STATISTICAL ANALYSIS PLAN

STUDY TITLE:	**A PHASE II, RANDOMIZED, OPEN-LABEL, MULTICENTER STUDY EVALUATING THE EFFICACY AND SAFETY OF GDC-9545 COMPARED WITH PHYSICIAN'S CHOICE OF ENDOCRINE MONOTHERAPY IN PATIENTS WITH PREVIOUSLY TREATED ESTROGEN RECEPTOR-POSITIVE, HER2-NEGATIVE LOCALLY ADVANCED OR METASTATIC BREAST CANCER**
STUDY NUMBER:	WO42312
STUDY NAME:	acelERA Breast Cancer
VERSION NUMBER:	2
ROCHE COMPOUND(S):	GDC-9545 (RO7197597)
EUDRACT NUMBER:	2020-001984-10
IND NUMBER:	132673
NCT NUMBER:	NCT04576455
PLAN PREPARED BY:	[redacted]

STATISTICAL ANALYSIS PLAN APPROVAL

Date and Time(UTC)	Reason for Signing	Name
27-Jan-2022 03:32:10	Company Signatory	[redacted]

SPONSOR:	F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS:	Grenzacherstrasse 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp above

STATISTICAL ANALYSIS PLAN VERSION HISTORY

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
1	.16 August 2021	Version 3, 09-July 2021
2	See electronic date stamp on title page	Version 3, 09-July 2021

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

This Statistical Analysis Plan (SAP) for Study WO42312 (acelERA Breast Cancer) has been amended to incorporate the following changes.

VERSION 2:

- Objective Reponses Rate (ORR) population removed in Section 4. Inclusion criteria for study is patients with measurable disease hence additional ORR population with measurable disease was deemed redundant.
- Analysis population for ORR in section 5.4.2 changed from ORR population to full analysis set.

Additional minor changes have been made to improve clarity and consistency.

TABLE OF CONTENTS

1. INTRODUCTION 8
   - 1.1 Objectives AND Estimands 8
   - 1.2 Study Design 12
   - 1.2.1 Treatment Assignment and Blinding 12
   - 1.2.2 Independent Review Facility 13
   - 1.2.3 Data Monitoring 13
2. STATISTICAL HYPOTHESES 14
   - 2.1 Multiplicity Adjustment 14
3. SAMPLE SIZE DETERMINATION 14
4. ANALYSIS SETS 15
5. STATISTICAL ANALYSES 15
   - 5.1 General Consideration 16
   - 5.2 Participant Disposition 16
   - 5.3 Primary Endpoint Analysis 16
   - 5.3.1 Definition of Primary Estimand 16
   - 5.3.2 Main Analytical Approach for Primary Estimand 17
   - 5.3.3 Handling of Missing Data 17
   - 5.3.4 Sensitivity Analyses for Primary Endpoint 18
   - 5.3.5 Supplementary Analyses for Primary Endpoint 19
   - 5.3.5.1 Subgroup Analyses for Primary Endpoint 19
   - 5.4 Secondary Endpoints Analyses 20
   - 5.4.1 Overall Survival 20
   - 5.4.2 Overall Response Rate 21
   - 5.4.3 CBR 22
   - 5.4.4 DOR 22
   - 5.4.5 PFS by ESR1 Mutation Status 23
   - 5.4.6 Patient Reported Outcomes 23
   - 5.5 Exploratory Endpoint Analysis 24
   - 5.6 Safety Analyses 25
   - 5.6.1 Extent of Exposure 25

5.6.2 Adverse Events 25
5.6.3 Laboratory Data 26
5.6.4 Exploratory Safety Analyses 26
5.7 Other Analyses 27
5.7.1 Summaries of Conduct of Study 27
5.7.2 Summaries of Treatment Group Comparability 27
5.7.3 Pharmacokinetic Analyses 27
5.7.4 Biomarker Analyses 28
5.7.5 Health Status Utility 28
5.7.6 Analyses of China Subpopulation 28
5.8 Interim Analyses 29
5.8.1 Interim Analyses for Secondary Efficacy Endpoint: OS 29
6. SUPPORTING DOCUMENTATION 30
7. REFERENCES 30

LIST OF TABLES

Table 1: Objectives and Corresponding Endpoints 8
Table 2: Recruitment intensity 15
Table 3: Assumptions and Characteristics for the Interim and Final Analyses of OS 29

LIST OF FIGURES

Figure 1: Study Schema 12

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
AE	adverse event
aBC	advance breast cancer
BPI-SF	Brief Pain Inventory - Short Form
CI	confidence interval
CSR	clinical study report
CCOD	clinical cut-off date
CBR	clinical benefit rate
ctDNA	circulating tumor DNA
DMC	Data Monitoring Committee
DOR	duration of response
eBC	early breast cancer
ER	estrogen receptor
ESR1	estrogen receptor 1
EORTC	European Organisation for Research and Treatment of Cancer
GHS	Global Health Status
GP5	General Population, Question 5
HR	hazard ratio
IA	interim analysis
ICE	intercurrent event
ICH	International Council on Harmonization
IMC	Internal Monitoring Committee
IRC	Independent Review Committee
ITT	intent to treat
IxRS	interactive voice/web-based response system
MBC	metastatic breast cancer
MDD	minimally detectable difference
MedDRA	Medical Dictionary for Regulatory Activities
NCI CTCAE v5.0	National Cancer Institute Common Terminology Criteria for Adverse Events, Version 5.0
NPT	non-protocol anti-cancer treatment
OS	overall survival
ORR	objective response rate
PFS	progression-free survival

PRO	patient-reported outcomes
PRO-CTCAE	Patient-Reported Outcomes Common Terminology Criteria for Adverse Events
PK	pharmacokinetic
QoL	Quality of Life
RECIST v1.1	Response Evaluation Criteria in Solid Tumors, Version 1.1
SAE	serious adverse events
SAP	Statistical Analysis Plan
SMQs	standardized MedDRA queries
TTD	time to deterioration
VAS	visual analog scale

1. INTRODUCTION

This Statistical Analysis Plan (SAP) provides details of the planned analyses and statistical methods for study WO42312, acelERA breast cancer. For detail background information on the study, refer to Protocol section 1.

