]{\frac{RR}{1000}}}$$

Where QT and RR is in msec.

Creatinine clearance

Creatinine clearance will be calculated using the Cockcroft and Gault formula below, from creatinine:

For creatinine values in mol/L

Men: [(140 – age) x weight (kg) x 1.23] / creatinine (µmol/L)

Women: [(140 – age) x weight (kg) x 1.04] / creatinine (µmol/L)

Corrected calcium will be calculated using the formula below, from total calcium and albumin:

Corr. Calcium = Total Calcium (mmol/L) + ([40 – Albumin (g/L)] x 0.02)

Baseline for laboratory variables will be defined as last non missing measurement prior to the first dose of medication and change from baseline will be calculated for each scheduled visit using the formula:

Change = [(Post dose measurement - baseline measurement)]

And percentage change from baseline using,

Percentage change = [(Post dose measurement - baseline measurement) / Baseline measurement] *100

3.4.6 Visit windows for other safety assessments

For physical examination, weight, ECG, vital signs, LVEF, lipids, glycosylated haemoglobin, glucose, insulin and insulin c-peptide the following conventions should apply to define visit response.

The time windows should be exhaustive so that data recorded at any time point has the potential to be summarised. 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.
- 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).
- Listings should display all values contributing to a time point for a patient.
- For visit based summaries:
    - If there is more than one value per patient within a time window then the closest value should be summarised, or the earlier in the event the values are equidistant from the nominal visit date.
- For summaries at a patient level, all values should be included, regardless of whether they appear in a corresponding visit based summary, when deriving a patient level statistic such as a maximum.

Unless otherwise stated, baseline will be defined as the last non-missing measurement prior to dosing with study treatment (AZD5363/placebo or Paclitaxel). For laboratory data and vital signs data, any assessments made on day 1 will be considered pre-dose. Where safety data are

summarised over time, study day will be calculated in relation to date of first treatment (AZD5363 or placebo).

Handling missing safety data

Missing safety data will generally 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.

3.4.7 Duration of Exposure

Total duration of exposure will be calculated as

For AZD5363 drug:

Date (last day of week of last dose of AZD5363) – Date (first dose of AZD5363) + 1

For paclitaxel:

For patients who discontinued paclitaxel therapy during the 1st or 2nd week of a dosing cycle

Date (last day of week of last dose of paclitaxel) – Date (first dose of paclitaxel) + 1

Otherwise

Date (last day of cycle of last dose of paclitaxel) – Date (first dose of paclitaxel) + 1

3.4.8 Dose Intensity

Dose intensity of AZD5363 and paclitaxel will be addressed by considering relative dose intensity (RDI) and Percent intended dose (PID), defined as follows:

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 plus the protocol-defined post-dose rest period.

PID = 100% * d1/D1

Where d1 is the actual cumulative dose delivered up to progression (or a censoring event) and D1 is the intended cumulative dose up to progression (or a censoring event). D1 is the total dose that would be delivered, if there were no modifications to dose or schedule.

RDI and PID for AZD5363 and paclitaxel will be calculated for the entire intended treatment period (censored at data cut-off).

Examples of dose intensity:

[redacted]

[redacted]

3.5 Derivation of pharmacokinetics variables:

PK concentration data will be summarised and displayed graphically in the CSR. Details of sample collections are provided in CSP section 6.6.

The pharmacokinetic analysis, i.e. derivation of pharmacokinetic variables will be performed by Quantitative Clinical Pharmacology, AstraZeneca using Population PK methods. These analyses will be described in a separate modeling analysis plan, and the results will not be reported in the CSR.

3.6 Derivation of pharmacodynamic variables

pPRAS40, total PRAS40, pAKT, pGSK3β, CTC count from platelet rich plasma samples will be reported in the Clinical Study Report (CSR). Other exploratory biomarker research will be reported separately from the CSR.

3.7 Derivation of other exploratory biomarker research variables

Results from the other exploratory biomarker and pharmacogenetic research, including matched tumour biopsy samples and investigating the relationship between AZD5363 exposure and these exploratory biomarkers, are out of scope of this SAP, and will be reported separately from the CSR.

4. ANALYSIS METHODS

4.1 General principles

The data cut off for primary analysis will be performed when recruitment to Part B has completed and 76 PFS events in the overall population, and 38 events in the *PIK3CA* mutation detected subgroup, have occurred.

The secondary outcome variable of OS will be formally analysed if there is sufficient data to warrant formal statistical analysis (ie enough OS events at the primary analysis to analyse OS).

In each of the following groups and sub-groups (ITT population):

i. overall (regardless of *PIK3CA* status)

ii. *PIK3CA* mutation detected stratum

iii. *PIK3CA* mutation not detected stratum

the hypotheses described below will be tested. The assignment of mutation status is described in more detail in section 1.2.

