NCT04059484

STATISTICAL ANALYSIS PLAN

An open-label randomized Phase 2 trial of amcenestrant (SAR439859), versus endocrine monotherapy as per physician's choice in patients with estrogen receptor-positive, HER2-negative locally advanced or metastatic breast cancer with prior exposure to hormonal therapies

SAR439859-ACT16105

STATISTICIAN: [redacted]

Statistical Project Leader: [redacted]

DATE OF ISSUE: 10-Dec-2021

Total number of pages: 83

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AEs:	adverse events
AESI:	adverse event of special interest
AI:	aromatase inhibitor
ALT:	alanine aminotransferase
AST:	aspartate aminotransferase
ATC:	anatomic category
BOR:	best overall response
CBR:	clinical benefit rate
cfDNA:	cell-free deoxyribonucleic acid
CI:	confidence interval
COD:	cut-off date
CTCAE:	Common Terminology Criteria for Adverse Events
DCR:	disease control rate
DI:	dose intensity
DOR:	duration of response
ECG:	electrocardiogram
ECOG:	Eastern Cooperative Oncology Group
eCRF:	electronic case report form
EDR:	early discrepancy rate
EORTC:	European Organisation for Research and Treatment of Cancer
EOT:	end of treatment
EQ-5D-5L:	EuroQoL questionnaire with 5 dimensions and 5 levels per dimension
ESR1:	estrogen receptor 1 gene
FWER:	family-wise error rate
GGT:	gamma-glutamyltransferase
GHS/QoL:	Global Health Status and Quality-of-Life Scale
HLGT:	high-level group term
HLT:	high-level term
HR:	hazard ratio
HRQL:	health-related quality of life
ICF:	informed consent form
ICR:	Independent Comitee Review
IHC:	immunohistochemistry
IMP:	investigational medicinal product
IRT:	interactive response technology
ITT:	intent-to-treat
LDR:	late discrepancy rate
LLT:	lower-level term
MedDRA:	Medical Dictionary for Regulatory Activities
NCI:	National Cancer Institute
ODR:	outcome discrepancy rate
OS:	overall survival

PCSA:	potentially clinically significant abnormalities
PD:	progressive disease
PFS:	progression free survival
PS:	performance status
PT:	preferred term
QLQ:	Quality of Life Questionnaire
RD:	relative dose
RECIST:	Response Evaluation Criteria in Solid Tumors
SDF:	survival distribution function
SERD:	selective estrogen receptor down-regulator
SERM:	selective estrogen receptor modulator
SOC:	system organ class
TEAE:	treatment-emergent adverse event
VAS:	visual analogue scale
WHO-DD:	World Health Organization-Drug Dictionary

1 OVERVIEW AND INVESTIGATIONAL PLAN

1.1 STUDY DESIGN AND RANDOMIZATION

This is an international, prospective, open-label, Phase 2 randomized study. Men, postmenopausal women and premenopausal women on a gonadotropin-releasing hormone analog with locally advanced or metastatic breast cancer will be randomly (1:1) assigned to one of the following treatment arms: amcenestrant (SAR439859) or an endocrine monotherapy of the choice of the physician.

The population will be stratified according to the presence of visceral metastasis (defined by at least 1 liver or lung metastasis) (Yes or No), prior treatment with CDK4/6 inhibitors (Yes or No), and Eastern Cooperative Oncology Group status (0 or 1).

Overall, 282 participants will be randomly assigned to study intervention with a balanced randomization ratio of 141 participants randomized per treatment arm from approximately 110 sites. The number of participants naïve to CDK4/6 inhibitors should not be higher than 20% of the overall sample size.

1.2 OBJECTIVES

1.2.1 Primary objectives

Table 1 - Primary objective and endpoint

Objective	Endpoint
Primary	
To determine whether SAR439859 400 mg per os improves progression-free survival (PFS) when compared with an endocrine monotherapy of the choice of the physician, in participants with metastatic or locally advanced breast cancer.	Progression-free survival is defined as the time interval from the date of randomization to the date of first documented tumor progression as per Response Evaluation Criteria in Solid Tumors (RECIST 1.1) assessed by independent central review (ICR) or death (due to any cause), whichever comes first.

1.2.2 Secondary objectives

Table 2 - Secondary objectives and endpoints

Objectives	Endpoints
Secondary	
To compare the overall survival in the 2 treatment arms.	Overall survival is defined as the time interval from the date of randomization to the date of documented death (due to any cause).
To assess the objective response rate in the 2 treatment arms.	Objective response rate is defined as the proportion of participants who have a confirmed complete response (CR) or partial response (PR), as best overall response (BOR) derived from overall response determined by ICR as per RECIST 1.1, from the date of randomization to the date of end of treatment.

Objectives	Endpoints
To evaluate the disease control rate in the 2 treatment arms.	Disease control rate is defined as the proportion of participants who have a confirmed CR, PR, stable disease (SD), or Non-CR/ Non-PD as BOR determined by ICR as per RECIST 1.1 from the date of randomization to the date of end of treatment.
To evaluate the clinical benefit rate in the 2 treatment arms.	Clinical benefit rate is defined as the proportion of participants who have a confirmed CR, PR, SD, or Non-CR/ Non-PD for at least 24 weeks determined by ICR as per RECIST 1.1, from the date of randomization to the date of end of treatment.
To evaluate the duration of response in the 2 treatment arms.	Duration of response is defined as the time from first documented evidence of CR or PR until progressive disease (PD) as determined by ICR as per RECIST 1.1 or death from any cause, whichever occurs first.
To evaluate the PFS according to the estrogen receptor 1 gene (ESR1) mutation status in the 2 treatment arms.	Progression-free survival as per ESR1 status determined at Cycle 1 Day 1.
To evaluate the pharmacokinetics of amcenestrant as single agent.	Amcenestrant plasma concentrations during the treatment period.
To evaluate health-related quality of life in the 2 treatment arms.	Disease-specific and generic health-related quality of life, disease and treatment-related symptoms, health state utility, and health status will be evaluated using the European Organisation for Research and Treatment of Cancer (EORTC) core quality of life questionnaire (QLQ-C30), the EORTC-QLQ breast cancer specific module (BR23) and the EuroQoL questionnaire with 5◦dimensions and 5 levels per dimension (EQ-5D-5L), from Cycle 1 Day 1 until 90 days after last dose of the study treatment.
To evaluate the overall safety profile in the 2 treatment arms.	Adverse events/serious adverse events and laboratory abnormalities.

1.2.3 Tertiary/exploratory objectives

Table 3 - Tertiary/exploratory objectives and endpoints

Objectives	Endpoints
Tertiary/exploratory	
To evaluate in participants the gene mutation profile of the tumor over time (baseline and end of treatment) by cell-free deoxyribonucleic acid (cfDNA) analysis, and ESR1 mutation analysis (pre- and on-treatment) by cfDNA.	The gene mutation profile of the tumor over time (Cycle 1 Day 1 [pre-treatment] and end of treatment) by cfDNA analysis, and ESR1 mutation analysis by cfDNA analysis pre-treatment and at Cycle 3 Day 1.
To evaluate in participants tumor biomarkers over time such as estrogen receptor (ER), Ki67, Bcl-2, and progesterone receptor (PgR) protein, and ribonucleic acid (RNA) gene expression profiles (for participants with tumor sites accessible for biopsy who accept biopsies at study entry and on treatment).	Tumor ER, Ki67, Bcl-2, and PgR protein, and RNA gene expression profiles in optional paired biopsies at Cycle 1 Day 1 (pre-treatment) and at Cycle 3, Day 1 (allowed up to Day 15; or within 14 days after first radiological tumor assessment after baseline).
To explore pharmacokinetic/pharmacodynamics relationship of amcenestrant.	Pharmacokinetic exposure and response, or safety endpoints during the treatment period.
To evaluate the time to first use of chemotherapy after disease progression in the 2 treatment arms.	Time to first use of chemotherapy after disease progression is defined as the time from randomization to start of the first use of chemotherapy after disease progression.

1.3 DETERMINATION OF SAMPLE SIZE

The sample size for this study is determined based on following assumptions: the median PFS for the control arm is assumed to be 4.5 months (1, 2, 3, 4, 5, 6) and an improvement of 53% to a median PFS of 6.9 months (corresponding to a hazard ratio [HR] = 0.65) would be considered clinically meaningful.

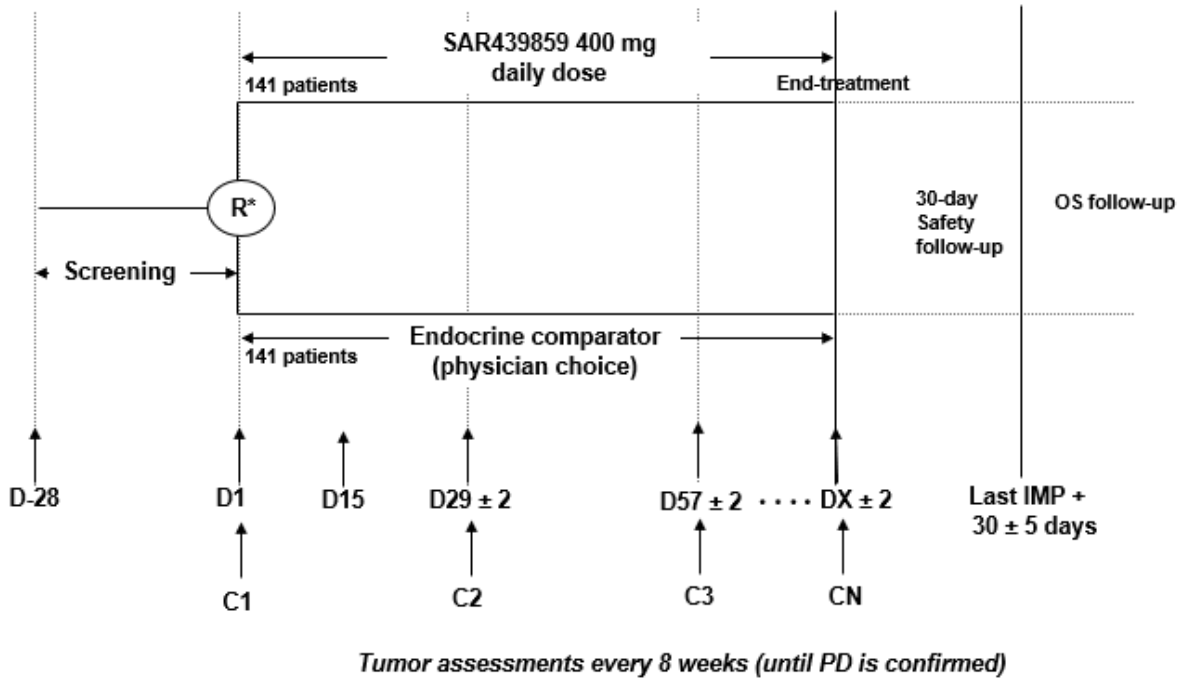
A total of 201 ICR-assessed PFS events will be required in the 2 treatment arms for the study to have approximately 85% power to detect an increase in PFS, assuming a true HR of 0.65 (representing a 53% increase in median PFS from 4.5 to 6.9 months), tested at a one-sided significance level of alpha = 0.025.

For the global part of the study, assuming a uniform accrual rate of 23.5 participants per month accomplished over a period of about 12 months and an annual dropout rate of 10%, a total sample size of approximately 282 participants (randomized in a 1:1 ratio) is required. The sample size calculation considers 1 futility interim analysis at 50% of the planned number of events. Under this current assumption, the COD for final PFS is approximately 18 months after first participant randomized.

Sample size calculation was performed using East 6.5.

1.4 STUDY PLAN

Figure 1 - Graphical study design



C = Cycle; D = day; IM = intramuscular; IMP = investigational medicinal product; OS = overall survival; PD = progressive disease; PFS = progression-free survival; R = randomization.

The complete schedule of activities (SoA) is presented in Section 1.3 of the protocol.

1.5 MODIFICATIONS TO THE STATISTICAL SECTION OF THE PROTOCOL

The protocol amendment history Table 4 gives the version number, description and rationale of major changes to the protocol statistical section. All changes were performed after the first patient randomized (12-Nov-2019) and before the planned primary PFS analysis (early 2022).

Table 4 - Protocol amendment statistical changes

Amendment #	Description of Change	Brief Rationale
02	Overall survival has been updated as the key secondary objective. The analysis methods for OS have been updated to include a formal comparison between the 2 treatment arms (stratified logrank for statistical testing).	Overall survival has been considered the key secondary endpoint, in response to the Food and Drug Administration recommendations.
	The objective response rate (ORR) is no longer key secondary endpoint. Its definition has been updated to clarify that BOR will be derived from overall response determined by ICR.	Updated for accuracy, as the best overall response will be derived from overall response determined by ICR

Amendment #	Description of Change	Brief Rationale
02	The sample size calculation has been updated. The total number of required ICR assessed PFS events has been updated to 201 from 195. The COD for final PFS has been updated to 18 months after first randomized participant. The power calculations for OS analysis have been included. The following will be applied for this endpoint: • The interim analysis of OS is planned at the time of PFS analysis, ie, approximately 18 months from the start of study randomization. • The final OS analysis is planned to be performed when 196 randomized participants have died (ie, at approximately 70% OS data maturity). • The COD for OS is approximately 64 months after the first randomized participant.	The sample size calculations and COD projections have been updated following the newly added futility interim analysis for PFS and the inclusion of OS as the key secondary endpoint.
02	A new section has been included for estradiol assessments, which have been added in this amendment.	Estradiol assessments have been included to explore the possible influence of circulating levels of estradiol on the efficacy of SAR439859.
02	The response-evaluable population has been removed.	This population has been removed from the study in consistency with the other studies from the same program. The analyses that were to be performed in this population will be performed in the intent-to-treat (ITT) participants with measurable disease at study entry, regardless of postbaseline assessment availability. This ITT analysis will provide a more accurate estimation of the ORR.
02	An interim analysis for futility has been planned to be carried out at 50% of the planned total number of PFS events. An interim analysis on OS is also planned at the time of final PFS analysis, if PFS is statistically significant.	To clarify the methodology used for the interim analyses.
03	The following text has been added: "The DMC will also oversee the interim analyses on PFS detailed in Section 9.5.1".	To clarify that the DMC will be overseeing the PFS interim analyses planned for this study.
04	The definitions of Disease Control Rate (DCR) and Clinical Benefit Rate (BCR) have been updated to add the "Non-Complete Response/Non-Progressive Disease" in the endpoint's descriptions.	To take in account the participants with only non-measurable disease in the definitions of DCR and CBR.
04	The stopping boundary for futility has been updated to: based on the observed HR based on Cox proportional hazard model, ie, an HR>1.1.	To revise the futility analysis strategy in order to recommend stopping the study only in case of observed increase risk of PFS events in the experimental arm compared to the control arm.

Amendment #	Description of Change	Brief Rationale
04	Photosensitivity has been added as AESI	Preclinical studies using amcenestrant indicate a potential risk for phototoxicity. Photosensitivity events has been added as an AESI in order to collect relevant information.
06	The definition of the COD for the final PFS and OS analysis has been changed	To allow more flexibility for the COD definition for final PFS and OS analyses.
06	The definition of the censoring and event scheme for the PFS analysis has been changed.	To align AMEERA-3 with the general FDA recommendation to censor patients after initiation of further anti-cancer therapy or if the documented progression or death occurred after two or more non-evaluable tumor assessments.

1.6 STATISTICAL MODIFICATIONS MADE IN THE STATISTICAL ANALYSIS PLAN

The statistical analysis plan history Table 5 below gives the version number, description and rationale for major changes to the statistical analysis features in the statistical analysis plan.

The initial SAP was approved on 23-Oct-2019. Interim analysis was performed on 24-Mar-2021. The primary PFS analysis is planned for early 2022.

Table 5 - Major changes in statistical analysis plan

SAP Version	Approval Date	Description of Change	Brief Rationale
1.0	23-Oct-2019	Not Applicable	Original version
2.0	11-Jan-2021	Overall survival has been updated as the key secondary objective. The analysis methods for OS have been updated to include a formal comparison between the 2 treatment arms (stratified logrank for statistical testing).	Overall survival has been considered the key secondary endpoint, in response to the Food and Drug Administration recommendations.
		The sample size calculation has been updated. The total number of required ICR assessed PFS events has been updated to 201 from 195. The COD for final PFS has been updated to 18 months after first randomized participant. The power calculations for OS analysis have been included.	The sample size calculations and COD projections have been updated following the newly added futility interim analysis for PFS and the inclusion of OS as the key secondary endpoint.
		The analysis plan for interim PFS futility and interim OS at the time of final PFS analysis (if PFS is statistically significant) has been described.	To clarify the methodology used for the interim analyses.
		Prior anti-cancer therapies endpoints are updated and reorganized.	To better organize the summary of prior anti-cancer treatment.
		Response-evaluable population has been removed from the analysis population.	Change as per protocol amendment 02.

