SPONSOR:

GTx Inc

PROTOCOL NUMBER:

G200802

STATISTICAL ANALYSIS PLAN

Author:	Sarah Lilley
Version:	1.0
Date:	05 JAN 2018

1 Cover and signature pages

Sponsor:	GTx Inc
Protocol Number:	G200802
Study Title:	A Phase 2 Open Label, Multi-Center, Multinational, Randomized, Parallel Design Study Investigating The Efficacy and Safety Of GTx-024 On Metastatic or Locally Advanced ER+/AR+ Breast Cancer (BC) in Postmenopausal Women
Document Version No	FINAL 1.0

We, the undersigned, confirm that we have read, understood and agree to the content of this document and hereby authorise its approval.

Name **Title, Organisation**	**Signature**	**Date**
Sarah Lilley Principal Biostatistician Cmed CRS Ltd.		
Jean-Luc Beffy Principal Biostatistician Cmed CRS Ltd.		
Mayzie Johnston Vice President, Clinical Development GTx. Inc.		

Table of Contents

2 List of Abbreviations and Definition of Terms

AE	Adverse Event/Adverse Experience
ALT	Alanine Transaminase
AR	Androgen Receptor
AR+	Androgen Receptor Positive
BC	Breast Cancer
BMI	Body Mass Index
BOR	Best Overall Response
CB	Clinical Benefit
CBR	Clinical Benefit Rate
CR	Complete Response
CRA	Clinical Research Associate
CRO	Contract Research Organization
CT	Computerized Tomography
CTCs	Circulating Tumor Cells
EAS	Evaluable Subjects Analysis Set
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EOT	End Of Treatment
EQ-5D-5L	EQ-5D 5-level version
ER	Estrogen Receptor
ER+/AR+	Estrogen Receptor Positive and Androgen Receptor Positive
EU	European Union
FAS	Full Analysis Set
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
ICF	Informed Consent Form
ICH	International Conference on Harmonization
IHC	Immunohistochemistry

INN	International Nonproprietary Names for Pharmaceutical Substances
INR	International Normalized Ratio
LFT	Liver Function Test
MedDRA®	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
NCI-CTCAE	National Cancer Institute-Common Terminology Criteria for Adverse Events
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PFS	Progression Free Survival
PI	Principal Investigator
PO	Per Os (oral)
PPS	Per Protocol Set
PR	Partial Response
PSA	Prostate Specific Antigen
QoL	Quality of Life
RBC	Red Blood Cells
RECIST	Response Evaluation Criteria In Solid Tumors
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Safety Analysis Set
SD	Stable Disease
SMC	Safety Monitoring Committee
TFL	Tables Figures Listings
TTP	Time-to-Progression
ULN	Upper Limit of the Normal Range
US	United States
VAS	Visual analog scale
WBC	White Blood Cells

3 Introduction

This is a Phase 2, open label, multi-center, multinational, randomized, parallel design study investigating the efficacy and safety of GTx-024 on metastatic or locally advanced ER+/AR+ breast cancer (BC) in postmenopausal women. Subjects will be randomized to receive either GTx-024 9 mg or 18 mg given orally (PO) daily for up to 24 months. Simon's two-stage (optimal) design[1] will be used to assess primary efficacy of each dose arm and will require up to 88 evaluable subjects; i.e., subjects with centrally confirmed AR+ who receive at least one dose of study drug.

The preparation of this statistical analysis plan (SAP) has been based on International Conference on Harmonisation (ICH) E3 and E9 Guidelines and in reference to protocol G200802 Version 3.0, Protocol Amendment 2, 25 July 2016.

4 Study Objectives

4.1 PRIMARY EFFICACY OBJECTIVE

The **primary efficacy objective** of this trial is to estimate the clinical benefit rate (CBR) at 24 weeks (defined as complete response [CR], partial response [PR], or stable disease [SD]) (by RECIST 1.1) of GTx-024 9 mg and of GTx-024 18 mg given PO daily in subjects with estrogen receptor positive and androgen receptor positive (ER+/AR+) BC who have centrally confirmed AR+ status.

4.2 SECONDARY EFFICACY OBJECTIVES

Secondary efficacy objectives:

- Estimate the CBR at 24 weeks (by RECIST 1.1) of GTx-024 9 mg and 18 mg in all subjects randomized who receive at least one dose of study medication (the full analysis set [FAS]) regardless of AR status as determined by the central laboratory

The following secondary efficacy objectives apply to both centrally confirmed AR+ subjects (the evaluable subset of the FAS) as well as to all subjects in the FAS:

- Estimate the objective response rate (ORR; defined as CR or PR) (by RECIST 1.1) of GTx-024 9 mg and 18 mg at 24 weeks
- Estimate the best overall response rate (BOR) of GTx-024 9 mg and 18 mg
- Estimate the progression free survival (PFS) of subjects receiving GTx-024 9 mg and 18 mg

- Estimate the time to progression (TTP) of subjects receiving GTx-024 9 mg and 18 mg
- Estimate duration of response (time from documentation of tumor response to disease progression or death) of subjects receiving GTx-024 9 mg and 18 mg
- Estimate overall survival (OS; time from randomization until death or last follow up date with a maximum of 12 months post last dose) of subjects receiving GTx-024 9 mg and 18 mg

4.3 TERTIARY OBJECTIVES

The following tertiary efficacy objectives apply to both centrally confirmed AR+ subjects (the evaluable subset of the FAS) as well as to all subjects in the FAS:

- Assess the effect of GTx-024 9 mg and 18 mg on serum prostate specific antigen (PSA)
- Assess the effect of GTx-024 9 mg and 18 mg on Quality of Life (QoL) as measured by EQ-5D-5L [3]
- Assess the effect of GTx-024 9 mg and 18 mg on circulating tumor cells (CTCs)
- Assess the impact of ER mutations at baseline on outcome
- Assess the impact of ER mutations at Week 24 and/or EOT on outcome
- Assess the impact of duration of prior CB on outcome
- Assess the impact of time from diagnosis of metastases to randomization on outcome
- Describe the effect of GTx-024 9 mg and 18 mg on tumor volumetrics
- Assess the effect of plasma concentrations of GTx-024 and GTx-024 glucuronide on CBR at 24 weeks

In addition to the tertiary efficacy objectives mentioned in the protocol, the following objectives will be considered and applied to the same populations:

- Assess the effect of PSA levels at baseline on outcome
- Assess the effect of change in PSA levels on outcome

4.4 SAFETY OBJECTIVE

To describe the safety profile of GTx-024 9 mg and 18 mg PO daily in subjects with ER+/AR+ BC with centrally confirmed AR+ as well as in all subjects randomized and treated.

4.5 PHARMACOKINETIC OBJECTIVE

To describe the plasma concentrations of GTx-024 and GTx-024 glucuronide at each of the assessed time points.

5 Study Design

5.1 STUDY DESIGN AND POPULATION

This is a Phase 2, open label, multi-center, multinational, randomized, parallel design study investigating the efficacy and safety of GTx-024 on metastatic or locally advanced ER+/AR+ breast cancer (BC) in postmenopausal women. Subjects will be randomized to receive either GTx-024 9 mg or 18 mg given orally (PO) daily for up to 24 months. Simon's two-stage (optimal) design[1] will be used to assess primary efficacy of each dose arm and will require up to 88 evaluable subjects; i.e., subjects with centrally confirmed AR+ who receive at least one dose of study drug. A total of 44 evaluable subjects is planned for each dose arm. The goal of the design is to determine whether each of the two dose arms has sufficient biological activity against breast cancer to warrant further clinical development.

In order to obtain this number of evaluable subjects, thirty-six to one hundred and eighteen (36–118) subjects, including replacement subjects (thirty of the aforementioned subjects may be considered replacement subjects to account for lack of centrally confirmed AR+ status or for the rare subject who is randomized but does not receive study drug (assumes 25% of enrolled subjects are not evaluable for the primary efficacy analysis)), will be randomized in a 1:1 fashion to receive a daily PO dose of either GTx-024 9 mg or 18 mg. Thirty-six to 88 (36–88) subjects with centrally confirmed AR+ who receive at least one dose of study drug (evaluable subjects) will be needed for primary efficacy analysis purposes and will be a subset of the FAS. The trial will test for an unacceptably low clinical benefit rate (CBR) of ≤ 10% versus a CBR more consistent with ≥ 30%. The first stage in each study arm will be assessed among the first 18 evaluable subjects. If at least 3/18 subjects achieve CB defined as complete response [CR], partial response [PR], or stable disease [SD], per Response Evaluation Criteria in Solid Tumors [RECIST], Version 1.1[2]) at week 24, the arm will proceed to the second stage of recruitment up to a total of 44 evaluable subjects per arm. Otherwise, the arm will be discontinued for lack of efficacy.

Subjects who are not confirmed AR+ may remain on the trial, but will not be part of the primary efficacy analysis – these subjects will contribute to secondary and tertiary analyses.

It is planned to enroll these subjects primarily in approximately 45 sites in the United States, Europe and Australia.