H_0: No treatment difference in terms of PFS between AZD5363 plus paclitaxel and AZD5363 matching placebo plus paclitaxel (i.e. HR =1)

H_1: There exists a treatment difference in term of PFS between AZD5363 plus paclitaxel and AZD5363 matching placebo plus paclitaxel (i.e. HR ≠1)

Each of i, ii and iii will be tested at the 1 sided 10% significance level and the corresponding confidence interval (2-sided 80% CI) will be provided. No adjustments for multiplicity will be performed for multiple subgroups or interim analyses. The above hypotheses are considered as the primary analyses. [redacted]

AZD5363 + paclitaxel is hereafter referred to as AZD5363, and AZD5363 matching placebo + paclitaxel is hereafter referred to as placebo. Since patients were randomised into this study based on the stratification factor *PIK3CA* mutation status (detected or not detected), all statistical analysis modelling in the overall population will include a term for this stratification covariate.

For stratification based on *PIK3CA* mutation, changes to stratum allocation after randomization will be discussed in the CSR. Analysis will be based on confirmed stratum allocation (for example if the baseline status was '*PIK3CA* not detected', but the patient's tissue/blood(ctDNA) was subsequently determined to be '*PIK3CA* detected' , analysis will proceed based on the patients last known *PIK3CA* status or that determined from the allocations recorded in the IVRS randomization procedure).

Summaries of the safety and efficacy data will be produced using standard summary statistics. For continuous variables, these statistics will include (but not be limited to) the mean, standard deviation, median, minimum and maximum. For categorical data, these statistics will consist of frequencies and associated percentages. Missing safety and efficacy data will not be included in the numerator of the percentage calculation.

Measures of location (minimum and maximum) will be reported to the same degree of precision as the raw data, other measure of location (mean, median) will be reported to one further degree of precision. Measures of spread (standard deviation, standard error) will be reported to two further degree of precision. Where confidence intervals are provided, unless otherwise stated, they will have been calculated using an assumption of normal distribution.

All safety and efficacy data will be listed.

4.2 Analyses

Table 9 summarises all formal statistical analyses planned for this study:

Table 9 Formal statistical analyses planned for the primary and secondary endpoints

Endpoint Analysed	Analyses
Progression free survival (PFS)	Primary Analysis: Cox proportional hazards model with covariates for *PIK3CA* mutation status in the overall population and repeated for subgroup of *PIK3CA* mutation detected patients and *PIK3CA* not detected patients separately. Sensitivity analysis: Evaluation-time bias assessment Attrition bias assessment

Table 9 Formal statistical analyses planned for the primary and secondary endpoints

Endpoint Analysed	Analyses
Overall Survival (OS)	Cox proportional hazards model with covariates for *PIK3CA* mutation status in the overall population and repeated for subgroup of *PIK3CA* mutation detected patients and *PIK3CA* not detected patients separately.
Tumour Size (TS)	ANCOVA with covariates for *PIK3CA* mutation status, time from randomisation to baseline tumour assessment and baseline tumour size in the overall population and repeated for subgroup of *PIK3CA* mutation detected patients and *PIK3CA* not detected patients separately.
Duration response rate (DRR)	Logistic regression with covariates for *PIK3CA* mutation status in the overall population and repeated for subgroup of *PIK3CA* mutation detected patients and *PIK3CA* not detected patients separately.
Objective response rate (ORR)	Main Analysis: Logistic regression with covariates for*PIK3CA* mutation status in the overall population and repeated for subgroup of *PIK3CA* mutation detected patients and *PIK3CA* not detected patients separately.

* Some analyses specified above will be repeated based on determination of PIK3CA status via ctDNA samples only and via tissue samples only.

4.2.1 Subgroup analyses

The following subgroups have been defined for the primary analysis:

- PIK3CA mutation status (detected/not detected in ctDNA or tissue samples)
- PIK3CA mutation (detected/not detected in ctDNA samples only)
- PIK3CA mutation (detected/not detected in tissue only)
- Age (<=65, >65)
- Region (Asian/Not Asian)
- Visceral disease (Yes/No)
- Prior taxane use in adjuvant/neoadjuvant setting (Yes/No)

Where variables are not categorical (e.g. age), cut-offs have been defined to determine subgroup classification. Statistical analysis will be performed when there are a minimum of 10 patients/events in a subgroup.

Mutation *PIK3CA* status (detected/not detected) and sample type (tissue/ctDNA) will be summarised and listed. To investigate the concordance between the results of the two sample types, an analysis will be done using kappa statistics with 95% confidence intervals.