SAP Version	Approval Date	Description of Change	Brief Rationale
		RMST in case of non-proportional hazard has been added.	To clarify supportive analysis that may be performed in case of non-proportional hazard.
		Subgroup analyses for OS and sensitivity analyses adjusting OS for switch to subsequent anti-cancer treatment have been added.	To clarify subgroup analysis and supportive analysis that may be performed for OS.
		The algorithm of PFS censoring under the condition of "documented progression (or death) after two or more non-evaluable tumor assessment" has been updated to "An event occurring at least 18 weeks (excluded) after the last evaluable tumor assessment".	To cover the 7-day window before and after the scheduled tumor assessment.
		Updated censoring rule for the primary PFS analysis, sensitivity analysis #1 and sensitivity analysis for investigator assessment.	Updated based on feedback from health authorities on AMEERA-5.
		The wording in 2.5.2 has been modified for Date of tumor assessment.	Rephrased the wording for clarification.
3.0	10-Dec-2021	Wording of the censoring rule for the primary PFS and OS analyses has been updated.	To clarify the censoring rule for all possible scenarios for the primary PFS analyses.
		The definition of the COD for the final PFS and OS analysis has been changed.	To allow more flexibility for the COD definition for final PFS and OS analyses.
.		Baseline definition has been added for biomarker data and clinical laboratory/vital sign/ECG data that, "baseline is defined as the last sample collection prior to the first administration of the IMP."	To clarify the baseline definition for biomarker and clinical laboratory/vital sign/ECG data.
		Photosensitivity has been added as AESI (after protocol amendment 04).	AESI added following protocol amendment 04.
		Analysis visit for PRO analyses has been added and longitudinal repeated measures analysis has been clarified.	To estimate the treatment effect on quality of life according to the real duration of exposure.
		Analyses of the time to sustained deterioration for all EORTC QLQ-C30 and EORTC QLQ-BR23 scales have been added.	To better characterize the effect of SAR439859 on Quality of Life.
		The definition of "as-treated" has been changed to "analyzed according to the randomized arm if the participant has received at least one administration of the as-randomized intervention".	Participants will be classified to the randomized arm if receiving at least one administration of as-randomized intervention. This definition is aligned with AMEERA-5 study for project level consistency.

2 STATISTICAL AND ANALYTICAL PROCEDURES

2.1 ANALYSIS ENDPOINTS

2.1.1 Demographic and baseline characteristics

The baseline value is defined as the last value or measurement taken up to the date of randomization. This definition applies for all variables unless otherwise specified.

All baseline safety and efficacy parameters (apart from those listed below) are presented along with the on-treatment summary statistics in the safety and efficacy sections (Section 2.4.6 and Section 2.4.5).

Demographic characteristics

Demographic variables are age in years (quantitative and qualitative variable : 18 to 64, 65 to 84, ≥85), gender, race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Island, Unknown), ethnicity (Hispanic or Latino, not Hispanic or Latino, Not reported, Unknown), weight (kg), eastern cooperative oncology group (ECOG) performance status (PS) and menopausal status for females (peri/premenopausal, postmenopausal) at baseline.

Medical or surgical history

Medical or surgical history includes relevant history of previous or associated pathologies other than the tumor.

This information will be coded using the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at Sanofi at the time of database lock.

Disease characteristics at diagnosis

The following disease characteristics at initial diagnosis will be described:

- Time from initial diagnosis of breast cancer to randomization date (in years)
- Histology (diagnosis type as collected in eCRF)
- Disease location and laterality (as collected in eCRF)
- Histopathology type (undifferentiated, poorly differentiated, moderately differentiated, well differentiated and other)
- Stage of the disease (as collected in eCRF)

Disease status at study entry

The following disease characteristics at study entry will be described:

- Extent of the disease (metastatic, locally advanced, other)

- Number of organ(s) involved (as collected in eCRF and per ICR)
- Type of organ(s) involved (as collected in eCRF and per ICR)
- Type of disease: Visceral metastases (ie, liver or lung) or not visceral metastases (as collected in eCRF and per ICR)
- HER2 status (as collected in eCRF)
- ER status (as collected in eCRF)
- PgR status (as collected in eCRF)
- Disease status (measurable disease, non-measurable disease), as collected in eCRF and per ICR)

Prior anticancer therapies

Prior anti-cancer treatments are collected by regimen in the eCRF. The following variables will be collected/derived:

- Intent of prior anti-cancer therapy according to the following categories: neoadjuvant only, adjuvant only, advanced only, both neoadjuvant and adjuvant, both neoadjuvant and advanced, both adjuvant and advanced, neoadjuvant and adjuvant and advanced.
- Time from last relapse/progression to randomization date
- Number of prior lines of treatment in advanced setting descriptive statistics and by class 0, 1, 2 or ≥3. A line of therapy in the advanced setting consists of a single agent, combination or a sequential therapeutic strategy with several drugs, given until a PD is documented. It corresponds to a regimen in advanced setting.
- Intent of the last prior anti-cancer therapy according to the following categories: no prior treatment, neoadjuvant or adjuvant, advanced and unknown.
- Endocrine resistance status according to the following categories:
  - Primary resistance, defined as relapse <24 months after the start of adjuvant hormonotherapy, for patients without advanced hormonotherapy; progression <6 months after the start of the last prior advanced hormonotherapy, for patients with advanced hormonotherapy.
  - Secondary resistance, defined as relapse >= 24 months after the start and <12 months after the end of adjuvant hormonotherapy, for patients without advanced hormonotherapy; progression >= 6 months after the start of the last prior advanced hormonotherapy, for patients with advanced hormonotherapy.
  - Sensitive, defined as relapse >= 12 months after the end of adjuvant hormonotherapy and hormonotherapy-naïve in advanced therapy.
- Type of prior anticancer therapy in neoadjuvant setting (targeted therapy, hormonotherapy, chemotherapy, immunotherapy or other as collected in eCRF),
- Type of prior anticancer therapy in adjuvant setting (targeted therapy, hormonotherapy, chemotherapy, immunotherapy or other as collected in eCRF),

- Time from start of adjuvant therapy to relapse in adjuvant setting (years), for participants who relapsed during adjuvant therapy
- Time from end of adjuvant therapy to relapse in adjuvant setting (years), for participants who relapsed after completion of adjuvant therapy
- Duration of adjuvant therapy (years)
- Among patients with prior adjuvant hormonotherapy,
  - Number of patients with relapse <24 months after the start of adjuvant hormonotherapy
  - Number of patients with relapse ≥24 months after the start and <12 months after the end of adjuvant hormonotherapy
  - Number of patients with relapse ≥12 months after the end of adjuvant hormonotherapy
- Type of prior anticancer therapy in advanced setting (targeted therapy, hormonotherapy, chemotherapy, immunotherapy or other as collected in eCRF),
- Reason for discontinuation of the last prior line in advanced setting,
- Best response to the last prior line in advanced setting,
- Time to progression of last prior line in advanced settings (ie, start date of the first drug within the last prior line of treatment up to last progression) (months)
- Duration of last prior line in advanced settings (last line end date - last line start date + 1) (months)
- Among prior hormonotherapy:
  - Number of patients with intent: neoadjuvant only, adjuvant only, advanced only, both neoadjuvant and adjuvant, both neoadjuvant and advanced, both adjuvant and advanced, neoadjuvant and adjuvant and advanced
  - Number of prior hormonotherapy-based lines in advanced setting by category (0, 1, 2 or ≥3)
  - Type of prior hormonotherapy in neoadjuvant or adjuvant settings:
    - Aromatase inhibitor (AI)
    - SERM (eg, Tamoxifen)
    - SERD (eg, Fulvestrant)
    - Other
  - Type of prior hormonotherapy in advanced setting:
    - AI
    - SERM (eg Tamoxifen)
    - SERD (eg Fulvestrant)
    - Other

- Among prior chemotherapy:
  - Number of patients with intent: neoadjuvant only, adjuvant only, advanced only, both neoadjuvant and adjuvant, both neoadjuvant and advanced, both adjuvant and advanced, neoadjuvant and adjuvant and advanced,
  - Number of prior chemotherapy lines in advanced setting by category (0, 1, 2 or ≥3)
  - Type of prior chemotherapy in neoadjuvant or adjuvant settings and Type of prior chemotherapy in advanced settings:
    - Anthracycline
    - Taxane
    - Capecitabine
    - Other
- Among prior targeted therapy:
  - Number of patients with intent: neoadjuvant only, adjuvant only, advanced only, both neoadjuvant and adjuvant, both neoadjuvant and advanced, both adjuvant and advanced, neoadjuvant and adjuvant and advanced,
  - Number of prior targeted-based lines in advanced setting by category (0, 1, 2 or ≥3)
  - Type of prior targeted therapy in neoadjuvant or adjuvant settings and Type of prior targeted therapy in advanced settings:
    - Anti-HER2
    - CDK4/6 inhibitor
    - mTOR inhibitor
    - PI3K inhibitor
    - PARP inhibitor
    - Other
- Among prior immunotherapy:
  - Number of patients with intent: neoadjuvant only, adjuvant only, advanced only, both neoadjuvant and adjuvant, both neoadjuvant and advanced, both adjuvant and advanced, neoadjuvant and adjuvant and advanced,
  - Number of prior immunotherapy lines in advanced setting by category (0, 1, 2 or ≥3)
  - Type of prior immunotherapy in neoadjuvant or adjuvant settings and Type of prior targeted therapy in advanced settings:
    - Anti PD-1
    - Anti PD-L1
    - Other
- Prior anti-cancer therapies in combination with endocrine therapy
  - Type of prior endocrine-based combinations in advanced settings:

- CDK4/6 inhibitor + AI
- CDK4/6 inhibitor + SERD
- AI + SERD
- AI + SERM
- PI3K inhibitor + SERD
- mTOR inhibitor + AI
- mTOR inhibitor + SERD
- Other

- Prior surgery: number (n, %) of patients with any prior surgery related to breast cancer, type of procedure (Preferred Term) and time from the last surgery to the randomization date (months).
- Prior radiotherapy: number (n, %) of patients with any prior external radiotherapy related to breast cancer, intent, intent of last prior external radiotherapy, time from the last external radiotherapy to the randomization date (months) overall and by intent (curative and palliative) and location of prior external radiation therapy by intent. Prior internal radiotherapy will be listed.

Any technical details related to computation, dates, and imputation for missing dates are described in Section 2.5.4.

2.1.2 Prior or concomitant medications (other than anticancer therapies)

All medications taken from the signed informed consent date up to the randomization date and until 30 days after the end of treatment are to be reported in the eCRF.

All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD) using the version currently in effect at Sanofi at the time of database lock.

- Prior medications are those the patient used from the date of informed consent until first study treatment administration intake. Prior medications can be discontinued before first administration or can be ongoing during treatment phase.
- Concomitant medications are any treatments received by the patient concomitantly to the IMP, from first study treatment administration to the end of treatment + 30 days. A given medication can be classified both as a prior medication and as a concomitant medication. Concomitant medications do not include medications started during the posttreatment period (as defined in the observation period in Section 2.1.4).

Any technical details related to computation, dates, imputation for missing dates are described in Section 2.5.4.

2.1.3 Efficacy endpoints

2.1.3.1 Primary efficacy endpoint(s)

The primary endpoint is Progression Free Survival (PFS) defined as the time from the date of randomization to the first documented date of the first documentation of objective PD according to RECIST 1.1 definitions (see Appendix D) or death due to any cause, whichever comes first.

The following censoring rules will be used:

- If documented progression or death are not observed before the analysis cut-off or the date of initiation of further anticancer therapy, PFS will be censored at the date of the last evaluable tumor assessment not showing documented progression prior to the initiation of a further anticancer therapy (if any).
- A participant without PFS event (death or documented progression) and without any evaluable post-baseline tumor assessments will be censored at the day of randomization (Day 1).
- A participant with an event documented after two or more non-evaluable tumor assessments will be censored at the date of the last evaluable tumor assessment documenting no progression prior to the initiation of a further anticancer therapy.

Clinical/non-radiological progression (as collected in the eCRF) will not be considered as documented progression for primary PFS analysis.

An independent central review (ICR) blinded to the randomization arm and to the patient characteristics will be conducted to evaluate overall response at each tumor assessment and determine disease response including progression status per RECIST 1.1 definitions (Appendix D). The full details regarding the determination of the progressive disease are provided in protocol and the ICR charter. The ICR will stop after the cut-off date (COD) for PFS analysis because no more efficacy assessment will be performed except collection of survival status.

Additional details regarding the definition of PFS and handling of events and censoring are given in Section 2.5.2 and Appendix E.

2.1.3.2 Secondary efficacy endpoint(s)

Overall Survival

Overall survival is defined as the time from date of randomization to date of death due to any cause. In the absence of observation of death, survival time will be censored to last date the participant is known to be alive (last contact date as defined in Section 2.5.3).

Objective response rate

The Best Overall Response will be determined by ICR according to RECIST v1.1 (see ICR charter for definition). Confirmation of responses (CR or PR) is necessary. As a supportive analysis, BOR according to the investigator assessment will also be determined.

The ORR on each randomized treatment arm will be estimated by dividing the number of participants with objective response (confirmed CR or PR as BOR assessed by ICR, according to RECIST 1.1) by the number of participants from the analysis population of the respective treatment arm.

As a supportive analysis, ORR according to the investigator assessment will also be determined.

In addition, in order to evaluate the tumor shrinkage of the target lesions, best relative change from baseline in tumor size will be assessed for each patient with measurable disease. Additional details are given in Section 2.5.3.

Disease control rate

The DCR on each randomized treatment arm will be estimated by dividing the number of participants with disease control (confirmed CR or PR, or SD or Non-CR/Non-PD as BOR assessed by ICR, according to RECIST 1.1) by the number of participants from the analysis population of the respective treatment arm.

As a supportive analysis, DCR according to the investigator assessment will also be determined.

Clinical benefit rate

The CBR on each randomized treatment arm will be estimated by dividing the number of participants considered as clinical benefit responders based on ICR assessment, by the number of participants from the analysis population of the respective treatment arm. For patients with measurable disease at baseline, they will be considered as clinical benefit responders if they achieve a CR or PR as BOR, or SD with an overall response recorded as SD or better at 24-1=23 weeks or later from randomization, allowing for the ±7 days visit window for tumor assessment. For patients with non-measurable disease at baseline, they will be considered as clinical benefit responders if they achieve a CR as BOR or Non-CR/Non-PD with an overall response recorded as Non-CR/Non-PD or better at least 23 weeks after randomization.

As a supportive analysis, CBR according to the investigator assessment will also be determined.

Duration of response

The DOR is defined as the time from the date of the first tumor assessment at which the overall response was recorded as CR or PR that is subsequently confirmed to the first date of tumor assessment at which the overall response was recorded as PD or death, whichever comes first. For participants with ongoing response at the time of the analysis, DOR will be censored at the date of the last valid disease assessment not showing documented progression performed before the initiation of a new anticancer treatment (if any).

DOR is determined only for patients who have achieved a BOR of PR or better.

As a supportive analysis, DOR according to the investigator assessment will also be determined.

2.1.4 Safety endpoints

The safety analysis will be based on the reported adverse events (AEs) and other safety information, such as clinical laboratory data, vital signs, weight, electrocardiogram (ECG) and Eastern Cooperative Oncology Group (ECOG) performance status (PS).

Observation period

The observation period will be divided into 3 epochs:

- The **pretreatment** period is defined as the time from when the participants give informed consent to the first administration of the IMP.
- The **on-treatment** period is defined as the time from the first dose of IMP up to 30 days after the last dose of IMP.
- The **post-treatment** period is defined as the time starting 31 days after the last dose of IMP to study closure.

2.1.4.1 Adverse events variables

Adverse event observation period

- Pretreatment AEs are defined as any AE occurring during the pretreatment period.
- Treatment-emergent AEs are defined as AEs that develop, worsen (according to the Investigator's opinion), or become serious during the on-treatment period.
- Post-treatment AEs are defined as AEs that are reported during the post-treatment period.

All adverse events (including serious adverse events and adverse events of special interest) will be graded according to National cancer institute common terminology for adverse events (NCI-CTCAE) v5.0 and coded to a lower-level term (LLT), preferred term (PT), high-level term (HLT), high-level group term (HLGT), and associated primary system organ class (SOC) using the version of MedDRA currently in effect at Sanofi at the time of database lock.

Adverse events of special interest

Specific analyses will be performed for the following AEs (see Section 2.1.4.1):

- Pregnancy of female participant entered in a study with IMP,
- Symptomatic overdose (serious or non-serious) with IMP.
- Increase in alanine aminotransferase (ALT) ≥ Grade 3.
- Photosensitivity (when occurred after local approval of protocol amendment 4)

2.1.4.2 Deaths

The deaths will be summarized as follows:

- Deaths on-study: includes all deaths occurring from the first IMP up to study closure:
  - Death on-treatment: includes all deaths occurring from the first IMP up to the end of treatment + 30 days,
  - Death post-treatment: includes all deaths occurring after the end of treatment + 30 days up to study closure.

2.1.4.3 Laboratory safety variables

Clinical laboratory data consists of blood analysis, including hematology, clinical chemistry, and urinalysis. Clinical laboratory values after conversion will be analyzed into standard international units and international units will be used in all listings and tables.

Blood samples for clinical laboratories will be taken as defined in the SoA (Section 1.3 of the protocol) and as clinically indicated. Urinalysis tests on dipstick will be assessed at baseline and at End of treatment (EOT) and if clinically relevant. Patients with 3+ or greater urine protein dipstick reading should undergo further assessment with a 24-hour urine collection for determination of proteinuria.

The laboratory parameters will be classified as follows:

- Hematology
  - **Red blood cells and platelets and coagulation**: hemoglobin, hematocrit, red blood cell count, platelet count, prothrombin time and international normalized ratio (INR)
  - **White blood cells:** white blood cell count, neutrophils, lymphocytes, monocytes, basophils, eosinophils
- Clinical chemistry
  - **Metabolism:** glucose, albumin
  - **Electrolytes**: sodium, potassium, chloride, calcium, phosphate, magnesium
  - **Renal function**: creatinine, eGFR (as collected in the eCRF at baseline), blood urea nitrogen, urea
  - **Liver function**: alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase, lactate dehydrogenase, total bilirubin, gamma-glutamyltransferase (GGT)
- Urinalysis
  - **Qualitative urinalysis by dipstick:** proteins, glucose, ketones, leukocytes, erythrocytes

2.1.4.4 Vital signs variables

Vital signs include heart rate, systolic and diastolic blood pressure, weight and ECOG PS (0, 1, 2, 3, 4).

For a given parameter, a patient will be considered as evaluable if at least one measure of this parameter is available during the on-treatment period.

