Appendix 16.1.9 Documentation of statistical methods

History of changes	
Version	**Summary of changes**
1.0	Original version

NOVARTIS

Clinical Development

BYL719/Alpelisib

CBYL719H12301 / NCT04251533

EPIK-B3: A Phase III, multicenter, randomized, double-blind, placebo-controlled study to assess the efficacy and safety of alpelisib (BYL719) in combination with nab-paclitaxel in patients with advanced triple negative breast cancer with either phosphoinositide-3-kinase catalytic subunit alpha (PIK3CA) mutation or phosphatase and tensin homolog protein (PTEN) loss without PIK3CA mutation

Statistical Analysis Plan (SAP) – Amendment 2

Author: Trial Statistician, [redacted]

Document type: SAP Documentation

Document status: Version 3.0

Release date: 03-Jul-2023

Number of pages: 69

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
08-Jun-2020	Prior to FPFV	Creation of final version	N/A - First version (v 1.0)	NA
28-Sep-2022	After protocol Am 1	More consistent with protocol Am1 and BYL program rules.	Version 2.0	Section 1, updated to align with protocol amendment 1 Section 1.2.1and 1.2.2, updated to align with the latest estimand framework Section 2.3.1, added summary of protocol deviations related to Covid-19. Section 2.8.1.4 clarified exclusion of wrong pre-dose insulin records in the analysis; updated Table 2-7 to show different analysis time windows for different laboratory assessments; clarified the use of fasting vs. non-fasting glucose values in summary output. Original Section 2.8.3 (becomes Section 2.9.1 after Am 2) updated rules for censoring about PRO assessments after the discontinuation of study treatment or the start of any new antineoplastic therapy. Section 5.2 deleted old wordings about Grade 5 AE, since it has become applicable after the database upgrade for this study. Section 5.3 clarified imputation rules for bilirubin value in the format of '<x'. All sections: replaced patient(s) /subject(s) with participant(s); updated wordings to align within BYL program.
03-Jul-2023	After protocol Am 2	To align with protocol Am2, update of the analysis plan according to study design change.	Version 3.0	General: For all subsections for Part B2, clarified that Part B2 was never initiated. Section 1, updated protocol version to align with protocol amendment 2 (PA2) Section 1.1, for Part A, removed interim analysis (IA) to align with PA2 and clarified key changes of study design per PA2 Section 1.2, for Part A, moved OS to secondary endpoint instead of key secondary endpoint, [redacted] All subsequent sections were adjusted accordingly. Section 2, adjusted text for IA to align with PA2; clarified timing of primary analysis for Part A.

	Table 2-1, added 30-days safety follow up for the derivation of last contact date per new discussion. Section 2.2.3, removed some subgroups due to limited sample size and change in clinical interests; clarified that only unstratified model will be used for subgroup analysis. Section 2.3.1, removed 'survival follow-up' from post-treatment follow-up disposition to be consistent with other BYL studies. Section 2.5.1, removed BIRC, which is not applicable after PA2. Section 2.5.2, updated analysis strategy for Part A to align with PA2 that no formal testing, but only descriptive analysis will be performed. Section 2.5.5, due to decreased samples, removed supplementary analyses for 'missing TA PFS' , 'backdating PFS', 'clinical PD', removed multivariate Cox regression model and any analyses based on BIRC review. Sections 2.6, 2.7, 2.7.1, and 2.7.2 clarified that the key secondary objective is not applicable after PA2 and instead OS has been moved to secondary objectives. All subsections for OS definition and analysis methods are adjusted accordingly. [redacted] Section 2.7.2, for ORR/CBR, removed supplementary analysis based on central review, also removed analysis of confirmed or unfirmed response for measurable disease only; For DOR, removed supplementary analysis based on unconfirmed response. Section 2.8.1.2, removed time to first occurrence for AESI Section 2.8.1.4, removed shift table for lab parameter, removed plot for HbA1c and lab box plot, removed time to first occurrence and time to resolution analyses for lab parameter Section 2.8.1.5, removed summary for change of ECG by timepoint Section 2.8.1.6, removed cardiac imaging shift table and summary of LVEF change

	Section 2.8.1.7, removed summary for change of vital sign, but kept summary for notable vital sign and the corresponding listing. [redacted] Section 2.9.1, removed PFS2. Section 2.10, clarified that DMC safety review are the only interim looks for this study and also clarified the timing of primary PFS analysis for Part A; In addition, removed all other originally planned interim analyses that are not applicable after PA2. Section 3, clarified the strategic change for recruitment halt, while the original sample size calculations still remain in the document for reference; Audit size for BIRC is removed. Section 5, removed subsections that are not applicable after PA2, including hypothesis and test statistic, audit based BIRC assessment of PFS and group sequential design.

Table of contents

List of tables

List of figures

List of abbreviations

AE	Adverse event
AESI	Adverse event of Special Interest
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BIRC	Blinded Independent Review Committee
BMI	Body mass index
BOR	Best overall response
BPI-SF	Brief Pain Inventory – short form
BSA	Body surface area
CBR	Clinical benefit rate
CI	Confidence Interval
CIS	Commonwealth of Independent States
CR	Complete response
CRF	Case Report Form
CRS	Case retrieval strategy
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
ctDNA	Circulating Tumor Deoxyribonucleic Acid
DB	Double-blind
DI	Dose Intensity
DMC	Data Monitoring Committee
DOR	Duration Of Response
DRL	Drug Reference Listing
ECHO	Echocardiogram
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EDR	Early discrepancy rate
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer's core quality of life questionnaire
EQ-5D-5L	EuroQol 5 Dimension questionnaire
FAS	Full analysis set
FPFV	First patient first visit
GI	Gastrointestinal
HGLT	High level group term
HLT	High level terms
HR	Hormone Receptor
HR.	Hazard Ratio
IA	Interim analysis
IHC	Immunohistochemistry

IRT	Interactive Response Technology
ITT	Intent To Treat
LDR	Late discrepancy rate
LLOQ	Lower limits of quantitation
LPLV	Last patient last visit
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
MUGA	Multigated Acquisition Scan
NCI	National Cancer Institute
NGS	Next Generation Sequencing
NMQ	Novartis MedDRA queries
OL	Open-label
ORR	Overall response rate
OS	Overall survival
PD	Progressive disease
PD.	Pharmacodynamic(s)
PDI	Planned dose intensity
PFS	Progression-free survival
PIK3CA	Phosphoinositide-3-kinase catalytic subunit alpha
PK	Pharmacokinetic
PPOS	predictive probability of success
PR	Partial response
PRO	Patient Reported Outcome
PT	Preferred term
PTEN	Phosphatase and tensin homolog protein
QoL	Quality of Life
QTcF	Corrected QT interval (corrected by Fridericia's formula)
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria In Solid Tumors
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
SMQ	Standardized MedDRA queries
SOC	System organ class
SUSAR	Suspected Unexpected Serious Adverse Reactions
TA	Tumor assessment
TBIL	Total bilirubin
TNBC/mTNBC	Triple-Negative Breast Cancer / metastatic Triple-Negative Breast Cancer
TTR	Time To Response
UNK	Unknown
WHO	World Health Organization

1 Introduction

This document describes the detailed statistical methodology to be used for the clinical study report (CSR) for the primary analysis of study CBYL719H12301, a Phase III, multicenter, randomized, double-blind, placebo-controlled study to assess the efficacy and safety of alpelisib (BYL719) in combination with nab-paclitaxel in patients with advanced triple negative breast cancer with either phosphoinositide-3-kinase catalytic subunit alpha (PIK3CA) mutation or phosphatase and tensin homolog protein (PTEN) loss without PIK3CA mutation.

The content of this SAP is based on the CBYL719H12301 protocol amendment 2 released on Mar 14, 2023. All decisions regarding primary PFS analysis, as defined in this document, have been made prior to the database lock and unblinding of the study data.

CSR deliverables (shells for tables, figures, listings) and further programming specifications are described in the Tables, Figures & Listings (TFL) shells and Programming Datasets Specification (PDS), respectively.

1.1 Study design

Study CBYL719H12301 has been designed as a three part, Phase III, multicenter, international trial evaluating the efficacy and safety of alpelisib in combination with nab-paclitaxel in 1st or 2nd line therapy of advanced (loco-regionally recurrent or metastatic) TNBC with a PIK3CA mutation or PTEN loss (without a PIK3CA mutation).

- Part A: randomized, double-blind (DB) and placebo-controlled.
- Part B1: single-arm, open-label (OL).
- Part B2: randomized, double-blind and placebo-controlled.

An independent Data Monitoring Committee (DMC) will monitor unblinded safety and efficacy data accruing during the trial. The interim PFS analysis (see Section 2.10 for details) will be performed by an independent statistician external to Novartis and the results will be provided to the DMC by the independent statistician. A separate DMC SAP specifies the analyses to be performed for the DMC reviews.

Part A: Advanced TNBC with PIK3CA mutation

The purpose of study Part A is to assess whether treatment with alpelisib in combination with nab-paclitaxel prolongs PFS compared to placebo in combination with nab-paclitaxel in advanced TNBC participants with a PIK3CA mutation (with or without concurrent PTEN loss). Approximately 252 participants will be randomly assigned to receive either alpelisib plus nab-paclitaxel or placebo plus nab-paclitaxel.

Randomization to treatment will follow a 1:1 ratio and will be stratified by three factors of prognostic value: i) line of therapy in metastatic setting (1st line versus 2nd line), ii) hormone receptor status at initial breast cancer diagnosis (HR+ versus HR-), and iii) prior therapy with a checkpoint inhibitor (Yes versus No).

With protocol amendment 02, no interim efficacy analyses will be performed.

Part B1: Advanced TNBC with PTEN loss (without PIK3CA mutation or PIK3CA status unknown)

The purpose of study part B1 is to determine whether alpelisib in combination with nab-paclitaxel in 32 participants with advanced TNBC and PTEN loss (without concurrent PIK3CA mutation or PIK3CA status unknown) warrants further development. ORR and safety data after 6 months of follow-up will be used to make the decision on whether or not to initiate part B2. No interim efficacy analysis is planned for part B1.

Note: Participants with a PIK3CA mutation and concurrent PTEN loss are only eligible to be randomized into Part A of this study.

Part B2: Advanced TNBC with PTEN loss (without PIK3CA mutation)

If review of part B1 results leads to the decision to open part B2 (refer to Section 2.5.1), the purpose of Part B2 is to determine whether treatment with alpelisib in combination with nab-paclitaxel prolongs PFS compared to placebo in combination with nab-paclitaxel in advanced TNBC participants with PTEN loss (without a concurrent PIK3CA mutation). Approximately 282 participants will be randomly assigned to receive either alpelisib plus nab-paclitaxel or placebo plus nab-paclitaxel.

Randomization to treatment will follow a 1:1 ratio and will be stratified by three factors of prognostic value: i) line of therapy in metastatic setting (1st line versus 2nd line) , ii) hormone receptor status at initial breast cancer diagnosis (HR+ versus HR-), and iii) prior therapy with a checkpoint inhibitor (Yes versus No). The stratification factors and definition of line of therapy for Parts A and B2 are identical.

Enrollment in Part B1 was completed in February 2022. In November 2022, Novartis took the decision to halt the enrollment to the Part A due to slow recruitment. Additionally, Part B2 was not initiated since Part B1 didn't meet its primary endpoint of ORR.

The reference to Part B2 remains in the SAP to reflect the initial study design and avoid confusion. Part B2 will not be conducted.