1.1 OBJECTIVES AND ESTIMANDS

This study will evaluate the efficacy and safety of GDC-9545 30 mg PO QD administered on Days 1-28 of each 28-day cycle, beginning on Day 1 of Cycle 1 compared with physician's choice of endocrine monotherapy in patients with previously treated estrogen receptor (ER)-positive, HER2 negative locally advanced or metastatic breast cancer who have received one or two prior lines of systemic therapy in the locally advanced or metastatic setting. Endocrine monotherapy is defined as either fulvestrant or an aromatase inhibitor.

The primary comparison of interest is the hazard ratio of progression-free survival. The primary trial objective is to demonstrate superiority of the GDC-9545 arm over the physician's choice of endocrine monotherapy treatment arm.

Specific objectives and corresponding endpoints for the study are outlined below.

Table 1: Objectives and Corresponding Endpoints

Primary Efficacy Objective	Corresponding Endpoint
• The primary efficacy objective for this study is to evaluate the efficacy of GDC-9545 compared with physician's choice of endocrine monotherapy on the basis of Progression-free survival (PFS).	• **Progression-free survival (PFS)**, defined as the time from randomization to the first occurrence of disease progression or death from any cause (whichever occurs first), as determined by the investigator according to Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST v1.1)

Definition of Primary Estimand:

Following the estimand framework introduced in the ICH-E9 addendum (ICH 2020), the attributes of the estimand built around the primary endpoint are defined as follows:

- **Population**: Patients with previously treated estrogen receptor (ER)-positive, HER2-negative locally advanced or metastatic breast cancer who have received one or two prior lines of systemic therapy in the locally advanced or metastatic setting, as defined by the inclusion / exclusion criteria in study protocol (refer to Protocol section 4.1)
- **Variable**: Progression Free Survival (PFS), defined as time from randomization to the first occurrence of disease progression or death from any cause (whichever occurs first), as determined by the investigator according to RECIST v1.1
- **Treatment**: GDC-9545 or physician's choice of endocrine monotherapy. During the conduct of the study, participants may also receive concomitant medications as detailed in Protocol section 4.4.

- **Intercurrent Events (ICE) and Handling Strategy**:
  - Use of any non-protocol anti-cancer treatment (NPT) prior to disease progression as detailed in Protocol section 4.4.3.
  - Discontinuation of study treatment prior to disease progression

  Handling of ICE: Following treatment policy, all the above ICE will be ignored and tumor assessment data collected after the ICE will be included in the primary PFS analysis.
- **Population-level summary**: Hazard ratio

Secondary Efficacy Objectives #	Corresponding Endpoints
• The secondary efficacy objective for this study is to evaluate the efficacy of GDC-9545 compared with physician's choice of endocrine monotherapy on the basis of:	• **Overall survival (OS)**, defined as the time from randomization to death from any cause • **Objective Response Rate (ORR)**, defined as the proportion of patients with a complete response (CR) or PR on two consecutive occasions ≥ 4 weeks apart, as determined by the investigator according to RECIST v1.1 • **Duration of response (DOR)**, defined as the time from the first occurrence of a documented objective response to disease progression or death from any cause (whichever occurs first), as determined by the investigator according to RECIST v1.1 • **Clinical Benefit Rate (CBR)**, defined as the proportion of patients with stable disease for ≥ 24 weeks or a CR or PR, as determined by the investigator according to RECIST v1.1 • **Investigator-assessed PFS, in subgroups categorized by baseline *ESR1* mutation status** • **Time to deterioration (TTD) in pain severity after randomization**, defined as the time from randomization to the first documentation of a ≥ 2-point increase from baseline on the "worst pain" item score from the Brief Pain Inventory–Short Form (BPI-SF)

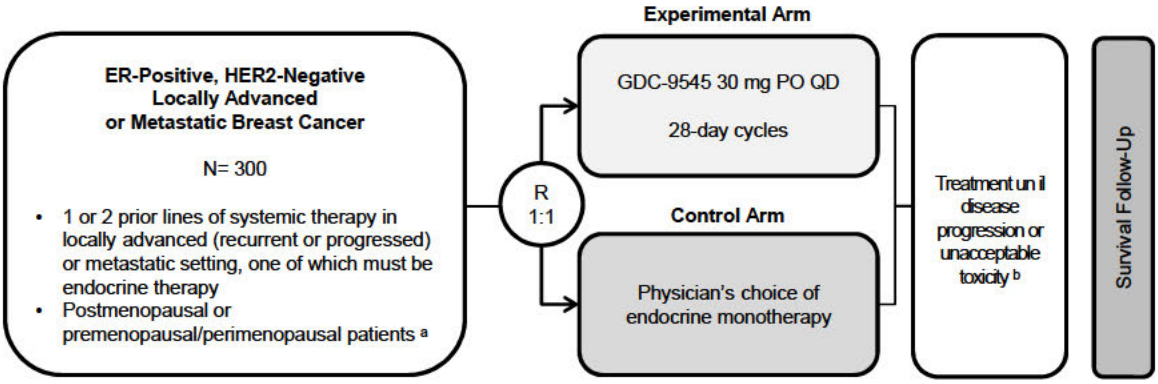
	• **TTD in pain presence and interference after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point increase from baseline in the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 linearly transformed pain scale score • **TTD in Physical Functioning (PF) after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point decrease from baseline in the EORTC QLQ-C30 linearly transformed PF scale score • **TTD in Role Functioning (RF) after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point decrease from baseline in the EORTC QLQ-C30 linearly transformed RF scale score • **TTD in global health status (GHS) and quality of life (QoL) after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point decrease from baseline in the EORTC QLQ-C30 linearly transformed GHS/QoL scale score
Safety Objectives	**Corresponding Endpoints**
• The safety objective for this study is to evaluate the safety of GDC-9545 compared with physician's choice of endocrine monotherapy on the basis of	• Incidence and severity of adverse events, with severity determined according to National Cancer Institute Common Terminology Criteria for Adverse Events, Version 5.0 (NCI CTCAE v5.0) • Change from baseline in targeted vital signs • Change from baseline in targeted clinical laboratory test results
Exploratory Safety Objectives	**Corresponding Endpoints**
• To evaluate the tolerability of GDC-9545 compared with physician's choice of endocrine monotherapy	• Presence, frequency of occurrence, severity, and/or degree of interference with daily function of selected symptomatic treatment toxicities, as assessed through use of the NCI PRO-CTCAE • Overall tolerability (i.e., bother experienced due to side effects of treatment), as assessed through the GP5 item from the FACT-G questionnaire