2.1.4.5 Electrocardiogram variables

The incidence of patients with at least 1 abnormal ECG at any time during the on-treatment period will be summarized irrespective of the baseline level and/or according to the following baseline status categories:

- Normal/missing
- Abnormal

2.1.5 Pharmacokinetic variables

SAR439859 plasma concentrations will be determined at the predefined time point according to the SoA presented in Section 1.3 of the protocol. In addition, population PK approaches will be used to calculate individual PK estimates. The population PK analysis will be described in a separate report.

PK estimates may also be used to conduct the exploratory exposure-response analyses for safety (eg, incidence of AEs) and efficacy (eg, BOR, CBR, DOR).

2.1.6 Pharmacodynamic, genomics and biomarker endpoints

Estrogen receptor 1 gene mutation status in circulating free DNA

The mutation status (wild type, mutant) of twelve specific mutations of the ESR1 gene will be determined at baseline and on Cycle 3 Day 1 by multiplex droplet digital PCR (ddPCR), including their mutant frequency and concentration. For each ESR1 mutation the type of mutation will be also specified (missense, frameshift, inframe, etc).

Mutational profiling in circulating free DNA

The mutations of a panel of 77 cancer genes (Roche AVENIO panel) will be determined at baseline. The type of mutation (missense, frameshift, inframe, etc), the mutant frequency and the mutant concentration will be also specified. In addition, other genomic aberrations such as the copy number variants or the fusion genes may be highlighted.

Estrogen receptor degradation and tumor biopsy biomarkers

The expression of ER will be determined by immunohistochemistry (IHC) at baseline and on treatment at Cycle 3 in tumor biopsy samples in both treatment arms. Change in ER expression will be used to assess ER degradation in patients with accessible tumor.

In addition, expression levels of biomarkers such as Ki67, Bcl-2, and PgR will also be evaluated by IHC. From tumor biopsies, tumor gene expression profiles related to ER degradation (and other gene signature and pathways) will be also obtained. These analyses will be performed on transcriptome (mRNA).

Estradiol

Circulating level of estradiol will be measured at baseline and at Cycle 3 Day 1.

2.1.7 Quality-of-life endpoints

Health-related quality of life (HRQL) will be assessed using:

- The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire – Cancer specific module with 30 items (EORTC QLQ-C30) (7, 8, 9).
- The EORTC QLQ-BR23 breast cancer-specific module with 23 items (EORTC QLQ-BR23) (10).
- The EuroQol measure with 5-dimensions and 5-levels per dimension EQ-5D-5L (11).

EORTC QLQ-C30

The EORTC core quality of life questionnaire (QLQ-C30) is a cancer-specific instrument that contains 30 items and provides a multidimensional assessment of HRQL (7, 8, 9) evaluating symptoms as well as functioning.

The QLQ-C30 is composed of both multi-item scales and single-item measures. These include 5 functional scales (physical, role, emotional, cognitive, and social functioning), 9 symptom scales (fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties), and a Global Health Status/quality of life scale (GHS/QoL). The recall period is the past week. All items are scored 1 (“not at all”) to 4 (“very much”) except for the items contributing to the global health status (GHS)/quality of life (QoL), which are scored 1 (“very poor”) to 7 (“excellent”). All scales are transformed from raw scores to linear scales ranging from 0 to 100 (see Appendix F). A higher score for functional and GHS/QoL scales represents a higher/healthier level of functioning/HRQL; whereas a higher score for symptoms scales represents a higher symptom burden. A 10-point change is often used as a conservative indicator of clinically meaningful change from baseline (12).

The following endpoints will be derived:

- Change from baseline in scores for each scale, calculated at each visit.
- The time to sustained deterioration for each scale, defined as the first time from first IMP date to the time of a worsening from baseline ≥10 in on-treatment (including EOT) post-baseline scores before cutoff date and start of further anticancer therapy, among patients for whom all subsequent scores meet the 10-point or greater post-baseline threshold (ie, an increase from baseline for symptom scales, and a decrease from baseline for the functional scales/GHS/QoL). In the absence of a sustained worsening from baseline in scales, the

time to sustained deterioration will be censored to the date of the last EORTC QLQ-C30 evaluation before the start of further anticancer therapy, last contact date, lost to follow-up, death or withdrawal of consent or the cutoff date, whichever occurs first. In case of sustained deterioration after start of further anticancer therapy or cutoff date, the patient will be censored using the same rules previously described.

EORTC QLQ-BR23

The EORTC QLQ breast cancer specific module (QLQ-BR23) is a disease-specific HRQL measure that is used in conjunction with the QLQ-C30. The QLQ-BR23 assesses the impact breast cancer and the side effects of treatment in breast cancer patients.

The QLQ-BR23 contains 23 items. It is composed of both multi-item scales and single-item measures. 4 functional scales (body image, sexual functioning, sexual enjoyment, future perspective) and 4 scales related to symptoms of disease or treatment (arm symptoms, breast symptoms, systemic therapy side effects, and upset by hair loss). The recall period is the past week. All items are scored 1 ("not at all") to 4 ("very much"). All scales are transformed from raw scores to linear scales ranging from 0 to 100 (see Appendix F). A higher score for functional scales represents a better outcome; whereas a higher score for symptoms scales represents a higher symptom burden for symptoms of disease or treatment. Currently, there are no robust published estimates for clinically meaningful thresholds of within-patient change in EORTC QLQ-BR23 scores. A 10-point change from baseline is commonly applied in this setting (13, 14, 15).

Rules of handling with missing items are detailed in Appendix F for the EORTC questionnaires.

The following endpoints will be derived:

- Change from baseline in scores for each scale, calculated at each visit.
- The times to sustained deterioration for each scale, defined as the first time from first IMP date to the time of a worsening from baseline ≥10 in on-treatment (including EOT) post-baseline scores before cutoff date and start of further anticancer therapy, among patients for whom all subsequent scores meet the 10-point or greater post-baseline threshold (ie, an increase from baseline for symptom scales, and a decrease from baseline for the functional scales). In the absence of a sustained worsening from baseline in scales, the time to sustained deterioration will be censored to the date of the last EORTC QLQ-BR23 evaluation before the start of further anticancer therapy, last contact date, lost to follow-up, death or withdrawal of consent or the cutoff date, whichever occurs first. In case of sustained deterioration after start of further anticancer therapy or cutoff date, the patient will be censored using the same rules previously described.

EQ-5D-5L

The EuroQoL questionnaire with 5 dimensions and 5 levels per dimension (EQ-5D-5L) is a standardized measure of health status, designed for self-completion by the participants, that provides a simple, generic measure of health for clinical and economic appraisal, and consists of 2 sections: descriptive system and visual analogue scale (VAS). The descriptive section consists

of 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The respondent is asked to indicate his/her health state by ticking (or placing a cross) in the box against the most appropriate statement in each of the 5 dimensions; this results in a 1-digit number expressing the level for that dimension. The digits for 5 dimensions can be combined in a 5-digit number describing the respondent's health state. The VAS records the participant's self-rated health on a 20 cm vertical VAS ranging from "the worst health you can imagine (0)" to "the best health state you can imagine (100)". This information can be used as a quantitative measure of health as judged by the individual participants. The recall period is today.

VAS will be described as given by the participant. The health state will be converted into a single utility index value by using the value sets based on UK population and according to the crosswalk algorithm developed by B. van Hout (11).

Change from baseline in VAS and utility index will be calculated at each visit.

2.1.8 Further therapy after discontinuation of investigational medicinal product administration during the study

Further therapies after discontinuation of IMP include further anti-cancer therapy, surgery and radiotherapy.

Time to first use of chemotherapy after disease progression or treatment discontinuation (without disease progression)

The time to first use of chemotherapy after disease progression or treatment discontinuation (without disease progression) is defined as the time from date of randomization to start of the first use of chemotherapy after disease progression or treatment discontinuation (without disease progression). In the absence of chemotherapy after disease progression or treatment discontinuation (without disease progression), time will be censored to last contact date or the cutoff date, whichever occurs first. In the unlikely event if a patient starts chemotherapy while on study treatment without disease progression, it will be considered as an event.

2.1.9 Impact of COVID-19pandemic

The impact of Covid-19 pandemic may be evaluated with below endpoints:

- Premature EOT due to COVID-19
- Premature EOS due to COVID-19
- Critical or major protocol deviations related to COVID-19
- Adverse events related to COVID-19
- Sensitivity analysis on primary PFS may be performed, with patients diagnosed with COVID-19 censored at the date of diagnosis.

2.2 DISPOSITION OF PATIENTS

This section describes patient disposition for both patient study status and the patient analysis populations.

Screened patients are defined as any patients who signed the informed consent.

Randomized patients (ie, Intent-to-treat (ITT) population) consist of all patients from the enrolled population (ie, with a signed informed consent form) who have been allocated a randomization number by the Interactive Response Technology (IRT), regardless of whether the patient was treated or not.

For patient study status, the total number of patients in each of the following categories will be presented in the clinical study report using a summary table:

- Nonrandomized but treated patients
- Randomized patients
- Randomized but not treated patients
- Randomized and treated patients
- Patients who discontinued study treatment with main reason for permanent treatment discontinuation
- Patients still on treatment
- Status at the cut-off date (Alive/Death)

Number of screened patients, number of screen failure patients and reasons for screen failure will be summarized in a separate table. In addition, number and percentage of screened patients by country and sites will be presented.

For all categories of patients (except for the screened and nonrandomized categories) percentages will be calculated using the number of randomized patients as the denominator.

All critical or major deviations potentially impacting efficacy analyses, randomization, and drug-dispensing irregularities, and other major or critical deviations will be summarized in tables giving numbers and percentages of deviations by treatment group.

Additionally, the following analysis populations will be summarized.

- Intent-to-treat (ITT) population
- Safety population
- Pharmacokinetic-evaluable population

Definition of study populations are provided in Section 2.3.

2.2.1 Randomization and drug dispensing irregularities

Randomization and drug-dispensing irregularities occur whenever:

1. A randomization is not in accordance with the protocol-defined randomization method, such as a patient is randomized based on an incorrect stratum or a patient is randomized twice

   OR

2. A patient is dispensed an IMP kit not allocated by the protocol-defined randomization, such as a patient at any time in the study is dispensed a different treatment kit than as randomized (which may or may not contain the correct-as-randomized IMP), or a nonrandomized patient is treated with IMP reserved for randomized patients. This irregularity is applicable only for patients in SAR439859 arm and for patients randomized in the control arm for countries dispensed by the IRT.

Randomization and drug-dispensing irregularities will be monitored throughout the study and reviewed on an ongoing basis.

All randomization and drug-dispensing irregularities will be documented in the clinical study report. If the number of irregularities is large enough to make a tabular summary useful, the irregularities will be categorized and summarized among randomized patients (number and percentages). Nonrandomized, treated patients will be described separately.

Randomization and drug-dispensing irregularities to be prospectively identified include but are not limited to:

Randomization and drug allocation irregularities
Kit dispensation without IRT transaction
Erroneous kit dispensation
Kit not available
Randomization by error
Patient randomized twice
Stratification error

Number of patients (%) in the stratification factors from IRT and derived from eCRF data will be displayed in the same table and the number of patients (%) with at least one wrong stratum will be calculated, by treatment and overall.

2.3 ANALYSIS POPULATIONS

Patients treated without being randomized will not be considered randomized and will not be included in any efficacy population.

The safety experience of patients treated and not randomized will be reported separately, and these patients will not be in the safety population.

2.3.1 Intent-to-treat population

All participants from the enrolled population (ie, who sign the ICF) and for whom there is a confirmation of successful allocation of a randomization number by IRT. Participants will be analyzed according to the treatment arm assigned at randomization.

2.3.2 Safety population

All participants randomly assigned to study intervention and who took at least 1 dose of study intervention. For participants receiving more than one study intervention during the study, they will be analyzed according to the randomized arm if the participant has received at least one administration of the as-randomized intervention.

In addition:

- Nonrandomized but treated patients will not be part of the safety population; however, their safety data will be presented separately
- Randomized patients for whom it is unclear whether they took the IMP will be included in the safety population as randomized

2.3.3 Pharmacokinetic-evaluable population

All participants from the safety population who received at least one dose of SAR439859 and with at least 1 available plasma concentration post treatment with adequate documentation of date and time of dosing and date and time of sampling.

2.4 STATISTICAL METHODS

Continuous data will be summarized using the number of available data, mean, standard deviation (SD), median, minimum, and maximum for each treatment group. Categorical and ordinal data will be summarized using the number and percentage of patients in each treatment group. One-sided tests will be used in this study for the primary endpoint (PFS) and key secondary endpoint (OS) at the 1-sided alpha of 2.5%. Confidence intervals will be estimated as two-sided and will be used for descriptive purposes only.

2.4.1 Demographics and baseline characteristics

Parameters will be summarized on the ITT population analyzed in the treatment group to which they were randomized. Analyses for the safety population will be included in the appendices if the size of the safety population is different (>10%) from the size of that in the primary analysis population for any treatment group.

Parameters described in Section 2.1.1 may be summarized by treatment group, according to the class of treatment (AI, SERM and SERD) for the control arm and overall using descriptive statistics, depending on the number of patients by class of treatment (AI, SERM, SERD).

The medical and surgical history will be summarized according to the SOC and PT (SOC will be sorted according to the internationally agreed order (see Appendix C) and PT by overall decreasing frequency).

2.4.2 Prior or concomitant medications (other than anticancer therapies)

The prior and concomitant medications will be presented for the ITT population.

Medications will be summarized by treatment group and according to the class of treatment (AI, SERM and SERD) for the control arm according to the WHO-DD dictionary, considering the first digit of the anatomic category (ATC) class (anatomic category) and the first 3 digits of the ATC class (therapeutic category). All ATC codes corresponding to a medication will be summarized, and patients will be counted once in each ATC category (anatomic or therapeutic) linked to the medication. Therefore, patients may be counted several times for the same medication.

The table for prior medications will be sorted decreasing frequency of ATC followed by all other therapeutic classes based on the overall incidence across treatment groups. In case of equal frequency regarding ATCs (anatomic or therapeutic categories), alphabetical order will be used.

The tables for concomitant medications will be sorted by decreasing frequency of ATC followed by all other therapeutic classes based on the incidence in the SAR439859 group. In case of equal frequency regarding ATCs (anatomic or therapeutic categories), alphabetical order will be used.

2.4.3 Prior anticancer therapies

In addition, the following specific medications will be summarized:

- A table for prior anti-cancer therapies will be provided using the ATC code (chemical class) and the standardized medication name. This table will be sorted by decreasing frequency of ATC followed by medication names based on the overall incidence. In case of equal frequency regarding ATCs, the alphabetical order will be used.
- The same table will be provided for prior anti-cancer therapies in advanced setting.

Summary tables will be displayed by treatment group, according to the class of treatment (AI, SERM and SERD) for the control arm and overall using descriptive statistics.

2.4.4 Extent of investigational medicinal product exposure and compliance

The extent of IMP exposure will be assessed and summarized by actual treatment within the safety population (Section 2.3.2), by treatment group, by class of treatment (AI, SERM, SERD) and according to each drug for the control arm.

In addition, class of treatment and drug will be summarized in the safety population. A shift table of planned drug versus actual taken drug will be provided.

2.4.4.1 Extent of investigational medicinal product exposure

The extent of IMP exposure will be assessed by the duration of IMP exposure and actual dose information, by class of treatment (AI, SERM, SERD) and for each drug for the control arm.

The extent of exposure will be assessed for each patient as:

- Number of cycles started, defined by the number of cycles in which at least one dose of any study treatments is administered.
- Duration of exposure (week) is defined as [(Last day of exposure – first day of exposure + 1)/7].
- The first date of exposure is defined as the first administration date with non-zero dose whichever the IMP (SAR439859, Anastrozole, Letrozole, Exemestane, Tamoxifen or Fulvestrant).

  The last date of exposure is defined as:

  - The last administration date for SAR439859, Anastrozole, Letrozole, Exemestane or Tamoxifen.
  - For Fulvestrant, min(last administration date + 14 - 1 days if last cycle is Cycle 1 or last administration date + 28 - 1 days if last cycle is Cycle 2 or later ; date of death)

Total number of cycles started, number of cycles started by patient as a quantitative variable and by category (ie, number (%) of patients having started at least 1 cycle, at least 2 cycles, etc), and duration of exposure will be summarized by descriptive statistics.

The following parameters will be calculated for each drug:

- Cumulative dose (mg): the cumulative dose is defined as the sum of all doses during the study treatment exposure.
- Dose intensity (DI, in mg/day) is defined as

$$\text{DI (mg/day)} = \frac{\text{Cumulative dose (mg)}}{\text{Duration of exposure (week)} \times 7}$$

- Total planned dose (mg):
  - For SAR439859, Anastrozole, Letrozole, Exemestane or Tamoxifen: Total Planned Dose = Planned dose at Cycle 1 Day 1 × 7 × duration of exposure (week)
  - For Fulvestrant: Total Planned Dose = Planned dose at Cycle 1 Day 1 × Theoretical total number of doses during started cycles (2 for Cycle 1 and 1 for Cycle 2 or later)
- The relative dose (RD, in %) is defined as:

$$\text{Relative Dose} = \frac{\text{Cumulative dose}}{\text{Total planned Dose}} \times 100$$

  Note: for SAR439859, Anastrozole, Letrozole, Exemestane or Tamoxifen the RD corresponds to a Relative Dose Intensity.

Summary statistics will be provided for DI (for SERD, SERM, and by each drug for the control arm) and RD. For RD, categorical summaries will also be presented according to the following categories: 0-80%, 80-100%, >100%.