Figure 1-1 Study design

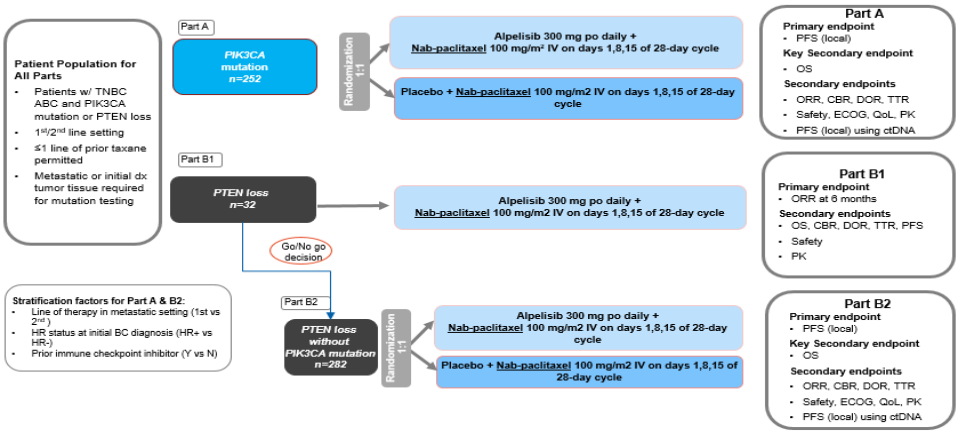
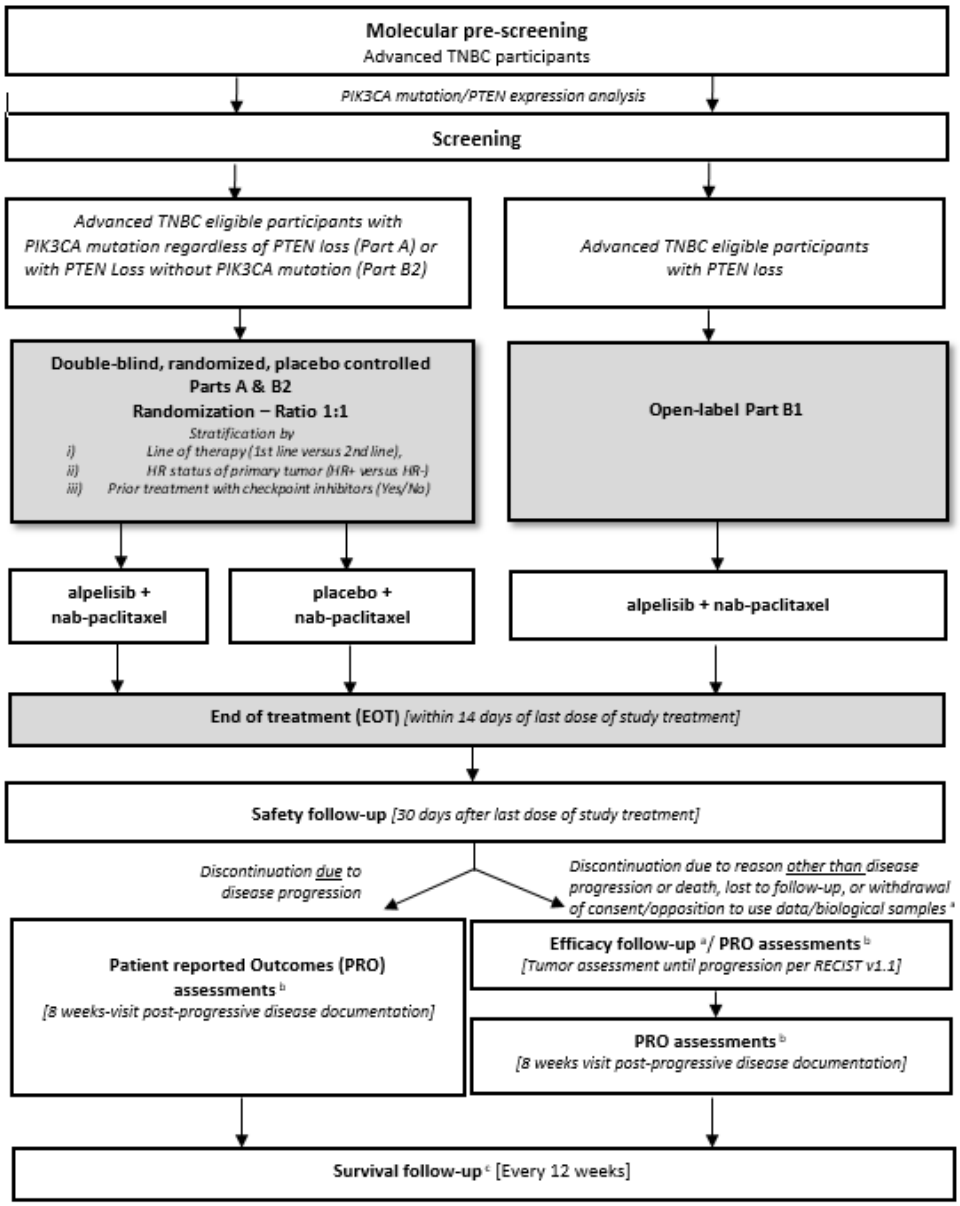


Figure 1-2 Study flow



[a] until documented disease progression, death, lost to follow-up, or withdrawal of consent /opposition to use data/biological samples

[b] ONLY for patients enrolled in Parts A and B2

[c] except if patient is lost to follow-up or in case of withdrawal of consent /opposition to use data/biological samples

1.2 Study objectives, endpoints and estimands

Objectives and related endpoints for study Part A, Part B1, and Part B2 are described in Table 1-1, Table 1-2, and Table 1-3 respectively below.

Table 1-1 Objectives and related endpoints for Part A (TNBC with PIK3CA mutation)

	Objective	Endpoint (s)
Primary	To assess whether treatment with alpelisib in combination with nab-paclitaxel prolongs PFS compared to placebo in combination with nab-paclitaxel	PFS based on investigator assessment using RECIST 1.1 criteria (refer to Section 1.2.1 for the primary estimand)
Secondary	To assess whether treatment with alpelisib in combination with nab-paclitaxel prolongs OS compared to placebo in combination with nab-paclitaxel	Overall Survival (OS)
	To assess safety and tolerability of alpelisib in combination with nab-paclitaxel	**Safety:** Incidence, type, and severity of adverse events per CTCAE v4.03 criteria including changes in laboratory values, vital signs, liver assessments and cardiac assessments **Tolerability:** dose interruptions, reductions, dose intensity, and duration of exposure for all drug components
	To assess additional efficacy parameters	ORR with confirmed response, CBR with confirmed response, Duration Of Response (DOR) with confirmed response, TTR based on local radiology assessments and using Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 criteria
	To evaluate the association between PIK3CA mutation status as measured in ctDNA at baseline with PFS upon treatment with alpelisib	PFS based on local radiology assessments using RECIST 1.1 criteria for participants by PIK3CA mutation status measured in baseline ctDNA
[redacted]		

[redacted]

Table 1-2 Objectives and related endpoints for Part B1 (TNBC with PTEN loss without PIK3CA mutation or PIK3CA status unknown)

	Objective	Endpoint (s)
Primary	To assess the antitumor activity of alpelisib in combination with nab-paclitaxel	ORR with confirmed response using RECIST 1.1 criteria (refer to Section 1.2.1 for the primary estimand)
Secondary	To determine the safety and tolerability of alpelisib in combination with nab-paclitaxel	**Safety:** Incidence, type, and severity of adverse events per CTCAE v4.03 criteria including changes in laboratory values, vital signs, liver assessments, and cardiac assessments **Tolerability:** dose interruptions, reductions, dose intensity, and duration of exposure for all drug components
	To evaluate additional efficacy parameters of alpelisib in combination with nab-paclitaxel	CBR with confirmed response, DOR with confirmed response, TTR, and PFS (based on local radiology assessments and using RECIST 1.1 criteria), OS
[redacted]		

Table 1-3 Objectives and related endpoints for Part B2 (TNBC with PTEN loss without PIK3CA mutation) – Not initiated

	Objective	Endpoint (s)
Primary	To determine whether treatment with alpelisib in combination with nab-paclitaxel prolongs PFS compared to placebo in combination with nab-paclitaxel	PFS based on investigator assessment using RECIST 1.1 criteria (refer to Section 1.2.1 for the primary estimand)
Key Secondary	To determine whether treatment with alpelisib in combination with nab-paclitaxel prolongs OS compared to placebo in combination with nab-paclitaxel	OS (refer to Section 1.2.2 for the key secondary estimand)
Secondary	To assess safety and tolerability of alpelisib in combination with nab-paclitaxel	**Safety**: Incidence, type, and severity of adverse events per CTCAE v4.03 criteria including changes in laboratory values, vital signs, liver assessments, and cardiac assessments **Tolerability**: dose interruptions, reductions, dose intensity, and duration of exposure for all drug components
	To assess additional efficacy parameters	ORR with confirmed response, CBR with confirmed response, DOR with confirmed response, TTR based on local radiology assessments and using RECIST 1.1 criteria
	To characterize exposure of alpelisib when administered in combination with nab-paclitaxel	Summary statistics of plasma alpelisib concentrations by time point
	To evaluate patient-reported outcomes (PRO) of alpelisib in combination with nab-paclitaxel versus placebo with nab-paclitaxel	1) Change from baseline in the global health status/QoL scale score of the EORTC QLQ-C30 2) Time to 10% definitive deterioration in the global health status/QoL scale score of the EORTC QLQ-C30
	To evaluate the association between PIK3CA mutation status as measured in ctDNA at baseline with PFS upon treatment with alpelisib	PFS based on local radiology assessments using RECIST 1.1 criteria for participants by PIK3CA mutation status measured in baseline ctDNA
	To evaluate alpelisib in combination with nab-paclitaxel versus placebo with nab-paclitaxel with respect to time to deterioration of ECOG performance status	Time to definitive deterioration of the ECOG performance status from baseline

1.2.1 Primary estimand

Part A

The primary scientific question of interest is: what is the treatment effect based on PFS for alpelisib in combination with nab-paclitaxel versus placebo in combination with nab-paclitaxel in men and women with advanced (locally recurrent or metastatic) TNBC with a PIK3CA mutation (with or without concurrent PTEN loss), regardless of study treatment discontinuation or start of new anti-neoplastic therapy?

The justification for targeting this treatment effect is based on the desire to assess the treatment effect not only during the on-treatment period but also after the discontinuation of study treatment, during the entire course of the study; and compare not just alpelisib+nab-paclitaxel vs placebo+nab-paclitaxel, but alpelisib+nab-paclitaxel followed by any new anti-neoplastic therapy vs. placebo+nab-paclitaxel followed by any new anti-neoplastic therapy, i.e. any subsequent anti-neoplastic therapy is part of treatment attribute.

The primary estimand is characterized by the following attributes:

1. Population: all participants randomized with advanced (locally recurrent or metastatic) TNBC with a PIK3CA mutation. Further details on the population are provided in Section 5 of the study protocol.

2. Treatment: the investigational treatment is alpelisib in combination with nab-paclitaxel plus any subsequent anti-neoplastic therapy as needed. The control treatment is placebo in combination with nab-paclitaxel plus any subsequent anti-neoplastic therapy as needed. Further details about the investigational treatment and control treatment are provided in Section 6 of the study protocol.

3. Variable: PFS based on local investigator assessment and using RECIST 1.1 criteria. Further details on PFS are provided in Section 2.5.1.