4.2.2 Primary variable – Progression Free Survival (PFS)

Primary analysis

PFS will be summarised using Kaplan-Meier (KM) methods. Estimates of median PFS with CIs and proportion of patients progression-free at 6 months and 1 year will be presented by treatment group for the overall analysis and in the analysis of *PIK3CA* mutation detected patients and *PIK3CA* not detected group. In addition, an analysis of PFS will be conducted on the subset of patients who have *PIK3CA* mutation detected by blood (ctDNA) only.

Kaplan Meier plots of PFS will be presented for the following scenarios by treatment (AZD5363 or placebo):

a. overall population by treatment regardless of *PIK3CA* mutation status,

b. by *PIK3CA* mutation (detected/not detected) as specified in the protocol (any method) and treatment,

c. by *PIK3CA* mutation (detected/not detected in ctDNA samples only) and treatment,

d. by *PIK3CA* mutation (detected/not detected in tissue samples only) and treatment,

e. by *PIK3CA* mutation (detected only) in ctDNA and tissue samples,

f. by *PIK3CA* mutation (not detected only) in ctDNA and tissue samples.

g. by *PIK3CA* mutation (detected/not detected) by protocol definition and by overall treatment regardless of *PIK3CA* mutation status (i.e. overlay of a and b above)

The primary analysis of progression free survival for all randomised patients (ITT) will be performed by Cox proportional hazard model using SAS® PROC PHREG with the Efron method to control for ties. PFS will be analysed for the overall population and for the subgroup of *PIK3CA* mutation detected patients and *PIK3CA* mutation not detected patients, using the protocol defined definition and based on ctDNA only. PFS will be analysed using separate Cox proportional-hazards models allowing for the effect of treatment and including a term for *PIK3CA* mutation status in the overall analysis (mutation detected or not detected).

The hazard ratios in the overall population and in the *PIK3CA* mutation detected and *PIK3CA* mutation not detected subgroups (HR; AZD5363 + paclitaxel: placebo + paclitaxel) for treatment will be estimated [redacted]

A forest plot will be used to display (on a single plot) the HRs and 2 sided 80% confidence intervals for the overall population subset by *PIK3CA* mutation status and the subgroups detailed in section 4.2.1.

Proportionality assumption

The assumption of proportionality, for all time to event analyses, will be assessed firstly by examining plots of complementary log-log (event times) versus log (time) and, if this raises concerns, subsequently fitting a time dependent covariate (adding a treatment-by-time or treatment-by-ln(time) interaction term) to the model to assess the extent to which this represents random variation. If a lack of proportionality is evident then the HR can still be meaningfully interpreted as an average HR over time unless there is extensive crossing of thetime to event curves.

Sensitivity analysis

The following sensitivity analyses will be performed on PFS data using the ITT.

Evaluation-time bias assessment

For PFS, a supportive analysis will be performed to assess time assessment bias via an Evaluation-time bias approach (Sellke et al. 1983, Sun et al. 2010). This will be performed using a validated SAS program for generation of the p-value. The most pragmatic and preferred approach to assessing evaluation-time bias is to analyse, using standard log-rank , the midpoint between the time of progression and the previous visit, this approach has been shown to be robust to even highly asymmetric assessment schedules.

To support this analysis, the mean of patient-level average inter-assessment times will be tabulated for each treatment.

Attrition bias assessment

A further analysis of the primary endpoint will be conducted, whereby patients who received subsequent anti-cancer therapies prior to progression (or death in the absence of progression) will be censored at the time patients received subsequent therapies, while patients who progress or die following two or more missed visits will not be censored. The analysis will use the same approach as the primary analysis model (outlined within Section 4.2.2). A list of relevant anti-cancer therapies will be provided by the study team prior to database lock (DBL). In support of this analysis, a Kaplan-Meier plot where the roles of events and censored observations are reversed will be presented to assess the impact of rate and nature of censoring on final results.

Additional supportive summaries/graphs

A summary of the number of patients who have incomplete PFS follow-up, i.e. were alive and progression free at the date of data cut-off and had not had a RECIST assessment within 12 weeks of the data cut-off will be provided by treatment group. Additionally, summary statistics will be given for the number of days from censoring to data cut-off for all censored patients.

A summary of the median duration of follow-up will also be presented. Each patient's duration of follow-up will be defined as the number of days from randomisation to the date of last contact. Duration of follow-up will be summarised using medians:

- In censored (not progressed) patients only: Time from randomisation to date of censoring (date last known to be non-progressor) and
- In all patients: Time from randomisation to the date of progression/date of death or to the date of censoring for censored patients.

Duration of follow-up will be presented for the overall population, and the PIK3CA detected and not detected strata (protocol defined). All of the collected RECIST data will be listed for all randomised patients.