Dose or cycle modifications

For each drug, the following variables will be calculated:

- Dose reduction: a dose is deemed to have been reduced if the dose taken by a patient is lower than the dose taken on the previous day or the day before dose(s) omitted. For C1D1, a dose is deemed to have been reduced if the dose taken is lower than the planned dose.
- Dose omission (per os treatment only): one omission corresponds to a dose not taken or equal to 0 mg/day between two non-zero doses. For SAR439859, Anastrozole, Letrozole, Exemestane or Tamoxifen, several consecutive dose omissions will be counted as one episode of omission. Dose omission is not applicable for Fulvestrant.

For Fulvestrant only, the following variable will be calculated:

- Cycle delays: a cycle is deemed to have been delayed if start date of current cycle – start date of previous cycle >30 days (28+2 days).

Dose modification variables will be summarized descriptively. Analyses will be performed based on the number of patients. The number of patients with at least one dose modification with the following details will be provided:

- Number (%) of patients with at least one dose modification (reduction, omission or cycle delay)
- Number (%) of patients with at least one dose reduction
- Number (%) of patients with at least one dose omission for SAR439859, Anastrozole, Letrozole, Exemestane or Tamoxifen
- Number (%) of dose reductions by patient according to the following categories: 0, 1, >1
- Number (%) of episodes of dose omissions by patient for SAR439859, Anastrozole, Letrozole, Exemestane or Tamoxifen according to the following categories: 0, 1, >1
- Number (%) of patients with at least 7 consecutive days of dose omission for SAR439859, Anastrozole, Letrozole, Exemestane or Tamoxifen
- Number (%) of patients with at least one cycle delay for fulvestrant

2.4.5 Analyses of efficacy endpoints

All efficacy analyses will be performed on the ITT population unless stated otherwise. Analyses by class of treatment will be performed according to the planned class of treatment for participants randomized to the control arm. All primary and secondary efficacy endpoints based on radiological assessments of tumor burden (ie, PFS, BOR, ORR, DCR, CBR, and DOR) will be

derived using the ICR tumor assessment. Analysis based on local radiologist's/Investigator's assessment will also be performed but considered as supportive analyses.

2.4.5.1 Analysis of primary efficacy endpoint(s)

Primary efficacy analysis will consist of PFS comparison between the SAR439859 arm and the control arm through a logrank test procedure stratified by the stratification factors as entered in the IRT (ie, presence of visceral metastasis, prior treatment with CDK4/6 inhibitors and ECOG). A one-sided Type I error rate of 2.5% will be used for statistical testing. Following hypotheses will be tested:

- The null hypothesis that the survival distribution functions (SDF) for the PFS of the SAR439859 arm is lower than or equal to the SDF of the control arm.
  - H_0: SDF (SAR439859) ≤ SDF (Control)

versus

- The alternative hypothesis that the SDF for the PFS of the SAR439859 arm is superior to the SDF of the control arm.
  - H_1: SDF (SAR439859) > SDF (Control)

The cut-off date for the analysis of PFS is the actual date when approximately 201 PFS events (first occurrence of either documented progression assessed by ICR or death due to any cause) have been observed or when all participants from the global cohort have been followed-up for at least 10 months (or discontinued treatment), whichever is earlier. It is approximately estimated 18 months after first randomized participant assuming a 12 months uniform enrollment and a median PFS of 4.5 months in the control arm and a hazard ratio of 0.65.

The actual cut-off date will be prospectively determined based on the timeline prediction to reach 201 ICR-assessed PFS events or the requirement of minimum of 10-month follow-up since the last participant randomized. The actual number of events may be slightly different from 201 at the data cut-off date for the primary PFS analysis.

In addition, the following estimates will be provided:

- PFS data will be analyzed using the Kaplan-Meier method by treatment group:
  - Kaplan-Meier estimates of the 25th, 50th and 75th percentiles and their associated 95% confidence interval (CI) will be provided. The 95% CIs will be constructed using a log-log transformation of the survival function and the methods of Brookmeyer and Crowley.
  - Number of patients at risk as well as the probabilities of being event-free at least 2, 4, 6, 8, 10 and 12 months with 95% CIs will be estimated for each treatment group using the Kaplan-Meier method and a log-log approach based on a normal approximation following the Greenwood's formula. Additional timepoints may be considered based on the actual data.

- Kaplan-Meier curves will be plotted. These plots will include the number of patients at risk at key time points by treatment group.

- The hazard ratio (HR) and its 95% CIs will be estimated using the Cox proportional hazards model stratified by the same stratification factors as those used for the log-rank test described above. Ties will be handled using the exact method. Underlying assumptions of the Cox Proportional hazards model will be assessed by graphical methods (ie, log-log graphical methods).
- For patients with events, the type of event (documented progression or death) will be summarized by treatment group using counts and percentages.
- For patients who died without documented progression, the time from the last evaluable disease assessment to the death will be summarized by treatment group using descriptive statistics.
- The number (%) of censored patients, the reason and timing of their censoring (ie, censored at randomization, censored at the last evaluable tumor assessment before the initiation of further anti-cancer therapy, or censored at last evaluable tumor assessment before the cut-off date), and the time from the last evaluable disease assessment to the cut-off date will be summarized by treatment group. For each censoring reason, when applicable, distinction will be made between cases where no event was observed and cases where an event was observed after the censoring.
- Follow-up duration (months) will be estimated using the reverse Kaplan-Meier approach, where censored data are treated as events and events are treated as censored data. Kaplan-Meier estimates of the 25th, 50th and 75th percentiles will be provided.

2.4.5.1.1 Sensitivity analyses

Different censoring and events rules

The same statistical methods used in the primary analysis will be applied using different censoring and event rules as defined below.

The sensitivity analyses will include the following censoring rules:

- Ignoring further anti-cancer therapy.
- Ignoring further anti-cancer therapy and considering events (documented progression or death) occurring after two or more non-evaluable tumor assessment as event.
- Ignoring further anti-cancer therapy and events (documented progression or death) occurring after exactly one or two or more non-evaluable tumor assessment are considered as event and back-dated to the next schedule assessment.

Additional details are provided in Appendix E.

Sensitivity analysis #1 (Ignoring further anti-cancer therapy)

PFS will be analyzed based on ICR assessment, ignoring further anti-cancer therapy.

Sensitivity analysis #2 (PFS considering events occurring after two or more non-evaluable tumor assessment as event and ignoring further anti-cancer therapy)

PFS will be analyzed based on ICR assessment, including events (documented progression or death) occurring after two or more non-evaluable tumor assessment as event. The date of progression (or death) will be used for date of outcome. Further anti-cancer therapy will be ignored.

Sensitivity analysis #3 (PFS considering events occurring after two or more non-evaluable tumor assessment as event and back-dating at the next schedule assessment and ignoring further anti-cancer therapy)

If more than 10% patients have two or more consecutive non-evaluable assessments prior to PFS event, PFS will be analyzed based on ICR assessment, including events (documented progression or death) occurring after two or more non-evaluable tumor assessment as event. The date of the next schedule assessment will be used for date of outcome. Further anti-cancer therapy will be ignored.

Stratification factors

The same statistical methods used in the primary analysis will be applied using different stratification rules as defined below.

Sensitivity analysis #4 (PFS analysis using stratification factors derived from eCRF data)

If more than 10% patients have a discordance between the strata as entered in the IRT system and as derived from eCRF data, PFS will be analyzed based on ICR assessment stratified by the stratification factors as derived from eCRF data.

Non-proportional hazard

As mentioned above, underlying assumptions of the Cox Proportional hazards model will be evaluated to assess the relevance of estimating the treatment effect from hazard ratio. In case the proportional hazards assumption may not be valid, Restricted Mean Survival Time (RMST) method (16) may be conducted for PFS with the primary analysis censoring rule. The RMST methodology is valid under any distribution of the time to event in the treatment groups and provide an estimate of the expected PFS between randomization and a common timepoint denoted by τ. The timepoint τ should be limited to the largest event time:

- τ_{max} = minimum of (largest observed PFS event time for SAR43985 arm, largest observed PFS event time for control arm).

The RMST estimate up to τ_{max} and associated 95% CI for each treatment arm will be provided. The treatment effect will be estimated based on the difference between the two treatment arms in RMST up to τ_{max}. Similar analysis will be performed using clinical meaningful truncation points (eg, up to 12 months if possible based on the largest event time within each treatment arm). Additionally, the RMST estimate within each arm will be plotted against time τ, with τ varying

from 0 to the largest observed PFS event time for the respective arm. Treatment effect based on the difference in RMST between the two treatment arms and associated 95% CI will also be provided against time τ, varying from 0 to τ_{max}.

Similar analyses may be performed for the comparison of SAR439859 and each class of treatment for the control arm. For these analyses, τ_{max} will be recalculated based on the minimum of the largest observed PFS event time for SAR43985 arm and largest observed PFS event time for the comparator.

2.4.5.1.2 Supportive analyses

PFS analysis will be supported by the following analysis based on the investigator/local radiologist's assessment:

- Supportive analysis #1: Same event/censoring rules as for the PFS primary analysis (see Table 17)
- Supportive analysis #2: Same event/censoring rules as for the sensitivity analysis #2 (see Table 19)
- Supportive analysis #3: Same event/censoring rules as for the supportive analysis #1 but considering the clinical/non-radiological progression as event, with the date of clinical/non-radiological progression as date of outcome (see Table 21)

Additional supportive analyses based on the investigator/local radiologist's assessment could be performed using different censoring and event rules or stratification rules as defined in Section 2.4.5.1.1, if relevant.

Concordance of PFS outcome

A comparison of PFS outcome (ie, "Event", "Censored") between the ICR and the investigator assessments will be summarized for each treatment group.

Table 6 - Cross-tabulation of ICR and investigator assessments of PFS outcome

Investigator/local radiologist	**ICR**	
	Event	Censored
Event	n_{11}	n_{12}
Censored	n_{21}	n_{22}

The PFS Outcome Discrepancy Rate (PFS ODR) will be calculated for each treatment arm as follow.

$$\text{PFS ODR} = \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}}$$

Concordance of tumor assessment evaluation

The differential discordance between the ICR and the investigator assessment of documented progression will be assessed using the Pharmaceutical Research and Manufacturers of America (PhRMA) method (17). The early discrepancy rate (EDR) and late discrepancy rate (LDR) differences between the two treatment arms will be calculated as follows:

Table 7 - Cross-tabulation of ICR and investigator assessments of documented PD

Investigator/local radiologist	ICR	
	Documented PD	No documented PD
Documented PD	$a = a_1 + a_2 + a_3$	b
No documented PD	c	d

Only documented PD component of PFS is considered; if death occurs without prior PD, a subject is counted under 'No documented PD'
a_1: number of agreements on timing and occurrence of documented PD
a_2: number of times investigator declares documented PD later than ICR
a_3: number of times investigator declares documented PD earlier than ICR

The timing of investigator/local radiologist and ICR (for participants with agreement on documented PD) will be considered to agree if they occur within ±7 days of each other, aligned with the protocol-specified window for tumor assessments.

The EDR is defined as:

$$EDR = \frac{b + a_3}{a + b}$$

The EDR quantifies the frequency with which the investigator declares progression early relative to ICR as a proportion of the total number of investigator assessed PDs, within each arm.

The LDR is defined as:

$$LDR = \frac{c + a_2}{b + c + a_2 + a_3}$$

The LDR quantifies the frequency with which the investigator declares progression later than ICR as a proportion of the total number of discrepancies, within each arm.

If the distribution of discrepancies is similar between the arms then this suggests the absence of evaluation bias favoring a particular arm.

The EDR and LDR will be calculated for each treatment arm and the differential discordance (DD) for each measure will be summarized as the rate on the SAR439859 arm minus the rate on the control arm. A negative differential discordance for the EDR and/or positive differential discordance for the LDR are suggestive of a bias in the investigator favoring the experimental arm.

2.4.5.1.3 Subgroups analyses

Evaluation of consistency

The consistency of the results from the primary analysis will be evaluated across pre-defined subgroups in patients available in the subgroup of consideration. The definition of each subgroup is defined in Table 8. Depending upon the study results, additional subgroups may be examined, and subgroups with small sample sizes may be pooled to create a larger meaningful subgroup. For each subgroup, Kaplan-Meier estimates of the median and its associated 95% CI will be provided for each treatment arm along with the HR and its 95% CI estimated using the unstratified Cox proportional hazards model. A forest plot summarizing the results for each subgroup will be provided.

Table 8 - Subgroups analyses: covariates investigated

Subgroup	Description
Presence of visceral metastasis as per IRT	Yes or No
Prior treatment with CDK4/6 inhibitors as per IRT	Yes or No
ECOG as per IRT	0 or 1
Age	<65 years or ≥65 years
Race	Asian or White or Other
Geographical region	Europe or North America or Asia or Other
Intent of last anti-cancer therapy	Neoadjuvant/Adjuvant or Advanced
Number of prior hormonotherapy lines in advanced setting	0 or 1 or 2 or ≥3
Prior chemotherapy	No prior chemotherapy or early breast cancer only or advanced metastatic
PgR status	PgR+ or PgR-
ESR1 mutational status	Wild type or mutated
Number of organs involved based on ICR assessment at baseline	<3 or ≥3
Menopausal status at study entry	Peri/premenopausal or Postmenopausal
Bone-only metastases based on ICR assessment at baseline	Yes or No

Note: In case of discrepancy between IRT and eCRF stratification factors, subgroup analyses will also be performed according to the stratification factors as per eCRF. Some subgroups may be pooled based on the number of patients within each subgroup.

Evaluation of interactions

For each pre-defined factor defined in Table 8, PFS will be analyzed using an unstratified Cox proportional hazards model with terms for the factor, treatment and their interaction. The 2-sided p-value of the test of interaction will be provided for descriptive purposes.

Evaluation of confounding

Since the results from the primary analysis could be impacted by confounding factors, any potential issues will be examined and, if confirmed, exploratory analysis of the primary endpoint will be done accordingly. A multivariate Cox proportional hazards model will be used to identify prognostic factors among the demographic and baseline characteristics factors described in the Table 8 using a stepwise selection procedure with a 15% 2-sided significance level for removing effects. For significant prognostic factors identified in the multivariate model, the balance between treatment groups will be assessed. If major confounding is identified through screening for treatment group imbalances in a prognostic factor at baseline, an exploratory analysis of PFS will be done after adjusting for the prognostic factors in the multivariate Cox proportional hazards model. Differences between the adjusted and unadjusted models will be discussed in the clinical study report.

2.4.5.1.4 Exploratory analysis

An exploratory analysis on the PFS according to ICR and to investigator/local radiologist's assessment will be performed by the class of treatment (AI, SERM, SERD) of the control group. Same statistical method as the primary analysis will be used, except that formal statistical testing will not be performed.

2.4.5.2 Analyses of secondary efficacy endpoints

Analysis of response-based endpoints (ie, ORR, DCR, CBR, and DOR) will be performed primarily on the ITT population and supported by the analyses based on the ITT population with measurable disease at study entry (for analysis based on ICR, measurability will be based on ICR assessment of lesions; for analysis based on investigator, measurability will be based on investigator assessment of lesions). The other secondary analyses will be performed on ITT population only.

The results will be presented by treatment group, and also displayed for the control arm according to the class of treatment (AI, SERM and SERD).

Overall survival (OS)

Overall survival (OS) will be evaluated as a key secondary efficacy endpoint. A hierarchical testing strategy will be used to ensure a strong control of the overall Type I error rate at one-sided 2.5%. In other words, comparison between arms on the OS will be performed only if the primary analysis of the PFS is statistically significant.

In case of statistically significant PFS, OS will be compared between treatment arms through a logrank test procedure stratified by the stratification factors as entered in the IRT. Following hypotheses will be tested:

- The null hypothesis that the survival distribution functions (SDF) for the OS of the SAR439859 arm is lower than or equal to the SDF of the control arm.
  - H_0: SDF (SAR439859) $\leq$ SDF (Control)

versus

- The alternative hypothesis that the SDF for the OS of the SAR439859 arm is superior to the SDF of the control arm.
  - H_1: SDF (SAR439859) > SDF (Control)

Otherwise, descriptive statistics of OS will be provided at the time of final PFS analysis.

The COD for final OS analysis will be the date when approximately 196 death events have been observed (approximately 70% of the participants have died).

The following estimates will be provided for OS:

- OS data will be analyzed using the Kaplan-Meier method by treatment group in the ITT population:
  - Kaplan-Meier estimates of the 25th, 50th and 75th percentiles and their associated 95% CIs will be provided. The 95% Cis will be constructed using a log-log transformation of the survival function and the methods of Brookmeyer and Crowley.
  - Number of patients at risk as well as the probabilities of surviving at least 6, 12, 18 and 24 months with 95% CI will be estimated for each treatment group using the Kaplan-Meier method and a log-log approach based on a normal approximation following the Greenwood's formula. Additional timepoints may be considered based on the actual data.
  - Kaplan-Meier curves will be plotted. These plots will include the number of patients at risk at key time points by treatment group.
- The hazard ratio (HR) and its 95% CI will be estimated using the Cox proportional hazards model stratified with the stratification factors as entered in the IRT. Ties will be handled using the exact method. Underlying assumptions of the Cox proportional hazard model will be assessed by graphical methods (ie, log-log graphical methods).
- The number of censored patients, the reasons for their censoring (ie, alive at the cut-off date, alive at the last contact before the cut-off date, and lost to follow-up) and the time between the date of last contact and the cut-off date will be summarized by treatment group.
- Follow-up duration (months) will be defined as the time interval from the date of randomization to the date of last contact (see Section 2.5.3) with the patient. Patients who have died will be censored on their date of death. Kaplan-Meier estimates of the 25th, 50th and 75th percentiles will be provided.

A multivariate Cox proportional hazards model may be used to identify prognostic factors among the demographic and baseline characteristics factors described in the Table 8. In addition, sensitivity analyses adjusting OS for switch to subsequent anti-cancer treatment could also be performed at interim and/or final analyses (eg, using inverse probability of censoring weighting (IPCW) method) (18).