	• Change from baseline in symptomatic treatment toxicities and overall tolerability/side-effect burden, as assessed through use of the PRO-CTCAE and the GP5 item, respectively
Pharmacokinetic Objective	**Corresponding Endpoints**
• The Pharmacokinetic (PK) objective for this study is to characterize the GDC-9545 PK profile (± LHRH agonist) on the basis of	• Plasma concentration of GDC-9545 at specified time points.
Exploratory Biomarker Objective	**Corresponding Endpoints**
• To identify and/or evaluate biomarkers that are predictive of response to GDC-9545 (i.e., predictive biomarkers), are early surrogates of efficacy, are associated with progression to a more severe disease state (i.e., prognostic biomarkers), are associated with acquired resistance to GDC-9545, are associated with susceptibility to developing adverse events or can lead to improved adverse event monitoring or investigation (i.e., safety biomarkers), can provide evidence of GDC-9545 activity (i.e., PD biomarkers), or can increase the knowledge and understanding of disease biology and drug safety.	• Relationship between biomarkers in blood, plasma, and tumor tissue and efficacy, safety, PK, disease biology, or other biomarker endpoints
Exploratory Health Status Utility Objective	**Corresponding Endpoints**
• To evaluate health status utility scores of patients treated with GDC-9545 compared with physician's choice of endocrine monotherapy	• Health utility and visual analog scale score of the EQ-5D-5L for pharmacoeconomic modeling at specified time points.
[#] *Refer to Section 5.4 for description of estimand for the non-PRO secondary efficacy endpoints of OS, ORR, DOR, CBR and PFS by ESR1 mutation.*	

1.2 STUDY DESIGN

The study schema is shown in **Figure 1**.

Figure 1: Study Schema



ER =estrogen receptor; HER2 =human epidermal growth factor receptor 2; LHRH =luteinizing hormone-releasing hormone; PO =orally; QD=once a day; R =randomization.

[a] *Premenopausal/perimenopausal patients and male patients in both arms will receive an LHRH agonist. To minimize the potential of exposure to the medication decreasing to sub-therapeutic levels towards the end of the treatment cycle, monthly injections of LHRH agonist are preferred and to be synchronized with Day 1 of each 28-day cycle (or according to clinical practice for the selected agent).*

[b] *An exception will be made for patients who have developed isolated brain metastases that are treatable with radiation, provided the patients have experienced a partial response, complete response, or stable disease for ≥ 24 weeks (see Protocol Section 4.4.3).*

This study will initially enroll approximately 300 patients across all sites in a global enrollment phase. After completion of the global enrollment phase, additional patients may be enrolled in an extended China enrollment phase at sites in mainland China and/or Taiwan. The global population will include all patients enrolled during the global enrollment phase (including patients enrolled in mainland China and/or Taiwan during that phase), and the China subpopulation will include all patients enrolled in mainland China and/or Taiwan (i.e., during both the global enrollment phase and the extended China enrollment phase). *Separate analyses will be performed for the global population and the China subpopulation (see section 5.7.6 for China subpopulation analyses).*

1.2.1 Treatment Assignment and Blinding

This is a randomized, open-label study. After initial written informed consent has been obtained, all screening procedures and assessments have been completed, and eligibility has been established for a patient, the study site will obtain the patient's identification number and treatment assignment from an interactive voice or web-based response system (IxRS).

Patients will be randomly assigned to one of two treatment arms to receive either GDC-9545 or physician's choice of endocrine monotherapy. Randomization will occur in a 1:1 ratio through use of a permuted-block randomization method to ensure a balanced assignment to each treatment arm. The randomization method implemented in the China extension cohort will be the same as that implemented in the global population. Randomization will be stratified according to the following factors:

- Site of disease (visceral vs. non-visceral), locally assessed
  - *Visceral is defined as any lung and/or liver involvement.*
  - *Non-visceral is defined as absence of any lung and/or liver involvement.*
- Prior treatment with a CDK4/6 inhibitor (*yes vs. no*)
- Prior treatment with Fulvestrant (*yes vs. no*)

The number of premenopausal/perimenopausal patients and male patients enrolled will be limited to approximately 20% of the study population. The cap of 20% has been chosen to ensure that the mix of patients in this study approximates global clinical practice patterns. Crossover between treatment arms will not be permitted in the study.

1.2.2 Independent Review Facility

All radiological data (e.g., computed tomography [CT] scan, magnetic resonance imaging [MRI], bone scan), photographs of skin lesions, and any additional clinical information required will be sent to a blinded, independent, core imaging laboratory (contracted by the Sponsor) to facilitate a retrospective evaluation of disease response and progression by an IRC. Details about IRC membership and procedures (e.g., tumor assessments) are outlined in a separate IRC Charter.

1.2.3 Data Monitoring

An Independent Monitoring Committee (IMC) will provide additional safety oversight throughout the study. The IMC will include representatives from Clinical Science, Safety Science and Biostatistics. In addition to the ongoing assessment of the incidence, nature, and severity of adverse events, serious adverse events, deaths, and vital sign and laboratory abnormalities performed by the Medical Monitor, the IMC will periodically review cumulative safety data during the study. No interim efficacy analyses are planned for WO42312 study. The IMC may review efficacy data if safety concerns necessitate risk-benefit assessments. Specific operational details such as the committee's composition, frequency and timing of meetings, and members' roles and responsibilities are detailed in a separate IMC Charter.

2. STATISTICAL HYPOTHESES

The primary analysis will be a comparison of the Investigator-assessed PFS between the two treatment arms using a stratified log-rank test at an overall 0.05 significance level (two-sided).

The statistical hypothesis of this study is as follows:

- H_0: *PFS (Arm A) = PFS (Arm B)*
- H_1: *PFS (Arm A) ≠ PFS (Arm B)*

PFS (Arm A) represents the progression free survival function in the GDC-9545 arm, and PFS (Arm B) represents the progression free survival function in the physician's choice of endocrine monotherapy arm.

The null and alternative hypotheses will be tested at a two-sided 0.05 significance level. The primary trial objective is to demonstrate superiority of the experimental over the control treatment.

2.1 MULTIPLICITY ADJUSTMENT

The hypothesis testing pertaining to investigator assessed PFS at randomization will take place using a two-sided 5% significance level. A hierarchical approach for testing PFS followed by OS will be used for multiplicity adjustments. Thus, the key secondary endpoint of Overall Survival (OS) would only be tested if the primary endpoint, PFS, were statistically significant at the primary analysis.