With the exception of elevated liver function tests (LFTs) and hypercalcemia, both of which are further addressed below, subjects on the 18 mg treatment arm who experience an AE with Grade ≥ 3 intensity (National Cancer Institute-Common Terminology Criteria for Adverse Events [NCI-CTCAE], Version 4.0[4]) and/or intolerance may have a dose reduction from 18 mg to 9 mg per day or a drug interruption based on the medical judgment of the Investigator and after confirmation by the study Medical Monitor. The drug interruption may last for a period of up to 7 days after which the subject must be rechallenged with study drug (18 mg or 9 mg) or discontinued from the study. In the case of a dose reduction, once the AE has resolved or reduced in intensity to Grade 1, the subject may be rechallenged with 18 mg or maintained at 9 mg at the discretion of the Investigator.

Subjects on the 9 mg treatment arm who experience an AE with Grade ≥ 3 intensity (NCI-CTCAE 4.0[4]) and/or intolerance may have a drug interruption based on the medical judgment of the Investigator and after confirmation by the study Medical Monitor. The drug interruption may last for a period of up to 7 days after which the subject must be rechallenged with study drug (9 mg) or discontinued from the study.

For ≥ Grade 3 elevations in LFTs at the 18mg dose, study drug should be withheld. A recheck of LFTs should be done within 48 hours. Study drug may be restarted at the 18 mg dose or reduced to 9 mg dose when LFTs have resolved to Grade 1 or better.

For ≥ Grade 3 elevations in LFTs at the 9mg dose, study drug should be withheld. A recheck of LFTs should be done within 48 hours. Study drug may be restarted at the 9 mg dose when LFTs have resolved to Grade 1 or better.

In the event of ≥ Grade 2 hypercalcemia, study drug should be withheld until hypercalcemia resolves to Grade 1 or better.

The subjects who demonstrate CB will be treated for up to 24 months from the date of randomization (as long as they continue to demonstrate CB from the treatment during these 24 months). Subjects who continue to demonstrate a CB from the study treatment at 24 months

will be offered to continue in a safety extension study under a separate protocol.

All subjects will be followed-up at one month after the last dose of GTx-024 is received, for safety purposes, and for vital (survival) status every 60 days, thereafter, for up to 12 additional months.

The implementation of Simon's two-Stage design could cause a pause in enrollment of the trial prior to proceeding to the second stage. For example, if the required 18 evaluable subjects in the first stage of an arm are enrolled and 15 have either terminated the trial or been assessed as not achieving CBR at 24 weeks, it will be necessary to pause enrollment in that arm until the remaining three evaluable subjects are fully assessed for CB, i.e., either withdraw prior to week 24 or have the week 24 RECIST 1.1 assessment.

In order to protect the safety of the subjects, a Safety Monitoring Committee (SMC) will be established for the study to review the safety data on an ongoing basis. See the SMC charter for further details.

5.2 STUDY TREATMENTS AND ASSESSMENTS

As not all assessments are performed at every visit, please see APPENDIX A: SCHEDULE OF EVENTS of the protocol for an outline of the timing of the study procedures/evaluations.

Screening assessments must occur within 28 days prior to randomization for determination of subject's overall eligibility. Informed consent will be obtained from the subjects before performing any other study procedure or test. AEs will be collected from the signing of the informed consent form (ICF). Other assessments performed at screening (Visit 1) include: eligibility criteria check, medical history (current and past), demographics, diagnosis and extent of cancer disease history (including prior anticancer hormonal treatment), prior/concomitant medications including previous and current anticancer treatment, physical examination (including height, weight and vital signs), ECG, ECOG performance status, laboratory data, radiological evaluation of tumor (as applicable CT/MRI/bone scan), hormone receptor status review (ER, PR, HER-2), eye examination, assessment of AR status (confirmed at a central laboratory), screening for HIV, HBV, HCV, coagulation status, PSA, serum pregnancy test and urine analysis .

Subjects who do not meet the inclusion/exclusion criteria at screening will report all screening

assessment data collected.

Baseline/Randomization assessments

The assessments performed, or results checked and confirmed, at the baseline/randomization visit (Visit 2) include: eligibility criteria check (prior to randomization), review of prior/concomitant medications including previous and current anticancer treatment, serum lipid profile, serum hormones, CTC enumeration, CTC gene expression, ctDNA analysis of ER mutations, pharmacokinetic sampling (prior to first dose), review of adverse events, randomization, administration of the first dose, and quality of life questionnaire EQ-5D-5L (before initiation of treatment). The following medical procedures and assessments will only be repeated if not done at screening within 7 days before the first dose of study treatment: vital signs, ECOG performance status, laboratory data and PSA. The urine analysis will be performed only if clinically indicated.

GTx-024 dispensing and administration of first dose of study drug will take place once baseline assessments have been completed.

Treatment phase

A safety check involving a comprehensive metabolic panel (Chem 14), including serum calcium and liver function tests, will be performed at Weeks 1, 3, and 5 (±3 days). The Week 3 safety check should include a clinic visit; at Weeks 1 and 5 blood draws only are required. The investigator may repeat these checks at Week 2 (±3 days), and Week 4 (±3 days) if clinically indicated. During the treatment phase at Visits 3, 4, 5, 6, 8, 9, 10, 11, 12 and 13 (weeks 6, 12, 18, 24, 36, 48, 60, 72, 84 and 96) the following medical procedures and assessments will be performed: review of concomitant medications, physical examination (including weight and vital signs), ECOG performance status, laboratory data, serum lipid profile, coagulation status, serum hormones, PSA, collection of adverse events, GTx-024 dispensing, GTx-024 accountability and quality of life questionnaire EQ-5D-5L. The urine analysis will be performed only if clinically indicated. At any time point during the treatment period, if clinically indicated an eye exam will be performed based on medical judgment. Pharmacokinetic sampling will be performed at visits 3, 5 and 6 only.

Radiological evaluation of tumor (as applicable CT/MRI/bone scan) will be performed at visits 4,

6, 7, 8, 9, 10, 11, 12, 13 and/or EOT and CTC enumeration will be performed at visits 4, 6, 8, 9, 10, 11, 12 and/or EOT. ctDNA analysis of ER mutations will be performed at Visit 6 and/or EOT. Should the subject have CR or PR at week 24 (visit 6), then a confirmation scan will be performed at week 28 (visit 7).

Dispensing of GTx-024 will be performed at all study visits (except Visit 7) during the treatment phase (every 6 weeks at Visits 3, 4, and 5). At Visits 6, 8, 9, 10, 11, 12, and 13, in order to accommodate the visit schedule of every 12 weeks (± 7 days), the subjects will receive two carton boxes of study drug (each containing 7 blisters) to cover study treatment for 14 weeks.

Not all assessments are performed at every visit within the treatment phase, see APPENDIX A: SCHEDULE OF EVENTS of the protocol for details.

End of treatment

End of treatment assessments must occur within 3 days after the last dose of study treatment. Assessments performed at end of treatment include: review of concomitant medications, physical examination (including weight and vital signs), ECOG performance status, laboratory data, serum lipid profile, coagulation status, serum hormones, PSA, CTC enumeration, CTC gene expression, ctDNA analysis of ER mutations, radiological evaluation of tumor (as applicable CT/MRI/bone scan), collection of adverse events, GTx-024 accountability, quality of life questionnaire EQ-5D-5L and eye examination. The urine analysis will be performed only if clinically indicated.

The assessments performed at end of treatment apply whether the subject completed 24 months of treatment or terminated early.

Post-treatment safety follow-up

Subjects will be followed up for 1 month after the last dose of GTx-024. Assessments performed during the follow-up period include: review of concomitant medications, physical examination (including weight and vital signs), ECOG performance status, laboratory data, PSA and collection of adverse events. The urine analysis will be performed only if clinically indicated.

Long term follow-up

Vital (survival) status follow-up assessments will be performed every 60 days after last dose for up to 12 months after last dose. For subjects who die during this follow-up period, the date of death will be collected.

5.3 RANDOMIZATION AND BLINDING

Eligible subjects will be randomized in a 1:1 ratio to GTx-024 9 mg daily or GTx-024 18 mg daily. If, for any reason, the 18 mg arm is terminated, randomization will end and all future enrollees will be enrolled on the 9 mg arm. Randomization will be stratified by the following factors:

- Bone only metastases (yes/no)
- Setting of immediately preceding therapy (adjuvant/metastatic)

The study is open-label and only the central imaging facility will be blinded to the GTx-024 dose received.

5.4 SAMPLE SIZE JUSTIFICATION

This trial will employ a Simon's two-stage (optimal) design independently for each dose arm, 9 mg and 18 mg. The assumptions for the design are as follows:

- H_0: CBR ≤ 0.10
- H_1: CBR ≥ 0.30
- α = 0.025 (one sided)
- Power = 90%

Based on the above assumptions, a sample size of N = 44 ER+/AR+ BC subjects with centrally confirmed AR+ status who have received at least one dose of study medication (evaluable subjects) are needed for each arm to proceed to completion of the second stage. An arm will proceed to the second stage if at least 3 subjects among the first 18 evaluable subjects randomized in stage 1 achieve CB at week 24, defined as CR, PR, or SD, as per RECIST 1.1 as determined by central review. Otherwise, the arm will be closed due to lack of efficacy. If an arm proceeds to the second stage, a statistical success that favors further evaluation of the dose arm(s) in future trials will require at least 9/44 subjects to achieve CB at week 24, i.e., the null hypothesis of an unacceptably low rate of CB, ≤ 10.0%, can be rejected in favor of the

alternative hypothesis that indicates the higher rate, ≥ 30.0%, is more likely. The lower limit of the exact 95% confidence interval at exactly 9 CBs, 11.0%, exceeds 10.0%. At any time during the first stage, if 3 evaluable subjects in any one arm achieve CB at week 24, that arm will proceed to the second stage; otherwise, it will be halted for lack of efficacy. If at any time, 9 evaluable subjects in an arm achieve CB at week 24, the efficacy criteria for that arm has been met; however, the arm should proceed to full accrual of 44 evaluable subjects in the successful arm(s) in order to better characterize the CB rate, evaluate secondary endpoints, and describe the safety profile of the dose level(s). The intent of the trial is not to statistically compare the CB rate between the two dose arms and the trial is not powered to do so.