Response-based endpoints

ORR, DCR and CBR according to ICR and investigator assessments (supportive analysis) will be summarized by treatment arm with descriptive statistics at the time of the primary analysis on PFS (based on data collected up to the PFS analysis cut-off date). In addition, 95% two-sided CIs will be computed using the Clopper-Pearson method.

The DOR will only be summarized on the subgroup of participants who have achieved objective response (confirmed CR or PR as BOR). Duration of response by treatment arm will be summarized according to ICR and investigator assessments (supportive analysis), using Kaplan-Meier methods and displayed graphically, if appropriate. The median DOR and associated 95% CI will be provided for each treatment group based on the Kaplan-Meier method using the log-log transformation of the survival function and the methods of Brookmeyer and Crowley. Percentage of responders with DOR greater or equal to different timepoints (eg, 2 months, 4 months, 6 months, etc) and associated 95% CI will also be estimated for each treatment group using the Kaplan-Meier method and a log-log approach based on a normal approximation following the Greenwood's formula.

Of note, the BOR for each participant will also be summarized according to ICR and investigator assessments.

Exploratory analyses of response-based endpoints may be provided according to subgroups of interest (Table 8).

Waterfall plots will be used to display the best relative change from baseline in tumor size observed for each patient. It will be sorted by decreasing order of best relative change from baseline in patients with measurable disease at study entry. BOR will be displayed. Only patients with post-baseline target lesions measurement will be included in the plot.

Progression-free survival according to ESR1 mutation status at baseline

Progression-free survival according to ESR1 mutation status at baseline will be assessed with the same censoring rules as the primary analysis of PFS, and same statistical methods will be used with the exception that no statistical test will be made. Supportive analysis based on the investigator assessments will be performed.

2.4.5.3 Multiplicity issues

Hypothesis testing of the key secondary efficacy endpoint will be carried out. In order to ensure a strong control of the overall Type I error rate at a one-sided 2.5%, a hierarchical testing strategy will be used. In other words, comparison between arms on the OS will be performed only if the primary analysis of the PFS is statistically significant.

2.4.6 Analyses of safety data

The summary of safety results will be presented by treatment group, and also displayed for the control arm according to the class of treatment (AI, SERM and SERD).

General common rules

All safety analyses will be performed on the safety population as defined in Section 2.3.2, unless otherwise specified, using the following common rules:

- Safety data in patients who do not belong to the safety population (eg, exposed but not randomized) will be listed separately
- The analysis of the safety variables will be essentially descriptive and no systematic testing is planned. Relative risks versus control and their 95% confidence intervals may be provided, if relevant

2.4.6.1 Analyses of adverse events

Generalities

The primary focus of adverse event reporting will be on treatment-emergent adverse events. Pretreatment and posttreatment adverse events will be described separately.

The severity grade will be taken into account in the summary. For patients with multiple occurrences of the same adverse event, the maximum (worst) grade by period of observation is used. Summaries will be provided for all grades and for grade ≥3 (including Grade 5). Missing grades, if any, will be included in the "all grades" category.

Sorting within tables ensures the same presentation for the set of all adverse events within the observation period (pretreatment, treatment-emergent, and posttreatment). For that purpose, the table of all treatment-emergent adverse events presented by SOC and PT sorted by the internationally agreed SOC order and decreasing frequency of PTs within SOCs will define the presentation order for all other tables unless otherwise specified. Sorting will be based on results for the SAR439859 arm.

Analysis of all treatment-emergent adverse events

The following treatment-emergent adverse event summaries will be generated for the safety population.

- Overview of treatment-emergent adverse events, summarizing number (%) of patients with any
  - Treatment-emergent AEs (TEAE)
  - Grade ≥3 TEAEs
  - Grade 5 TEAEs (any TEAE with a fatal outcome during the on-treatment period)
  - Serious TEAEs
  - TEAEs leading to treatment discontinuation
  - Treatment-related TEAEs
  - AESI

- All TEAEs by primary SOC, HLGT, HLT, and PT, showing number (%) of patients with at least 1 treatment-emergent adverse event sorted by the SOC internationally agreed order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.
- All TEAEs by primary SOC and PT, showing the number (%) of patients with at least 1 treatment-emergent adverse event, sorted by the internationally agreed SOC order and by decreasing incidence of PTs within each SOC in the SAR439859 arm. This sorting order will be applied to all other tables, unless otherwise specified.
- Most frequent (≥5% of patients in any treatment arm, on PT) TEAE by primary SOC and PT, sorted by the sorting order defined in the TEAEs summary table.
- All treatment-related TEAEs by primary SOC and PT, showing the number (%) of patients with at least 1 treatment-related TEAE, sorted by the sorting order defined in the TEAEs summary table.

Analysis of all treatment emergent serious adverse event(s)

- All serious TEAEs by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients with at least 1 serious TEAEs, sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.
- All serious TEAEs by primary SOC and PT, showing the number (%) of patients with at least 1 serious TEAE, sorted by the order defined in the TEAEs summary table.
- All serious treatment-related TEAEs, by primary SOC and PT, showing the number (%) of patients with at least 1 serious related TEAE, sorted by the order defined in the TEAEs summary table.

Analysis of all treatment-emergent adverse event(s) leading to treatment discontinuation

- All TEAE leading to treatment discontinuation by primary SOC and PT, showing the number (%) of patients with at least 1 TEAE leading to treatment discontinuation, sorted by the order defined in the TEAEs summary table.

Analysis of all treatment-emergent adverse event(s) leading to dose modification

- All TEAEs leading to dose modification by primary SOC and PT, showing the number (%) of patients with at least 1 TEAE leading to dose reduction and/or dose delay/drug interruption (for Fulvestrant), sorted by the sorting order defined in the TEAEs summary table.
- All TEAEs leading to dose reduction by primary SOC and PT, showing the number (%) of patients with at least 1 TEAE leading to dose reduction, sorted by the sorting order defined in the TEAEs summary table.
- A listing of TEAEs leading to dose delay by primary SOC and PT will be provided for patients who take Fulvestrant.
- A listing of TEAEs leading to drug interruption by primary SOC and PT will be provided for patients who take Fulvestrant.

Analysis of adverse events of special interest (AESI)

A listing of patients with at least one AESI mentioned in Section 2.1.4.1 will be provided. This listing will include the category of AESI, the cycle of occurrence, the seriousness and the outcome.

Analysis of pretreatment and posttreatment adverse events

- All pre-treatment AEs by primary SOC and PT, showing the number (%) of patients with at least 1 pre-treatment AE, sorted by the internationally agreed SOC order and decreasing frequency of PT in the SAR439859 arm.
- All post-treatment AEs by primary SOC and PT, showing the number (%) of patients with at least 1 post-treatment AE, sorted by the internationally agreed SOC order and decreasing frequency of PT in the SAR439859 arm.

2.4.6.2 Deaths

The following summaries of deaths will be generated for the safety population.

- Number (%) of patients who died by study period (on-treatment period and post-treatment) and reasons for death (disease progression, AE and other).
- Listing of deaths in non-randomized patients or randomized but not treated patients (this listing will be generated on the screened patients).
- All TEAEs leading to death (regardless of the date of death/period) by primary SOC and PT, sorted by the order defined in the TEAEs summary table.
- Summary of AEs leading to death including fatal TEAEs (Grade 5 during treatment or any grade AE leading to death post-treatment) and Grade 5 post-treatment AEs, presented by primary SOC and PT. These tables will be provided for death occurring:
  - in the context of disease progression (death within 30 days from last study treatment administration and the cause of death is disease progression),
  - in the context other than disease progression (death within 30 days from last study treatment administration and for whom cause of death is not disease progression or the death occurred more than 30 days from last study treatment administration and the cause of death is adverse event).
- Overview of Grade 5 AEs will be provided with the following categories:
  - Grade 5 AE (TEAE and post-treatment).
  - TEAE leading to death (regardless date of death/period).
    - Grade 5 TEAE with a fatal outcome during the treatment period,
    - Any Grade TEAE with a fatal outcome during the post-treatment period.
  - Post-treatment Grade 5 AE (excluding a TEAE that worsened to Grade 5 during the post-treatment period).
- Listing of deaths

2.4.6.3 Analyses of laboratory variables

For clinical laboratory data, baseline is defined as the last sample collection prior to the first administration of the IMP.

Hematology and biochemistry

Hematological and clinical biochemistry toxicities will be assessed from laboratory test parameters defined in Section 2.1.4.3. Each test result will be graded by NCI-CTCAE version 5.0, whenever applicable.

For hematological parameters and for some selected biochemistry parameters, Sanofi sponsor generic ranges (LLN, ULN) are defined and will be used for grading. For other biochemistry parameters (eg, for hepatic enzymes ALT, AST, Alkaline phosphatase, total bilirubin), grading will be derived using local laboratory normal ranges.

The number of patients with abnormal laboratory tests at baseline will be presented by grade and all grades together. The frequency of patients in each grade and all grades of laboratory abnormalities during treatment will be summarized. For patients with multiple occurrences of the same laboratory variable during the treatment, the maximum grade (worst) per patient will be used. The denominator used for percentage calculation is the number of patients with at least 1 evaluation of the laboratory test during the considered observation period. When appropriate, the summary table will present the frequency of patients with any grade of abnormal laboratory tests and with Grade 3-4 abnormal laboratory tests.

In addition, for hematology and biochemistry toxicities, shift tables showing the number of patients in each grade at baseline by worst grade during the on-treatment period will be provided.

Urinalysis

For dipstick analyses, a frequency table of results for each parameter (leukocytes, erythrocytes, proteins, glucose, and ketones) will be provided using the worst value observed on-treatment.

A summary of baseline results will also be provided for dipstick analyses.

For laboratory tests for which NCI-CTCAE V5.0 scale is not applicable, potentially clinically significant abnormalities (PCSA) values are defined as abnormal values considered medically important by the Sponsor according to predefined criteria/thresholds based on literature review.

PCSA criteria will determine which patients had at least 1 PCSA during the on-treatment period, taking into account all evaluations performed during the on-treatment period, including nonscheduled or repeated evaluations. The incidence of PCSA any time during the on-treatment period will be summarized by treatment group irrespective of the baseline level.

2.4.6.4 Analyses of vital sign variables

Vital signs parameters are described in Section 2.1.4.4. Baseline is defined as the last sample collection prior to the first administration of the IMP.

A shift table of baseline ECOG PS versus last and worst ECOG PS on treatment will be provided.

For blood pressure, heart rate and weight parameters, potentially clinically significant abnormality (PCSA) values are defined as abnormal values considered medically important by the Sponsor according to predefined criteria/thresholds based on literature review (Appendix A). PCSA criteria will determine which patients had at least 1 PCSA during the on-treatment period, taking into account all evaluations performed during the on-treatment period, including nonscheduled or repeated evaluations. The number of all such patients will be the numerator for the on-treatment PCSA percentage. The treatment-emergent PCSA denominator by group for a given parameter will be based on the number of patients assessed for that given parameter in the treatment-emergent adverse event period by treatment group on the safety population

The incidence of PCSAs at any cycle during the on-treatment period will be summarized by treatment group irrespective of the baseline and/or according to the following baseline categories:

- Normal/missing.
- Abnormal according to PCSA criterion or criteria.

In addition, for blood pressure and heart rate a graph describing mean changes from baseline and associated +/- SEM will also be done throughout the on-treatment period.

2.4.6.5 Analyses of electrocardiogram variables

For ECG, baseline is defined as the last sample collection prior to the first administration of the IMP. The incidence of patients with at least 1 abnormal ECG at any time during the on-treatment period will be summarized irrespective of the baseline level and/or according to the following baseline status categories:

- Normal/missing
- Abnormal

2.4.7 Analyses of pharmacokinetic and pharmacodynamic variables

2.4.7.1 Analysis of pharmacokinetic variables

Plasma concentrations will be summarized for SAR439859 treatment arm by time point using descriptive statistics (such as the number of observations, arithmetic and geometric mean, median, standard deviation (SD), coefficient of variation (CV)%, minimum, and maximum).

In addition, steady state C_{trough} concentrations, defined as the median C_{trough} of individual patient across study duration (plasma concentrations of predose samples at Day 15 Cycle 1 and Day 1 Cycles 2, 3, 4 and 6), will be summarized using standard descriptive statistics.

Plasma concentrations will be included in the descriptive statistics if actual sampling occurs in the following time windows: within 2 hours before administration for predose samples on Day 15 Cycle 1 and Day 1 Cycles 2, 3, 4 and 6; [1-2h] and [3-5h] for samples at 1.5 and 4h post dose, respectively on Day 1 Cycle 1 and Day 1 Cycle 2; [7-9h] for samples at 8h post dose on Day 1 Cycle 2.

In addition, a listing of individual concentration data by cycle and visit day will be provided.

Concentrations reported as below the limit of quantification (BLQ) will be replaced by 0 for descriptive statistics. Geometric mean will not be computed in case at least one concentration is below LLOQ.

Descriptive statistics and graphical presentation of SAR439859 PK parameters of interest may be done in regards of safety/efficacy endpoints.

2.4.7.2 Analysis of Pharmacodynamics/genomic and biomarkers variables

The main analyses of pharmacodynamics/genomics and biomarker variables are provided in this section. Additional exploratory biomarker analyses may be conducted depending on information obtained during SAR439859 studies or from the literature. This may be described in a separated SAP.

Baseline for biomarker data is defined as the last sample collection prior to the first administration of the IMP.

Estrogen receptor 1 gene mutation status in circulating free DNA

Descriptive statistics of the ESR1 mutational status (wild-type or mutated: at least one ESR1 mutation detected) and each of the twelve mutations will be provided at baseline and over time by comparing baseline and Cycle 3 Day 1 for each treatment arm. The evolution of mutant allele frequency and concentration will be also described.

In addition to the association with the PFS (Section 2.4.5), the ESR1 mutational status and/or the presence of some specific mutations may be described for other efficacy endpoints in the two treatment arms. The ESR1 mutational status may be also described according to prior anti-cancer therapies. On-treatment changes in ESR1 mutation may be also associated with efficacy endpoints for exploratory purpose.

Mutational profiling in circulating free DNA

Descriptive statistics of the genomic aberrations of the 77 genes will be provided at baseline for each treatment arm and by prior anti-cancer therapies. The baseline mutant allele frequency and concentration for some genes of interest may be also described.

The relationship between some genes of interest and efficacy endpoints (eg, PFS) may be explored in the two treatment arms.

Estrogen receptor degradation and tumor biopsy biomarkers

Descriptive statistics of the baseline and the change over time of tumor biomarkers such as ER expression, Ki67, Bcl-2, and PgR will be provided. The relationship between the change of each protein expression and efficacy endpoints (eg, PFS) may be explored in the two treatment arms. Some tumor gene signatures (such as the ER activation signature and other gene signature of interest) may be also derived and described.

Estradiol

Descriptive statistics of the circulating level of estradiol will be provided at baseline and over time by comparing baseline and Cycle 3 Day 1 for each treatment arm. The relationship between baseline circulating levels of estradiol and clinical response may be explored in the two treatment arms. On-treatment changes in circulating level of estradiol may be also associated with efficacy endpoints for exploratory purpose.

2.4.8 Analyses of quality of life

Patient reported outcomes endpoints for each of the 3 selected PRO/HRQL and health utility instruments (EORTC QLQ-C30, QLQ-BR23, and EQ-5D-5L) will be analyzed in participants from the safety population. The baseline value is defined as the last questionnaire assessment before or on the first IMP date.

Compliance rate

For each questionnaire the compliance profile over time will be summarized based on the cycle reported in the eCRF on the safety population (number and percentage of forms received versus expected, and number and percentage of forms evaluable versus expected). A questionnaire is considered received if at least one item on the form is completed. A questionnaire is expected as defined in the SoA (Section 1.3 of the protocol) and based on the number of cycles started by the patient. A questionnaire is evaluable if at least one scale is complete according to scale developer scoring algorithms (Appendix F).

Change from baseline analysis

For the QLQ-C30 (15 total scales), QLQ-BR23 (8 scales), and EQ-5D-5L (health index and visual analogue scale) instruments, descriptive statistics on the absolute value and changes from baseline will be done for each treatment arm at each analysis visit, at EOT and follow-up based on the safety population. Analysis visit will be derived based on the study day of assessment, defined as date of assessment – first IMP date + 1.

Table 9 - Analysis visit for PRO analyses

Planned visit	Planned study day	Analysis visit study day window	Analysis visit
C1D1	1	Before or on first IMP date	Baseline
C2D1	29	[2;42]	C2D1
C3D1	57	[43;70]	C3D1
C4D1	85	[71;112]	C4D1
C6D1	141	[113;168]	C6D1
C8D1	197	[169;224]	C8D1
CxD1	(x-1)*28+1	[(x-1)*28+1-28; (x-1)*28+1+27]	CxD1
EOT	IMP date + 30	[last IMP date+1; min(last IMP date+35; start of further anticancer therapy)]	EOT
FUP	IMP date + 90	[min(last IMP date+35; start of further anticancer therapy)+1; last IMP date+95]	FUP

Note: only assessments performed up to and including the last IMP can be mapped into a CxD1 analysis visit.

In the case of multiple assessments falling in the same analysis visit window, the assessment closest to the planned study day will be used. If 2 assessments falling within the same analysis visit window are equidistant from the planned study day, then the assessment obtained prior to target date will be used.