3. SAMPLE SIZE DETERMINATION

The primary analysis of PFS will be conducted when approximately 166 PFS events from both arms are observed. The study is designed with 80% power at the 5% (two-sided) level of significance to detect hazard ratio of 0.647, which corresponds to an improvement in median PFS from 5.5 months to approximately 8.5 months. The largest hazard ratio (smallest effect size, minimal detectable difference, MDD) determined to be statistically significant at the 5% level will be approximately 0.738 (which corresponds to median improvement in PFS from 5.5 to 7.5 months).

Approximately 300 patients will be enrolled and randomized in a 1:1 ratio to receive either GDC-9545 (experimental arm) or physician's choice of endocrine monotherapy (control arm). Given assumptions on recruitment intensity (see Table 2 below); the enrollment is projected to finish approximately 15 months after randomization of the first patient. For both the GDC-9545 and control arms, an annual loss to follow-up rate of 10% is assumed. On that basis, it is projected that the primary PFS analysis will occur approximately 18 months after the first patient is enrolled.

Table 2: Recruitment intensity

	Recruitment month from ramp-up to peak														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Participants per month	1	4	9	14	19	23	26	29	28	29	28	29	28	29	4

Sample Size for the China Subpopulation

This study will initially enroll approximately 300 participants across all sites during the global enrollment phase. After completion of the global enrollment phase, additional participants may be enrolled in an extended China enrollment phase at sites in mainland China and/or Taiwan.

4. ANALYSIS SETS

The following analysis sets are defined:

Population	Definition
Full Analysis Set	All participants assigned to treatment groups as randomized by the IxRS.
Safety-evaluable	All participants randomly assigned to study treatment and who received at least one dose of study treatment.
ctDNA-evaluable	All participants assigned to treatment groups as randomized by the IxRS with evaluable plasma ctDNA at baseline.
PRO-evaluable	All randomized participants who have a baseline and at least one post-baseline PRO assessment.
PK-evaluable	All randomized participants in the GDC-9545 arm who have at least one evaluable GDC-9545 plasma concentration.

ctDNA = circulating tumor DNA, PRO = patient reported outcome, PK = pharmacokinetic.

5. STATISTICAL ANALYSES

This section is a summary of the planned statistical analyses of the most important endpoints, including primary and key secondary endpoints. It also describes the sensitivity analysis related to the primary study estimand and any pre-specified subgroup analyses.

The analyses described in this SAP will supersede those specified in Protocol Version-3, 09-July-21 or thereafter.

Separate analyses will be performed for the global population and the China subpopulation. The global population will include all patients enrolled during the global enrollment phase (including patients enrolled in mainland China and/or Taiwan during that phase), and the China subpopulation will include all patients enrolled in mainland China and/or Taiwan (i.e., during both the global enrollment phase and the extended China enrollment phase).

5.1 GENERAL CONSIDERATION

All efficacy analyses will be performed on the full analysis set population, unless otherwise specified. All safety analyses will be performed in the safety-evaluable population, unless otherwise specified.

Analyses of demographics and other baseline information will be based on full analysis set population. The baseline value of any variable will be defined as the last available data point prior to the first administration of study drug.

5.2 PARTICIPANT DISPOSITION

The total number of participants screened and the reasons for screen failure will be summarized. A detailed study disposition of participants including the number/percentage of participants who have completed the study vs. number/percentage of participants who have prematurely withdrawn from the study, as well as the primary reasons for withdrawal will be summarized overall by treatment arm. Additionally, a summary table of survival follow-up period for participants remaining in the study until the time of final overall survival will be provided.

5.3 PRIMARY ENDPOINT ANALYSIS

The primary comparison of interest is the hazard ratio of progression-free survival (PFS). The primary trial objective is to demonstrate superiority of the GDC-9545 arm over the physician's choice of endocrine monotherapy treatment arm.

5.3.1 Definition of Primary Estimand

Following the estimand framework introduced in the ICH-E9 addendum (ICH 2020), the attributes of the estimand built around the primary endpoint are defined as follows:

- **Population**: Patients with previously treated estrogen receptor (ER)-positive, HER2-negative locally advanced or metastatic breast cancer who have received one or two prior lines of systemic therapy in the locally advanced or metastatic setting, as defined by the inclusion / exclusion criteria in study protocol (refer to Protocol section 4.1)
- **Variable**: Progression Free Survival (PFS), defined as time after randomization to the first occurrence of disease progression or death from any cause (whichever occurs first), as determined by the investigator according to RECIST v1.1
- **Treatment**: GDC-9545 or physician's choice of endocrine monotherapy. During the conduct of the study, participants may also receive concomitant medications as detailed in Protocol section 4.4.

- **Intercurrent events and Handling Strategy**:
    - Use of any non-protocol anti-cancer treatment (NPT) prior to disease progression as detailed in Protocol section 4.4.3. → Following treatment policy the ICE will be ignored and tumor assessment data collected after the ICE will be included in the PFS analysis
    - Discontinuation of study treatment prior to disease progression → Following treatment policy the ICE will be ignored and tumor assessment data collected after the ICE will be included in the PFS analysis.
- **Population-level summary:** Hazard ratio

5.3.2 Main Analytical Approach for Primary Estimand

If participants have any intercurrent event(s), then the strategies defined as defined in section 5.3.1 to handle the intercurrent events will be implemented. Otherwise, data for participants without the occurrence of disease progression or death as of the clinical cutoff date (CCOD) will be censored at the time of the last tumor assessment prior to the CCOD (or at the time of randomization plus 1 day if no tumor assessment was performed after the baseline visit). PFS will be compared between treatment arms using the stratified log-rank test. The hazard ratio will be estimated using a stratified Cox proportional hazards model. The 95% CI for the hazard ratio will be provided. The stratification factors used will be the same as the randomization stratification factors (as entered in IxRS, see Protocol section 4.2.1). Results from an unstratified analysis (see section 5.3.4) will also be provided. For each treatment arm, Kaplan-Meier methodology will be used to estimate the median PFS, and the Brookmeyer-Crowley method will be used to construct the 95% CI for the median PFS. Kaplan-Meier curves will also be produced.

5.3.3 Handling of Missing Data

The impact of missing scheduled tumor assessments on PFS will be assessed depending on the number of participants who missed assessments scheduled immediately prior to the date of disease progression per RECIST v1.1 or the data cutoff. If > 5% of participants missed two or more assessments scheduled immediately prior to the date of disease progression per RECIST v1.1 or the data cutoff in any treatment arm then the impact of missing scheduled tumor assessments on PFS will be assessed by performing a sensitivity analysis based on the interval censoring method.