4. Intercurrent event:

- discontinuation of study treatment for any reason

Details on how to handle intercurrent events are provided in Section 2.5.3.

5. Summary measure: PFS hazard ratio (alpelisib versus placebo) and its 95% confidence interval, estimated using a Cox proportional hazard model stratified by the randomization stratification factors. Further details on how the summary measure will be tested are provided in Section 2.5.2.

Part B1

The scientific question of interest is: what is the treatment effect based on ORR for alpelisib in combination with nab-paclitaxel in men and women with advanced (locally recurrent or metastatic) TNBC with PTEN loss (without PIK3CA mutation or PIK3CA status unknown) prior to treatment discontinuation date+30 days or prior to start of new anti-neoplastic therapy, whichever occurs first.

The justification for targeting this treatment effect is based on the desire to assess the treatment effect only during the on-treatment period defined as from the start of treatment until treatment discontinuation+30 days and before any new anti-neoplastic therapy.

The primary estimand is characterized by the following attributes:

1. Population: all participants with advanced (locally recurrent or metastatic) TNBC with PTEN loss (without PIK3CA mutation or PIK3CA status unknown). Further details on the population are provided in Section 5 of the study protocol.

2. Treatment: the investigational treatment is alpelisib in combination with nab-paclitaxel . Further details about the investigational treatment are provided in Section 6 of the study protocol.

3. Variable: Best overall response with confirmed response based on local investigator assessment and using RECIST 1.1 criteria. Further details on ORR are provided in Section 2.5.1.

4. Intercurrent events:

- discontinuation of study treatment for any reason
- start of new anti-neoplastic therapy

Details on how to handle intercurrent events are provided in Section 2.5.3.

5. Summary measure: ORR with confirmed response and its 95% confidence interval. Further details on how the summary measure will be tested are provided in Section 2.5.2.

Part B2

The primary scientific question of interest is: what is the treatment effect based on PFS for alpelisib in combination with nab-paclitaxel versus placebo in combination with nab-paclitaxel in participants with advanced (locally recurrent or metastatic) TNBC with PTEN loss (without PIK3CA mutation), regardless of study treatment discontinuation or start of new anti-neoplastic therapy?

The justification for targeting this treatment effect is based on the desire to assess the treatment effect not only during the on-treatment period but also after the discontinuation of study treatment, during the entire course of the study; and compare not just alpelisib+nab-paclitaxel vs placebo+nab-paclitaxel, but alpelisib+nab-paclitaxel followed by any new anti-neoplastic therapy vs. placebo+nab-paclitaxel followed by any new anti-neoplastic therapy, i.e. any subsequent anti-neoplastic therapy is part of treatment attribute.

The primary estimand is characterized by the following attributes:

1. Population: all participants randomized with advanced (locally recurrent or metastatic) TNBC with PTEN loss (without PIK3CA mutation). Further details on the population are provided in Section 5 of the study protocol.

2. Treatment: the investigational treatment is alpelisib in combination with nab-paclitaxel plus any subsequent anti-neoplastic therapy as needed. The control treatment is placebo in combination with nab-paclitaxel plus any subsequent anti-neoplastic therapy as needed. Further details about the investigational treatment and control treatment are provided in Section 6 of the study protocol.

3. Variable: PFS based on local investigator assessment and using RECIST 1.1 criteria. Further details on PFS are provided in Section 2.5.1.

4. Intercurrent event:

- discontinuation of study treatment for any reason

Details on how to handle intercurrent events are provided in Section 2.5.3.

5. Summary measure: PFS hazard ratio (alpelisib versus placebo) and its 95% confidence interval, estimated using a Cox proportional hazard model stratified by the randomization stratification factors. Further details on how the summary measure will be tested are provided in Section 2.5.2.

Note Part B2 was never initiated.

1.2.2 Key secondary estimand

Not applicable following protocol amendment 02.

2 Statistical methods

2.1 Data analysis general information

The analysis for open-label Part B1 and the final PFS analysis for Part A will be performed by Novartis. The interim safety analysis for Part A before the primary PFS analysis will be performed by an independent statistician external to Novartis for review by a DMC. SAS version 9.4 or later will be used to perform all data analyses and to generate tables, figures, and listings.

Study results will be presented by treatment group and study part, unless stated otherwise.

Data included in the analysis

Part A

The primary descriptive efficacy and safety analyses for Part A will be performed when all participants have completed 6 months of study treatment or have been discontinued from study treatment. Any additional data for participants continuing to receive study treatment past this time and for participants continuing for efficacy follow-up (PFS, OS), as allowed by the protocol, will be further summarized in the final study close-out report. Refer to Section 2.10 for timing of interim analyses (safety).

Part B1

The primary efficacy and safety analyses for Part B1 will be performed after all enrolled participants have completed 6 months treatment or have been discontinued from study treatment whichever occurs earlier. Any additional data for participants continuing to receive study treatment past this time and for participants continuing for efficacy follow-up (PFS, OS), as allowed by the protocol, will be further summarized in the subsequent study report.

Part B2

Not applicable following protocol amendment 02.

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis in each Study part. Due to expected small number of participants enrolled at centers, no center effect will be assessed.
Qualitative data (e.g., gender, race, etc.) will be summarized by means of contingency tables by treatment group; a missing category will be included as applicable. Percentages will be calculated using the number of participants in the relevant population or subgroup as the denominator.
Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (i.e. mean, standard deviation, median, minimum, and maximum) by treatment group.

2.1.1 General definitions

2.1.1.1 Investigational drug and study treatment

Investigational drug, will refer to alpelisib only. **Control drug** will refer to alpelisib matching placebo only. Whereas, **study treatment** will refer to alpelisib +/- nab-paclitaxel and placebo +/- nab-paclitaxel.

The term investigational treatment and alpelisib matching placebo may also be referred to as "**study treatment"** and placebo, respectively which are used throughout this document.

2.1.1.2 Date of first administration of study treatment component

The date of first administration is defined as the first date when a non-zero dose of that component of study treatment is administered and recorded on the "Study Treatment Summary" eCRF (for alpelisib/placebo) or "Study Treatment" eCRF for (nab-paclitaxel).

2.1.1.3 Date of last administration of study treatment component

The date of last administration is defined as the last date when a non-zero dose of that component of study treatment is administered and recorded on the "Study Treatment Summary" eCRF (for alpelisib/placebo) or "Study Treatment" eCRF (for nab-paclitaxel).

2.1.1.4 Date of first administration of study treatment

The date of first administration of study treatment is defined as the first date when a non-zero dose of any component of study treatment is administered. The date of first administration of study treatment will also be referred to as the start of study treatment.

Note: Dates from eCRFs relating to PK sampling will not be used for this derivation. For example: if the first dose of alpelisib/placebo is administered on 05-Jan-2020, and first dose of nab-paclitaxel is taken on 03-Jan-2020, then the date of first administration of study treatment is 03-Jan-2020.

2.1.1.5 Date of last administration of study treatment

The date of last administration of study treatment is defined as the last date when a non-zero dose of any component of the study treatment (alpelisib/placebo or nab-paclitaxel) is administered.

Note: Dates from eCRFs relating to PK sampling will not be used for this derivation. For example: if the last dose of alpelisib/placebo is administered on 15-Apr-2020, and the last dose of nab-paclitaxel is taken on 17-Apr-2020, then the date of last administration of study treatment is on 17-Apr-2020.

2.1.1.6 Study day

The study day describes the day of the event or assessment date, relative to the reference start date.

The reference start date is designated as Study Day 1. Study Day –1 is the day that precedes Day 1. Study Day 0 is not defined. Study day is not to be used in numerical computations, for example in calculating exposure.

The study day will be calculated as follows:

- The date of the event (visit date, onset date of an event, assessment date etc.) - reference start date + 1 if the event is on or after the reference start date;
- The date of the event (visit date, onset date of an event, assessment date etc.) - reference start date if the event precedes the reference start date.

The reference start date for safety assessments (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption, PK etc.) is the start of study treatment.

The reference start date for all other, non-safety assessments (i.e., tumor assessment, death, disease progression, tumor response, ECOG performance status, and patient reported outcomes (PRO)) is the date of randomization. (Example: if randomization date is 15-DEC-2020, start of study treatment is on 18-DEC-2020, and the date of death is 28-DEC-2020, then the study day when the death occurred is 14).

The study day will be displayed in data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

2.1.1.7 Time Unit

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

2.1.1.8 Baseline

For efficacy evaluations, the last available assessment, including unscheduled assessments, on or before the date of randomization is taken as the "baseline" value or "baseline" assessment. In the context of baseline definition, the efficacy evaluations also include PRO and performance status. For RECIST-based endpoints only, including PFS, ORR, CBR, time to response and duration of response, a window of 7 days after the start of study treatment will be allowed, i.e. the investigator responses will be considered as a candidate for the baseline assessment if the assessment is within 7 days after the treatment start date. For biomarker assessments please see Section 2.9.4.

For safety evaluations (i.e. laboratory assessments, ECGs and vital signs), the last available assessment, including unscheduled assessments, on or before the start date of study treatment as described in Section 2.1.1.4 is taken as "baseline" assessment. Assessments specified to be collected post-dose on the first date of treatment are not considered as baseline values.

If participants have no value as defined above, the baseline result will be considered missing.

2.1.1.9 On-treatment assessment/event

Safety summaries and selected summaries of deaths will summarize only on-treatment assessments/events.

An on-treatment assessment/ event is defined as any assessment event in the following time interval:

- [date of first administration of study treatment; date of last administration of study treatment + 30 days], i.e. including the lower and upper limits

Note: the calculation of study treatment duration will use different rules as specified in Section 2.4.1.1.

An AE starting in the screening phase and ongoing in the on-treatment phase will not be considered an on-treatment AE unless it has worsened in severity.

If the last date of study treatment is missing, any assessment/event occurring after the start date of study treatment will be considered as on-treatment.

Data listings will include all assessments/events, flagging those which are not on-treatment assessments/events.

Note: The date of first administration of study treatment and the date of last administration of study treatment are defined in Section 2.1.1.4 and Section 2.1.1.5, respectively.

2.1.1.10 Windows for multiple assessments

In order to summarize ECOG, laboratory, PRO, and other data collected over time (including unscheduled visits), the assessments will be time slotted. The following general rule will be applied in creating the assessment windows (except for ECOG, see Section 2.7.1 for ECOG-specific rules): If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, then the earlier of the 2 assessments will be used. If multiple assessments on the same date then the worst case will be used. Data from all assessments (scheduled and unscheduled), including multiple assessments, will be listed.

Time windows for ECOG, lab and PRO assessments are listed in Table 2-12, Table 2-7 and Table 2-10, respectively.

2.1.1.11 Last contact date

The last contact date will be derived for participants not known to have died at the analysis cut-off date based on the last complete date among the criteria in Table 2-1:

Table 2-1 Last contact date data sources

Source data	Conditions
Date of Randomization	No Condition
Last date participant was known to be alive from Survival eCRF	Participant status is reported to be alive or unknown.
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term.
Start/End dates from “Study Treatment Summary” eCRF (for alpelisib/placebo) or “Study Treatment” eCRF (for nab-paclitaxel)	Non-missing dose. Doses of 0 are allowed.
End of treatment date from End Of Treatment Disposition eCRF/end date of post-treatment follow-up disposition CRF	No condition.
End date of 30 days safety follow up	No condition
Date of PRO assessment	At least one non-missing answer to questionnaire
Tumor (RECIST) assessment date	For non-target lesion: non-missing lesion status For target lesion: non-missing lesion diameter

Source data	Conditions
	For new lesion: "Is there a new lesion?" = yes
Laboratory collection dates	At least one non-missing parameter value
PK collection dates	Was sample taken? = Yes
Vital signs or ECG assessment date	At least one non-missing parameter value
ECOG Performance Status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term
Concomitant medication date	Non-missing medication term
Biomarker collection date	At least one non-missing biomarker measurement
Hospitalization admission/discharge date	Non-missing verbatim term
Cardiac imaging date	Non-missing LVEF or overall interpretation

The last contact date is defined as the latest complete date from the above list on or before the data cut-off date.