For each scale of each instrument, comparison between the two treatment arms will be based on a longitudinal repeated measures analysis using a mixed effects model on the change from baseline in patients from the safety population who have completed the baseline and at least 1 postbaseline on-treatment assessment. Only on-treatment assessments will be included in the model (ie, up to but excluding EOT). Fewer analysis visits (eg, visits with at least 10 participants in each arm, visits with at least 15 participants in each arm, etc) may be considered in case of model convergence issue after unblinding. The variables included in the model will be treatment, time, treatment-by-time, with baseline value and stratification factors used as covariates. Time will be considered as categorical variable based on the analysis visit defined in Table 9. Parameter estimates will be initially based on a restricted maximum likelihood method and an unstructured covariance matrix will be used. The structure of the correlation matrix will be investigated and simplified using likelihood ratio tests if appropriate. The following covariance structure may be investigated: Toeplitz, Autoregressive(1). The Kenward-Roger approximation will be used to estimate denominator degrees of freedom. The least square means (LS means) and associated standard error (SE) for each treatment arm and differences in LS means with the corresponding 2-sided 95% CI and 2-sided p-values will be presented at each analysis visit (for descriptive purposes only). A graphical display of the LS means (SE) of the changes from baseline for each treatment arm over time will also be provided for each scale of each instrument. Overall LS means and associated SE for each treatment arm, overall difference in LS means (95% CI) and associated

2-sided p-value (for descriptive purposes only) will be provided along with the 2-sided p-value of the interaction term treatment-by-time.

For each of the EOT and FUP visit, comparison between the two treatment arms will be based on a linear model for the change from baseline with treatment, baseline value and stratification factors used as covariates. LS means (SE) for each treatment arm and differences in LS means (95% CI) will be presented. Two-sided p-values will be presented (for descriptive purposes only).

As p-values will be presented for descriptive purposes only, no statistical hypotheses will be tested and no adjustments for multiple comparisons will be made.

Time to sustained deterioration in EORTC QLQ-C30 and EORTC QLQ-BR23 scales

The time to sustained deterioration for each EORTC QLQ-C30 and QLQ-BR23 scale will be analyzed using the Kaplan-Meier method by treatment group in patients from the safety population who have completed the baseline and at least 1 postbaseline on-treatment assessment, and who have baseline scales scores that allow for a 10-point deterioration (ie, no greater than 90 for symptom scales, and no lower than 10 for the functional scales and GHS/QoL):

- Kaplan-Meier estimates of the 25th, 50th and 75th percentiles and their associated 95% CI) will be provided. The 95% CIs will be constructed using a log-log transformation of the survival function and the methods of Brookmeyer and Crowley.
- Number of patients at risk as well as the probabilities of having no deterioration at least 6, 12, 18 and 24 months with 95% CIs will be estimated for each treatment group using the Kaplan-Meier method and a log-log approach based on a normal approximation following the Greenwood's formula.

The number (%) of censored patients and the reasons for their censoring will be summarized by treatment group.

The Kaplan-Meier curves will be plotted. These plots will include the number of patients at risk at key time points by treatment group.

The hazard ratio (HR) for sustained deterioration and its 95% CIs may be estimated using the Cox proportional hazards model stratified by the stratification factors as entered in the IRT. Ties will be handled using the exact method.

2.4.9 Further therapy after discontinuation of investigational medicinal product administration during the study

A summary table will be provided for further therapies based on WHO-DD coding. Similar analysis will be performed for further radiotherapy and further surgery.

Time to first use of chemotherapy after disease progression or treatment discontinuation (without disease progression)

The following estimates will be provided for the time to first use of chemotherapy after disease progression or treatment discontinuation (without disease progression):

- Time to first use of chemotherapy after disease progression or treatment discontinuation (without disease progression) data will be analyzed using the Kaplan-Meier method by treatment group in the ITT population:
  - Kaplan-Meier estimates of the 25th, 50th and 75th percentiles and their associated 95% CIs will be provided. The 95% CIs will be constructed using a log-log transformation of the survival function and the methods of Brookmeyer and Crowley.
  - Number of patients at risk as well as the probabilities of being free of use of chemotherapy at least 6, 12, 18 and 24 months with 95% CI will be estimated for each treatment group using the Kaplan-Meier method.
- The number of censored patients, the reasons for their censoring (ie, alive without use of chemotherapy, death, and lost to follow-up) will be summarized by treatment group.
- Kaplan-Meier curves will be plotted.

2.5 DATA HANDLING CONVENTIONS

2.5.1 General conventions

The following formulas will be used for computation of parameters.

Time unit

A month length is 30.4375 days (365.25 / 12). If duration is to be reported in months, duration in days is divided by 30.4375. If duration is to be reported in years, duration in days will be divided by 365.25.

Duration

Unless otherwise specified, difference between two dates (Date A ≤ Date B) will be calculated as follows:

$$\text{Duration (days)} = \text{Date B} - \text{Date A} + 1$$

2.5.2 Data handling conventions for primary efficacy variables

The following formulas will be used for computation of PFS endpoint.

Date of tumor assessment

It is acknowledged that an assessment may include several methods of evaluation performed over a period of several days within a window of time around an expected assessment date. For each tumor assessment, a single date will be derived according to the overall response of that assessment.

The date of target lesion(s) assessment will be derived based on the date of the last assessment of target lesion(s).

When the overall response is different from PD, the date of tumor assessment is defined as the date of the last evaluation included in the series of evaluations performed within that time point (ie, target lesion(s), non-target lesions(s) and new lesion(s) assessments).

When the overall response is PD, the date of tumor assessment is the date when progression was first demonstrated according to the target lesion(s), non-target lesion(s) and new lesion(s), as specified below:

- For progression based on new lesion(s) the date of progression is the earliest date a new lesion has been detected.
- For progression based on non-target lesion(s), the date of progression is the earliest date a non-target lesion was considered as PD.
- For progression based on target lesion(s), the date of progression is the date of the last assessment of target lesion(s).

If progression is based on several events within the same tumor assessment (eg, new lesion(s) seen along with target lesion(s) progression), the earliest date of progression, according to the rules listed above, will be the date of assessment.

Evaluable tumor assessment

An evaluable tumor assessment is defined as a tumor assessment with an overall response different of non-evaluable (NE).

Date of documented progression

The date of documented progression is defined as the first date of tumor assessment at which the overall response was recorded as progressive disease.

Date of death

The date of death is defined as the date of death recorded in the eCRF.

Date of next scheduled assessment

The date of next scheduled assessment is the date of the next scheduled tumor assessment as per protocol schedule of assessment. This date may be used if back-dating is considered when the event occurred beyond the acceptable time window for the next tumor assessment (eg, 1 or 2 non-evaluable tumor assessment).

Date of last evaluable tumor assessment

The date of the last evaluable tumor assessment is defined as the last date of tumor assessment at which the overall response was recorded as CR, PR, SD or Non-CR/Non-PD before a censoring reason occurred.

2.5.3 Data handling conventions for secondary efficacy variables

The following formulas will be used for computation of secondary endpoint.

Best relative change from baseline in tumor size

Tumor size is defined as the sum of the longest diameters of the target lesions as per RECIST 1.1. It can be calculated only for measurable patients.

Relative change from baseline in tumor size at tumor assessment *t* will be calculated as follows:

> Relative change (%) from baseline in tumor size (*t*) = 100*(Tumor Size(*t*) – Tumor Size(*baseline*))/ Tumor Size(*baseline*)

Best relative change from baseline in tumor size will be the smallest relative change from baseline in tumor size.

Date of first documented response

The date of the first documented response is defined as the date of the first tumor assessment at which the overall response was recorded as CR or PR that is subsequently confirmed.

Date of last contact

The last contact date is derived for patients not known to have died at the analysis cut-off date based on the latest date among the following:

- Date of visits
- Assessment dates (eg, laboratory, vital signs, ECOG performance status, ECG, tumor assessment, PK assessment, EOT completion etc).
- Medication and procedures dates including study medication, concomitant medications, surgical and medical procedures, further anti-cancer therapies administered after treatment discontinuation.
- Adverse event start and end dates
- "Date of Last Available Information" collected on the "Subject status" page
- Study treatment start/end date
- Randomization date

The last contact date is defined as the latest date from the above list or the cut-off date, whichever comes first. The last contact date could be used for censoring of patients in time to event analysis.

2.5.4 Missing data

The analyses and summaries of continuous and categorical variables will be based on observed data only. Percentages will be calculated using as denominator the number of patients with non-

missing observation in the considered population. When relevant, the number of patients with missing data is presented.

Handling of disease characteristics missing/partial dates

- If the day is missing, it will be imputed by 1.
- If the month is missing, it will be imputed by 1 (only for medical history variables).
- If the year is missing, no imputation will be performed.

Incomplete date of cancer diagnosis:

- If the day of the cancer diagnosis is missing, the date will be imputed to the first day of the month.
- If day and month of the cancer diagnosis are missing, no imputation will be done.

Handling of medication missing/partial dates

No imputation of medication (other than anti-cancer therapies) start/end dates or times will be performed. If a medication date or time is missing or partially missing and it cannot be determined whether it was taken prior or concomitantly, it will be considered a prior, concomitant, and post-treatment medication.

For prior anti-cancer therapies, following rules will be applied:

- Incomplete start date of prior anti-cancer therapy: if the day of the start date of the prior regimen is missing, the date will be imputed to the first day of the month; if the month is missing, the date will be imputed to the first month of the year.
- Incomplete end date of prior anti-cancer therapy: if the day of the end date of last prior regimen is missing, the date will be imputed to the last day of the month; if the month is missing, the date will be imputed to the last month of the year.

Imputation of incomplete date for post anti-cancer treatment start date

For post anti-cancer treatments, if the medication start date is missing, it will be imputed as follows:

- If the medication start day and month are missing and the medication start year is the same as treatment end year, the medication start date will be set equal to treatment end date + 1.
- If the medication start day and month are missing and the medication start year is after the treatment end year, the medication start day and month will each be set to 01.
- If the medication start day is missing and medication start year and month is the same as the treatment end year and month, the medication start date will be set to the treatment end date + 1.
- If the medication start day is missing and medication start month is before the treatment end month and the medication start year is the same as treatment end year, the medication start day will be set to 01.

- If the medication start day is missing and the medication start month is after the treatment end month and the medication start year is the same as treatment end year, the medication start day will be set to 01.
- If the medication start day is missing and the medication start month is not missing and the medication start year is after the treatment end year, the medication start day will be set to 01.
- If the medication start day, start month and start year is missing, the medication start date will be set equal to the treatment end date + 1.

Handling of other missing dates

Incomplete date of progression for the last prior regimen:

- If the day of the progression for the last prior regimen is missing, the date will be imputed to the to the earliest of the end day of the month and the randomization date
- If day and month of the progression for the last prior regimen are missing, no imputation will be done.

Incomplete date of prior surgery:

- If the day of the prior surgery is missing, the date will be imputed to the earliest of the end day of the month and the randomization date.
- If day and month of the prior surgery are missing, no imputation will be done.

Incomplete end date of prior radiotherapy:

- If the day of the end date of the prior radiotherapy is missing, the date will be imputed to the earliest of the end day of the month and the randomization date.
- If the day and month of the end date of the prior radiotherapy are missing, no imputation will be done.

Handling of questionnaires dates

If the date is incomplete, it will be imputed by:

- If the day is missing, it will be imputed by the day of the visit date if year and month are the same. Otherwise, it will be imputed by 1 if same year but a later month compared to the visit date or to the end day of the month if same year but an earlier month.
- If the day and month are missing, it will be imputed by the day and the month of the visit date if years are the same. Otherwise, if the year of visit is later, missing date will be imputed by the last day of the year of the questionnaire; if the year of visit is earlier, missing date will be imputed by the first day of the year of the questionnaire.
- In other cases of incomplete date or missing date, it will be imputed by the visit date.

Handling of computation of treatment duration if investigational medicinal product end of treatment date is missing

For the calculation of the treatment duration, the date of the last dose of IMP is equal to the date of last administration reported on the end-of-treatment case report form page. If this date is missing, the exposure duration should be left as missing.

Handling of adverse events when date and time of first investigational medicinal product administration is missing

When the date and time of the first IMP administration is missing, all adverse events that occurred on or after the day of randomization should be considered as treatment-emergent adverse events.

Missing grade

If the grade is missing for one of the treatment emergent occurrences of an AE, the maximal severity on the remaining occurrences will be considered. If the severity is missing for all the occurrences, no imputation will be done and missing grades will be summarized in the "all grades" category.

Handling of missing assessment of relationship of adverse events to investigational medicinal product

If the assessment of the relationship to IMP is missing, then the relationship to IMP has to be assumed and the adverse event considered as such in the frequency tables of possibly related adverse events, but no imputation should be done at the data level.

Handling of potentially clinically significant abnormalities

If a patient has a missing baseline, he/she will be grouped in the category "normal/missing at baseline."

For PCSAs with 2 conditions, one based on a change from baseline value or a normal range and the other on a threshold value, with the first condition being missing, the PCSA will be based only on the second condition.

For a PCSA defined on a threshold and/or a normal range, this PCSA will be derived using this threshold if the normal range is missing; eg, for eosinophils the PCSA is >0.5 GIGA/L or > ULN if ULN ≥0.5 GIGA/L. When ULN is missing, the value 0.5 should be used.

Measurements flagged as invalid by the laboratory will not be summarized or taken into account in the computation of PCSA values.

Handling of Estradiol data below quantification limit

For Estradiol data with values below the quantification limit, the data will be imputed to half of the quantification limit.

2.5.5 Unscheduled visits

Unscheduled visit measurements of laboratory data, vital signs, and ECG will be used for computation of baseline and worst values and/or grades.

2.5.6 Pooling of centers for statistical analyses

Data from all sites will be pooled together for analyses.

2.5.7 Statistical technical issues

Not applicable.

3 INTERIM ANALYSIS

3.1.1 Interim analysis for progression-free survival

An interim analysis on PFS is planned for futility at 50% of the planned total number of PFS events (ie, approximately 101 PFS events). Without possible claim for overwhelming evidence of efficacy, no Type I error rate adjustment will be performed. The stopping boundary for futility is based on the observed HR based on Cox proportional hazard model, ie, an HR >1.1. If the observed HR >1.1, then the study may be stopped for futility.

A summary of the PFS analyses is provided in Table 10. The observed HR based on Cox proportional hazard model stratified by the stratification factors as entered in the IRT system will be compared with the futility boundary. As described in Section 2.4.5.1, PFS data will be analyzed using the Kaplan-Meier method by treatment group; HR and its 95% CIs will be estimated; the type of PFS event will be summarized; for patients who died without documented progression, the time from the last evaluable disease assessment to the death will be summarized; the number of censored patients, the reason and timing of censoring and the time from the last evaluable disease assessment to the cut-off date will be summarized; median follow-up duration will be estimated. Supportive PFS analysis according to investigator/local radiologist's assessment, with the same censoring rules as for the PFS primary analysis, will be provided.

Demographics and baseline characteristics, prior or concomitant medication, extent of IMP exposure and compliance, AEs (TEAE, death, SAE, TEAE leading to discontinuation, AESI), and laboratory variables (abnormality of hematological and chemistry test) will be analyzed at interim analysis. If needed, more data will be analyzed to inform the futility decision including additional efficacy, PK, and biomarker.

Table 10 - Progression-free survival analyses

Analysis	Months after FPI (approx. under PFS HR=0.65)	Planned accrual	Number of events (under PFS HR=0.65)	Information fraction	Cumulative Power (under PFS HR=0.65)	Futility boundary	Efficacy boundary
PFS IA (futility only)	10	239	101	50%		HR>1.1	NA
PFS Final analysis	18	282	201	100%	86%	p >0.025 (HR[a] >0.758)	p ≤0.025 (HR[a] ≤0.758)

a HR is provided only for information purposes. The final decisions will be based on p-values.
Note: numbers have been rounded. Calculations were made using East 6.5 software.
FPI = first participant in; HR = hazard ratio; IA = interim analysis; NA = not applicable; PFS = progression-free survival.

3.1.2 Interim analysis for overall survival

Comparison between arms on the OS will be performed only if the primary analysis of the PFS is statistically significant. Therefore, a maximum of 2 analyses are planned for OS: at the time of the primary analysis of PFS and at the final OS analysis.

A gamma error spending function ($\gamma = -8$) will be used, along with the hierarchical testing strategy to strongly control the family-wise error rate (FWER; overall Type I error rate). If the value of the test statistic exceeds the efficacy boundary ($z \leq -3.716$, $p \leq 0.0001$, assuming 63 OS events are observed), superiority of OS will be claimed but survival data will continue to be collected until the end of study when approximately 196 death events have been observed. The actual efficacy boundary will be adjusted according to the observed number of events at interim analysis.

A summary of the OS analyses is provided in Table 11. The test statistic obtained from a logrank test procedure stratified by the stratification factors as entered in the IRT system will be compared with the futility boundary. As described in Section 2.4.5.2, OS data will be analyzed using the Kaplan-Meier method by treatment group; HR and its 95% CIs will be estimated; the type of OS event will be summarized; the number of censored patients, reasons for censoring (ie, alive at the cut-off date, alive at the last contact before the cut-off date, and lost to follow-up) and the time between the date of last contact and the cut-off date will be summarized by treatment group.

Table 11 - Overall survival analyses

Analysis	Months after FPI (approx.)	Planned accrual	Number of deaths (approx.)	Information fraction	Cumulative Power[a] (under HR=0.75)	Futility boundary	Efficacy boundary
OS IA (at PFS final analysis)	18	282	63	32%	0.5%	NA	p ≤0.0001 (HR[b] ≤0.392)
Final analysis	64	282	196	100%	52%	p >0.0249 (HR[b] >0.756)	p ≤0.0249 (HR[b] ≤0.756)

a Marginal power conditional to statistically significance of PFS.
b HR is provided only for information purposes. The interim and final decisions will be based on p-values.
Note: number have been rounded. Calculations were made using East 6.5 software. Assume a 5% dropout rate at 64 months after first participant randomized.
FPI = first participant in; HR = hazard ratio; IA = interim analysis; NA = not applicable; OS = overall survival.