Subjects whose AR+ status is not centrally confirmed will be replaced in order to accrue the necessary number of evaluable AR+ subjects to be included in the primary analysis. Subjects who are not confirmed AR+ may remain on the trial, but will not be part of the primary efficacy analysis – these subjects will contribute to secondary and tertiary analyses. Subjects who are randomized but never receive study drug will be replaced. Assuming that 25% of subjects enrolled are not evaluable for the primary efficacy analysis for the reasons noted above, up to 30 additional subjects may need to be randomized.

6 Statistical Considerations

For the primary endpoint, CBR, this trial will employ a Simon's two-stage (optimal) design (see section 5.4 sample size for assumptions).

As the study is non-comparative, for binary endpoints, the exact 95% confidence interval about the endpoint will be constructed. These will be 2-sided with the lower bound observed applying the Clopper-Pearson confidence interval method. **The 95% Confidence interval lower bound is equivalent to the one-sided α = 0.025.**

For the time to event endpoints (PFS, TTP, duration of response and OS), the median time to event and 2-sided 95% Confidence intervals will be constructed using the log-log transformation.

Where change from baseline is analyzed using the Wilcoxon Signed Rank Test, the associated 95% confidence intervals for the median change will be based on the Hodges-Lehmann estimator.

The date of randomization will be taken from IVRS.

6.1 Missing data handling

No evaluable subject will have missing data for the primary efficacy analysis because subjects who have the week 24 assessment will be classified as having CB or not, and those who do not have a week 24 assessment will classified as not having CB.

No further imputation for missing data will be carried out other than to complete partial dates using standard imputation techniques as described below.

See section 8.7.1 for details of determining most severe grade of AE and study drug relationship in case of missing data.

6.2 Partial date imputation

The following rules should be used when modifying partial or missing dates for reporting purposes such as defining on treatment flags.

A permanent new date variable should be created if there is a requirement to present it in a table, listing or figure; otherwise, the same principles are used in determining flags, sort and other derived variables. Imputed data variable names will be defined in the derived dataset specifications.

Original (raw) date variables must not be overwritten. Imputed dates will not be displayed in the listings.

For the time to event variables censoring rules will apply, so there should be no missing data.

General rules

Disease history, medical history or concomitant medications are considered to have started at the earliest possible date and end at the latest possible date.

Some examples are given below (DDMMMYYYY)

In most cases, start dates are imputed as first day of the month or first of January.

Data Type	Start Date	Imputed Start Date	First Dose date	Last Dose Date	End Date	Imputed End Date

Medical History	--FEB2012	01FEB2012	01JAN2012	01MAR2012	--FEB2012	29FEB2012
Medical History	--FEB2012	01FEB2012	03FEB2012	01MAR2012	--MAR2012	31MAR2012
Adverse Event, Prior/Concomitant Meds	--FEB2012	01FEB2012	01JAN2012	01MAR2012	--FEB2012	29FEB2012
Adverse Event, Prior/Concomitant Meds	--FEB2012	03FEB2012	03FEB2012	01MAR2012	--MAR2012	31MAR2012
Adverse Event, Prior/Concomitant Meds	-----2012	03FEB2012	03FEB2012	01MAR2012	-----2012	16MAR2012 £
Adverse Event, Prior/Concomitant Meds		03FEB2012	03FEB2012	01MAR2012	--MAR2012	31MAR2012
Adverse Event, Prior/Concomitant Meds	--JAN2012	01JAN2012	03FEB2012	01MAR2012		01MAR2012 *

£ Subject discontinued on 16MAR2012; * Subject discontinued on 01MAR2012.

For adverse events or concomitant medications, any partial start date during the month of first dosing would be imputed at the date of first dose, taking the worst case scenario. For any AE or concomitant medication starting after the month of first dosing, the start date would be imputed at the first day of the month. Partial adverse event or concomitant medication end dates would be imputed at the last day of the month.

Partial dates are not expected for tumor assessment data and dates of death. However, should partial dates be present for death or on treatment tumor assessments, the dates would be imputed as the first day of the month if the month is present or first of January if only the year is present.

To calculate the age where only year of birth is recorded as per local regulations (e.g. Hungary), the year of birth will be subtracted from the year of screening. For age at randomization required as a covariate for the exploratory analyses, where only year of birth is recorded the year of birth will be subtracted from the year of randomization.

6.3 Visit windowing

In general there will be no programmatic windowing of visits. Post-baseline data will be

presented according to the visit at which it was collected on the eCRF. The only exception is the tumor assessment for centrally read assessments and investigators' assessments for which a window of ± 7 days will be used at Week 24 for EOT or unscheduled visits. For example, if an EOT assessment falls in the window ± 7 days of the Week 24 assessment (day 169 ± 7 days gives a visit window of day 162 to day 176) then it will be assigned as Week 24 in the tables and as Week 24/EOT in the listings. All other tumor assessments will be recorded at the visit at which they are entered in the eCRF.

6.4 Baseline

Baseline is defined as the last non-missing value/result where assessment date is less than or equal to the date of first study treatment, unless otherwise specified for individual assessments.

6.5 Reporting guidelines

The following guidelines will be followed:

- **Page Orientation:** Landscape
- Post-text tables and listings will be generated in .lst and converted to rtf. No in-text outputs are planned.
- **Font** – Courier New font with 8 point font size
- Paper size – US letter size paper, wider margin at top.
- Columns header will be left aligned.
- **Treatment labels:** unless otherwise stated, the subjects will be summarized under 'GTx-024 9 mg' and 'GTx-024 18 mg' treatment groups.
- **Visit labels**: the following visit labels will be used as required.

CRF Visit	Tables, Figures and Listings Label	Visit Number/ID
Screening	SCR	1/v1
Randomization/Baseline	BAS	2/v2
Treatment Phase: Day 43 (+/- 7 Days)	Week 6	3/v3
Treatment Phase: Day 85 (+/- 7 Days)	Week 12	4/v4
Treatment Phase: Day 127 (+/- 7 Days)	Week 18	5/v5

CRF Visit	Tables, Figures and Listings Label	Visit Number/ID
Treatment Phase: Day 169 (+/- 7 Days)	Week 24	6/v6
Treatment Phase: Day 197 (+/- 7 Days)	Week 28	7/v7
Treatment Phase: Day 253 (+/- 7 Days)	Week 36	8/v8
Treatment Phase: Day 337 (+/- 7 Days)	Week 48	9/v9
Treatment Phase: Day 421 (+/- 7 Days)	Week 60	10/v10
Treatment Phase: Day 505 (+/- 7 Days)	Week 72	11/v11
Treatment Phase: Day 589 (+/- 7 Days)	Week 84	12/v12
Treatment Phase: Day 673 (+/- 7 Days)	Week 96	13/v13
End of Treatment	EOT	777/eot
Safety Follow-Up	VFU	778/fup
Long Term Follow-Up	LTFU	779/ltfup (note: this is a summary eCRF page, and there will be multiple contact dates for a subject within the long term follow-up)
Unscheduled visit(s) will be mapped to the visit to which they correspond to and will be flagged in listing with "R" following the visit name.		uns/uns

- Unscheduled visit / repeat assessments: Data obtained at unscheduled assessments will only be included in time to event analyses and in the calculation of BOR/OOR and CBR (excluding the ones at specific timepoints), and it will be presented on data listings. No visit windows will be applied, except for tumor assessments as specified in Section 6.3.
- **Continuous data** will be summarized using n, mean, median, standard deviation, minimum value, maximum value and number of missing values (if there are any).
- **Categorical data** will be summarized using n and percentage based on number of non-missing values.
    - The number of missing values will be presented as a separate category with no percentage, but only if 1 or more subjects have missing data for the summary. Otherwise, all categories will be presented (even if no subjects are counted in the category).
    - Counts of zero in any category will be presented without percentage.
    - Except otherwise specified, percentages will be calculated based on the number of subjects in the population.
- **Precision of summary statistics:**

- Integer – Sample size (n, N) and number of missing responses (if displayed)
- One additional decimal place than reported/collected – mean, geometric mean, median, other percentile, confidence interval
- Two additional decimal places than reported/collected – standard deviation
- Same number of decimal places as reported/collected – minimum, maximum
- Percentages –one decimal place

- Data will be presented on listings in order of unique subject ID (study, center, subject, e.g. G200802-XXXX-XXXX), visit, assessment date / time and assessment (in order collected on eCRF, unless specified otherwise), by treatment arm.
- File naming: Each TFL output file will be named with a t, l or f to denote the output type and then according to its table numbering in the following way: Table 14.2-1.1 would be t14_2_1_1, Table 14.2-11 would be t14_2_11, Listing 16.2.7-1.1 would be l16_2_7_1_1, and Figure 14.2-2.1 would be f14_2_2_1.