4.2.3 Secondary variables

4.2.3.1 Change in tumour size at 12 weeks

The absolute values and percentage change in target lesion tumour size from baseline will be summarised using descriptive statistics and presented at each time point and by randomised treatment group.

The number and percentage of subjects in each treatment group whose week 12 data is imputed will also be presented.

At blind review, plots of the residuals from fitting an ANCOVA model (including covariates for baseline tumour size, time from randomisation to baseline assessment and *PIK3CA* mutation status) to the percentage change from baseline at week 12 will be produced. If the residuals appear to follow a normal distribution then the same model will be used in the final analysis with an additional term for treatment. The model will be repeated in each of the PIK3CA mutation subgroup (protocol definition) analyses. In each case, the results to be presented include: the adjusted least square means for each treatment, the difference in least square means between treatments, [redacted]

[redacted] In each case, the results to be presented include: the adjusted geometric least square means for each treatment (back transformed to the original scale), the ratio of the geometric least square means, [redacted]

Should neither of the above parametric models produce normally distributed residuals, a non-parametric approach will be applied instead. The recommended method is an ANCOVA model on the ranked percentage change in tumour size, including a covariate for baseline tumour size, time from randomisation to baseline assessment, treatment and *PIK3CA* status. The ranking will be of the analysis data set following appropriate imputation. The subject with the greatest reduction in target lesion tumour size will be assigned the lowest rank, with smaller changes and increases in tumour sizes taking increasing ranks. Deaths will be assigned the highest rank. The model will be repeated in each of the *PIK3CA* mutation subgroup analyses.

If a non-parametric analysis is performed, the p-value from the ANCOVA model will be presented together with the Hodges Lehmann estimate of the median difference and [redacted] The median percentage change and range will be presented for each treatment group, together with the number of subjects and percentage of subjects in each treatment group whose 12 week data is imputed (imputation as described above or using a regression method) in the non-parametric analysis.

Tumour size is also to be presented graphically using waterfall plots for each treatment group, to present each subject's week 12 percentage change in tumour size as a separate bar, with the bars ordered from the largest increase to the largest decrease. The waterfall plot will be repeated separately by *PIK3CA* mutation status at entry (protocol definition) and treatment group. A reference line at the –30% change in tumour size level will be added to the plots, which correspond with the definitions of 'complete or partial' response. In these waterfall plots, the subjects whose week 12 percentage change in tumour size is based on an imputation (imputation as described above or using a regression method), due to missing target lesion data, will be clearly identified (by different coloured bars/symbols on the waterfall plot). A waterfall plot by treatment group of best percentage change will also be reported, repeated for mutation subgroups (*PIK3CA* detected and not detected, based on the protocol definition).

4.2.3.2 Overall Survival

The analysis population for OS will be the ITT. OS on AZD5363 in combination with weekly paclitaxel compared to weekly paclitaxel plus placebo will also be analysed using the same methods as for PFS. Estimates of median OS with CIs will be presented by treatment group

for the overall analysis and in the analysis of *PIK3CA* mutation detected patients and *PIK3CA* not detected group. In addition, an analysis of PFS will be conducted on the subset of patients who have *PIK3CA* mutation detected/not detected by blood (ctDNA) only.

Cox proportional hazard models will be conducted on the overall analysis and in the analysis of *PIK3CA* mutation detected patients and *PIK3CA* not detected group (protocol definition) only.

Analysis will occur when a sufficient number of events are available in the overall population, i.e. when 50% of patients have died. An additional analysis after 75% of patients have died may also be carried out. If fewer than 50% of patients have died at the time of the primary PFS analysis, then the OS data will only be summarised at that time. Analyses of PFS and OS at other time points may be performed as deemed appropriate.

Kaplan Meier plots of OS will be presented for the same scenarios that are outlined in points a b and g for the primary analysis above.

4.2.3.3 Objective Response Rate (ORR)

Objective response rate (ORR) as defined in section 3.2.2.1 will be reported and analyzed over the whole study and reported at week 12 in the ITT.

A tabulation of response rate at 12 weeks and proportion progression free at 12 weeks will be presented for the overall population, *PIK3CA* detected and *PIK3CA* not detected.

Best overall response will be presented by treatment group at time of PFS analysis. Frequencies and percentages will be presented for complete responses and partial response, stable disease, progressive disease and not evaluable. Summaries will be repeated based on the overall population, *PIK3CA* detected, *PIK3CA* not detected, *PIK3CA* detected (blood (ctDNA)) and *PIK3CA* not detected (blood (ctDNA)).