3.1.3 Data Monitoring Committee

This study will use an independent DMC. The first DMC meeting will be set up to review early safety results (eg, after approximately 25 participants have completed at least 2 cycles in the SAR439859 arm, or after 6 months after first participant randomized), and then periodically. Ad hoc DMC meetings may also be held if a significant safety issue or an issue deemed important for discussion arises on this or other SAR439859 studies. After each meeting, the DMC will make recommendations to the Sponsor's representatives regarding the continued safety of treating ongoing and future study participants, as well as the course of action regarding the conduct of the study. The DMC will also oversee the interim analyses on PFS detailed in Section 3.1.1.

4 DATABASE LOCK

Estimated COD will be approximately 18 months after first randomized participant for the PFS analysis. No pharmacokinetic sample will be taken after Cycle 6 or PFS COD, whichever comes first. After COD for PFS analysis, no more efficacy assessment will be performed except collection of survival status. The COD for analysis of OS will be approximately 64 months after the first patient randomization.

5 SOFTWARE DOCUMENTATION

All summaries and statistical analyses will be generated using SAS version 9.4 or higher.

Biomarkers analyses will be generated using R version 3.3.2.

6 REFERENCES

1. Cristofanilli M, Turner NC, Bondarenko I, Ro J, Im SA, Masuda N, et al. Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. Lancet Oncol. 2016;17(4):425-39.

2. Kornblum N, Zhao F, Manola J, Klein P, Ramaswamy B, Brufsky A, et al. Randomized phase II trial of fulvestrant plus everolimus or placebo in postmenopausal women with hormone receptor-positive, human epidermal growth factor receptor 2-negative metastatic breast cancer resistant to aromatase inhibitor therapy: results of PrE0102. J Clin Oncol. 2018;36(16):1556-63.

3. Chia S, Gradishar W, Mauriac L, Bines J, Amant F, Federico M, et al. Double-blind, randomized placebo controlled trial of fulvestrant compared with exemestane after prior nonsteroidal aromatase inhibitor therapy in postmenopausal women with hormone receptor-positive, advanced breast cancer: Results from EFECT. J Clin Oncol. 2008;26(10):1664-70.

4. Yardley DA, Noguchi S, Pritchard KI, Burris HA 3rd, Baselga J, Gnant M, et al. Everolimus plus exemestane in postmenopausal patients with HR(+) breast cancer: BOLERO-2 final progression-free survival analysis. Adv Ther. 2013;30:870-84.

5. Piccart M, Hortobagyi GN, Campone M, Pritchard KI, Lebrun F, Ito Y, et al. Everolimus plus exemestane for hormone-receptor-positive, human epidermal growth factor receptor-2-negative advanced breast cancer: overall survival results from BOLERO-2. Ann Oncol. 2014;25:2357-62.

6. Johnston SR, Kilburn LS, Ellis P, Dodwell D, Cameron D, Hayward L, et al. Fulvestrant plus anastrozole or placebo versus exemestane alone after progression on non-steroidal aromatase inhibitors in postmenopausal patients with hormone-receptor-positive locally advanced or metastatic breast cancer (SoFEA): a composite, multicentre, phase 3 randomised trial. Lancet Oncol. 2013;14:989-98.

7. Aaronson NK, Ahmedzai S, Bullinger M, Crabeels D, Estape J, Filiberti AN, et al. The EORTC core quality of life questionnaire: interim results of an international field study. In: OSOBA D, editor. Effect of cancer on quality of life. Vancouver. 1991;185-202.

8. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365-7.

9. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 Scoring Manual (3rd edition). Brussels: EORTC, 2001.

10. Sprangers MA, Groenvold M, Arraras JI, Franklin J, te Velde A, Muller M, et al. The European Organization for Research and Treatment of Cancer breast cancer-specific quality-of-life questionnaire module: first results from a three-country field study. J Clin Oncol. 1996;14(10):2756-68.

11. Van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15:708-15.

12. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol. 1998;16(1):139-44.

13. Harbeck N, Iyer S, Turner N, Cristofanilli M, Ro J, Andre F, et al. Quality of life with palbociclib plus fulvestrant in previously treated hormone receptor-positive, HER2-negative metastatic breast cancer: patient-reported outcomes from the PALOMA-3 trial. Annals of Oncology. 2016;27(6):1047-54.

14. Kaufman PA, Toi M, Neven P, Sohn J, Grischke EM, Andre V, et al. Health‐related quality of life in MONARCH 2: Abemaciclib plus fulvestrant in hormone receptor‐positive, HER2‐negative advanced breast cancer after endocrine therapy. The Oncologist. 2020;25(2):e243.

15. Verma S, O'Shaughnessy J, Burris HA, Campone M, Alba E, Chandiwana D, et al. Health-related quality of life of postmenopausal women with hormone receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer treated with ribociclib+ letrozole: results from MONALEESA-2. Breast cancer research and treatment. 2018;170(3):535-45.

16. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol. 2013;13(1):152.

17. Amit O, Mannino F, Stone AM, Bushnell W, Denne J, Helterbrand J, et al. Blinded independent central review of progression in cancer clinical trials: results from a metaanalysis. Eur J Cancer. 2011;47(12):1772-8.

18. Watkins C, Huang X, Latimer N, Tang Y, Wright EJ. Adjusting overall survival for treatment switches: commonly used methods and practical application. Pharm Stat. 2013;12(6):348-57.

19. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). 2009;45(2):228-47.

7 LIST OF APPENDICES

Appendix A Potentially clinically significant abnormalities criteria

Parameter	PCSA	Comments
Vital signs		
HR	≤50 bpm and decrease from baseline ≥20 bpm ≥120 bpm and increase from baseline≥20 bpm	To be applied for all positions (including missing) except STANDING.
SBP	≤95 mmHg and decrease from baseline ≥20 mmHg ≥160 mmHg and increase from baseline ≥20 mmHg	To be applied for all positions (including missing) except STANDING.
DBP	≤45 mmHg and decrease from baseline ≥10 mmHg ≥110 mmHg and increase from baseline ≥10 mmHg	To be applied for all positions (including missing) except STANDING.
Weight	≥5% increase from baseline ≥5% decrease from baseline	FDA Feb 2007.
Laboratory		
Blood Urea Nitrogen	≥17 mmol/L	
Chloride	<80 mmol/L >115 mmol/L	
Basophils	>0.1 Giga/L	
Monocytes	>0.7 Giga/L	
RBC	≥6 Tera/L	
Hematocrit	≤0.37 v/v (Male); ≤0.32 v/v (Female) ≥0.55 v/v (Male); ≥0.5 v/v (Female)	
pH	≤4.6 ≥8	

Appendix B Summary of statistical analyses

EFFICACY ANALYSIS

Endpoint	Analysis population	Primary analysis	Sensitivity analysis	Supportive analysis	Subgroup analysis	Other analyses
Primary endpoint						
PFS according to ICR	ITT	Stratified Log-rank test Kaplan-Meier Cox proportional hazards model	#1 Ignoring further anti-cancer therapy #2 PFS considering events occurring after two or more non-evaluable tumor assessment as event and ignoring further anti-cancer therapy #3 PFS considering events occurring after two or more non-evaluable tumor assessment as event and back-dating at the next schedule assessment and ignoring further anti-cancer therapy #4 PFS analysis using stratification factors derived from eCRF data	- PFS according to investigator assessment (same censoring rules as for the PFS primary analysis, sensitivity analysis #2, PFS considering the clinical/non-radiological progression as event, and same rules as the sensitivity analyses if relevant) - Concordance of PFS outcome (PFS ODR) - Concordance of tumor assessment evaluation (EDR and LDR) - RMST if non-proportional hazard assumption is suspected	Yes	Multivariate Cox proportional hazards model
Secondary endpoints						
OS	ITT	Stratified Log-rank test Kaplan-Meier Cox proportional hazards model	No	IPCW to account for informative censoring due to further anti-cancer therapy, if relevant	Yes if relevant based on subgroups of primary endpoint	Multivariate Cox proportional hazards model

Endpoint	Analysis population	Primary analysis	Sensitivity analysis	Supportive analysis	Subgroup analysis	Other analyses
ORR according to ICR	ITT ITT measurable	Clopper-Pearson 95% CI	No	ORR according to investigator	Yes if relevant based on subgroups of primary endpoint	- Best Overall Response - Best relative change from baseline in tumor size (waterfall plot)
DCR according to ICR	ITT ITT measurable	Clopper-Pearson 95% CI	No	DCR according to investigator	Yes if relevant based on subgroups of primary endpoint	
CBR according to ICR	ITT ITT measurable	Clopper-Pearson 95% CI	No	CBR according to investigator	Yes if relevant based on subgroups of primary endpoint	
DOR according to ICR	ITT, for patients who have achieved confirmed CR or PR as BOR ITT measurable	Kaplan-Meier	No	DOR according to investigator	Yes if relevant based on subgroups of primary endpoint	
PFS according to ESR1 mutation status at baseline based on ICR assessment	ITT	Kaplan-Meier Cox proportional hazards model	No	PFS according to ESR1 mutation status at baseline based on investigator assessment	No	
Pharmacokinetics of SAR439859	Pharmacokinetic-evaluable population	Descriptive statistics	No	No	No	PK parameters according to safety/efficacy endpoints

Endpoint	Analysis population	Primary analysis	Sensitivity analysis	Supportive analysis	Subgroup analysis	Other analyses
Tertiary/exploratory endpoints						
ESR1 mutation status	ITT	Descriptive statistics				
Mutation profile	ITT	Descriptive statistics				Relationship between the mutation status and clinical response
ER degradation and tumor biopsy biomarkers	ITT	Descriptive statistics				Relationship between the protein expression and clinical response
Transcriptome profile (RNA)	ITT	Descriptive statistics				
Time to first use of chemotherapy after disease progression or treatment discontinuation (without disease progression)	ITT	Kaplan-Meier				

BOR: Best overall response, CI: Confidence interval, CR: Complete response, EDR: Early discrepancy rate, ITT: intent-to-treat, LDR: Late discrepancy rate, ODR: Outcome discrepancy rate, ORR: Overall response rate, OS: overall survival, PFS: Progression free survival, PR: Partial response
Stratified analyses are performed with the stratification factors as entered in the IRT ie, presence of visceral metastasis, prior treatment with CDK4/6 inhibitors and ECOG.

SAFETY ANALYSES

Endpoint	***Analysis population***	***Primary analysis***	***Supportive analysis***	***Subgroup analysis***
Adverse events	Safety	Descriptive statistics by treatment group	Descriptive statistics by treatment arm and according to the class of therapy of the control arm	No
Deaths	Safety	Descriptive statistics by treatment group	Descriptive statistics by treatment arm and according to the class of therapy of the control arm	No
Laboratory	Safety	Descriptive statistics by treatment group	Descriptive statistics by treatment arm and according to the class of therapy of the control arm	No
Vital signs	Safety	Descriptive statistics by treatment group	Descriptive statistics by treatment arm and according to the class of therapy of the control arm	No
ECG	Safety	Descriptive statistics by treatment group	Descriptive statistics by treatment arm and according to the class of therapy of the control arm	No
Health-related quality of life (HRQL): EORTC QLQ-C30, EORTC QLQ-BR23, EQ-5D-5L	Safety	Descriptive statistics of compliance rate by treatment group Descriptive statistics by treatment group	No	No
Health-related quality of life (HRQL): EORTC QLQ-C30 and EORTC QLQ-BR23 scales, EQ-5D-5L index score and VAS	Safety patients who have completed the baseline and at least 1 post-baseline assessment	Mixed model for repeated measures on the change from baseline in each scale	No	No
Health-related quality of life (HRQL): Time to sustained deterioration in EORTC QLQ-C30 and EORTC QLQ-BR23 scales	Safety patients who have completed the baseline and at least 1 post-baseline assessment and who have baseline scales scores that allow for a 10-point deterioration	Kaplan-Meier	No	No

Appendix C Internationally agreed SOC order

The internationally agreed order (Guideline on summary of product characteristics, December 1999, European commission) for SOC:

1. Infections and infestations
2. Neoplasms benign and malignant (including cysts and polyps)
3. Blood and the lymphatic system disorders
4. Immune system disorders
5. Endocrine disorders
6. Metabolism and nutrition disorders
7. Psychiatric disorders
8. Nervous system disorders
9. Eye disorders
10. Ear and labyrinth disorders
11. Cardiac disorders
12. Vascular disorders
13. Respiratory, thoracic and mediastinal disorders
14. Gastrointestinal disorders
15. Hepato-biliary disorders
16. Skin and subcutaneous tissue disorders
17. Musculoskeletal, connective tissue and bone disorders
18. Renal and urinary disorders
19. Pregnancy, puerperium and perinatal conditions
20. Reproductive system and breast disorders
21. Congenital and familial/genetic disorders
22. General disorders and administration site conditions
23. Investigations
24. Injury and poisoning
25. Surgical and medical procedures
26. Social circumstances
27. Product Issues

The other terms are sorted by dictionary code order.

Appendix D Response Evaluation Criteria in Solid Tumors (recist 1.1)

Details provided in bibliographic reference (19).

Definitions

At baseline, tumor lesions/lymph nodes will be categorized measurable or nonmeasurable as follows:

Measurable

- Tumor lesions: Must be accurately measured in at least one dimension (longest diameter in the plane of measurement is to be recorded) with a minimum size of:
  - Ten millimeters by computed tomography (CT) scan (CT scan slice thickness no greater than 5 mm).
  - Ten-millimeter caliper measurement by clinical examination (lesions which cannot be accurately measured with calipers should be recorded as nonmeasurable).
  - Twenty millimeters by chest X-ray.
- Malignant lymph nodes: To be considered pathologically enlarged and measurable, a lymph node must be ≥15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow-up, only the short axis will be measured and followed (see Special Issue 15 [19]). See also notes below on “Baseline documentation of target and nontarget lesions” for information on lymph node measurement.

Non-measurable

- All other lesions, including small lesions (longest diameter <10 mm or pathological lymph nodes with 10 to <15 mm short axis) as well as truly nonmeasurable lesions. Lesions considered truly nonmeasurable include: leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses/abdominal organomegaly identified by physical examination that is not measurable by reproducible imaging techniques.

Special considerations regarding lesion measurability

Bone lesions, cystic lesions, and lesions previously treated with local therapy require particular comment:

Bone lesions:

- Bone scan, positron emission tomography (PET) scan or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions.
- Lytic bone lesions or mixed lytic-blastic lesions, with identifiable soft tissue components, that can be evaluated by cross sectional imaging techniques such as CT or magnetic

resonance imaging (MRI) can be considered as measurable lesions if the soft tissue component meets the definition of measurability described above.

- Blastic bone lesions are nonmeasurable.

Cystic lesions:

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor nonmeasurable) since they are, by definition, simple cysts.
- 'Cystic lesions' thought to represent cystic metastases can be considered as measurable lesions, if they meet the definition of measurability described above. However, if noncystic lesions are present in the same participant, these are preferred for selection as target lesions.

Lesions with prior local treatment:

- Tumor lesions situated in a previously irradiated area, or in an area subjected to other locoregional therapy, are usually not considered measurable unless there has been demonstrated progression in the lesion. Study protocols should detail the conditions under which such lesions would be considered measurable.

Methods of measurement

- Measurement of lesions:
  All measurements should be recorded in metric notation, using calipers if clinically assessed. All baseline evaluations should be performed as close as possible to the treatment start and never more than 4 weeks before the beginning of the treatment.
- Method of assessment:
  The same method of assessment and the same technique should be used to characterize each identified and reported lesion at baseline and during follow-up. Imaging based evaluation should always be done rather than clinical examination unless the lesion(s) being followed cannot be imaged but are assessable by clinical examination. Clinical lesions: Clinical lesions will only be considered measurable when they are superficial and 10 mm diameter as assessed using calipers (eg, skin nodules). For the case of skin lesions, documentation by color photography including a ruler to estimate the size of the lesion is suggested. As noted above, when lesions can be evaluated by both clinical examination and imaging, imaging evaluation should be undertaken since it is more objective and may also be reviewed at the end of the study.
- Chest X-ray: Chest CT is preferred over chest X-ray, particularly when progression is an important endpoint, since CT is more sensitive than X-ray, particularly in identifying new lesions. However, lesions on chest X-ray may be considered measurable if they are clearly defined and surrounded by aerated lung.
- Computed tomography, MRI: CT is the best currently available and reproducible method to measure lesions selected for response assessment. This guideline has defined measurability of lesions on CT scan based on the assumption that CT slice thickness is 5 mm or less. When CT scans have slice thickness greater than 5 mm, the minimum size

for a measurable lesion should be twice the slice thickness. Magnetic resonance imaging is also acceptable in certain situations (eg, for body scans).

- Ultrasound: Ultrasound is not useful in assessment of lesion size and should not be used as a method of measurement.
- Endoscopy, laparoscopy: The utilization of these techniques for objective tumor evaluation is not advised. However, they can be useful to confirm complete pathological response when biopsies are obtained or to determine relapse in trials where recurrence following complete response or surgical resection is an endpoint.
- Tumor markers: Tumor markers alone cannot be used to assess objective tumor response. If markers are initially above the upper normal limit, however, they must normalize for a participant to be considered in complete response. Specific guidelines for both cancer antigen 125 response (in recurrent ovarian cancer) and prostate-specific antigen response (in recurrent prostate cancer), have been published. In addition, the Gynecologic Cancer Intergroup has developed cancer antigen 125 progression criteria which are to be integrated with objective tumor assessment for use in first-line trials in ovarian cancer.
- Cytology, histology: These techniques can be used to differentiate between partial response (PR) and complete response (CR) in rare cases if required by protocol (eg, residual lesions in tumor types such as germ cell tumors, where known residual benign tumors can remain). When effusions are known to be a potential adverse effect of treatment (eg, with certain taxane compounds or angiogenesis inhibitors), the cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment can be considered if the measurable tumor has met criteria for response or stable disease (SD) in order to differentiate between response (or SD) and progressive disease.