The cut-off date will not be used for last contact date, unless the participant was seen or contacted on that date. No date post cut-off date will be used. Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. For participants who died, the last contact date is defined as the date of death. Partial date imputation is allowed for event (death)/censoring if it is recorded on the 'Survival' eCRF.

The last contact date will be used for censoring of participants in the analysis of overall survival.

2.2 Analysis sets

Parts A and B2

The Full Analysis Set (FAS) comprises all participants to whom study treatment has been assigned by randomization. According to the intent to treat principle, participants will be analyzed according to the treatment and strata they have been assigned to during the randomization procedure. The FAS will be the primary population for all efficacy analyses.

The Safety Set includes all participants who received at least one dose of study treatment (i.e. at least one dose of any component of alpelisib/placebo, nab-paclitaxel). Participants will be analyzed according to the study treatment received, where treatment received is defined as the randomized treatment if the participant took at least one dose of that treatment or the first treatment received if the randomized treatment was never received.

The Pharmacokinetic Analysis Set (PAS) consists of all participants who receive at least one dose of alpelisib/placebo or nab-paclitaxel and provide at least one evaluable PK concentration.

Note Part B2 was never initiated.

Part B1

The Full Analysis Set (FAS) comprises all participants to whom study treatment has been assigned and who received at least one dose of study treatment. Participants will be analyzed according to the treatment they have been assigned to.

The Safety Set includes all participants who received at least one dose of any study treatment (i.e. at least one dose of any component of alpelisib or nab-paclitaxel). Participants will be analyzed according to the study treatment received, where treatment received is defined as the assigned treatment if the participant took at least one dose of that treatment or the first treatment received if the assigned treatment was never received.

[redacted]

2.2.1 Participant classification

Participants may be excluded from the analysis sets defined above based on the protocol deviations entered in the database and/or on specific participant classification rules defined in Table 2-2.

Table 2-2 Participant classification based on protocol deviations and non-PD criteria

Analysis set	Protocol deviation criteria leading to exclusion	Non-protocol deviation criteria leading to exclusion
FAS	No written informed consent	Not applicable
Safety set	No written informed consent	No dose of study treatment
PAS	No written informed consent	No dose of study treatment or no evaluable PK concentration

2.2.2 Withdrawal of informed consent

Any data collected in the clinical database after a participant withdraws informed consent from all further participation in the trial, will not be included in any analysis. The date on which a participant withdraws full consent is recorded in the eCRF. When an end of treatment visit occurs after withdrawal of informed consent, the end of treatment disposition status and reasons are retained.

Death events may be used in the analysis of PFS/overall survival if captured from public records (registers), local law and participant informed consent permitting.

Additional data for which there is a separate informed consent, e.g. PK, biomarker collection etc., collected in the clinical database without having obtained that consent will not be included in the analysis. These data will be excluded by the presence of the appropriate protocol deviation criterion.

2.2.3 Subgroups of interest

In Parts A and B2, subgroup analyses will be performed for efficacy and safety as outlined below. Subgroups will be formed using eCRF data (with the exception of some biomarker subgroups, for which the applicable third party biomarker data will be used). This includes variables related to stratification factors (line of therapy in metastatic setting [1st line vs. 2nd line], hormone receptor status at initial breast cancer diagnosis [HR+ vs. HR-], and prior therapy with a checkpoint inhibitor [Yes vs. No]). Analyses by stratum/by stratification factor based on IRT data are covered by the analyses described in Section 2.5.

Race subgroup will be defined as follows (as reported on Demography eCRF):

- Asian: race = Asian
- Other: Any other race except participants with missing race information, who will be excluded from the subgroup analyses.

Efficacy

The primary efficacy endpoint will be summarized by the following subgroups to examine the homogeneity of treatment effect:

- Line of therapy in metastatic setting per clinical database (1st line vs. 2nd line)
- Hormone receptor status at initial breast cancer diagnosis per clinical database (HR+ vs. HR-)
- Prior therapy with checkpoint inhibitors per clinical database (Yes vs. No)
- Age (<65 years vs. >=65 years)
- Race (Asian vs. Other)
- Menopausal status (Premenopausal vs. Postmenopausal)
- Number of metastatic sites (<3 vs. ≥3)
- Presence of lung metastases (Yes vs. No)
- Presence of liver metastases (Yes vs. No)
- Presence of brain metastases (Yes vs. No)
- Prior taxane-based chemotherapy use (Yes vs. No)
- PTEN loss of expression (Loss of expression [H-score < 10] vs No loss of expression [H-score ≥ 10]) (Part A only).

For each of the subgroups, the following analyses will be performed:

- Median Kaplan-Meier estimates of the survival distribution of PFS
- Hazard ratio with 95% CI using unstratified Cox proportional hazards model.

[redacted] Forest plot (HR, 95% CI) will be produced to graphically depict the treatment effect estimates in different subgroups. No inferential statistics (p-values) will be produced for the subgroups.

Descriptive subgroup analyses of OS will also be performed for the subgroups listed above.

Safety

The main safety analyses will be repeated on the safety set in the following subgroups:

- Age (<65 years vs. >=65 years)
- Race (Asian vs. Other)

These main safety analyses include:

- AEs by system organ class, preferred term and maximum grade
- Treatment-related AEs by system organ class, preferred term and maximum grade
- Serious AEs by system organ class and preferred term and maximum grade

For the AESI of Hyperglycemia only, a subgroup analysis by hyperglycemia diagnosis status at baseline per American Diabetes Association (ADA) 2017 will be presented:

- Diabetic: FPG ≥ 7.0 mmol/l or 126 mg/dl or HbA1c ≥ 6.5 % vs.
- Pre-diabetic: FPG 5.6- <7.0 mmol/l or 100-125 mg/dl or HbA1c 5.7- <6.5% vs.
- Normal: [FPG <5.6 mmol/l or <100 mg/dl] and HbA1c <5.7%

The objective for carrying out these subgroup analyses is to identify potential safety issues that may be limited to a subgroup of participants, or safety issues that are more commonly observed in a subgroup of participants.

2.3 Participant disposition, demographics and other baseline characteristics

Demographic and other baseline data including disease characteristics/prognostic data will be summarized descriptively by treatment group for the FAS and Safety Set (for DSUR). Relevant medical history and current medical conditions at baseline will be summarized separately by system organ class and preferred term, by treatment group.

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviation, median, minimum, and maximum will be presented. For selected parameters, 25th and 75th percentiles will also be presented.

2.3.1 Participant disposition

Enrollment by country and center will be summarized for all screened participants and also by treatment group for the randomized participants. The number (%) of randomized and treated participants included in the FAS will be presented overall and by treatment group. The number (%) of screened but not-randomized participants and the reasons for screening failure will also be displayed. The number (%) of participants in the FAS who are still on treatment, who discontinued the study phases and the reason for discontinuation will be presented overall and by treatment group.

The following summaries will be provided (with % based on the total number of FAS participants):

- Number (%) of participants who were randomized (based on data from IRT system)

- Number (%) of participants who were randomized but not treated (based on 'Study Treatment' eCRF not completed for any study treatment component)
    - Primary reason for not being treated (based on 'Treatment disposition' eCRF)
    - Number (%) of participants who were treated (based on 'Study Treatment' eCRF of each study treatment component completed with non-zero dose administered)
    - Number (%) of participants who are still on-treatment (based on the 'Treatment disposition' eCRF not completed);
    - Number (%) of participants who discontinued the study treatment phase (based on the 'Treatment disposition' eCRF)
- Primary reason for study treatment phase discontinuation(based on the 'Treatment disposition' and 'Participant status' eCRFs)
- Number (%) of participants who have entered the post-treatment follow-up (based on the 'Treatment disposition' and 'Participant status' eCRFs));
    - Number (%) of participants who have discontinued from the post-treatment follow-up (based on the 'Post-treatment follow-up disposition' eCRF);
    - Reasons for discontinuation from the post-treatment follow-up (based on Post-treatment follow-up disposition' eCRF);

Protocol deviations

The number (%) of participants in the FAS with any protocol deviation will be tabulated by deviation category (as specified in the Edit Check Specification Document) overall and by treatment group for the FAS. Moreover, protocol deviations related to Covid-19 will also be summarized separately by deviation category overall and by treatment group. Major protocol deviations leading to exclusion from analysis sets will be tabulated separately overall and by treatment group. All protocol deviations will be listed.

Analysis sets

The number and percentage of participants in each analysis set will be summarized by treatment group and randomization stratum.

2.3.2 Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed for the FAS by treatment group and study part. Categorical data (e.g. race, ECOG performance status, menopausal status) will be summarized by frequency counts and percentages; the number and percentage of participants with missing data will be provided. Continuous data (e.g. age, weight, etc.) will be summarized by descriptive statistics (N, mean, standard deviation, median, minimum and maximum). BMI (kg/m^2) will be calculated as weight[kg] / (height[m]^2) using weight at baseline.

2.3.3 Baseline stratification factors

The number (%) of participants in each stratum based on data obtained from the IRT system will be summarized overall and by treatment group for the FAS in randomized parts.

Discordances between the stratum recorded in IRT at the time of randomization and the actual stratum recorded in the clinical database through the data collected on eCRF will be cross-tabulated and listed.

Unless otherwise specified, stratified analyses and analyses "by stratum" will be based on IRT stratification data, while data for other subgroup analyses (Section 2.2.3) will be based on eCRF data.

2.3.4 Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. This analysis will include the following: primary site of cancer, details of tumor histology/cytology, histological grade, stage at initial diagnosis, time since initial diagnosis, time from initial diagnosis to first recurrence/progression (in months), time since most recent relapse/progression to randomization (in months), stage at time of study entry, presence/absence of target and non-target lesions, number and type of metastatic sites involved (including visceral involvement, bone only lesions), hormone receptor status at initial diagnosis, hormone receptor status at study entry. The presence/absence of target and non-target lesions will be based on the data collected on RECIST target/non-target lesion assessment eCRFs. Metastatic sites will be based on Diagnosis and Extent of Cancer eCRF.

Time since initial diagnosis will be summarized in months. A month is defined as 365.25/12=30.4375 days.

2.3.5 Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on the eCRF, will be summarized and listed by treatment group and study part. Separate summaries will be presented for ongoing and historical medical conditions. The summaries will be presented by system organ class (SOC), preferred term (PT) and treatment group. Medical history and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology. The MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings.

2.3.6 Other

All data collected at baseline, including participants' referrals, child bearing potential and details of informed consent for biomarker data collection, will be listed.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

Duration of exposure, cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by each component of study treatment by treatment group for each study part. The number of participants who have dose reductions or interruptions, and the reasons, will be summarized by treatment group. Details of the derivations and summaries are provided

in the following sections. Participants with no exposure to the study treatment component will be excluded from the corresponding tabular summaries.