For each participant in both treatment arms, the left and the right boundaries of the intervals will be derived based on the following rules:

Situations	Left Boundary	Right Boundary
Participants who had disease progression prior to death	The date of the last assessment that showed a progression-free*1 status	The date of the first assessment that showed disease progression
Participants who died without disease progression	The date of the last assessment that showed a progression-free*1 status	Death date
Participants who did not die nor had disease progression*2	The date of the last assessment that showed a progression-free*1 status	Not applicable (Missing)

*1 *For participants who did not have any post-baseline assessment with progression-free status, the left boundary is the date of randomization.*

*2 *Including participant withdrawing consent prior to disease progression*

The PFS survival curves will be estimated using the nonparametric maximum likelihood estimate (NPMLE, Turnbull 1974) for each treatment arm. The median PFS of each treatment arm will be reported and its 95% confidence interval will be constructed based on the Brookmeyer-Crowley method (Brookmeyer and Crowley 1982).

Hypothesis testing will be performed based on the stratified log-rank test proposed by Sun (Sun 1996) to compare the PFS in the treatment arms. The treatment effect will be estimated using a stratified proportional hazard regression model (Finkelstein 1986) with a parametric assumption of piecewise exponential distribution for the baseline hazard function (Friedman 1982, Royston and Parmar 2002). Results from an unstratified analysis will also be provided.

The analyses listed above will be performed using the procedures PROC ICLIEFTEST and PROC ICPHREG in SAS version 9.4 or a later version.

5.3.4 Sensitivity Analyses for Primary Endpoint

Following sensitivity analysis to the primary estimand of the primary PFS after randomization will be performed:

i. **PFS assessed by investigator based on unstratified analysis:** To assess the impact of stratification (as entered in IxRS, see Protocol section 4.2.1) the main analysis described in section 5.3.2 will be repeated without stratification factors.

ii. **PFS assessed by Independent Review committee (IRC):** To assess the concordance of PFS assessment by investigators PFS analysis, as mentioned in section 5.3.2, will be repeated based on IRC assessment. The IRC cannot evaluate patients if the baseline scans were missing hence in such cases data will be censored at randomization plus 1 day.

iii. **PFS assessed by investigator based on interval censoring approach:** The impact of missing scheduled tumor assessments on PFS will be assessed by performing sensitivity analysis based on the interval censoring method as described in section 5.3.3.

5.3.5 Supplementary Analyses for Primary Endpoint

Two supplementary analysis as outlined below based on different strategy of handling ICE are planned for PFS. Note that the attributes of population, variables and population level summary will remain the same as the primary estimand. The analysis for the supplementary estimands will be conducted only on the primary endpoint investigator assessed PFS.

1. **PFS assessed by investigator based on hypothetical strategy for use of any NPT prior to disease progression:** → To assess the impact of use of any NPT prior to disease progression, the primary analysis of investigator assessed PFS will be repeated with the ICE of use of any NPT handled using a hypothetical strategy. According to this strategy, patients who start NPT prior to disease progression will be censored at the time of the last disease status assessment before the initiation of NPT. If patients start any NPT before starting study treatment then the data of those patients will be censored at the time of randomization plus day 1. Approaches to handle other ICE (Discontinuation of study treatment prior to disease progression) and analysis method will be the same as mentioned in section 5.3.2.

2. **PFS assessed by investigator based on composite strategy for use of any NPT prior to disease progression:** → In addition to above estimand, here supplementary estimand for PFS will be estimate by following composite strategy. According to composite strategy, use of any NPT prior to disease progression will be considered as PFS event (progressed) at the time of initiation of NPT. If patients start any NPT before starting study treatment then the data of those patients will be censored at the time of randomization plus day 1. Approaches to handle other ICE (Discontinuation of study treatment prior to disease progression) and analysis method will be the same as mentioned in section 5.3.2.

5.3.5.1 Subgroup Analyses for Primary Endpoint

The generalizability of PFS after randomization results when comparing GDC-9545 compared to physician's choice of endocrine monotherapy will be investigated as a part of exploratory analysis by estimating the treatment effect in subgroups based on factors such as but not limited to:

- Age (<65, ≥65 years)

- Race (American Indian/Alaska native, Asian, Black/African American, Native Hawaiian or Other Pacific Islander, White, Unknown),
- ECOG Performance at baseline (0, 1)
- Menopausal status at baseline (Pre/Peri- Menopausal, Post-Menopausal)
- Hormonal Status (ER+PR+, ER+PR-)
- PFS by ESR1 mutation at baseline (mutation detected, no mutation detected)
- Site of disease assessed locally (visceral, non-visceral)
- Prior treatment with CDK4/6 inhibitor (yes, no),
- Prior treatment with Fulvestrant (yes/no)
- Prior treatment with Aromatase Inhibitor (yes, no)
- Prior treatment with Tamoxifen (yes, no)
- Disease status (Measurable, Non-measurable)
- Number of Organ sites (1, 2, >3)
- Central Nervous System (CNS) involvement (yes, no)
- Bone only involvement (yes, no)
- Previous lines of therapy in the aBC/MBC setting (1, 2)
- Prior Chemo in eBC (yes, no)
- Prior Chemo in aBC/MBC (yes, no)
- Disease Free Interval (≤12 months, >12 months)
- Endocrine Resistance (Primary, Secondary)

Further prognostic factors may be considered for subgroup analyses as deemed appropriate. Un-stratified analysis results will be presented for subgroup analyses due to the potentially limited number of patients in each subgroup. Summaries of PFS after randomization by above subgroups will be provided in forest plots including estimates for HR and 95% CIs from unstratified Cox proportional hazard models.

5.4 SECONDARY ENDPOINTS ANALYSES

The secondary efficacy endpoints for this study Overall Survival (OS), Objective response rate (ORR), Duration of response (DOR), Clinical Benefit Rate (CBR) and PFS in subgroups characterized by baseline ESR1 mutation status and the PROs Time to deterioration (TTD) in pain severity, pain presence/interference, physical functioning, role functioning, global health status and quality of life, are defined in Protocol Section 2.1.2. The subset of corresponding secondary efficacy endpoints which are not derived from patient-reported outcome (PRO) measures, are expressed using the estimand framework. Analysis method for all the secondary endpoints are described in following sub-section.

5.4.1 Overall Survival

The secondary comparison of interest is the hazard ratio of Overall Survival (OS). The secondary trial objective is to demonstrate superiority of the GDC-9545 arm over the physician's choice of endocrine monotherapy treatment arm.