7 Analysis Sets and Subgroups

7.1 Analysis Sets

Full Analysis Set (FAS): All subjects who are randomized and receive at least one dose of study drug irrespective of AR+ status.

Evaluable Subjects Analysis Set (EAS): Subjects in the FAS who have centrally confirmed AR+ status; this analysis set is used for the primary efficacy analysis.

Per Protocol Set (PPS): All subjects with centrally confirmed AR+ who have baseline tumor scans, complete the week 24 RECIST 1.1 evaluation, receive at least 80% of the anticipated doses, and have no major protocol deviations.

Safety Analysis Set (SAS): All subjects who are randomized and receive at least one dose of study drug.

Safety analysis subjects will be analyzed in the treatment arm in which they are initially dosed. Efficacy analysis subjects will be analyzed according to the treatment arm to which they were randomized.

The Safety Analysis Set and Full Analysis Set will be the same population.

The analysis sets applicable for each endpoint are summarized below:

Endpoint	Key population	Supporting population
Primary	EAS	FAS, PPS
Secondary	EAS	FAS, PPS
Tertiary	EAS	FAS
PK	EAS	FAS
Safety	SAS	EAS

7.1.1 Protocol deviations

The full list of types of protocol deviations and their relation to the analysis sets, along with the method of identification of each protocol deviation, are detailed in the protocol deviation criteria form which is separate to this SAP. This will be used as a basis for identifying subjects with protocol deviations throughout the study.

Protocol deviations noted during the trial by the CRAs, Medical Monitors or Data managers will be tracked throughout the study. Additionally protocol deviations will be identified programmatically in SAS® using data from the clinical database.

At the protocol deviation review meetings, a consolidated list of protocol deviations will be reviewed. Prior to database lock, this review will include a review and agreement of the final analysis populations. List of subjects and summary of protocol deviations by deviation category will be provided.

Major protocol deviations will be summarized by deviation category. A listing of major protocol deviations by subject will also be provided. Major protocol deviations are identified as 'key' deviations on the protocol deviation criteria form. These deviations are a subset of deviations that require notification to the Project Manager (PM), Medical Monitor (MM) and/or Sponsor. These are significant deviations that may impact the safety of the subject or integrity of the study.

7.2 Subgroup Analyses

The following subgroup categories are defined:

- Bone/no bone metastases at baseline
    - Subjects with bone metastases
    - Subjects without bone metastases
- Regions
    - US
    - Europe and Australia
- Setting of most recent therapy (randomization stratification factor)
    - Adjuvant
    - Metastatic
- Previous chemotherapy for breast cancer
    - Yes
    - No
- Length of response to immediately preceding hormonal therapy
    - Less than or equal to 12 months
    - Greater than 12 months

    This will be determined from the length of response to most recent hormone therapy as recorded on the prior hormonal therapies eCRF. If there are any subjects with unknown time of response to prior hormonal therapy, they will not be included in the subgroup analysis. Subjects who did not respond to the prior hormonal therapy will have a length of response of <= 12 months.

Should any sites be added outside of US, Europe and Australia, then these groups may be updated.

If there are at least 5 subjects in each subgroup, a by subgroup summary of CBR at week 24 and time to event (PFS, TTP, duration of response and OS) analyses, based upon centrally read assessments will be presented. There will be no formal statistical comparison of the subgroups.

8 Methods of Analyses and Presentations

8.1 SUBJECT DISPOSITION

The number and percentage of subjects who were screened, with a randomization/baseline visit, treated, discontinued from study treatment, follow-up completed or withdrawn or completed 24 months of treatment will be presented for each treatment group, along with the reasons for withdrawal from treatment and/or follow-up. Information on study discontinuation and treatment withdrawal will also be listed.

The number of subjects in each analysis population will be summarized by treatment group. The number of subjects in the FAS and EAS will also be presented in a table by region, country and center.

8.2 DEMOGRAPHIC AND BASELINE CHARACTERISTICS

Continuous demographic data (i.e., age, weight, height, and body mass index [BMI]) for each treatment group will be listed and summarized by descriptive statistics. Categorical demographic data (i.e., gender, race, ethnicity, and ECOG performance status) will also be listed and tabulated by treatment groups as frequency counts and percentages. Summaries of demographic data for each treatment group will be provided for the FAS, EAS and PPS analysis sets.

8.3 MEDICAL HISTORY

Medical history will be summarized by treatment group for FAS and EAS. Frequency counts will be presented for each treatment group by MedDRA preferred term.

Information on history of breast cancer will be summarized by treatment group. For categorical variables including stage of cancer, site, presence of disease progression or disease recurrence, AR status confirmation, HER-2 status, ER status, PR status, EGFR status, BRCA1 Status, BRCA2 Status, measureable and bone only non-measureable disease, and TNM stage, frequency counts and percentages will be displayed for each treatment group. For continuous variables like percentage of cells staining AR+, ER+ and PR+, AR+ composite score, time since initial diagnosis of breast cancer, time since diagnosis of metastatic breast cancer and time since initial diagnosis, summary statistics will be displayed. Time since initial diagnosis and time since

diagnosis of metastatic breast cancer are calculated relative to date of randomization. Any partial dates should be imputed using the rules given in Section 6.

Prior breast cancer hormonal therapy including information on purpose of therapy (neo-adjuvant, adjuvant, metastatic and other), duration of response, and response to therapy will be summarized for each treatment group by means of frequency counts and percentages. More than one setting can be reported for each subject. The preferred terms for prior hormonal therapy will be summarized for each treatment group using the WHO-DDE dictionary.

8.4 EFFICACY DATA ENDPOINTS AND ANALYSES

Tumor response will be assessed by RECIST 1.1[2] criteria from centrally read results.

The table below (from RECIST 1.1) provides a summary of the overall response status determination at each time point for subjects who have measurable disease at baseline.

Time point response calculations for subjects with target (+/–non-target) disease

Target lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	Non-CR/non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD
CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = not evaluable.			

When subjects have non-measurable (therefore non-target) disease only, the following table is to be used.

Time point response calculations for subjects with non-target disease only

Non-target lesions	New lesions	Overall response

Non-target lesions	New lesions	Overall response
CR	No	CR
Non-CR/non-PD	No	Non-CR/non-PD[a]
Not all evaluated	No	NE
Unequivocal PD	Yes or No	PD
Any	Yes	PD

CR = complete response, PD = progressive disease, and NE = not evaluable.
[a] 'Non-CR/non-PD' is preferred over 'stable disease' for non-target disease since SD is increasingly used as endpoint for assessment of efficacy in some trials so to assign this category when no lesions can be measured is not advised.

Tumor response is judged relative to baseline tumor assessments. A subject will be considered to have confirmed clinical benefit (CB) if the post-baseline assessment indicates SD, PR, or CR. For subjects with non-measurable disease at baseline (only non-target lesions), SD is defined as those subjects with a response of non-CR/non-PD for determining their clinical benefit. The investigators assessment of overall response will be presented as supportive analyses only.

As per protocol, tumor assessments are collected up to the end of treatment/early termination, therefore all post-baseline assessments will be considered for the analyses (unless otherwise stated).

8.4.1 Primary Efficacy Endpoint and Analyses

The primary analyses will be based upon the centrally read results through week 24 based upon the subjects confirmed response at week 24.

Please see section 8.4.2 for the determination of confirmed response at week 24 in order to identify the CB at 24 weeks.

CB will be determined based on the following information supplied in the database (either central or local reader) and is defined as:

- CB at week 24 = a response of SD or confirmed PR or CR
- No CB = all other scenarios including if week 24 is missing an assessment (i.e. no other prior or post week 24 reads will be considered)

The exact 2-sided 95% confidence interval about the CBR at week 24 will be constructed using the Clopper-Pearson method programmatically in SAS®.

The analysis will be performed at following timings:

Stage 1: This assessment will be performed once the first consecutive 18 evaluable subjects (EAS) in each arm have been followed for at least 24 weeks and have their tumor assessment complete, or prior to week 24 have progressed, discontinued study medication, withdrawn from the study, or died. At this stage the requirement for 3 out of 18 evaluable subjects in each arm having CB in order to proceed to the next stage is assessed.

Stage 2: This assessment will be performed when the first consecutive 44 evaluable subjects (EAS) in each arm have progressed, discontinued study medication, withdrawn from the study, or died before 24 weeks, or have been followed for at least 24 weeks and have their tumor assessment complete. At this stage the requirement of 9 out of 44 evaluable subjects in each arm having CB in order to deem the efficacy criteria to have been met is assessed.

Should at the end of stage 1 and stage 2 there be more than 18 and 44 EAS subjects respectively in each arm at that timepoint, then the CBR for the primary analysis will be determined based upon the first 18 or 44 EAS subjects only (in each arm), identifying them based in order of their date of informed consent date from the eCRF. Separate tables for CBR for the primary analysis at week 24 will be presented for EAS Stage 1 (N=18 per arm), EAS Stage 2 (N=44 per arm) and all subjects in the EAS.

The CB rate at 24 weeks will be summarized for centrally read results separately for EAS, FAS and PPS analysis sets. Any tumor assessments after the subject starts any new prohibited anti-cancer therapy or surgery on target lesions should be excluded from the primary assessment of CBR at 24 weeks. For the centrally read results, the assessments of only the 'final read' (i.e. the reader selected as most appropriate by the adjudicator) will be included in the analysis of CB and other secondary efficacy endpoint. The results of both 'Reader 1 and 'Reader 2' will be listed.