Objective tumour response rates will be compared between the two treatment arms for scenarios i – iii outlined in the primary analysis above using a logistic regression model allowing for the effect of treatment and adjusted for *PIK3CA* mutation status in the overall population. The analysis will also be repeated based on PIK3CA mutation detected/not detected by blood (ctDNA) only.

The objective tumour response rate will be estimated for each arm. If there is an imbalance in the covariates, adjusted response rates will also be presented. These are the counterparts of lsmeans from linear models and require conditioning on one arm's observed response rate. An estimate statement would be used that weights log-odds in line with the prevalence of the covariates.

4.2.3.4 Duration of Response (DoR)

Duration of response (DoR) is defined in section 3.2.2.2 will be presented by using Kaplan Meier plots and KM estimates. Kaplan Meier plots of DoR will be presented for scenarios a, b and g detailed in the primary analysis description above. DoR in the responding patients will be presented as individual line plots for each patient, with lines indicating the DoR and symbols indicating whether the response had ended (progressed) or censored at the data cut off.

Supporting summary tables will be presented on the overall population and on the analysis of *PIK3CA* mutation detected patients and *PIK3CA* not detected group (protocol definition).

4.2.3.5 Durable Response Rate (DRR)

Durable response rate (DRR) as defined in section 3.2.2.3 will be reported and analyzed in the ITT population set. Frequencies and percentages will be presented for patient with durable response (complete or partial response lasting continuously for at least 24 weeks).

Durable response rate (DRR) will be compared between the two treatment arms in overall population, *PIK3CA* mutation detected, *PIK3CA* mutation not detected, detected in blood (ctDNA) only and not detected in blood (ctDNA) only subgroups using a logistic regression model allowing for the effect of treatment and adjusting for *PIK3CA* mutation status in the overall analysis [redacted]

4.2.3.6 Health related quality of life (HR QoL) data

Heath related quality of life (HR QoL) is defined in section 3.3. Tables of summary statistics will be provided for the 15 domains of the QLQC-30 and any further domains of the BR-23. Responses from individual items will be listed only. Where sufficient data are available then summary statistics (frequencies and percentages) will be presented by treatment group for the following:

- Summary of compliance of subjects to the PRO instruments.
- Summary of subjects best QoL response
- Summary statistics (n, mean, median, SD, minimum, maximum) for domain scores over time
- Summary statistics (n, mean median, SD, minimum, maximum, q1, q3) for change from baseline in domain scores and categories of change from baseline

Time-to-deterioration will only be reported and analysed, should initial review of the summary outputs for the QoL data indicate that it would be meaningful. In case that sufficient data is collected, and further analysis is required a log-rank test would be implemented to estimate the hazard ratios and corresponding confidence intervals of the effect of treatment.

PROC LIFETEST would be used to estimate these measures. Breslow method of handling of ties would be used. For further exploration, a repeated-measures ANOVA method may be implemented taking treatment effect as a factor and each of the PRO scores as the continuous response thus estimating the effect of treatment on the changes in score.

4.2.4 Safety

Safety data will be listed and summarised only. No formal statistical analyses will be performed on the safety data. All safety data will be summarised by initial treatment group (AZD5363 or placebo) including patients who have dose modification. However, some listings such as AEs listings will display the actual dose the patient received at onset of an AE.

A reduced set of safety outputs will be replicated for patients having *PIK3CA* mutation detected and mutation not detected status (using the protocol definition).

4.2.4.1 Adverse Events

All AEs occurred will be summarised (percentages and frequencies) by actual treatment group (AZD5363 and Placebo) using MedDRA system organ class, preferred term and maximum CTCAE grade (version 4.0).

Any AE occurring before the first dose of IP (i.e., before Cycle 1 Week 1 Day 2) will be included in the data listings but will not be included in the summary tables of adverse events.

Any AE occurring within the defined 28 day follow-up period after discontinuation of study therapy (AZD5363/placebo and paclitaxel) will be included in the AE listings and summaries. AEs occurring after the 28-day follow-up period after discontinuation of IP will be listed, but not included in the summaries. AE occurred before first dose of IP or after 28 days follow-up period will be flagged in AE listings.

All reported AEs will be listed along with the date of onset, date of resolution (if AE is resolved), investigator's assessment of severity and relationship to study drug. Frequencies and percentages of patients reporting each preferred term will be presented (i.e. multiple events per patient will not be accounted for apart from on the episode level summaries).