Following the estimand framework (ICH 2020), the attributes of the estimand for the secondary endpoint are defined as follows:

- **Population**: As defined for the primary estimand in section 5.3.1
- **Variable**: Overall Survival (OS), defined as time from randomization to death from any cause.
- **Treatment**: As defined for the primary estimand in section 5.3.1
- **Intercurrent events and Handling strategy**
    - Use of any non-protocol anti-cancer treatment (NPT) prior to disease progression as detailed in Protocol section 4.4.3. → Following treatment policy the ICE will be ignored and observations collected after the ICE will be used.
    - Discontinuation of study treatment prior to disease progression. → Following treatment policy the ICE will be ignored and observations collected after the ICE will be used.
- **Population-level summary:** Hazard ratio for OS.

As per protocol, crossover between treatment arms will not be permitted in the study hence crossover not anticipated as an ICE. If participants have any intercurrent events, then the strategies defined below to handle the intercurrent events will be implemented. Otherwise, data for patients who are alive at the time of the analysis data cutoff will be censored at the last date they were known to be alive. Data from patients without postbaseline information will be censored at the date of randomization plus 1 day.

OS will be hierarchically tested if the primary endpoint, PFS, is statistically significant at the primary PFS analysis. The OS analyses will be conducted at the time of the primary analysis of PFS, at least 12 months after the primary analysis, and at the end of the study, which is expected to occur at least 25 months after the last patient is enrolled in the global study, refer to section 5.8 for details on interim analysis for OS.

5.4.2 Overall Response Rate

The analysis population for ORR will be full analysis set. Patients not meeting the criteria for ORR, including patients without any postbaseline tumor assessment, will be considered non-responders.

Following the estimand framework (ICH 2020), the attributes of the estimand for the secondary endpoint are defined as follows:

- **Population**: As defined for the primary estimand in section 5.3.1.
- **Variable**: ORR, defined as the proportion of patients with a CR or PR on two consecutive occasions $\geqslant 4$ weeks apart, as determined by the investigator according to RECIST v1.1.
- **Treatment**: As defined for the primary estimand in section 5.3.1

- **Intercurrent events and Handling strategy**:
    - As described for the primary estimand in section 5.3.1.
    - **ICE Handling Strategy:→** Following treatment policy, all the ICE's will be ignored and tumor assessment data collected after the ICE will be included in the ORR analysis.
- **Population-level summary:** Difference in proportion.

An estimate of ORR and 95% CI will be calculated using the Clopper-Pearson method for each treatment arm. ORR will be compared between treatment arms using the stratified Cochran-Mantel-Haenszel test. The stratification factors to be used will be the same as those used for the analysis of the primary endpoint. The difference in ORR between treatment arms will be calculated, and 95% CI will be calculated using the Newcombe methodology (Newcombe 1998).

5.4.3 CBR

Following the estimand framework (ICH 2020), the attributes of the estimand for the secondary endpoint are defined as follows:

- **Population**: As defined for the primary estimand in section 5.3.1
- **Variable**: CBR, defined as the proportion of patients with stable disease for ⩾24 weeks or a CR or PR, as determined by the investigator according to RECIST v1.1.
- **Treatment**: As defined for the primary estimand in section 5.3.1
- **Intercurrent events and Handling strategy**:
    - As described for the primary estimand in section 5.3.1
    - **ICE Handling Strategy:→** Following treatment policy, all the ICE's will be ignored and tumor assessment data collected after the ICE will be included in the CBR analysis.
- **Population-level summary:** Difference in proportion.

CBR analysis including handling of ICE will follow the same methods as those used for ORR.

5.4.4 DOR

Analysis of DOR will include only patients who had an objective response.

Following the estimand framework (ICH 2020), the attributes of the estimand for the secondary endpoint are defined as follows:

- **Population**: All randomized patients who had an objective response.
- **Variable**: DOR, defined as the time from the first occurrence of a documented objective response to disease progression or death from any cause (whichever occurs first), as determined by the investigator according to RECIST v1.1.

- **Treatment**: As defined for the primary estimand in section 5.3.1
- **Intercurrent events and Handling strategy**
  - As described for the primary estimand in section 5.3.1.
  - **ICE Handling Strategy:→** Following treatment policy, all the ICE's will be ignored and tumor assessment data collected after the ICE will be included in the DOR analysis.
- **Population-level summary:** Hazard ratio.

Handling of ICE will follow the same methods as those used for PFS. Data for patients who have not progressed and who have not died at the time of analysis will be censored at date of the last tumor assessment. The Kaplan-Meier approach will be used to estimate the median DOR and to construct survival curves for each treatment arm for a visual description of the difference among arms. The Brookmeyer-Crowley methodology will be used to construct the 95% CI for the median (Brookmeyer and Crowley 1982).

5.4.5 PFS by ESR1 Mutation Status

The analysis population for PFS by ESR1 mutation will be all randomized patients with evaluable plasma ctDNA at baseline.

Following the estimand framework introduced in the ICH-E9 addendum (ICH 2020), the attributes of the estimand for the secondary endpoint are defined as follows:

- **Population**: All randomized patients with evaluable plasma ctDNA at baseline.
- **Variable**: Investigator-assessed PFS in subgroups categorized by ESR1 mutation status determined at baseline from plasma ctDNA.
- **Treatment**: As defined for the primary estimand in section 5.3.1
- **Intercurrent events and Handling strategy**
  - As described for the primary estimand in section 5.3.1.
  - **ICE Handling Strategy:→** Following treatment policy, the above ICE will be ignored and tumor assessment data collected after the ICE will be included in the analysis.
- **Population-level summary:** Hazard ratio.

PFS by ESR1 mutation status (mutation detected, no mutation detected) will be compared between treatment arms using the stratified log-rank test. Analytical approach and handling of ICE will follow the same methods as those used for PFS in section 5.3.2.