As a supportive analysis, the CB rate at 24 weeks will also be summarized for investigators' assessments using EAS and FAS populations. This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.

Analysis	Population	Derivation
CBR at week 24 (Main analysis)	EAS (Stage 1, Stage 2 and all EAS)	Confirmed response (clinical benefit – confirmed CR, confirmed PR, or SD) at week 24 using centrally read assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions.
CBR at week 24 (Secondary)	FAS	As for main analysis
CBR at week 24 (Secondary)	PPS	As for main analysis
CBR at week 24 (Supportive)	EAS/FAS	Confirmed response (CB- clinical benefit – confirmed CR, confirmed PR, or SD) at week 24 using Investigators' assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions. This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.

8.4.2 Secondary Efficacy Endpoints and Analyses

Unless otherwise stated, all main secondary analyses are based upon the centrally read tumor assessments. Analyses of the investigators' assessments are presented as supportive analyses only. A summary of the planned analyses is as follows:

Secondary analyses	Population	Derivation
CBR at week 24 (Main) ORR at week 24 (Main)	EAS/FAS/PPS	Confirmed response (CB - clinical benefit – confirmed CR, confirmed PR, or SD) at week 24 using centrally read assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions. For ORR, include only subjects with measurable disease at baseline. Response is subjects with confirmed CR or confirmed

Secondary analyses	Population	Derivation
		PR.
CBR at week 24 (Supportive) ORR at week 24 (Supportive)	EAS/FAS	Confirmed response at week 24 using Investigators' assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions For ORR, include only subjects with measurable disease at baseline. This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.
CBR at week 24 (Supportive)	EAS/FAS	Unconfirmed response (CB – CR, PR or SD) at week 24 using centrally read assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions.
CBR at week 24 (Supportive)	EAS/FAS	Unconfirmed response (CB – CR, PR or SD) at week 24 using Investigators' assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.
CBR at week 12 (Supportive)	EAS/FAS	Unconfirmed response (CB – CR, PR or SD) at week 12 using centrally read assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions.

Secondary analyses	Population	Derivation
CBR at week 12 (Supportive)	EAS/FAS	Unconfirmed response (CB – CR, PR or SD) at week 12 using Investigators' assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.
ORR over entire period (Main)	EAS/FAS/PPS	BOR over entire period using centrally read assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions For ORR, include only subjects with measurable disease at baseline.
ORR over entire period (Supportive)	EAS/FAS	BOR over entire period using Investigators' assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions. For ORR, include only subjects with measurable disease at baseline. This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.
PFS, TTP (Main)	EAS/FAS/PPS	Centrally read assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or

Secondary analyses	Population	Derivation
		surgery on target lesions. Include subjects with both measurable and non-measurable disease at baseline.
PFS, TTP (Supportive)	EAS/FAS	Investigators' assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions. Include subjects with both measurable and non-measurable disease at baseline. This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.
Duration of response (Main)	EAS/FAS/PPS	Centrally read assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions. Include subjects with measurable disease at baseline and with a confirmed response of CR or PR.
Duration of response (Supportive)	EAS/FAS	Investigators' assessments. Exclude tumor assessments after start of any new prohibited anti-cancer therapy or surgery on target lesions. Include subjects with measurable disease at baseline and with a confirmed response of CR or PR. This analysis will include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.
Overall survival (Main)	EAS/FAS/PPS	

Clinical Benefit Rate at 24 Weeks

Deriving the CBR as for week 24, the exact 95% confidence interval about the CBR at 24 weeks will be constructed using the Clopper-Pearson method. The lower bound of the 95% CI is the main measure of interest as it is equivalent to the one-side alpha=0.025. The CB rate at 24 weeks will be summarized for centrally read results EAS, FAS and PPS populations. The CB rate at 24 weeks will be summarized for investigators assessments using EAS and FAS populations as supportive analyses.

Any tumor assessments after the subject starts any new anti-cancer therapy or surgery on target lesions should be excluded from the assessment of CBR at 24 weeks.

Best Overall Response (BOR)/ Objective Response Rate (ORR)

The number and percentage of subjects in the following categories of BOR will be tabulated:

- Complete Response (CR)
- Partial Response (PR)
- Stable Disease (SD)
- Progressive Disease (PD)
- Not evaluable (NE)

The ORR is defined as the percentage of subjects, with measurable disease at baseline, having a confirmed response of either CR or PR at week 24 time point only. Only subjects with measurable disease (target lesions) at baseline will be included in the analysis of the ORR.

The ORR at 24 weeks will be summarized for centrally read results separately for EAS, FAS and PPS analysis sets. The ORR at 24 weeks will be summarized for investigators' assessments using EAS, FAS populations. The exact 95% confidence interval about the ORR (PR or CR) for the confirmed response will be constructed at week 24 using the Clopper-Pearson method.

Subjects need to have two consecutive assessments of PR or CR to be a responder (note: a sequence of CR-NE-CR or PR-NE-PR would be considered as confirmed CR and PR respectively). PR or CR has to be confirmed by 2 consecutive tumor evaluations spaced at least 4 weeks (28 days) apart.

Subjects who do not have any evaluable post-baseline assessments will be classified as having both confirmed response and BOR of Not evaluable.

The following table, based on Table 3 from the RECIST 1.1 guidelines[2], summarizes the algorithm describing how confirmed response and BOR are determined from the overall tumor assessments. The order used to determine BOR is CR>PR>SD>PD, ignoring visits with missing tumor assessments.

All assessments following an assessment of PD will be excluded from the calculation of BOR:

- For week 24 include assessments up to week 24, and up to week 28 if CR or PR need to be confirmed. The responses at week 24, or week 28 if CR or PR needs to be confirmed, are timepoint [b] in the table below.
- The BOR is determined once all the data for the subject is known. For BOR over the entire period, all tumor assessments are included (excluding those after intake of any new anti-cancer therapy or assessment of PD)

Case	Overall response first timepoint [a]	Overall response subsequent timepoint [b]	Confirmed response/BOR
1	CR	CR	CR (if assessments at least 28 days apart). (note: sequence of CR – NE – CR would be considered as confirmed CR)
2	CR	PR	SD, PD or PR • If CR truly met at first timepoint, any subsequent assessment of PR should make the disease PD at that point. That is, neither a PR nor SD may follow CR • Therefore, SD, if CR assessment >=11 weeks (77 days after date of randomization), otherwise PD. • However BOR may be PR if subsequent scans suggests small lesions were still present at first assessment (in which case first assessment of CR should be changed to PR)
3	CR	SD	SD or PD • SD, if CR or SD assessment >=11 weeks (77

Case	Overall response first timepoint [a]	Overall response subsequent timepoint [b]	Confirmed response/BOR
			days after date of randomization), • otherwise PD
4	CR	PD	SD or PD • SD, if CR assessment >=11 weeks (77 days after date of randomization), • otherwise PD
5	CR	NE	SD or NE • SD, if CR assessment >=11 weeks (77 days after date of randomization • otherwise NE
6	PR	CR	PR (if assessments at least 28 days apart).
7	PR	PR	PR (if assessments at least 28 days apart). (note: sequence of PR – NE – PR would be considered as confirmed PR) Where there are cases of more than one SD assessment between two PR assessments, then this should be discussed.
8	PR	SD	SD
9	PR	PD	SD or PD • SD, if PR assessment >=11 weeks (77 days after date of randomization), • otherwise PD
10	PR	NE	SD or NE • SD, if PR assessment >=11 weeks (77 days after date of randomization), • otherwise NE
11	SD	SD	SD
12	SD	PD	SD or PD • SD, if SD assessment >=11 weeks (77 days after date of randomization), • otherwise PD
13	SD	NE	SD or NE • SD, if SD assessment >=11 weeks (77 days after date of randomization), • otherwise NE

Case	Overall response first timepoint [a]	Overall response subsequent timepoint [b]	Confirmed response/BOR
14	NE, -	SD	SD
15	CR, PR, SD	-	SD or NE • SD, if assessment >=11 weeks (77 days after date of randomization) and does not qualify for CR or PR, • otherwise NE.
16	PD		PD. Ignore all assessments after initial overall response of PD.
17	NE	NE	NE Where all assessments are Not evaluable

[a] For week 24 response, this is the time point prior to week 24 for SD or week 24 to confirm PR or CR. For BOR over the entire period, this is the time point prior to the 'best' observed response.
[b] For week 24 response, this is the week 24 assessment or week 28 assessment to confirm PR or CR. For BOR over the entire period, this is the time point of the 'best' observed response.

The exact 95% confidence interval about the ORR (PR or CR) for the BOR during the whole treatment period will be constructed using the Clopper-Pearson method.

The BOR and ORR (PR or CR) over the entire treatment period will be summarized for centrally read results separately for EAS, FAS and PPS analysis sets. The ORR (PR or CR) for BOR over the entire treatment period will be summarized for investigators assessments using EAS and FAS populations. The exact 95% confidence interval about the ORR during the whole treatment period will be constructed using the Clopper-Pearson method.