Tumor response evaluation

To assess objective response or future progression, it is necessary to estimate the overall tumor burden at baseline and use this as a comparator for subsequent measurements. Only participants with measurable disease at baseline should be included in protocols where objective tumor response is the primary endpoint. Measurable disease is defined by the presence of at least 1 measurable lesion. In studies where the primary endpoint is tumor progression (either time to progression or proportion with progression at a fixed date), the protocol must specify if entry is restricted to those with measurable disease or whether participants having non measurable disease only are also eligible.

Response criteria

Table 12 - Response criteria, evaluation of target lesions

Complete Response (CR)	Disappearance of all target lesions. Any pathological lymph nodes (whether target or nontarget) must have reduction in short axis to <10 mm.
Partial Response (PR):	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.
Progressive Disease (PD):	At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).
Stable Disease (SD):	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study

Table 13 - Response criteria, evaluation of nontarget lesions

Complete Response (CR):	Disappearance of all nontarget lesions and normalization of tumor marker level. All lymph nodes must be nonpathological in size (<10 mm short axis).
Non-CR/Non-PD:	Persistence of one or more nontarget lesion(s) and/or maintenance of tumor marker level above the normal limits.
Progressive Disease (PD):	Unequivocal progression (see comments below)[a] of existing nontarget lesions. (Note: the appearance of one or more new lesions is also considered progression).

a Although a clear progression of "nontarget" lesions only is exceptional, in such circumstances, the opinion of the treating physician should prevail and the progression status should be confirmed later on by the review panel (or study chair).

Evaluation of best overall response

It is assumed that at each protocol specified time point, a response assessment occurs. The following table (Table 14) provides a summary of the overall response status calculation at each time point for participants who have measurable disease at baseline. When patients have non-measurable (therefore non-target) disease only, Table 15 is to be used.

Table 14 - A summary of overall response status for measurable 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; PD = progressive disease; PR = partial response; SD = stable disease.

Table 15 - A summary of overall response status for non-measurable disease

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; PR = partial response; SD = stable disease.

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.

Table 16 - A summary of overall response status for non-measurable disease

Overall response First time point	Overall response Subsequent time point	BEST overall response
CR	CR	CR
CR	PR	SD, PD or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE

a If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the participant had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

CR = complete response; NE = not evaluable; PD = progressive disease; PR = partial response; SD = stable disease

The best overall response is determined once all the data for the participant is known.

When SD is believed to be best response, it must also meet the protocol specified minimum time from baseline (defined as 42 days). If the minimum time is not met when SD is otherwise the best time point response, the participant's best response depends on the subsequent assessments. For example, a participant who has SD at first assessment, PD at second and does not meet minimum duration for SD, will have a best response of PD. The same participant lost to follow-up after the first SD assessment would be considered inevaluable. Complete or partial responses may be claimed only if the criteria for each are met at a subsequent time point as specified in the protocol (generally 4 weeks later). In this circumstance, the best overall response can be interpreted as in Table 16.

Appendix E Description of censoring and event rules for primary and sensitivity analyses of PFS

Table 17 - PFS Primary analysis

Situation	Date of outcome	Outcome	Category
No baseline tumor assessments[a]	Date of randomization	Censored	No baseline tumor assessments
No evaluable post-baseline tumor assessments[a]	Date of randomization	Censored	No evaluable post-baseline tumor assessments
Alive and no documented progression	Date of the last evaluable tumor assessment documenting no progression	Censored	Alive without progression
Documented progression (or Death) at or between scheduled visits	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after exactly one non-evaluable tumor assessment	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after two or more non-evaluable tumor assessment[b]	Date of the last evaluable tumor assessment documenting no progression	Censored	Event occurred after two or more missed tumor assessment
Clinical/non-radiological progression	Ignored	Ignored	
Initiation of further anti-cancer therapy	Date of the last evaluable tumor assessment before start date of further anti-cancer therapy (if any), date of randomization otherwise	Censored	Initiation of further anti-cancer therapy

a Except if the patient dies within 17 weeks after the date of randomization in which case a death event outcome will be considered for the PFS with date of outcome corresponding to the date of death

b An event occurring at least 18 weeks (excluded) after the last evaluable tumor assessment, documenting no progression or randomization date, whichever is last. 18 weeks corresponds to twice the time between two disease assessments per protocol (every 2*8 weeks), plus the 7-day window before and after.

Note: for patients falling in situation with discordant outcomes, the event outcome takes the priority to the censored outcome (eg, death without baseline tumor assessment)

Table 18 - PFS sensitivity analysis #1 (ignoring further anti-cancer therapy)

Situation	Date of outcome	Outcome	Category
No baseline tumor assessments[a]	Date of randomization	Censored	No baseline tumor assessments
No evaluable post-baseline tumor assessments[a]	Date of randomization	Censored	No evaluable post-baseline tumor assessments
Alive and no documented progression	Date of the last evaluable tumor assessment documenting no progression	Censored	Alive without progression
Documented progression (or Death) at or between scheduled visits	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after exactly one non-evaluable tumor assessment	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after two or more non-evaluable tumor assessment[b]	Date of the last evaluable tumor assessment documenting no progression	Censored	Event occurred after two or more missed tumor assessment
Clinical/non-radiological progression	Ignored	Ignored	
Initiation of further anti-cancer therapy	Ignored	Ignored	

a Except if the patient dies within 17 weeks after the date of randomization in which case a death event outcome will be considered for the PFS with date of outcome corresponding to the date of death

b An event occurring at least 18 weeks (excluded) after the last evaluable tumor assessment, documenting no progression or randomization date, whichever is last. 18 weeks corresponds to twice the time between two disease assessments per protocol (every 2*8 weeks), plus the 7-day window before and after.
Note: for patients falling in situation with discordant outcomes, the event outcome takes the priority to the censored outcome (eg, death without baseline tumor assessment)

Table 19 - PFS sensitivity analysis #2 (PFS considering events occurring after two or more non-evaluable tumor assessment as event and ignoring further anticancer therapy)

Situation	Date of outcome	Outcome	Category
No baseline tumor assessments[a]	Date of randomization	Censored	No baseline tumor assessments
No evaluable post-baseline tumor assessments[a]	Date of randomization	Censored	No evaluable post-baseline tumor assessments
Alive and no documented progression	Date of the last evaluable tumor assessment documenting no progression	Censored	Alive without progression
Documented progression (or Death) at or between scheduled visits	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after exactly one non-evaluable tumor assessment	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after two or more non-evaluable tumor assessment[b]	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Clinical/non-radiological progression	Ignored	Ignored	
Initiation of further anti-cancer therapy	Ignored	Ignored	

a Except if the patient dies in which case a death event outcome will be considered for the PFS with date of outcome corresponding to the date of death

b An event occurring at least 18 weeks (excluded) after the last evaluable tumor assessment, documenting no progression or randomization date, whichever is last. 18 weeks corresponds to twice the time between two disease assessments per protocol (every 2*8 weeks), plus the 7-day window before and after.

Note: for patients falling in situation with discordant outcomes, the event outcome takes the priority to the censored outcome (eg, death without baseline tumor assessment)

Table 20 - PFS sensitivity analysis #3 (PFS considering events occurring after two or more non-evaluable tumor assessment as event and back-dating at the next schedule assessment and ignoring further anticancer therapy)

Situation	Date of outcome	Outcome	Category
No baseline tumor assessments[a]	Date of randomization	Censored	No baseline tumor assessments
No evaluable post-baseline tumor assessments[a]	Date of randomization	Censored	No evaluable post-baseline tumor assessments
Alive and no documented progression	Date of the last evaluable tumor assessment documenting no progression	Censored	Alive without progression
Documented progression (or Death) at or between scheduled visits	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after exactly one non-evaluable tumor assessment	Date of the next schedule assessment	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after two or more non-evaluable tumor assessment[b]	Date of the next schedule assessment	Event	Documented progression (or Death without documented progression)
Clinical/non-radiological progression	Ignored	Ignored	
Initiation of further anti-cancer therapy	Ignored	Ignored	

a Except if the patient diesin which case a death event outcome will be considered for the PFS with date of outcome corresponding to the date of death

b An event occurring at least 18 weeks (excluded) after the last evaluable tumor assessment, documenting no progression or randomization date, whichever is last. 18 weeks corresponds to twice the time between two disease assessments per protocol (every 2*8 weeks), plus the 7-day window before and after.

Note: for patients falling in situation with discordant outcomes, the event outcome takes the priority to the censored outcome (eg, death without baseline tumor assessment)

Table 21 - PFS supportive #3 analysis for investigator assessment (taking clinical/non-radiological progression as event)

Situation	Date of outcome	Outcome	Category
No baseline tumor assessments[a]	Date of randomization	Censored	No baseline tumor assessments
No evaluable post-baseline tumor assessments[a]	Date of randomization	Censored	No evaluable post-baseline tumor assessments
Alive and no documented progression	Date of the last evaluable tumor assessment documenting no progression	Censored	Alive without progression
Documented progression (or Death) at or between scheduled visits	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after exactly one non-evaluable tumor assessment	Date of progression (or Date of death)	Event	Documented progression (or Death without documented progression)
Documented progression (or Death) occurring after two or more non-evaluable tumor assessment[b]	Date of the last evaluable tumor assessment documenting no progression	Censored	Event occurred after two or more missed tumor assessment
Clinical/non-radiological progression	Date of clinical/non-radiological progression	Event	Non-documented progression
Initiation of further anti-cancer therapy	Date of the last evaluable tumor assessment before start date of further anti-cancer therapy (if any), date of randomization otherwise	Censored	Initiation of further anticancer therapy

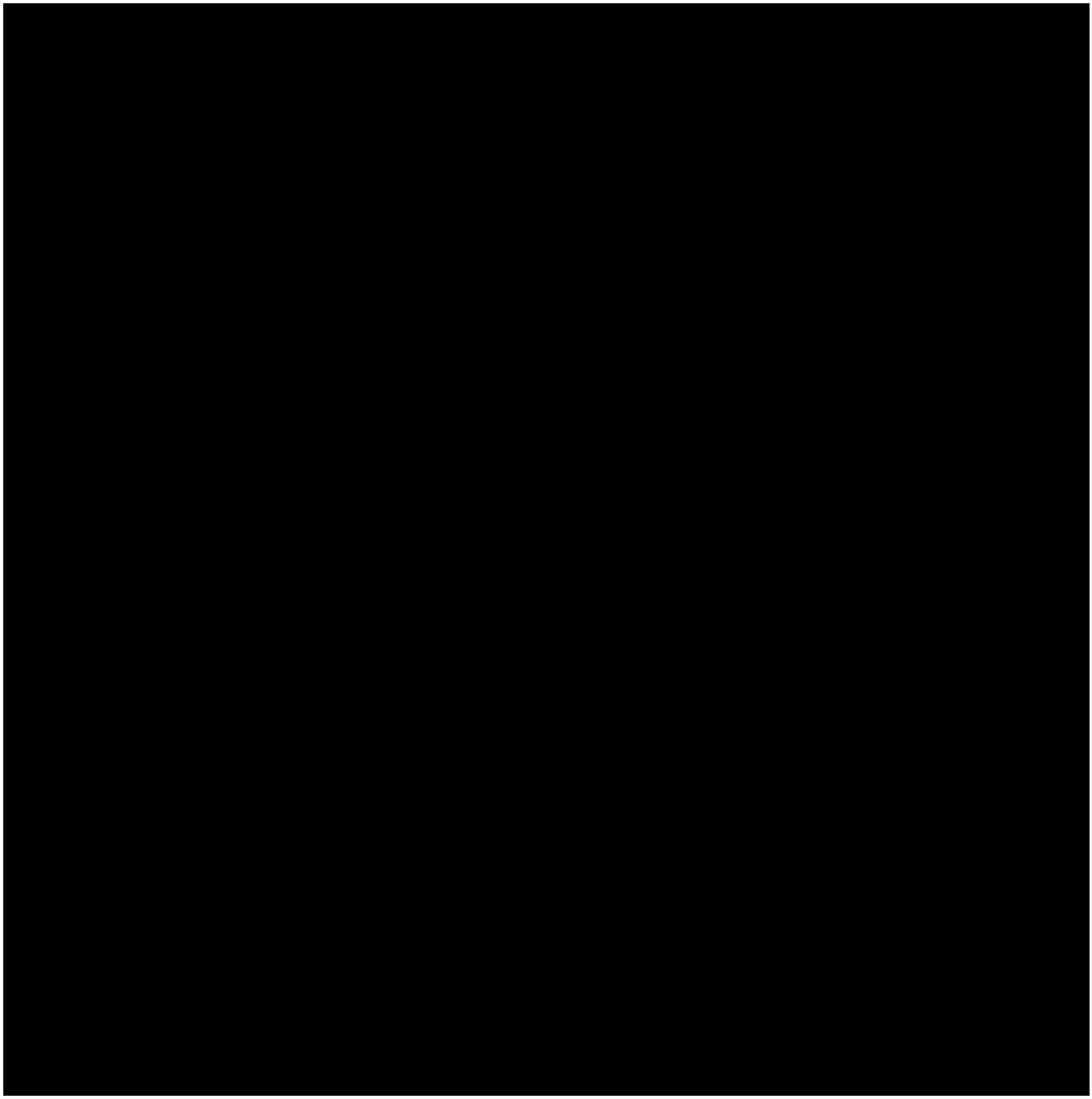
a Except if the patient dies within 17 weeks after the date of randomization in which case a death event outcome will be considered for the PFS with date of outcome corresponding to the date of death

b An event occurring at least 18 weeks (excluded) after the last evaluable tumor assessment, documenting no progression or randomization date, whichever is last. 18 weeks corresponds to twice the time between two disease assessments per protocol (every 2*8 weeks), plus the 7-day window before and after.

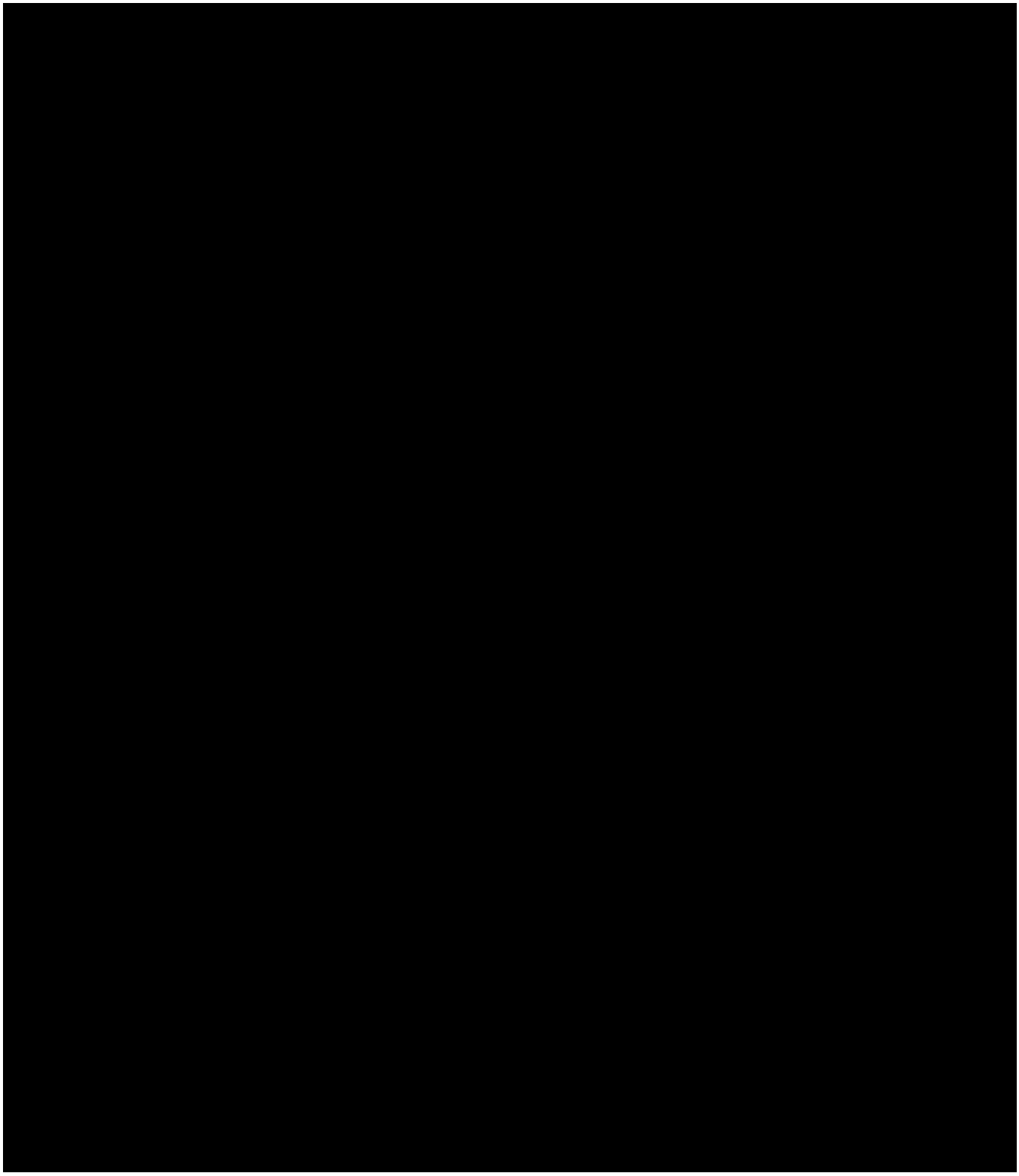
Note: for patients falling in situation with discordant outcomes, the event outcome takes the priority to the censored outcome (eg, death without baseline tumor assessment)

Appendix F EORTC QLQ-C30 and QLQ-BR23 items, scales and scores

For QLQ-C30:



For QLQ-BR23



Signature Page for VV-CLIN-0548794 v3.0
act16105-16-1-9-sap

Approve & eSign	[redacted]
Approve & eSign	