The following tables of summaries will also be produced:

- Summary information (the number and percent of patients by treatment) will be tabulated for all AEs and all AEs (by episode)
- All AEs
- All AEs with causality related to study medication (AZD5363/Paclitaxel/both)
- All serious adverse events (SAEs)
- All SAEs with causality related to study medication (AZD5363/Paclitaxel/both)

- AE leading to study medication (AZD5363/Paclitaxel/both) interrupted
- AEs leading to discontinuation of study medication (AZD5363/Paclitaxel/both)
- AEs leading to discontinuation of AZD5363/Paclitaxel/both, causally related to AZD5363/Paclitaxel/both
- AEs leading to dose reduction of study medication (AZD5363/Paclitaxel)
- AEs with an outcome of death
- AEs with an outcome of death, causally related to study medication (AZD5363/Paclitaxel/both)
- AEs with CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, causally related to study medication (AZD5363/Paclitaxel/both)
- Other significant AEs
- Other significant AEs causally related to study medication (AZD5363/Paclitaxel/both)

An overall summary of the number and percentage of patients in each category will be presented, as will an overall summary of the number of episodes in each category. With the exception of Other Significant AEs, all the above categories will be summarised for each initial treatment group by system organ class and preferred term. In addition, a truncated adverse event table of most common AEs, showing all events that occurs in at least 5% of patients overall will be summarised by preferred term, by decreasing frequency.

AEs will be assigned CTCAE grades (National cancer institute (NCI) CTCAE version 4.0) and summaries of the number and percentage of patients will be provided by maximum CTCAE grade, preferred term and initial treatment group. Fluctuations observed in CTCAE grades during study will be listed.

In addition, serious adverse events, other significant adverse events, adverse events leading to discontinuation of treatment, adverse events leading to dose interruption of AZD5363/placebo and paclitaxel adverse events with causality AZD5363/placebo and adverse events of special interest will be flagged in the relevant listing. Key patient information for serious adverse events and adverse events leading to death will be listed.

A summary of deaths (categorised as related to disease under investigation, AE outcome=death or both, AE with outcome=death >=28 days after last treatment and Patient with unknown reason for death) will be provided by number and percentage of patients by initial treatment group and a corresponding listing will also be produced. In this table, patients

who died more than 28 days following their last dose of AZD5363/Placebo and did not die from disease progression, it is assumed they died from an AE.

Burden of adverse event of diarrhoea

The burden of AE of diarrhoea will be determined for each patient in the safety population and summarised. [redacted]

Grouped adverse events of special interest

Preferred terms used to identify adverse events of special interest will be listed and documented in the Study Master File. Groupings of AE will be based on preferred terms provided by the medical team prior to unblinding.

Grouped summary tables of certain MedDRA preferred terms will be produced. For each 'grouped' term, the number (%) of patients experiencing any of the specified terms will be presented by maximum CTCAE grade. Time to onset of first AE for each grouped term and preferred term within it will also be produced. A listing of the preferred terms in each grouping will be provided. Changes observed in CTCAE grade during study will be listed only.

Additional summaries of the above-mentioned grouped AE categories will include number (%) of patients who have:

- At least one adverse event of special interest presented by outcome
- At least one adverse event of special interest by CTCAE grade
- At least one adverse event of special interest causally related to study medication (AZD5363/placebo, paclitaxel, AZD5363/placebo and paclitaxel)
- At least one adverse event of special interest leading to discontinuation of AZD5363/placebo.

Go/no-go Criteria

[redacted]

4.2.4.2 Laboratory parameters

All laboratory data (absolute value), change from baseline and percentage change from baseline values will be listed and summarised using descriptive statistics at each scheduled assessment time by treatment groups.

For all laboratory variables, which are included in the CTCAE version 4.0, the CTCAE grade will be calculated and summarised using frequency counts and percentages in the form of shifts from baseline to maximum grade post baseline.

For urinalysis, shift table comparing baseline to maximum value by treatment will be presented (i.e. using number of patients with results of negative, trace or positive).

Clinically significant laboratory values will be listed. Laboratory parameters will be summarised and listed by treatment groups.

Frequencies and percentages will be presented for:

- Elevated Alanine aminotransferase (ALT), Aspartate aminotransferase (AST) by maximum total bilirubin during the study.

  − ALT <= 1x , > 1x – 3x, > 3x – 5x, > 5x – 10x, > 10x – 20x and > 20x ULN during the study

  − AST <= 1x, > 1x – 3x, > 3x – 5x, > 5x – 10x, > 10x – 20x and > 20x ULN during the study

  − Total bilirubin <= 1x, > 1x - 1.5x, > 1.5x – 2x, > 2x ULN during the study

- Narratives will be provided for patients who have ALT >3x ULN plus Total Bilirubin >2x ULN or AST >3x ULN plus Total Bilirubin >2x ULN at any visit.

Individual patient data where ALT or AST plus bilirubin are elevated at any time will be listed.