Categorical data will be summarized as frequencies and percentages. For continuous data, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum will be presented.

The safety set will be used for all summaries and listings of study treatment.

2.4.1.1 Duration of study treatment exposure

The duration of exposure to study treatment will be calculated as:

Duration of exposure to study treatment (days) = (last date of exposure to any study treatment component) – (date of first administration of study treatment) + 1.

Duration of exposure to any single component of study treatment will be calculated as:

Duration of exposure (days) = (last date of exposure to study treatment component) – (date of first administration of study treatment component) + 1.

The ***last date of exposure*** is defined as follows for the ***study treatment*** components:

The ***last date of exposure to alpelisib/placebo*** is the date of last administration of alpelisib/placebo.

Due to the irregularly spaced nab-paclitaxel dose administration, the ***last date of exposure to nab-paclitaxel*** is calculated as last date of administration of nab-paclitaxel) + (length of time interval to next scheduled dose - 1).

- If a participant died or was lost to follow-up within last date of administration of nab-paclitaxel + (length of time interval to next scheduled dose - 1) days, then the last date of exposure to nab-paclitaxel is the date of death or the date of last contact, respectively.
- If the derived last date of exposure goes beyond the data cutoff date, it should be truncated to the date of data cutoff.

If the participant discontinues nab-paclitaxel between Cycle X Day 1 and Cycle X Day 14 inclusive, the length of time interval to the next scheduled dose is 7. If the participant discontinues nab-paclitaxel on or after Cycle X Day 15 and before Cycle X+1 Day 1, the length of time interval to the next scheduled dose is 14.

This duration of exposure includes the periods of temporary interruption (of any component of the study treatment for any reason). The duration of study treatment exposure and exposure to each treatment component will be summarized by Study Part, in months. In addition, the duration of exposure will be categorized into time intervals (e.g. <1 month; <3 months; 3-<6 months; 6-<9 months, etc.); frequency counts and percentages will be presented for the number of participants in each interval.

2.4.1.2 Cumulative dose

Cumulative dose for any component of study treatment is defined as the total dose of the medication given during the study treatment exposure.

Cumulative dose will be summarized using descriptive statistics for each component of study treatment by treatment group for each study part.

For alpelisib/placebo, the cumulative dose is the sum of the non-zero doses recorded over the dosing period.

For nab-paclitaxel, the cumulative dose should be defined based on the days when the participant is assumed to have taken a non-zero dose during dosing periods.

2.4.1.3 Dose intensity and relative dose intensity

Dose intensity (DI) for participants with non-zero duration of exposure to each study treatment component is defined as follows:

- DI (dosing unit / unit of time) = Cumulative dose (dosing unit) / Duration of exposure (unit of time).

For participants who did not take any drug, the DI is equal to zero. Planned dose intensity (PDI) is the assigned dose by unit of time planned to be given to participants as per protocol in the same dose unit and unit of time as that of the Dose Intensity. DI, PDI and Relative dose intensity (RDI) is defined as:

For alpelisib/placebo:

- DI (mg/day) = Cumulative dose (mg) / duration of exposure (days)
- PDI is 300 mg/day
- RDI (%) = DI (mg/day) / PDI (mg/day) *100

For nab-paclitaxel:

- DI (mg/m^2/cycle) = Cumulative dose (mg/m^2) / (*adjusted* duration of exposure (days) / 21)

To account for the irregularly spaced nab-paclitaxel dose administration, the *adjusted* duration of exposure (days) for nab-paclitaxel is 21*(# completed cycles) + min(21, [(date of last administration of nab-paclitaxel + 6) – (date of first administration of nab-paclitaxel) + 1] in the last incomplete cycle).

For example, if the duration of exposure to nab-paclitaxel is 63 days (corresponding to two completed cycles plus 7 days, i.e. the last date of exposure is on C3D1), then the adjusted duration of exposure to nab-paclitaxel is 21*2+min(21,7)=42+7=49 days.

- PDI is 300 mg/m2/cycle
- RDI (%) = DI (mg/m^2/cycle) / PDI (mg/m^2/cycle) * 100

Cumulative dose is expressed in mg/m^2 for nab-paclitaxel:

- The sum of ('dose administered' (mg) / BSA (m^2)) during the exposure to paclitaxel.

BSA (m^2) will be determined by the Mosteller formula: √(wt (kg)×ht (cm) / 3600).

Both categorical and continuous summaries of RDI (i.e. mean, standard deviation etc.) by treatment component will be presented.

2.4.2 Dose reductions, interruptions or permanent discontinuations

The number (%) of participants with dose reductions or interruptions and permanent discontinuations, and associated reasons, will be summarized separately for each component of study treatment (alpelisib/placebo and nab-paclitaxel).

Dose administered (mg) and dosing frequency from the "Study Treatment Summary" eCRF (for alpelisib/placebo) or "Study Treatment" eCRF (for nab-paclitaxel) will be used to determine the dose reductions and interruptions.

'Dose permanently discontinued' box ticked from the "Study Treatment Summary" eCRF (for alpelisib/placebo) or "Study Treatment" eCRF (for nab-paclitaxel) will be used to determine permanent discontinuation.

Dose reductions

For alpelisib/placebo, a dose reduction is defined as a decrease in dose from the previous non-zero dose to another non-zero dose less than the protocol planned starting dose, even if this decrease has been directly preceded by an interruption. For example, in the sequence 300 mg daily to 0 mg to 250 mg daily, the 250 mg dose will be counted as a reduction. On the other hand, if the dose decrease is followed by an interruption, with the dose resuming at the same level prior to the interruption (e.g. in the sequence 250 mg daily to 0 mg to 250 mg daily), the second dose decrease will not be counted as dose reduction.

If, due to a dosing error, a participant receives a higher than planned starting dose and moves down to the planned starting dose then this is not considered a dose reduction. However if the dose change is from a higher than planned starting dose down to a lower than protocol planned starting dose, then this is considered a dose reduction (e.g. in the sequence: 350 mg daily to 300 mg daily to 250 mg daily; 250 mg is considered a dose reduction).

If, due to a dosing error, a participant receives a lower than previous non-zero dose and resumes later at the protocol specified dose reduction, then the lower dose received due to dosing error and protocol specified dose reduction are dose reductions (e.g. in the sequence 300 mg daily to 200 mg daily to 250 mg daily, then 200 mg and 250 mg are considered dose reductions).

If, due to a dosing error, a participant receives a lower than previous non-zero dose and resumes later at a lower than previous non-zero dose, then 2 dose reductions will be counted (e.g. in the sequence 300 mg daily to 250 mg daily to 200 mg daily, 250 mg and 200 mg are dose reductions).

Table 2-3 Examples of Dose Reduction for alpelisib

Sequence	Reduction
With dose change	
300 mg daily to 250 mg daily to 0 mg to 250 mg daily	1 reduction (the 1st 250 mg)
300 mg daily to 300 mg daily to 0 mg to 250 mg daily	1 reduction (250 mg)
300 mg daily to 0 mg to 250 mg daily	1 reduction (250 mg)

Sequence	Reduction
With interruption	
300 mg daily to 0 mg to 300 mg daily	0 reductions
With dosing error	
300 mg daily to 250 mg daily to 200 mg daily*	2 reductions (250 mg, 200 mg)
300 mg daily to 200 mg daily* to 300 mg daily	1 reduction (200 mg)
300 mg daily to 200 mg daily* to 250 mg daily	2 reductions (200 mg, 250 mg)
300 mg daily to 400 mg daily* to 350 mg daily*	0 reductions since 400 mg and 350 mg are dose escalations not reduction
300 mg daily to 150 mg daily* to 300 mg daily	1 reduction (150 mg)
With dosing error at the 1st administration	
150 mg daily* to 300 mg daily	1 reduction (150 mg)
150 mg daily* to 0 mg to 150 mg daily* to 300 mg daily	1 reduction (150 mg)
150 mg daily* to 300 mg daily to 0 mg to 250 mg daily	2 reductions (150 mg and 250 mg)

*dosing error

For nab-paclitaxel, a dose reduction is defined as a decrease in dose (mg/m^2) from the protocol planned dose of 100 mg/m^2 at any administration. The reduction counting rules from above apply.

Dose interruptions

An interruption is defined as a 0mg dose on one or more days between two non-zero dosing records. Any two or more consecutive zero doses of alpelisib / placebo (e.g. in the sequence 300 mg daily to 0 mg to 0 mg to 300 mg daily) or nab-paclitaxel will be counted as 1 interruption if the reasons for these two consecutive dose interruption are the same. It will be counted as two different interruptions only if the reasons are different. For participants who have dose interruption checked in the eCRF but never resume with a non-zero dose, the dose interruption will not be counted. For example, in the sequence of 300 mg daily to 0 mg (dose interruption) to 0 mg (dose permanently discontinued), the 0 mg (dose interruption) will not be counted as a dose interruption. Interruptions will be summarized for each component of study treatment.

Note: The last zero dose of alpelisib/placebo or nab-paclitaxel at permanent discontinuation is not considered as a dose interruption.

The number (%) of dose interruptions along with reasons will be summarized for each study part.

Permanent discontinuations

In addition, the reasons for permanent discontinuation of each study treatment component will be summarized by treatment group, based on the information on the respective component-specific eCRF.

2.4.3 Prior, concomitant and post-treatment therapies

Prior anti-cancer therapy

The number and percentage of participants who received any prior anti-neoplastic medications, prior anti-neoplastic radiotherapy or prior anti-neoplastic surgery (biopsy and non-biopsy separately) will be summarized by treatment group using the FAS.

Prior anti-neoplastic medications will be summarized by therapy type (e.g. chemotherapy, hormonal therapy etc.), setting (e.g. adjuvant, neoadjuvant, metastatic, etc.) and also by lowest ATC class, preferred term and treatment. Summaries will include total number of regimens, best response, and time from treatment end date of the last therapy to progression for the last therapy. The medication therapy type of any combination therapy will be classified based on the following order: chemotherapy, biologic therapy, targeted therapy, hormonal therapy. For example, a combination therapy of chemotherapy and hormonal therapy will be classified as 'chemotherapy'. This classification will be based on medical review of observed trial data and the excel spreadsheet created will be stored in GPS.

For radiotherapy, the location and setting (e.g. adjuvant, palliative) for the last therapy will be summarized.

For surgery (excluding biopsies), procedure at last surgery and the time since last surgery will be summarized.

Concomitant therapy

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) other than the study treatment administered to a participant coinciding with the study treatment period. Concomitant therapy include medications (other than study treatment) starting on or after the start date of study treatment or medications starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system and summarized by lowest ATC class and preferred term using frequency counts and percentages. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and preferred term. These summaries will include:

1. Medications starting on or after the start of study treatment but no later than 30 days after the last dose of study treatment and
2. Medications starting prior to start of study treatment and continuing after the start of study treatment.

All concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of study treatment or starting more than 30 days after the last date of study treatment

will be flagged in the listing. The safety set will be used for all concomitant medication tables and listings.

Post treatment anti-cancer therapy

Anti-neoplastic therapies after discontinuation of study treatment will be listed and summarized by ATC class, preferred term, by treatment group; by means of frequency counts and percentages using the FAS.

2.5 Analysis of the primary objective

Part A:

The primary objective of Part A is to assess whether treatment with alpelisib in combination with nab-paclitaxel prolongs PFS compared to placebo in combination with nab-paclitaxel.

Part B1:

The primary objective of Part B1 is to assess the antitumor activity of alpelisib in combination with nab-paclitaxel.