5.4.6 Patient Reported Outcomes

The secondary efficacy objective of the study is based on the following Patient Reported Outcome (PRO) endpoints:

- **Time to deterioration (TTD) in pain severity after randomization**, defined as the time from randomization to the first documentation of a ≥ 2-point increase from baseline on the "worst pain" item score from the Brief Pain Inventory–Short Form (BPI-SF)
- **TTD in pain presence and interference after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point increase from baseline in the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 linearly transformed pain scale score
- **TTD in Physical Functioning (PF) after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point decrease from baseline in the EORTC QLQ-C30 linearly transformed PF scale score
- **TTD in Role Functioning (RF) after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point decrease from baseline in the EORTC QLQ-C30 linearly transformed RF scale score
- **TTD in global health status (GHS) and quality of life (QoL) after randomization**, defined as the time from randomization to the first documentation of a ≥ 10-point decrease from baseline in the EORTC QLQ-C30 linearly transformed GHS/QoL scale score

All randomized patients will be used for completion analyses and time-to-deterioration (TTD) analyses. Analyses will be performed based on the treatment arm assigned at randomization. The comparison of interest is the difference in time to deterioration between treatment arms expressed by the hazard ratio. The Kaplan-Meier analysis methods similar to those described for PFS analysis will be applied for TTD analysis. Patients who do not have an observed deterioration prior to discontinuation from study treatment or at the time of the clinical cut-off date (CCOD), will be censored at the last available assessment date prior to or at the time of discontinuation from study treatment or CCOD, whatever is earlier. Data for patients without a post baseline assessment will be censored at the time of randomization plus 1 day.

5.5 EXPLORATORY ENDPOINT ANALYSIS

The exploratory efficacy objective for this study is to evaluate the efficacy of GDC-9545 compared with physician's choice of endocrine monotherapy on the basis of the following endpoints:

- Mean scores and mean change from baseline in functional scores (physical, role, cognitive, emotional, and social), GHS/QoL, and disease- and treatment-related symptom scores,

This exploratory analysis will be based on PRO-evaluable population, which will include all randomized patients who have a baseline and at least one post-baseline assessment. Summary statistics (mean, standard deviation, median, 25th and 75th percentiles, and

range) of scores will be reported for the "worst pain" item of the BPI-SF, as well as all linear transformed scores for scales (symptoms, functional domains, and GHS/QoL) of the QLQ-C30 and QLQ-BR23 questionnaires for each assessment time point. The mean change of the linear transformed scores from baseline (and 95% CI using the normal approximation) will also be analyzed for each treatment arm. Line charts depicting the mean and mean changes from the baseline assessment (and 95% CIs) of items and scales over time will be provided for each treatment arm.

In the event of incomplete data for all questionnaire scales, if more than 50% of the constituent items are completed, a prorated score will be computed consistent with the scoring manuals and validation papers (see protocol). For scales with less than 50% of the items completed, the scale will be considered as missing in accordance with the EORTC scoring manual guidelines.

PRO completion, compliance rates, and reasons for missing data will be summarized at each time point by treatment arm for each measure in ITT population. The compliance rate will be based on the total number of patients expected to complete the questionnaire at a particular time point.

5.6 SAFETY ANALYSES

The safety analysis population will consist of all patients who received at least one dose of study drug (GDC-9545 or physician's choice of endocrine monotherapy), with patients grouped according to treatment received.

5.6.1 Extent of Exposure

Drug exposure will be summarized using descriptive statistics, including duration of treatment, cumulative dose, and dose intensity.

5.6.2 Adverse Events

Adverse events will be mapped to MedDRA thesaurus terms and appropriate MedDRA hierarchy. Adverse event severity will be graded according to NCI CTCAE v5.0. Multiple occurrences of the same event will be counted once at the maximum severity.

The frequency, nature, and severity of treatment-emergent adverse events, adverse events leading to death, adverse events leading to study drug discontinuation, serious adverse events, and adverse events of special interest will be summarized by treatment arm. All deaths will be summarized. Treatment-emergent adverse events are defined as adverse events that occur after the first dose of study treatment.

Relevant vital signs will be presented using summary statistics by treatment received. Selected laboratory and vital sign data will be summarized by treatment received, and

grade compared with baseline and measurements outside of the normal range will be identified.

5.6.3 Laboratory Data

Relevant laboratory values will be summarized by treatment arm over time, with NCI CTCAE v5.0 Grade 3 and Grade 4 values identified, where appropriate. Summary tables of clinically relevant shifts in NCI CTCAE v5.0 grades (Grades 1, 2) at baseline to the worst post-baseline (Grade ≥3) value will be presented.

A Hy's Law analysis will be provided: the finding of an elevated ALT or AST (> 3x upper limit of normal [ULN]) in combination with either an elevated total bilirubin (> 2x ULN) or clinical jaundice in the absence of cholestasis or other causes of hyperbilirubinemia is considered to be an indicator of severe liver injury (as defined by Hy's Law).

5.6.4 Exploratory Safety Analyses

The exploratory safety objective for this study is to evaluate the tolerability of GDC-9545 compared with physician's choice of endocrine monotherapy based on the following endpoints:

- Presence, frequency of occurrence, severity, and/or degree of interference with daily function of selected symptomatic treatment toxicities, as assessed through use of the NCI Patient-Reported Outcomes Common Terminology Criteria for Adverse Events (PRO-CTCAE).

- Overall tolerability (i.e., bother experienced due to side effects of treatment), as assessed through the General Population, Question 5 (GP5) item from the Functional Assessment of Cancer Therapy-General (FACT-G) questionnaire.

- Change from baseline in symptomatic treatment toxicities and overall tolerability/side-effect burden, as assessed through use of the PRO-CTCAE and the GP5 item, respectively.

In this study PRO-CTCAE is composed of seven adverse events selected from the NCI item bank; {diarrhea, nausea, vomiting, decreased appetite, fatigue, mouth sores, and rash}. The population to be used for the analysis of PRO-CTCAE is the safety evaluable population.

PRO-CTCAE analyses will be descriptive, with a focus on characterizing the pattern of symptomatic treatment toxicities over the course of the study. The numeric scores indicate an ordinal, rather than continuous outcome; therefore, analysis will focus on frequency counts and percentages. The number and percentage of patients reporting each symptom and the change from baseline by category (frequency of occurrence, severity, or

interference) will be summarized at each assessment time point by treatment arm. The worst post-baseline score will be summarized per treatment group, for each individual attribute and AE, and shift tables will be provided per treatment group. For items rated on a 5-point Likert scale, the maximum post-baseline score and change from baseline will be summarized by treatment arm.

Results from these exploratory analyses will be presented separately from the safety analyses. PRO-CTCAE data will be analyzed at the item level. Graphical representation of PRO-CTCAE data over time will also be provided. PRO-CTCAE data will be summarized over time. These analyses will also apply to the GP5 overall treatment burden item. A descriptive analysis of absolute scores and the proportion of patients selecting each response option at each assessment time point by treatment arm will be reported for the FACT-G single-item GP5. The proportion of missing data at each assessment time point will also be summarized to facilitate interpretation of data.