A summary of unconfirmed response at week 12 and 24 will also be presented as a supportive analysis. The unconfirmed response does not need to two consecutive CR or PR assessments described above for response. A single result of PR or CR will be sufficient to be determined as an unconfirmed response or clinical benefit.

PFS

The **PFS** is defined as the time from randomization until the first radiographically documented progression of disease or death from any cause, whichever occurs first. Subjects without measureable disease at baseline (i.e. no target lesions at baseline) will be included in the

evaluation of PFS.

The following are key events to consider during determination of PFS:

1. This will be measured only whilst subjects are on study treatment (up to their end of treatment visit), if at the time of treatment discontinuation, no PFS event has occurred, then the event will be censored.
2. Any tumor assessments after the subject starts any new anti-cancer therapy or surgery on target lesions should be excluded from the assessment of PFS.
3. Subjects with no PFS events will be censored at the time of the last evaluable (last assessment that is CR, PR or SD) tumor assessment. Subjects with no tumor assessment after the baseline visit will be censored at the time of randomization.

- **If a subject has disease progression observed then:**

  Progression free survival = (Date of first radiographically documented PD – Date of randomization) + 1

- **If a subject dies and does not have radiographically based disease progression observed then:**

  Progression free survival = (Date of death – Date of randomization) + 1

- **If a subject does not experience disease progression or does not die during the study** then a censored progression free survival will be calculated. This will be calculated as:

  Progression free survival = (Date of last available tumor assessment where subject is known to be PD free - Date of randomization) + 1

- **If a subject does not have any tumor assessments after randomization (i.e. post-baseline) then:**

  Progression free survival = (Date of randomization – Date of randomization) + 1 (i.e. Day 1)

The following table summarizes the potential scenarios for the main analysis of PFS (PFS) in

order of priority with 1 superseding any outcome for 2 onwards:

Order	PFS (includes documented progression only) Situation	Date of Progression or Censoring	Outcome
1	No baseline tumor assessments	Date of randomization	Censored
2	New anticancer treatment started	Date of last radiological assessment of measured lesions prior to start of new treatment	Censored
3	Radiographically documented progression at any visit	Earliest of date of radiological assessment showing progressive disease	Progressed
4	Death before first PD assessment	Date of death	Progressed
5	No progression	Date of last radiological assessment of measured lesions	Censored
6	Treatment discontinuation for reason other than death	Date of last radiological assessment of measured lesions	Censored
7	No post baseline tumor assessments	Date of randomization	Censored

Note the last radiological assessment is the last assessment that is either CR, PR or SD (or Non CR/Non PD for subjects with non-measurable disease).

PFS: Median PFS and 95% confidence intervals will be estimated by the Kaplan-Meier method and the survival function and associated 95% confidence intervals will be constructed at 12, 24, 36 and 48, 72 and 96 weeks only if there are sufficient data, however the final choice of time points may be updated dependent upon the data.

The PFS will be summarized for centrally read results separately for EAS, FAS and PPS populations. The PFS will be summarized using the investigators assessment of overall tumor response for EAS and FAS populations, and will be based on either radiological evidence or clinical information only.

The PFS will be presented graphically separately by centrally read results and investigators' assessments and for EAS, FAS and PPS populations (centrally read results) and for EAS and FAS populations (investigators assessments). The analysis of PFS based upon investigators' assessment will also include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.

TTP

Time to progression is defined as for PFS, except for subjects who did not have a radiographically documented progression and have died due to reason other than PD. For such subjects death will not be counted as an event, but they will be censored on the last available tumor assessment prior to the death date.

Order	TTP (includes documented progression only) Situation	Date of Progression or Censoring	Outcome
1	No baseline tumor assessments	Date of randomization	Censored
2	New anticancer treatment started	Date of last radiological assessment of measured lesions prior to start of new treatment	Censored
3	Radiographically documented progression at any visit	Earliest of date of radiological assessment showing progressive disease	Progressed
4	Death due to progressive disease	Date of death	Progressed
5	No progression	Date of last radiological assessment of measured lesions	Censored
6	Treatment discontinuation for reason other than death	Date of last radiological assessment of measured lesions	Censored
7	Death due to reason other than progressive disease	Date of last radiological assessment of measured lesions	Censored
8	No post baseline tumor assessments	Date of randomization	Censored

Any tumor assessments after the subject starts any new anti-cancer therapy or surgery on target lesions should be omitted in the assessment of TTP.

Median TTP and 95% confidence intervals will be estimated by the Kaplan-Meier method and the survival function and associated 95% confidence intervals will be constructed at 12, 24, 36, and 48, 72 and 96 weeks only if there are sufficient data, however the final choice of time points may be updated dependent upon the data.

The TTP will be summarized for centrally read results separately for EAS, FAS and PPS analysis sets. The TTP will be summarized using the investigators' assessment for EAS, FAS analysis sets and will be based on either radiological evidence or clinical information only.

The TTP will be presented graphically separately by centrally read results and investigators assessments and for EAS, FAS and PPS populations (centrally read results) and for EAS and FAS

populations (investigators' assessments). The analysis of TTP based upon investigators' assessment will also include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.

Duration of Response (in responders)

Duration of response, in responders, is defined as the period from the date of initial PR or CR until the date of progressive disease or death from any cause. Only subjects with BOR of CR or PR (i.e. responders) will be included in the analysis of duration of response.

Duration of response = (Date of disease progression/death - Date of first recorded CR/PR) + 1.

Subjects with no documented progression or death after CR or PR will be censored at the last date at which they are known to have had the CR or PR, respectively (regardless of the response at intermediate assessments). The method for handling censoring is the same as described for the PFS.

Duration of response = (Date of last CR/PR assessment - Date of first recorded CR/PR) + 1

Any tumor assessments after the subject starts any new anti-cancer therapy or surgery on target lesions should be discarded in the assessment of duration of response and the response information from the last radiological assessment prior to initiation of new therapy or surgery on target lesions should be used.

Duration of response: Median duration of response and 95% confidence intervals will be estimated by the Kaplan-Meier method and the survival function and associated 95% confidence intervals will be constructed at 12, 24, 36, and 48, 72 and 96 weeks only if there are sufficient data weeks for responders, however the final choice of timepoints may be updated dependent upon the data.

The duration of response will be summarized for centrally read results separately for EAS, FAS and PPS populations. The duration of response will be summarized using the investigators' reported assessment for EAS, FAS populations. The analysis of duration of response based upon investigators' assessment will also include investigators' assessments of progressive disease based on either radiological evidence or clinical information only.

The duration of response will be presented graphically separately for centrally read results and investigators assessments for EAS, FAS and PPS analysis sets (centrally read results) and for EAS and FAS analysis sets (investigators' assessments).

Overall survival

The OS is defined as the time from randomization to death or last follow up date with a maximum of 12 months post last dose. Subjects known to be alive at their last follow up will be censored at that time.

- **If a subject dies then:**

  Overall survival = (Date of death – Date of randomization) + 1

- **If a subject does not die during the study or in follow-up** then a censored overall survival time will be calculated as there will be no date of death. This will be calculated as:

  For subjects with follow-up assessment:
  Overall survival = (Date of last follow up date – Date of randomization) + 1
  For subjects without follow-up assessment:
  Overall Survival = (Date of last study treatment – Date of randomization) + 1

  For subjects with no post-baseline information:
  Overall Survival = (Date of randomization – Date of randomization) + 1

Overall survival: Median OS and 95% confidence intervals will be estimated by the Kaplan-Meier method and the survival function and associated 95% confidence intervals will be constructed at 12, 24, 36, and 48, 72 and 96 weeks only if there are sufficient data, however the final choice of time points may be updated dependent upon the data.

The OS will be summarized and presented graphically separately for EAS, FAS and PPS populations.

8.4.3 Tertiary Efficacy Endpoints and Analyses

The following tertiary efficacy endpoints will be assessed among the EAS and the FAS:

- PSA: PSA will be obtained during routine laboratory assessments and results at each scheduled assessment will be compared to baseline.

- QoL: QoL will be obtained using the EQ-5D-5L, with each scheduled assessment compared to baseline.

- CTCs: The number of CTCs (enumeration) will be obtained at each scheduled assessment and compared to baseline. The CTC gene expression will not be included on the clinical database. Should any subsequent analysis be performed on the CTC gene expression data, this will be covered by a separate statistical analysis plan.

- ER mutations: ctDNA analysis of ER mutations will determine the presence or absence as well as types of mutations at baseline, Week 24, and EOT and the impact of these mutations on outcome measures (CB response, PFS, TTP and duration of response) will be assessed.

- Duration of prior CB: the impact of duration of prior CB on outcome (CB response at 24 weeks, PFS, TTP and duration of response) will be assessed.

- Time from diagnosis of metastases to randomization: the impact of time from diagnosis of metastases on outcome (CB response at 24 weeks, PFS, TTP and duration of response) will be assessed.

- Tumor volume, change from baseline and percentage change from baseline in tumor volume.

- Plasma concentrations: GTx-024 and GTx-024 glucuronide plasma concentrations, and the effect of GTx-024 and GTx-024 glucuronide on outcome (CB response at week 24).

The following tertiary efficacy analyses apply to the evaluable subjects and the FAS:

- Change in the detectability (PSA above the limit of detectability/PSA undetectable) of PSA from baseline to each post-baseline visit will be summarized in shift tables, and the changes tested for significance using McNemar's test.