Box-plots of absolute values and change from baseline in all hematological and clinical chemistry variables will be presented for scenarios i – iii outlined in the primary analysis above.

For glucose, screening values (fasted only) will be listed. All other glucose assessments will be reported as separately regardless of the fasting/non-fasting status assigned to them. In the clinical chemistry shift tables, CTCAE grading will be used on all glucose assessments regardless of fasting status. It is acknowledged that while CTCAE grading is intended for fasted samples, the use of this grading on all assessments is a conservative approach which allows comparison of the placebo and active treatment groups.

4.2.4.3 Vital Signs

Vital signs (SBP, DBP and pulse) will be listed and summarised using descriptive statistics over time in terms of absolute values and changes from baseline at each scheduled time point by treatment group using descriptive statistics. Box-plots for absolute values and change from baseline for SBP, DBP and pulse will be presented graphically for scenarios i – iii outlined in the primary analysis above.

4.2.4.4 ECG parameters and echocardiogram data

Overall evaluation of ECG will be summarised by cycle using frequency counts and percentage of patients for each treatment group as normal or abnormal, and the relevance of the abnormality will be summarised by “clinically significant” or “not clinically significant”.

In addition, a shift table comparing baseline (normal, abnormal - not clinically significant, abnormal - clinically significant) to last observation on treatment will also be presented.

Continuous ECG measurements will be summarised over time in terms of absolute values and change from baseline at each scheduled measurement by dose group.

The baseline value for each parameter is defined as the last available value before the first dose of study treatment in cycle 1.

The number and percentage of patients with increases of QTcB and QTcF of >30msec and >60msec from baseline, and QTcB and QTcF values of >450msec, >480msec and >500msec will be produced.

Left ventricular ejection fraction (LVEF) will be summarised over time in terms of absolute values and changes from baseline at each scheduled measurement by treatment group.

Box and whisker plots of absolute values and change from baseline for all continuous ECG variables and LVEF will be presented for scenarios i – iii outlined in the primary analysis above. A listing of echocardiography will be provided by treatment group.

4.2.5 Treatment Exposure

The following summaries related to study treatment and dose reductions /interruptions/ discontinuation will be generated:

- Total exposure of AZD5363/placebo and paclitaxel
- Actual exposure of AZD5363/placebo and paclitaxel
- Duration of therapy at starting dose of AZD5363/placebo and paclitaxel
- Total number and duration of dose interruptions and modifications of AZD5363/placebo
- Total number and duration of dose interruptions, delays and reductions of paclitaxel
- Time to dose discontinuation of AZD5363/placebo and paclitaxel
- Time to first dose interruptions of AZD5363/placebo and paclitaxel
- Time to first dose reduction of AZD5363/placebo and paclitaxel
- Time to first dose delay of paclitaxel
- Reasons for dose reduction, dose interruptions, and dose modifications of AZD5363/placebo

- Reasons for dose reduction, dose interruptions, and dose delays of paclitaxel
- Number (%) of patients summarised by total number cycles received of AZD5363/Placebo and paclitaxel.
- PID and RDI of AZD5363/placebo (entire intended treatment period and according to defined treatment periods in weeks)
- PID and RDI of paclitaxel (entire intended treatment period and according to defined treatment periods in weeks)

For patients on study treatment at the time of the analysis, the data cut off (DCO) date will be used to calculate exposure.

All treatment information data will be listed by treatment group.

4.2.6 Demography

Information collected at baseline will be summarised by randomised treatment group. A subset of demography outputs will be repeated for subjects who have *PIK3CA* mutation detected/not detected status.

4.2.7 Pharmacokinetics

In the calculation of plasma concentration summary statistics, values below the lower limit of quantification (LLOQ) will be handled according to the following rules:

- If, at a given time point, 50% or less of the plasma concentrations are non-quantifiable (NQ), the geometric mean (gmean), coefficient of variation (CV), gmean ± standard deviation (SD), arithmetic mean, SD and median will be calculated by substituting the LLOQ for values which are NQ. The minimum at that time point will be reported as NQ.
- If more than 50%, but not all, of the concentrations are NQ, the gmean, CV, gmean ± SD, arithmetic mean and SD will be reported as not calculable (NC). The minimum and median at that time point will be reported as NQ.
- If all the concentrations are NQ, the gmean, arithmetic mean, median, minimum and maximum will be reported as NQ and the CV, gmean ± SD and SD as NC.
- If the calculation of the gmean - SD results in a value less than the LLOQ, NQ will be displayed.

The plasma concentrations determined using sparse PK sampling scheme, at sampling time points, will be summarised and listed. Box plots will be produced for the PK concentration data on the original raw scale. The geometric mean may also be displayed graphically on the linear and log linear scale.