Part B2:

The primary objective of Part B2 is to determine whether treatment with alpelisib in combination with nab-paclitaxel prolongs PFS compared to placebo in combination with nab-paclitaxel.

Note Part B2 was never initiated.

2.5.1 Primary endpoint(s) / estimand(s)

Parts A and B2

The primary endpoint (variable attribute of the primary estimand; refer to Section 1.2.1) is PFS, defined as the time from the date of randomization to the date of the first documented progression or death due to any cause. For the primary efficacy analysis, PFS will be based on local investigators' review of tumor assessments and using RECIST 1.1 criteria (see Section 16.1 of the study protocol). The primary analysis will be based on FAS and will include all data observed up-to the cut-off date. If a participant has not progressed or died at the analysis cut-off date, PFS will be censored at the date of the last adequate tumor assessment (see RECIST 1.1 criteria in Protocol Section 16.1 of the study protocol for further details) before the cut-off date. PFS events documented after the initiation of new anti-neoplastic therapy (i.e. RECIST 1.1. documented disease progression or death) will be considered for the primary analysis provided tumor assessments continue after initiation of new cancer therapy (see Section 2.5.4 for additional details regarding censoring rules and determination of date of last adequate tumor assessment). Discontinuation due to disease progression (collected on the 'Treatment disposition' and 'Post treatment follow up' disposition eCRFs without supporting objective evidence satisfying progression criteria per RECIST 1.1 will not be considered as disease progression for PFS derivation. Clinical deterioration without objective radiological evidence will not be considered as documented disease progression.

The tumor endpoint derivation is based on the sequence of overall lesion responses at each assessment/time point. However, the overall lesion response at a given assessment/time point may be provided from different sources as illustrated in Table 2-4.

Table 2-4 Sources for overall lesion response

Source 1	Investigator (local radiology) reported overall lesion response
Source 2	Novartis-calculated overall lesion response based on raw (i.e. individual lesion) measurements from investigator (local radiology)

The primary efficacy analysis will be based on the investigator/local radiology review. The investigator reported overall lesion response at each assessment/time point (Source 1 in Table 2-4) will be used to derive the efficacy endpoints.

The overall response at each assessment will also be calculated using the raw lesion measurements (Source 2 in Table 2-4). The calculated responses will be listed along with the responses given by the investigator.

Censoring conventions (i.e. handling of missing values/censoring/ discontinuations) are provided in Section 2.5.4.

Part B1

The primary endpoint is ORR with confirmed response, based on local radiology assessments in participants with measurable disease at baseline. The primary analysis for ORR will occur when all participants have completed treatment duration of at least 6 months or have discontinued treatment, whichever occurs first. All assessments up to the same cut-off date (when all patients have had at least 6 months treatment or discontinued treatment before 6 months) will be used in the ORR calculation. ORR is defined as the proportion of participants with best overall response (BOR) of confirmed complete response (CR) or confirmed partial response (PR) according to RECIST 1.1 (see Protocol Section 16.1 of the study protocol). If no valid post-baseline tumor assessments are available, the best overall response must be "Unknown" unless progression is reported. Participants whose BOR is unknown or missing will be determined to be non-responders. Details on the handling of missing values/ discontinuations are provided in Section 2.5.4.

2.5.2 Statistical hypothesis, model, and method of analysis

Part A:

The primary efficacy analysis is to compare and summarize PFS between two treatment groups descriptively.

The primary efficacy variable, PFS (variable attribute of the primary estimand; refer to Section 1.2.1) will be analyzed based on the FAS according to the randomized treatment group and strata assigned at randomization. The PFS distribution will be estimated using the Kaplan-Meier method, and Kaplan-Meier curves, medians and 95% confidence intervals of the medians will be presented for each treatment group. The hazard ratio for PFS will be calculated, along with its 95% confidence interval, from a stratified Cox model using the randomization stratification factors, i.e. (i) line of therapy in advanced/metastatic setting (1st line versus

2nd line), ii) hormone receptor status at initial breast cancer diagnosis (HR+ versus HR-), and iii) prior therapy with checkpoint inhibitors (Yes/No). Given no inferential analysis will be conducted, no p-value will be calculated.

Part B1

A response rate of 35% is considered a minimum clinically meaningful improvement in this study population based on data from first and second line mTNBC participants treated with either nab-paclitaxel or paclitaxel in the PAKT study (Schmid et al 2018), the LOTUS study (Kim et al 2017), the IMpassion130 study (Schmid et al 2018), as well as information provided in the Abraxane USPI (Nab-Paclitaxel FDA label 2015) the Abraxane EU label (Nab-Paclitaxel EU label 2018). Therefore, proof of preliminary efficacy of alpelisib in combination with nab-paclitaxel will be declared if both of the following conditions are met:

- the mean of the posterior distribution of ORR is at least 35%

and

- the posterior probability that the ORR is $\geq$ 25% is at least 0.9

The posterior distribution of ORR will be derived from the prior distribution and all available data from the participants included in the FAS. A minimally informative unimodal Beta prior (Neuenschwander et al 2008) will be used for ORR (see Section 5.4.3 of the study protocol for further details). Additionally, ORR will be summarized by descriptive statistics (N, %) along with a two-sided exact binomial 95% confidence interval (Clopper and Pearson 1934).

Part B2 (Not applicable as of protocol amendment 02):

The primary efficacy analysis will be the comparison of PFS between the two treatment groups using a stratified log-rank test at an overall one-sided 2.5% level of significance for Parts A and B2 individually. The stratification will be based on the randomization stratification factors, per IRT, i.e. i) line of therapy in metastatic setting (1st line versus 2nd line), ii) hormone receptor status at initial breast cancer diagnosis (HR+ versus HR-), and iii) prior therapy with checkpoint inhibitors (Yes/No).

Assuming proportional hazards model for PFS, the following statistical hypotheses will be tested to address the primary efficacy objective:

H_{01}: $\theta_1 \geq 0$ vs. H_{a1}: $\theta_1 < 0$

where θ_1 is the log-hazard ratio (alpelisib + nab-paclitaxel group vs. placebo + nab-paclitaxel) of PFS.

The primary efficacy variable, PFS (variable attribute of the primary estimand; refer to Section 1.2.1), will be analyzed at the interim analysis and final analysis of a group sequential design, using a Haybittle-Peto boundary. Refer to Section 2.10 for more details on interim analysis. Analyses will be based on the FAS according to the randomized treatment group and strata assigned at randomization. The PFS distribution will be estimated using the Kaplan-Meier method, and Kaplan-Meier curves, medians and 95% confidence intervals of the medians will be presented for each treatment group. The hazard ratio for PFS will be calculated, along with its 95% confidence interval, from a stratified Cox regression model using the same stratification factors as for the log-rank test.

Note Part B2 was never initiated.

2.5.3 Handling of intercurrent events of primary estimand

Parts A & B2

The primary estimand will account for different intercurrent events as explained in the following:

1. **Discontinuation of study treatment**: tumor assessment data collected after discontinuation of study treatment will be used for the primary analysis irrespective of the study treatment discontinuation reason (treatment policy strategy).

Note Part B2 was never initiated.

Part B1

The primary estimand will account for the first of the following two intercurrent events to occur:

1. **Discontinuation of study treatment for any reason**: tumor assessment data collected within 30 days after discontinuation of study treatment will be used in the analysis. Tumor assessment data collected more than 30 days after discontinuation of study treatment will be excluded from the analysis (While on treatment strategy).
2. **Start of new anti-neoplastic therapy**: tumor assessment data collected after the initiation of new anti-neoplastic therapy will be excluded from the analysis (While on treatment strategy).

2.5.4 Handling of missing values/censoring/discontinuations

Parts A and B2

In the primary analysis, PFS will be censored at the date of the last adequate tumor assessment if no PFS event is observed prior to the analysis cut-off date.

PFS events documented after the initiation of new anti-neoplastic therapy (i.e. RECIST 1.1. documented disease progression or death) will be considered for the primary analysis provided tumor assessments continue after initiation of new cancer therapy.

If a PFS event is observed after one or more missing or non-adequate tumor assessments, the actual date of event will be used (see RECIST 1.1 in Section 16.1 of the study protocol).

The term “missing adequate tumor assessment” is defined as a tumor assessment (TA) not performed or tumor assessment with overall lesion response of “UNK”.

The date of last adequate tumor assessment is the date of the last tumor assessment with overall lesion response of CR, PR or SD or non-CR/non-PD before an event or a censoring reason occurred. In this case the last tumor evaluation date at that assessment will be used. If no post-baseline assessments are available (before an event or a censoring reason occurred) then the date of randomization/start date of treatment will be used.

Refer to Table 2-5 for censoring and event date options and outcomes for PFS.

Table 2-5 Outcome and event/censor dates for primary PFS analysis

Situation	Date	Outcome
No baseline assessment	Date of randomization	Censored
Progression or death at or before next scheduled Assessment	Date of progression (or death)	Progressed
Progression or death after one or more missing assessments	Date of progression (or death)	Progressed
No progression (or death)	Date of last adequate assessment	Censored
Treatment discontinuation due to 'Disease progression' without documented progression, i.e. clinical progression based on investigator claim	Ignore the treatment discontinuation and follow situations above	As per above situations
New anticancer therapy given prior to protocol defined progression	Date of progression (or death) by ignoring the new anticancer therapy	Progressed
Death before first PD assessment	Date of death	Progressed

Note Part B2 was never initiated.

Part B1

Participants with unknown or missing BOR as per RECIST 1.1 will be counted as failures. If there is no baseline tumor assessment, all post-baseline overall lesion responses are expected to be 'Unknown'. If no valid post-baseline tumor assessments are available, the best overall response must be "Unknown" unless progression is reported. If a complete response is not confirmed, the best overall response must be "Unknown" unless it occurs after the first 7 weeks (stable disease), or confirmed partial response or progression is reported. If a partial response is not confirmed, the best overall response must be "Unknown" unless it occurs after the first 7 weeks (stable disease) or progression is reported. For the computation of ORR, these participants with unknown BOR will be included in the FAS and will be counted as 'failures'.

The BOR will be determined from response assessments undertaken while on treatment. Only tumor assessments performed on or before the start of a new antineoplastic treatment other than study drug(s) (not considering palliative radiotherapy) will be considered in the assessment of BOR.

2.5.5 Sensitivity and supplementary analyses

Sensitivity analyses

Parts A and B2

If there is a high rate of discrepancy (at least 5%) between the strata classifications constructed using the eCRF data and those obtained from the IRT, a sensitivity analysis will be performed in which a stratified Cox regression model will be used to estimate the treatment hazard ratio

and the associated 95% confidence interval based on the eCRF-derived strata. No other inferential statistics will be provided.

In addition, to assess the impact of stratification, the hazard ratio and 95% confidence interval for PFS per local investigator assessment will also be obtained from an un-stratified and covariate unadjusted Cox model.

Supplementary analyses

Parts A and B2

Kaplan-Meier estimates of median PFS along with 95% confidence intervals and hazard ratio and corresponding 95% confidence interval obtained using the Cox proportional hazards model will be provided for the below data. No other inferential statistics will be provided.

- using the primary analysis source (i.e., local investigator assessment) on the FAS and censoring PFS at the date of the last adequate tumor assessment before the start of new anticancer therapy if no PFS event is observed prior to the start of new antineoplastic therapy. In the summary tables, this approach is referred as 'new anticancer therapy PFS supplementary analysis'.