- QOL – see section 8.6

- CTCs: The number of CTCs will be summarized by visit. Changes in the number of CTCs (enumeration) from baseline to each scheduled assessment will be described and tested for a significant change from baseline using a Wilcoxon signed rank test. The calculation of the change from baseline value and the test used will be the same as for PSA. Note: Any analysis of CTC gene expression will be presented in a separate SAP.
- ER mutations: The impact of the presence and types of ER mutations found at baseline on the CB response at week 24 will be explored using logistic regression models with and without other covariates. The impact of the presence and types of ER mutations found at baseline on PFS, TTP, and duration of response outcomes will be assessed in a Cox proportional hazards model with and without other covariates. Each mutation will be classified as present/absent and each mutation will be considered a categorical variable in the logistic regression and Cox proportional hazards analyses. The following covariates will be considered: bone metastases (present/absent), time from diagnosis of metastases to randomization (dichotomized by median time), region (US, Europe and Australia), maximum change from baseline in tumor volume (dichotomized by median change), plasma concentrations of GTx-024 and of GTx-024 glucoronide, age at randomization, race, prior chemotherapy (yes/no), prior setting stratification factor (adjuvant/metastatic), AR % staining (low/high) variable created by cutting at the median of % staining AR+, AR composite score (low, high) variable created by cutting at the median of a computed composite score, PSA change at week 24 time point (increased, not increased) and PSA at baseline (undetectable/detectable).
- ER mutations: The impact of the presence and types of ER mutations found at Week 24 and EOT on outcomes of CB response at week 24, PFS, TTP and duration of response will be assessed as above.
- Duration of prior CB: the impact of duration of prior CB on the current CB response (CB = yes/no) will be explored using logistic regression models with and without other covariates (bone metastases (present/absent), time from diagnosis of metastases to randomization (dichotomized by median duration), region (US, Europe and Australia), maximum change in tumor volume (dichotomized by median change), plasma concentrations of GTx-024 and of GTx-024 glucuronide, age at randomization, race, prior chemotherapy (yes/no), prior setting stratification factor (adjuvant/metastatic), AR

% staining (lower than median/higher or equal to median) variable created by cutting at the median of % staining AR+, AR composite score (lower than median/higher or equal to median) variable created by cutting at the median of a computed composite score, PSA change at week 24 time point (increased, not increased), PSA at baseline (undetectable/detectable), and presence/absence of each of the various ER mutations at baseline and week 24/EOT. The odds ratio (OR) estimate together with the 95% confidence interval will be produced using the SAS® PROC LOGISTIC procedure for each potential covariate individually, and then for the best subsets model (using SELCTION=SCORE option) which will include duration of prior CB and two further covariates from the list above and their interaction terms.

  - Note: The composite score is defined as composite score=% staining AR+ * densiometric value.

- The impact of duration of prior CB on the PFS, TTP and duration of response outcomes will be assessed in a Cox proportional hazards model with and without the covariates mentioned above. The hazard ratio estimate together with the 95% confidence interval will be produce using the SAS® procedure PROC PHREG. As for the logistic regression analysis, the results will be presented for a model containing each potential covariate individually and for the best subsets model (using SELCTION=SCORE option) including duration of prior CB and two further covariates from the list above.

- Time from diagnosis of metastases to randomization: the impact of time from diagnosis of metastases on the current CB outcome will be explored using logistic regression models with and without other covariates (same covariates as in the paragraph above). The odds ratio (OR) estimate together with the 95% confidence interval will be produced using the SAS® PROC LOGISTIC procedure for each covariate individually, and then the best subsets model which will include time from diagnosis of metastases to randomization and two further covariates from the list above and their interaction terms in the same way as was done previously for duration of prior CB.

- The impact of time from diagnosis of metastases on the PFS, TTP and duration of response outcomes will be assessed in a Cox proportional hazards model with and without the covariates mentioned above. The hazard ratio estimate together with the

95% confidence interval will be produce using the SAS® procedure PROC PHREG. As for the duration of prior CB analysis, the results will be presented for a model containing each covariate individually and for the best subsets model including time from diagnosis of metastases to randomization and two further covariates from the list above.

- PSA levels at baseline (undetectable/detectable): the effect of PSA levels at baseline on the current CB outcome will be explored using logistic regression models with and without other covariates (same covariates as in the paragraph above). The odds ratio (OR) estimate together with the 95% confidence interval will be produced using the SAS® PROC LOGISTIC procedure for each covariate individually, and then the best subsets model which will include PSA levels at baseline and two further covariates from the list above and their interaction terms in the same way as was done previously for duration of prior CB.

- The effect of PSA levels at baseline (undetectable/detectable) on the PFS, TTP and duration of response outcomes will be assessed in a Cox proportional hazards model with and without the covariates mentioned above. The hazard ratio estimate together with the 95% confidence interval will be produce using the SAS® procedure PROC PHREG. As for the duration of prior CB analysis, the results will be presented for a model containing each covariate individually and for the best subsets model including PSA levels at baseline and two further covariates from the list above.

- Change in PSA levels at week 24/EOT (increased/not increased): the effect of change in PSA levels on the current CB outcome will be explored using logistic regression models with and without other covariates (same covariates as in the paragraph above). The odds ratio (OR) estimate together with the 95% confidence interval will be produced using the SAS® PROC LOGISTIC procedure for each covariate individually, and then the best subsets model which will include change in PSA levels and two further covariates from the list above and their interaction terms in the same way as was done previously for duration of prior CB.

- The effect of Change in PSA levels at week 24/EOT (increased/not increased) on the PFS, TTP and duration of response outcomes will be assessed in a Cox proportional hazards model with and without the covariates mentioned above. The hazard ratio estimate

together with the 95% confidence interval will be produce using the SAS® procedure PROC PHREG. As for the duration of prior CB analysis, the results will be presented for a model containing each covariate individually and for the best subsets model including change in PSA levels and two further covariates from the list above.

- Tumor volumetrics: the effect of GTx-024 18 mg on tumor volume will be described visually using waterfall plots. The largest decrease in total tumor volume from baseline will be presented overall from largest to smallest change (in order from largest to smallest). Total tumor volume, change from baseline in total tumor volume and percentage change from baseline in total tumor volume will be summarized by visit.

  The percentage change from baseline is calculated as follows:

  - Percentage change from baseline=100 x (post-baseline total tumor volume - baseline total tumor volume)/baseline total tumor volume

- Plasma concentrations: the effect of GTx-024 and GTx-024 glucuronide concentrations will be assessed for their association with CBR at weeks 24 using logistic regression by creating binary variables for concentrations by dichotomizing at the median concentration at week 24. The two groups will be 'Lower than median concentration' and 'Higher than or equal to median concentration'. The odds ratio (OR) estimate together with the 95% confidence interval will be produced using the SAS® PROC LOGISTIC procedure.

- The duration of clinical benefit, as assessed by centrally read results, amongst subjects who achieve CB at week 24 will be summarized. The duration of clinical benefit is calculated as the period in months during which the subject did not have PD or die. Essentially this is the same as the PFS but applied only to subjects who have CB at week 24. The duration of clinical benefit assessed by centrally read results will also be graphically presented.

  - If the subject progressed or died due to any cause, then the duration of clinical benefit will be:

    - Date of disease progression or death – Date randomization + 1

- If the subject did not progress or die due to any cause, then the duration of clinical benefit will be censored at the date of the last tumor assessment:
  - Date of last tumor assessment – Date of randomization + 1

8.4.4 Subgroup Analysis

Summaries of CBR at 24 weeks, PFS, TTP, duration of response and OS will be presented for the bone/no bone metastases, region, setting of most recent therapy, previous chemotherapy and the length of response to immediately preceding hormonal therapy subgroups using the EAS and FAS populations.

The subgroup categories are defined as follows:

- Bone/no bone metastases
  - Subjects with bone metastases
  - Subjects without bone metastases
- Region
  - US
  - Europe and Australia
- Setting of most recent therapy (randomization stratification factor)
  - Adjuvant
  - Metastatic
- Previous chemotherapy for breast cancer
  - Yes
  - No
- Length of response to immediately preceding hormonal therapy
  - Less than or equal to 12 months
  - Greater than 12 months

Forest plots will display PFS, TTP, duration of response and OS for each treatment arm by subgroup, including the EAS or FAS for comparison.

8.5 PHARMACOKINETIC ENDPOINTS AND ANALYSES

Plasma concentrations of GTx-024 and GTx-024 glucuronide will be listed and a summary provided by timepoint. These summaries will include mean concentration, standard deviation,

geometric mean, median, minimum and maximum. Pharmacokinetic samples are collected at a single timepoint pre-dose at baseline, week 6, week 18 and week 24 only.

8.6 QUALITY OF LIFE OR PHARMACOECONOMIC ENDPOINTS AND ANALYSES

QoL: The EQ visual analog scale (VAS) will be summarized for EAS and FAS by descriptive statistics at baseline and each assessment. The same summary will be presented split by confirmed response of SD or better at week 24. Changes from baseline will also be summarized using descriptive statistics and 95% confidence intervals and will be tested for a significant change from baseline in VAS using the Wilcoxon signed rank test.

Forest plots of the change from baseline by visit for EQ-5D VAS will be produced using the EAS and FAS populations. Cumulative distribution plot for change from baseline for EQ-5D VAS score will be presented for the EAS and FAS populations by subjects with a confirmed response of SD or better at week 24 and subjects with PD or Not evaluable at week 24.