5.7 OTHER ANALYSES

5.7.1 Summaries of Conduct of Study

Study enrollment, duration, study drug discontinuation, and study discontinuation, as well as reasons for study drug discontinuation and study discontinuation, will be listed and summarized overall and by treatment arm. Major protocol deviations, including major deviations with regard to inclusion and exclusion criteria, will also be listed and summarized overall and by treatment arm.

5.7.2 Summaries of Treatment Group Comparability

The evaluation of treatment group comparability between the treatment arms will include summaries of demographic and baseline characteristics, including stratification factors and patient treatment history. Continuous variables will be summarized using means, standard deviations, medians and ranges. Categorical variables will be summarized by counts and proportions.

5.7.3 Pharmacokinetic Analyses

The PK objective for this study is to characterize the GDC-9545 PK profile (±LHRH agonist) based on Plasma concentration of GDC-9545 at specified time points. This analysis will be based on PK analysis population, which will consist of all randomized participants in the GDC-9545 arm who have at least one evaluable GDC-9545 plasma concentration. PK parameters (e.g., area under the concentration-time curve, time to maximum concentration, maximum concentration observed, and half-life if appropriate) will be estimated from an intensive PK sample-subset patient population. Summary statistics

(mean, standard deviation, coefficient of variation, median, minimum, and maximum) will be presented for PK data. Plasma GDC-9545 concentration versus time data will be tabulated and plotted.

GDC-9545 PK data may be pooled and analyzed using a population PK analysis approach as appropriate and reported in a standalone report.

5.7.4 Biomarker Analyses

The exploratory biomarker objective is based on the relationship between biomarkers in blood, plasma, and tumor tissue and efficacy, safety, PK, disease biology, or other biomarker endpoints. No formal statistical analysis of exploratory biomarkers will be performed. Data may be analyzed in the context of this study and in aggregate with data from other studies. Results may be presented in a separate report.

5.7.5 Health Status Utility

The exploratory health status utility objective for this study is to evaluate health status utility scores of patients treated with GDC-9545 compared with physician's choice of endocrine monotherapy on the basis of the EQ-5D-5L. These data will be used in pharmacoeconomic models and reported separately from the CSR.

5.7.6 Analyses of China Subpopulation

A separate analysis will be performed for the China subpopulation, where data from all participants enrolled at mainland China and/or Taiwan (i.e., during both the global enrollment phase and the extended China enrollment phase) will be combined and summarized. Results from these analyses will be summarized in a separate Clinical Study Report.

The efficacy objective of the China subpopulation analyses is to evaluate whether the efficacy of GDC-9545 compared with physician's choice of endocrine monotherapy in the China subpopulation (enrolled during both the global enrollment phase and the extended China enrollment phase) is consistent with the efficacy observed in the global population enrolled during the global enrollment phase. Therefore, no formal hypothesis testing will be performed for China subpopulation.

The China subpopulation analyses will be conducted at the same time as global population. If the PFS data in the China subpopulation is not mature at the time of primary analysis, an additional PFS analysis in the China subpopulation may be conducted. Data for the

China subpopulation will be analyzed using the same statistical methods as described in Section 5.1–5.7 when data allow.

5.8 INTERIM ANALYSES

Currently there are no interim analysis planned for primary endpoint, PFS.

5.8.1 Interim Analyses for Secondary Efficacy Endpoint: OS

The study will incorporate three OS analyses (two interim analyses and one final analysis). The first OS interim analysis will be performed at the time of the primary PFS analysis. The second OS interim analysis will be performed after the occurrence of approximately 119 deaths and is projected to occur at approximately 30 months (~12 months after Primary Analysis) after the first patient is randomized. The final OS analysis will be performed after the occurrence of approximately 149 deaths (projected to occur approximately ~40 months after the first patient is randomized). The Lan-DeMets implementation of the O'Brien and Fleming use function will be used to control the overall type I error for the OS comparison at a two-sided 0.05 significance level.

The OS analyses were designed with an overall 5% (two-sided) level of significance to detect HR ratio of 0.893, which corresponds to an improvement in median OS from 25 months to approximately 28 months. However, this study is not designed to formally test OS. For the target number of OS events and anticipated study timeline, the OS analyses will be largely underpowered, i.e. ~10% statistical power. The details of the OS analyses are shown in Table 3. The exact efficacy boundary will be updated based on the actual number of observed OS events at the time of OS analysis.

Table 3: Assumptions and Characteristics for the Interim and Final Analyses of OS

Assumption	Findings
First interim analysis of OS (to be performed at time of final PFS analysis)	
Estimated cutoff date [a]	18 months
Projected number of events (% of final events)	60 (20%)
Projected MDD [b] (p-value)	0.418 (< 0.0008)
Second interim analysis of OS	
Estimated cutoff date [a]	30 months
Projected number of events (% of final events)	119 (39.7%)
Projected MDD [b] (p-value)	0.661 (< 0.0242)
Final analysis of OS	
Estimated cutoff date [a]	40 months
Projected number of events (% of final events)	149 (49.7%)
Projected MDD [b] (p-value)	0.717 (< 0.0428)

HR = hazard ratio; MDD = minimally detectable difference;
[a] Estimated data cutoff time from first randomization. Analysis results will be available after data cleaning.
[b] The largest observed HR that is projected to be statistically significant.

6. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document. For Synopsis, Schedule of assessments, PRO forms, etc. refer to study protocol.

7. REFERENCES

Brookmeyer R, Crowley J. (1982). A confidence interval for the median survival time. Biometrics 1982; 38:2941.

Finkelstein, D. M. (1986). A proportional hazards model for interval-censored failure time data. Biometrics 42, 845-854.

Friedman, M. (1982), “Piecewise Exponential Models for Survival Data with Covariates,” Annals of Statistics, 10, 101–113.

International Council for Harmonisation (ICH). E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2020.

Royston, P. and Parmar, M. K. B. (2002), “Flexible Parametric Proportional-Hazards and Proportional-Odds Models for Censored Survival Data, with Application to Prognostic Modelling and Estimation of Treatment Effects,” Statistics in Medicine, 21, 2175–2197.

Newcombe, R., “Interval Estimation for the Difference Between Independent Proportions: Comparison of Eleven Methods”, Statist. Med. 17, 873-890 (1998).

Sun, J. (1996). A non-parametric test for interval-censored failure time data with application to aids studies. Statistics in Medicine 15, 1387-1395.