Additionally, the following analyses will be performed:

- Number of participants and number of events by treatment group within each stratum (per IRT) will be presented along with the hazard ratio for treatment effect obtained using unstratified Cox proportional hazards regression models with corresponding confidence intervals. K-M plots of survival distributions will be presented by stratum.
- Timing of all tumor assessments will be depicted graphically separately for central radiology and investigator/local radiology and displayed by treatment group.
- Number of participants with a PFS event and number of participants censored for the PFS analysis will be summarized. In addition, a summary of reasons for PFS censoring will be provided by treatment group based on the reasons defined below. These summaries on censoring reasons will be produced for PFS by investigator radiology.

Censoring pattern of PFS

Number of participants with a PFS event and number of participants censored for the PFS analysis will be summarized. In addition, a summary of reasons for PFS censoring will be provided by treatment group based on the following reasons:

1: Ongoing without event
2: Lost to follow-up
3: Withdrew consent
4: Adequate assessment no longer available

The PFS censoring reasons are defined in the following way.

If the time interval between the last adequate TA date and the earliest of the following dates is smaller or equal to interval of 2 missing tumor assessments:

1. Analysis cut-off date,
2. Date of consent withdrawal,

3. Visit date of study treatment discontinuation or end of post-treatment follow-up discontinuation due to lost to follow-up.

Then the PFS censoring reason will be:

1. 'Ongoing',
2. 'Withdrew consent',
3. 'Lost to follow-up',

If the time interval is larger than the interval of 2 missing tumor assessments with no event observed, then the PFS censoring reason will always default to 'Adequate assessment no longer available'.

These summaries on censoring reasons will be produced for PFS by investigator. The censoring patterns will be compared between treatment groups .

Subgroup analyses for the primary endpoint

The primary endpoint of PFS will be summarized for the subgroups of stratification factors (per IRT) using an unstratified analysis. Additional subgroup analyses as specified in Section 2.2.3 will be performed to examine the homogeneity of treatment effect using the same conventions as for the primary analysis.

For each of the additional subgroups, the following analyses will be performed:

- Kaplan-Meier estimates of the median and its 95% CI
- Hazard ratio with 95% CI using unstratified Cox proportional hazards model.

Efficacy analyses in subgroups are intended to explore the consistency (homogeneity) of treatment effect. A forest plot (including sample size/number of PFS events and HR with 95% CI) will be produced to graphically depict the treatment effect estimates in different subgroups. No inferential statistics (p-values) will be produced for the subgroups.

Part B1

ORR with confirmed response as per blinded independent central review will be presented along with 95% confidence intervals.

Participants with 'unknown' BOR will be summarized by reason for having unknown status. The following reasons will be used:

- No valid post-baseline assessment
- All post-baseline assessments have overall lesion response UNK
- New anti-neoplastic therapy started before first post-baseline assessment
- SD too early
- PD too late

Note 1: A SD is considered as "SD too early" if the SD is documented within first 7 weeks after randomization date (i.e. until Day 49 included).

Note 2: A PD is considered as "PD too late" if the first documentation of PD is recorded more than 17 weeks after randomization date (i.e. from Day 119 included) with no qualifying CR, PR or SD in between.

Note 3: Special (and rare) cases where BOR is "unknown" due to both too early SD and too late PD will be classified as "SD too early".

Waterfall plot to depict anti-tumor activity

Waterfall graphs will be used to depict the anti-tumor activity. These plots will display the best percentage change from baseline in the sum of diameters of all target lesions for each participant. Only participants with measurable disease at baseline will be included in the waterfall graphs. Special consideration is needed for assessments where the target lesion response is CR, PR or SD, but the appearance of a new lesion or a worsening of non-target lesions results in an overall lesion response of PD. As a conservative approach, such assessments will not be considered for display as bars in the graph, since the percentage change in the sum of diameters of target lesions reflects the non-PD target lesion response, but the overall lesion response is PD. A participant with only such assessments will be represented by a special symbol (e.g. ✶) in the waterfall graph. Assessments with "unknown" target lesion response and assessments with unknown overall response will be excluded from the waterfall plots. Participants without any valid assessments will be completely excluded from the graphs.

The total number of participants displayed in the graph will be shown and this number will be used as the denominator for calculating the percentages of participants with tumor shrinkage and tumor growth. Footnote will explain the reason for excluding some participants (due to absence of any valid assessment).

All possible assessment scenarios are described in Table 2-6.

Table 2-6 Inclusion/exclusion of assessments used in waterfall graph

	Criteria for inclusion/exclusion			Possible sources of contradictions	
Case	Target response	Overall lesion response	Include in waterfall?	Non-target response	New lesion?
1	CR/PR/SD	PD	Yes but as ✶ only	PD	any
2	CR/PR/SD	PD	Yes but as ✶ only	any	Yes
3	UNK	UNK or PD	No	any	any
4	CR/PR/SD	UNK	No	UNK	No
5	CR/PR/SD	CR/PR/SD	Yes as a bar	SD/IR	No
6	PD	PD	Yes as a bar	any	any

Percentage change from baseline in the sum of diameters of all target lesions over time will be displayed for individual participants.

2.6 Analysis of the key secondary objective

Not applicable after protocol amendment 02.

2.7 Analysis of secondary efficacy objective(s)

Parts A and B2

The secondary efficacy objectives in Parts A and B2 are to compare the two treatment groups with respect to PFS for participants by PIK3CA mutation status as measured in baseline ctDNA, and to evaluate the overall survival (OS), overall response rate (ORR), clinical benefit rate (CBR), time to response (TTR), duration of response (DOR) and time to definitive deterioration in ECOG performance status.

Note Part B2 was never initiated.

Part B1

The secondary objectives in Part B1 are to evaluate additional efficacy parameters with respect to PFS and OS, and to evaluate clinical benefit rate (CBR), time to response (TTR) and duration of response (DOR).

No statistical testing of secondary endpoints in any Study Part will be undertaken.

2.7.1 Secondary endpoints

Overall survival (OS)

OS is defined as the time from date of randomization to date of death due to any cause. If a participant is not known to have died, then OS will be censored at the latest date the participant was known to be alive (on or before the cut-off date).

If a participant is not known to have died at the time of analysis cut-off, then OS will be censored at the last known date participant was alive, i.e., last contact date (see Section 2.1.1.11).

For rare cases when either day is missing or both month and day are missing for the date of death, the follow imputation rules will be implemented:

- If only day is missing, then impute max [(1 mmm-yyyy), min (last contact date +1, cutoff date)].
- If both day and month are missing, then impute max [(1 Jan-yyyy, min (last contact date +1, cutoff date)].

Overall response rate (ORR)

ORR with confirmed response is defined as the proportion of participants with BOR of confirmed complete response (CR) or confirmed partial response (PR), as per local review and according to RECIST 1.1 (see Section 16.1 of the study protocol) and Section 5.5.2 of this document.

ORR with confirmed response will be calculated based on the FAS. Participants with only non-measurable disease at baseline will be included in the numerator if they achieve a complete response. Only tumor assessments performed before the start of any further anti-neoplastic therapies (i.e. any additional anti-neoplastic therapy or surgery) and within 30 days after the

last administration of study treatment will be considered in the assessment of best overall response.

Further anti-neoplastic therapies will be identified from the data collected on 'Anti-neoplastic therapies since discontinuation of study treatment' eCRF. Palliative radiotherapy is the only setting of radiotherapy allowed during the study. Therefore, palliative radiotherapy will not be considered as an anti-neoplastic therapy for assessment of BOR. Continuation of combination partner therapy alone after end of study treatment will also not be considered as a new anti-neoplastic therapy.

Clinical benefit rate (CBR)

CBR with confirmed response is defined as the proportion of participants with a best overall response of confirmed complete response (CR) or confirmed partial response (PR), or an overall response of stable disease (SD) lasting for at least 24 weeks. CR, PR, and SD are defined as per local review according to RECIST 1.1 (see Section 16.1 of the study protocol). A participant will be considered to have SD for 24 weeks or longer if a SD response is recorded at 24-1=23 weeks or later from randomization, allowing for the ±1 week visit window for tumor assessments. Participants with only non-measurable disease at baseline will be included in the numerator only if they achieve a complete response or have a 'Non-CR/Non-PD' response at 24-1=23 weeks or later from randomization. CBR will be calculated using the FAS based on the local investigators' tumor assessments.

Time to response (TTR)

TTR (CR or PR) is defined as the time from the date of randomization to the first documented response of CR or PR, which must be subsequently confirmed (although date of initial response is used, not date of confirmation) using local investigators' review of tumor assessment data and according to RECIST 1.1 (see Section 16.1 of the study protocol).

All participants in the FAS will be included in TTR calculations. Participants without a confirmed CR or PR will be censored at the study-maximum follow-up time (i.e. LPLV-FPFV) for participants with a PFS event (i.e. disease progression or death due to any cause), or at the date of the last adequate tumor assessment for participants without a PFS event.

Duration of response (DOR)

DOR with confirmed response only applies to participants whose best overall response is confirmed CR or confirmed PR according to RECIST 1.1 based on tumor response data per local investigators' review. The start date is the date of first documented response of CR or PR (i.e. the start date of response, not the date when response was confirmed), and the end date is defined as the date of the first documented progression or death due to underlying cancer. Participants continuing without progression or death due to underlying cancer will be censored at the date of their last adequate tumor assessment.

2.7.2 Statistical hypothesis, model, and method of analysis

Secondary objectives fall outside of the statistical testing framework specified for the primary and key secondary objectives thus no statistical testing will be undertaken.

Overall Survival

OS will be analyzed in the FAS according to the treatment group and strata assigned at randomization (Parts A and B2) or at enrollment (Part B1). The OS distribution will be estimated using the Kaplan-Meier method, and the Kaplan-Meier curves, medians and 95% confidence intervals of the medians will be presented for each treatment group. The hazard ratio for OS will be calculated, along with its 95% confidence interval, using a stratified Cox model.

The pattern of censored data will be examined between the treatment groups: reasons for censoring ('Alive' or 'Lost to follow-up') and death cause will be summarized by treatment group. In addition, survival status, reason for censoring and death cause will be listed. Participants not known to have died will be censored for 'Lost to follow-up' if the time between their last contact date and the analysis cut-off date is longer than the protocol specified interval between the survival follow-up assessments plus 2 weeks, i.e., 12 + 2 weeks = 14 weeks for this study (i.e. the planned interval between two OS follow-up visits plus the 1 week window on either side).

Subgroup analysis of OS will be performed using subgroups as defined in Section 2.2.3. [redacted]

ORR/CBR

ORR/CBR with confirmed response per local review based on the FAS will be summarized using descriptive statistics (N, %) by treatment group along with two-sided exact binomial 95% CIs (Clopper and Pearson 1934).

As supplemental analyses in Parts A and B2, the following will also be calculated and presented by treatment group together with exact binomial 95% confidence intervals:

- ORR/CBR with unconfirmed response per local review based on the FAS

Time to response

Time to response will be listed and summarized by treatment group. The distribution of time to response will be estimated using the Kaplan-Meier method and the median time to response will be presented along with 95% confidence interval. A descriptive summary of time to response for the responders will also be presented. No inferential analysis that compares time to response between the two treatment groups will be performed.

Duration of Response

DOR with confirmed response will be listed and summarized by treatment group for all participants in the FAS with confirmed BOR of CR or PR. The distribution of duration of response will be estimated using the Kaplan-Meier method and the median duration of response will be presented along with 95% confidence interval. No inferential analysis that compares duration of response between the two treatment groups will be performed.