EQ-5D-5L single item scale scores (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) will be summarized by means of shift tables of baseline score verusus score at each assessment. The EQ-5D-5L single item scale score change from baseline will be presented as summary statistics by visit.

8.7 SAFETY DATA ENDPOINTS AND ANALYSES

Safety analyses will be conducted among the subjects in the SAS and EAS populations. AEs will be coded according to the current version of the Medical Dictionary for Regulatory Activities (MedDRA®). AEs, laboratory outcomes, and results from physical examinations will be graded according to NCI-CTCAE, Version 4.0[4].

Summaries of safety data will be presented for subjects in the SAS and EAS populations.

8.7.1 Adverse Events (AEs)

The version of the MedDRA dictionary used to code adverse events will be provided in the adverse events tables and listing footnotes.

Any adverse event which starts on or after the date of the first dose up to the date of last dose + 30 days is considered treatment emergent. Any adverse event that starts before the start of treatment but increases in severity after the start of treatment should be entered into the eCRF as different AE records with the differing severities recorded.

All information on adverse events (AEs) will be listed by subject. AEs will be collected starting from the ICF date but only Treatment-Emergent Adverse Event (TEAE), Related Treatment-Emergent Adverse Events and Serious Treatment-Emergent Adverse Events will be summarized. If the AE start date is partial or missing, the start date will be imputed as defined in section 6.

An overview of TEAE incidence rates for the SAS population and overall TEAE, related TEAE and serious TEAE by system organ class, preferred term and maximum severity for SAS and EAS populations will be provided.

If there are more than one adverse events within the same System Organ Class, the subject will be counted only once in that System Organ Class.

If there are more than one adverse events coded to the same Preferred Term within a System Organ Class, the subject will be counted only once in that System Organ Class and Preferred Term under the worst severity.

If the severity is missing for an AE its grade will be imputed as the worst grade reported for that subject (across all AEs) between mild, moderate, severe and life-threatening (i.e. excluding death). If the relationship for an adverse event is missing it will be assumed to be related to study drug in the summary tables.

All adverse events will be listed by subject; adverse events which start before the start of study treatment will be listed separately.

8.7.2 Prior and Concomitant Medication

Medications other than the study treatment will be classified with regards to the relationship between their administration dates and the start of treatment. The medications will be coded using WHO-DDE dictionary.

Medications will be defined as follows:

- Prior Medication: Any medication whose medication end date is before the start of study treatment.

- Concomitant Medication: Any medication whose medication start or end date is either the same as or after date of study treatment start.

Any medication with a missing medication end date will be assumed to be continuing.

The number of subjects with each medication will be tabulated by INN class and preferred term for SAS and EAS populations. Separate tables will be produced for Prior medications and Concomitant Medications. A subject will only be counted once within each preferred term and once within each INN class.

A separate summary of concomitant antiandrogen medications will be presented by INN class and preferred term. This summary will include medications from the antiandrogens group (code LO2BB for antiandrogens).

8.7.3 Study Drug Exposure and Compliance

Summary statistics will be produced for duration on treatment, cumulative dose received, and for overall percentage compliance for the SAS and EAS by means of mean, standard deviation, median and range. Where duration on treatment, cumulative dose received and overall percentage compliance are defined as:

- Duration on treatment (months) = (Date of last treatment – Date of first treatment + 1)/ 30.4375
- Compliance is calculated as follows, based upon the initially allocated treatment arm:

$$\text{Compliance } (\%) = \left(\frac{\text{Cumulative dose received (mg)}}{\text{Cumulative dose expected (mg)}} \right) \times 100$$

Where:

$$\text{Cumulative dose received (mg)} = \sum (\text{Tablets dispensed} - \text{Tablets returned} - \text{Tablets lost}) \times 3mg$$

For the 9 mg treatment arm:

$$\text{Cumulative dose expected (mg)} = (\text{Date of last treatment - Date of first treatment} + 1) \times 9mg$$

For the 18 mg treatment arm:

$$\text{Cumulative dose expected (mg)} = (\text{Date of last treatment - Date of first treatment} + 1) \times 18mg$$

The following rules will be applied when determining the compliance:

- If the number of tablets lost is missing it will be assumed to be 0.
- If the number of tablets returned is missing, the compliance will not be calculated.

Compliance will also be summarized using descriptive statistics, and the number (%) of subjects with compliance < 80 %, between 80% and 100%, between 100% and 120% and > 120% will also be reported.

Details of study drug administration, including dates and day of study drug interruptions and dose modifications and study drug compliance will be listed.

8.7.4 Clinical Laboratory Evaluations

Central laboratory parameters (hematology, blood chemistry, coagulation, serum lipid, serum hormones) results will be summarized descriptively by visit for the SAS and EAS, using the units as they are collected so no conversion is required.

Shift tables will be presented for selected biochemistry parameters with CTC criteria for baseline grade to worst CTC grade on treatment. Baseline will be the last observed value before the start of study drug. The worst on treatment value will be the most extreme CTC grade observed after the start of study treatment up to and including the one month follow-up period. Summary statistics of all safety laboratory parameters will be presented by visit.

All data will be listed in subject data listings and abnormal results will be flagged. The laboratory reference ranges will be presented in a listing.

A separate listing will be produced for liver function test parameters for any subject that has an ALT or AST result > 3x UNL at any time during the study. Furthermore a further listing will be produced of any laboratory result > 5 x UNL.

8.7.5 Vital signs

The following vital signs will be summarized for the SAS and EAS by visit:

- Heart rate
- Respiratory rate
- Sitting Systolic blood pressure
- Sitting Diastolic blood pressure
- Weight
- Body temperature

Vitals signs results will be listed and summarized by visit.

8.7.6 Eye exam

The results of the eye examinations will be listed as separate listings presenting the collected data by subject for fundus exam, intraocular pressure, slit lamp biomicroscopy and best corrected visual acuity results.

8.7.7 ECOG Performance Status

ECOG performance status shift tables for SAS and EAS populations will show the baseline ECOG performance status as compared to the end of treatment ECOG performance status. ECOG performance status will also be listed.

9 Interim Analyses

As per the Simon 2-stage design, the clinical benefit rate will be assessed in each study arm in the first 18 evaluable subjects at end of stage 1 before deciding whether to proceed with recruitment of the full 44 evaluable subjects for the stage 2 assessment of CBR. Separate summaries of CBR at week 24 will be presented for Stage 1 EAS (N=18 per arm) and Stage 2 EAS (N=44 per arm).

Safety will be assessed on a regular basis throughout the study by the SMC. The SMC will also review the number of subjects with clinical benefit at week 24, as per the Simon 2-stage design. The SMC will determine if the study should proceed to the second stage by reviewing how many events of interest happened in the first stage (no statistical inference), therefore will be no

multiplicity adjustment necessary.

10 Changes to Planned Analyses

In addition to those analyses stated in the protocol, the primary and main secondary analyses will be repeated using the investigators' assessment of tumor response.

Furthermore, an additional analysis will be the presented for the duration of clinical benefit amongst the subset of subjects who achieve CB at week 24, from start of study treatment until date of death or disease progression. If the subject does not die or have disease progression, they will be censored at the time of their last tumor assessment. Duration of clinical benefit will be derived based on the centrally read tumor assessments.

Summaries of CBR at 24 weeks, PFS, TTP, duration of response (using centrally read assessments) and OS will be presented by bone/no bone metastases, region, setting of most recent therapy, previous chemotherapy and length of response to immediately preceding hormonal therapy subgroups for the FAS and EAS analysis sets.

The protocol states that the change from baseline and percentage change from baseline in PSA will be reported at each assessment. However a high proportion of PSA results are below the level of detection. Therefore, PSA will instead be summarized at each assessment as a shift table of detectable/undetectable from the baseline of assessment.

A summary of CB rate based on unconfirmed response at weeks 12 and 24 will be presented for the EAS and FAS populations.

The following additional tertiary efficacy objectives will be applied to both centrally confirmed AR+ subjects (the evaluable subset of the FAS) as well as to all subjects in the FAS:

- Assess the effect of PSA levels at baseline on outcome
- Assess the effect of change in PSA levels on outcome

10.1 PROTOCOL CLARIFICATIONS

- According to the definition from the Protocol, the OS is defined as the time from

treatment initiation until death or last follow up date. This is not consistent with the PFS and TTP definitions which are using the time from randomization, hence OS defined as time from date of randomization.

- As per section 11.4.2 in the Protocol the impact of ER mutations on outcome measures (CB response and duration of CB) will be assessed. Only the CB response and duration of CB are mentioned here, but in the section 11.4.3.4 Tertiary efficacy analysis there are mentioned CB response, PFS, TTP, and duration of response.

- The Clopper-Pearson method for calculating exact 2-sided 95% confidence intervals will be used instead of the Blyth-Still-Casella method.

11 Document History

Date	Version	Modified by	Brief details of changes made to document
05 JAN 2018	1.0	Sarah Lilley	Initial final version

12 References

[1] Simon R. Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials* 1989; 10: 1 - 10.

[2] http://www.eortc.be/recist/documents/RECISTGuidelines.pdf

[3] http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L.pdf

[4] CTCAE v4 http://evs.nci.nih.gov/ftp1/CTCAE/About.html

13 Appendices

13.1 Tables, Figures and Listing shells