4.2.8 Exploratory outcome variables

The exploratory analyses discussed in the following sections will be included in the CSR, unless otherwise stated.

WHO performance status

WHO performance status will be listed only in the CSR.

Biomarkers Analysis

All biomarker analysis will be performed on the ITT analysis set patients for whom data is available. The following biomarkers will be reported in the CSR for all measured timepoints.

1. PRP data:

- phospo-PRAS40
- Total PRAS40
- pAKT
- pGSK3β

For each biomarker there will be two baseline results defined; one before the first dose of paclitaxel and one before the first dose of AZD5363. The absolute value and percentage change from baselines will be summarised at each time point, by dose and schedule. Standard summary statistics n, mean, standard deviation, median, minimum and maximum will be presented.

Box plots of percentage change from baseline (pre-paclitaxel) of each biomarker will be presented by treatment and timepoint. Line plots of mean absolute values over time split by treatment may be produced as well as line plots of individual patient's absolute values over time split by treatment.

2. CTC count

The absolute values will be summarised at each time point, by treatment. Standard summary statistics n, mean, standard deviation, median, minimum and maximum will be presented. The number of patients with < 5 vs ≥ 5 CTCs per 7.5 mL of whole blood will be presented.

Line plots of individual patient's absolute values will also be plotted by timepoint split by treatment.

If there are sufficient data, the association between CTC count at baseline (< 5 vs ≥ 5 CTCs per 7.5 mL) with outcome (PFS) by treatment may be explored. For the patients with increased baseline CTCs (≥ 5 CTCs per 7.5 mL), the association between CTC decrease to <5

per 7.5ml vs not at Cycle 2 Week 1 Day 1 and Cycle 3 Week 1 Day 1 with outcome (PFS) by treatment may be explored.

5. INTERIM ANALYSES

The first interim analysis will be done when the recruitment of 40 patients have been completed and followed up for 12 weeks (or progressed or died prior to 12 weeks). The interim analysis will be based on efficacy variables i.e. change in tumour size at 12 weeks. The description of statistical analysis at the interim will be same as section 4.2.

Additionally, the following variables will be summarised (frequency and percentage) by treatment groups for overall population.

- Response rate at week 12
- Best Objective response
- Percentage of patients without progression

The above stated analysis will be performed for overall population, *PIK3CA* mutation-positive patients and non-detected patient population separately.

The second interim analysis will be conducted when at least 38 PFS events across both *PIK3CA* mutation subgroups have been achieved and 30 *PIK3CA* mutation detected patients have completed 12 week follow up period (or progressed prior to 12 weeks).

The 2nd interim analysis will be conducted upon the following efficacy endpoints for the overall population, *PIK3CA* mutation detected and *PIK3CA* mutation not detected subgroups :

- Change in tumour size at week 12
- Progression-free survival
- Duration of response

Additionally the following variables will be summarised:

- Progression status
- Progression free survival
- Best objective response

More details on the interim analysis can be found in the interim SAP.

6. CHANGES OF ANALYSIS FROM PROTOCOL

The following changes have been made from the analysis outlined in the protocol:

1. The analysis of HR QoL C30/BR23 has been restricted to compliance, best QoL, summary of domain scores, change from baseline, and change from baseline categorical status. The rationale for this change is to enable the exploration of the data in the relevant populations of interest. From this initial exploration, further outputs outlined in the protocol may be produced should they be deemed necessary.

7. REFERENCES

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

Sun et al. 2010
Sun X, Chen C. Comparison of Finkelstein's Method With the Conventional Approach for Interval-Censored Data Analysis. Stat Biopharm Research 2010. 2:97-108.

8. APPENDIX (NOT APPLICABLE)

4.2.7 Biomarker analysis

All biomarker analysis will be performed on the Pharmacodynamic analysis set patients for whom data is available. The following biomarkers will be reported in CSR. Other exploratory biomarkers will be reported separately and will not be part of the CSR.

- phospo-PRAS40
- Total PRAS40
- pAKT
- pGSK3β

For each biomarker, the absolute value and the change from baseline will be summarised at each time point, by dose and schedule.

Standard summary statistics n, mean, standard deviation, median, minimum and maximum will be presented.

Box-plots and line plots will be produced for each biomarker by dose cohort, schedule and time. Line plots over time will be produced for each biomarker, for all patients along with an extra line reflecting median biomarker levels, for each cohort separately.

5. CHANGES OF ANALYSIS FROM PROTOCOL

Not applicable.

6. REFERENCES

Beal and Sheiner 1988-1998
Beal, SL, and Sheiner, LB (1988-1998). NONMEM Users Guides. Part I-VIII.NONMEM Project Group C255, University of California at San Francisco.