PFS by PIK3CA mutation status as measured in ctDNA at baseline

An analysis of PFS based on local radiology assessments and using RECIST 1.1 criteria by PIK3CA mutation status as measured in ctDNA at baseline will be conducted using the same analytical conventions as the primary analysis.

2.8 Analysis of other secondary objective(s)

2.8.1 Safety analyses

For all safety analyses, each Study Part will be summarized separately and the safety set will be used; no pooled safety analyses are planned. All listings and tables will be presented by treatment group for each Study Part.

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In addition, a separate summary for deaths including on treatment and post treatment deaths will be provided.

The on-treatment period lasts from the date of first administration of study treatment to 30 days after the date of the last actual administration of any study treatment.

The overall observation period will be divided into three mutually exclusive segments:

1. Pre-treatment period: from day of participant's informed consent to the day before first dose of study treatment
2. On-treatment period: from day of first dose of study treatment to 30 days after last dose of study treatment
3. Post-treatment period: starting at day 31 after last dose of study treatment.

2.8.1.1 Adverse events (AEs)

Summary tables for adverse events (AEs) will include only AEs that started or worsened during the on-treatment period, the treatment-emergent AEs. However, all safety data (including those from the pre and post-treatment periods) will be listed and those collected during the pre-treatment and post-treatment period are to be flagged. All information obtained on adverse events will be displayed by treatment group for each study part.

The number (and percentage) of participants with treatment emergent adverse events may be summarized in the following ways:

- by treatment, system organ class and preferred term.
- by treatment, system organ class, preferred term and maximum severity.
- by treatment, Standardized MedDRA Query (SMQ) and preferred term.

A participant with multiple adverse events within a system organ class is only counted once towards the total of the system organ class.

Separate summaries will be provided for study treatment related adverse events, death, serious adverse events, non-serious adverse events and other significant adverse events e.g. leading to discontinuation and adverse events leading to dose adjustment.

In addition, a summary of serious adverse events with number of occurrences will be produced (an occurrence is defined as >1 day between the start and end date of the same preferred term).

The incidence of treatment-emergent adverse events (new or worsening from baseline) will be summarized by system organ class and or preferred term, severity (based on CTCAE grades), type of adverse event, relation to study treatment. A participant with multiple grades for an AE will be summarized under the maximum grade recorded for the event. AEs with a missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries the system organ class will be presented alphabetically and the preferred terms will be sorted within SOC in descending frequency. The sort order for the preferred term will be based on their frequency in the alpelisib treatment group in each Study Part.

The frequency of grade 3 and above AEs will be summarized separately.

Any information collected (e.g. grades, relationship to study treatment, serious, action taken etc.) will be summarized as appropriate.

All AEs, deaths, and serious adverse events (including those from the pre and post-treatment periods) will be listed and those collected during the pre-treatment and post-treatment period will be flagged.

2.8.1.2 Adverse events of special interest / grouping of AEs

An adverse event of special interest (AESI) is a grouping of adverse events that are of scientific and medical concern specific to alpelisib. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HLGTs (high level group terms), HLTs (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. For each specified AESI, number and percentage of participants with at least one event of the AESI occurring during on-treatment period will be summarized.

Summaries of these AESIs will be provided by AESI grouping, maximum CTCAE grade and by treatment group for each part (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption etc.). A listing of all grouping levels down to the MedDRA preferred terms used to define each AESI will be generated.

An electronic Case Retrieval Sheet (eCRS) is to be used to map reported AEs to the AESI groupings. This file may be updated (i.e. it is a living document) based on review of accumulating trial data, and therefore the groupings are also subject to potential change. The most up-to-date version of the eCRS will be used at the time of the cut-off date for each analysis.

2.8.1.3 Deaths

Separate summaries for on-treatment and all deaths (including post-treatment death) will be produced by treatment group, system organ class and preferred term.

All deaths will be listed, post treatment deaths will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened participants.

2.8.1.4 Laboratory data

Regarding laboratory assessments, data from all sources (central and local laboratories) will be combined. For the purposes of handling multiple assessments at each visit, all scheduled/unscheduled assessments should be assigned to time windows (Table 2-7).

The summaries will include all laboratory assessments collected no later than 30 days after the last administration of study treatment. All laboratory assessments will be listed and those collected later than 30 days after the last treatment date will be flagged in the listings. Due to an error for insulin tubes (identified in January 2021), some pre-dose insulin samples were mistakenly collected instead of the 1 hour post-dose of alpelisib/placebo for post-baseline measurements. Such unexpected pre-dose insulin records will be flagged by comparing the insulin sample date/time with the chemistry sample date/time and excluded from the summary table.

Table 2-7 Time windows for laboratory assessments

Fasting Plasma Glucose (Baseline, C1D8, C1D15, C2D1, C2D8, C2D15, C3D1 and every CXD1 afterwards)

Assessment	Target day of assessment	Time Interval
Baseline		≤ Day 1
Cycle 1 Day 8	8	Day 2 to Day 11
Cycle 1 Day 15	15	Day 12 to Day 21
Cycle 2 Day 1	29	Day 22 to Day 32
Cycle 2 Day 8	36	Day 33 to Day 39
Cycle 2 Day 15	43	Day 40 to Day 49
Cycle 3 Day 1	57	Day 50 to day 70
Cycle k Day 1 (k≥4)	=(k-1)*28+1	Day d-14 to Day d+13
End of Treatment		Assessment taken at the EOT visit

HbA1C (Baseline, C2 D1 and every 3 cycles afterwards)

Assessment	Target day of assessment	Time Interval
Baseline		≤ Day 1
Cycle 2 Day 1	29	Day 2 to Day 70
Cycle k Day 1 (k=5,8,11,...)	=(k-1)*28+1	Day d-42 to Day d+41
End of Treatment		Assessment taken at the EOT visit

Insulin (Baseline, C2 D1 and every 2 cycles afterwards)

Assessment	Target day of assessment	Time Interval
Baseline		≤ Day 1
Cycle 2 Day 1	29	Day 2 to Day 56
Cycle k Day 1 (k=4,6,8,...)	=(k-1)*28+1	Day d-28 to Day d+27
End of Treatment		Assessment taken at the EOT visit

All laboratory data will be listed by treatment group, participant, and visit/time and if normal ranges are available abnormalities will be flagged. Summary statistics will be provided by treatment and visit/time.

Grading of laboratory values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher. The criteria to assign CTC grades are given in Appendix 5.3.

The following summaries will be produced for the laboratory data (by laboratory parameter):

- Number and percentage of participants with each CTC grade as their worst post-baseline value (regardless of the baseline status). Each participant will be counted only for the worst grade observed post baseline.
- HbA1c, fasting glucose and insulin data will be summarized in tables by treatment group and time point. Summary statistics include number of participants with available data, mean, standard deviation, median, minimum and maximum. Figures of mean glucose with two-sided 95% confidence intervals over time by treatment group may also be produced to view the trends over time (only fasting glucose values will be used in this plot).

The following listings will be produced for the laboratory data:

- Listing of participants with CTC grade 3 or 4 laboratory abnormalities;
- Listing of all laboratory data with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges

Note that for glucose, CTC criteria only considers fasting values for Grade 1 and 2. For CTC Grade 3 and 4, both fasting and non-fasting values are considered. For summary involving comparison between 2 visits, only fasting glucose will be used to be consistent between visits, e.g., shift table, plot for mean glucose results over time. Otherwise, both fasting and non-fasting glucose values will be used.

Liver function parameters

Liver function parameters of interest in this study are total bilirubin (TBIL), ALT, AST and alkaline phosphatase (ALP).

The number and percentage of participants meeting the following categorical liver function test criteria will be summarized:

- ALT > 3xULN, 5xULN, 10xULN, 20xULN
- AST > 3xULN, 5xULN, 10xULN, 20xULN
- ALT or AST > 3xULN, 5xULN, 8xULN, 10xULN, 20xULN
- TBIL > 2xULN, 3xULN

For the following combined categories, the assessments need not be concurrent, i.e. participants are counted based on their most extreme value for each parameter (highest in the case of ALT, AST and TBIL; lowest in the case of ALP). Further medical review can assess potential confounding factors such as, liver metastases, liver function at baseline etc.

- If AST and ALT ≤ ULN at baseline
  - ALT or AST >3x ULN & TBIL>2x ULN
  - ALT or AST >3x ULN & TBIL>2x ULN & ALP ≥2x ULN
  - ALT or AST >3x ULN & TBIL>2x ULN & ALP <2x ULN
- If AST and ALT > ULN at baseline
  - Elevated ALT or AST (>3x Baseline value or 8x ULN) & TBIL(>2x Baseline value and 2x ULN)
  - Elevated ALT or AST (>3x Baseline value or 8x ULN) & TBIL(>2x Baseline value and 2x ULN) & ALP ≥2x ULN
  - Elevated ALT or AST (>3x Baseline value or 8x ULN) & TBIL(>2x Baseline value and 2x ULN) & ALP <2x ULN

Additional categories may be added to the above list based on any updates to the internal guidelines on collection, analysis, and presentation of liver safety data.

2.8.1.5 ECG

Notable elevations of ECG summarize the number of participants meeting or exceeding predefined limits in terms of absolute QT/QTc interval data/PR/RR/QRS or changes from baseline as defined in Table 2-8. Notable elevations include only newly occurring ECG. A newly occurring ECG abnormality is defined as an abnormal post-baseline ECG finding that is not present at Baseline. The percentage of participants having notable ECG interval values is based on the number of participants at risk for the change with a value at baseline and post-baseline.

Table 2-8 Clinically notable ECG values

ECG parameter (unit)	Clinically notable criteria
QT, QTcF (ms)	New > 450 and ≤ 480
	New > 480 and ≤ 500
	New > 500
	Increase from Baseline > 30 to ≤ 60
	Increase from Baseline > 60
PR duration (ms)	Increase from baseline >25% and to a value > 200
	New > 200
QRS duration (ms)	Increase from baseline >25% and to a value > 120
	New > 120
Heart Rate (bpm)	< 50 and decrease from Baseline of > 25%
	> 100 and increase from Baseline of > 25%

All ECG data will also be listed by treatment group, participant and visit/time for each study part. Abnormalities will be flagged.

2.8.1.6 Cardiac imaging

A listing of participants with newly occurring clinically significant abnormality will be produced by study part and treatment group.

2.8.1.7 Vital signs

Vital signs assessments are performed in order to characterize basic body function. The parameters expected to be collected include: weight, body temperature, pulse rate, and systolic and diastolic blood pressure.

The criteria for clinically notable abnormalities are defined in Table 2-9 below.

Table 2-9 Clinically notable changes in vital signs

Vital sign (unit)	Clinically notable criteria	
	above normal value	**below normal value**
Systolic blood pressure (mmHg)	≥180 with increase from baseline of ≥20	≤90 with decrease from baseline of ≥20
Diastolic blood pressure (mmHg)	≥105 with increase from baseline of ≥15	≤50 with decrease from baseline of ≥15
Pulse rate (bpm)	≥100 with increase from baseline of >25%	≤50 with decrease from baseline of > 25%
Body temperature	≥ 39.1	-
Weight (kg)	increase ≥ 10% from Baseline	decrease ≥ 10% from Baseline

The following summary will be produced for each vital sign parameter:

- Number and percentage of participants with notable vital sign values during treatment period

In addition, the following listing will be produced by treatment group:

- Participants with clinically notable vital sign abnormalities.

In both listings, the clinically notable values will be flagged and also assessments collected later than 30 days after the last treatment date will be flagged.

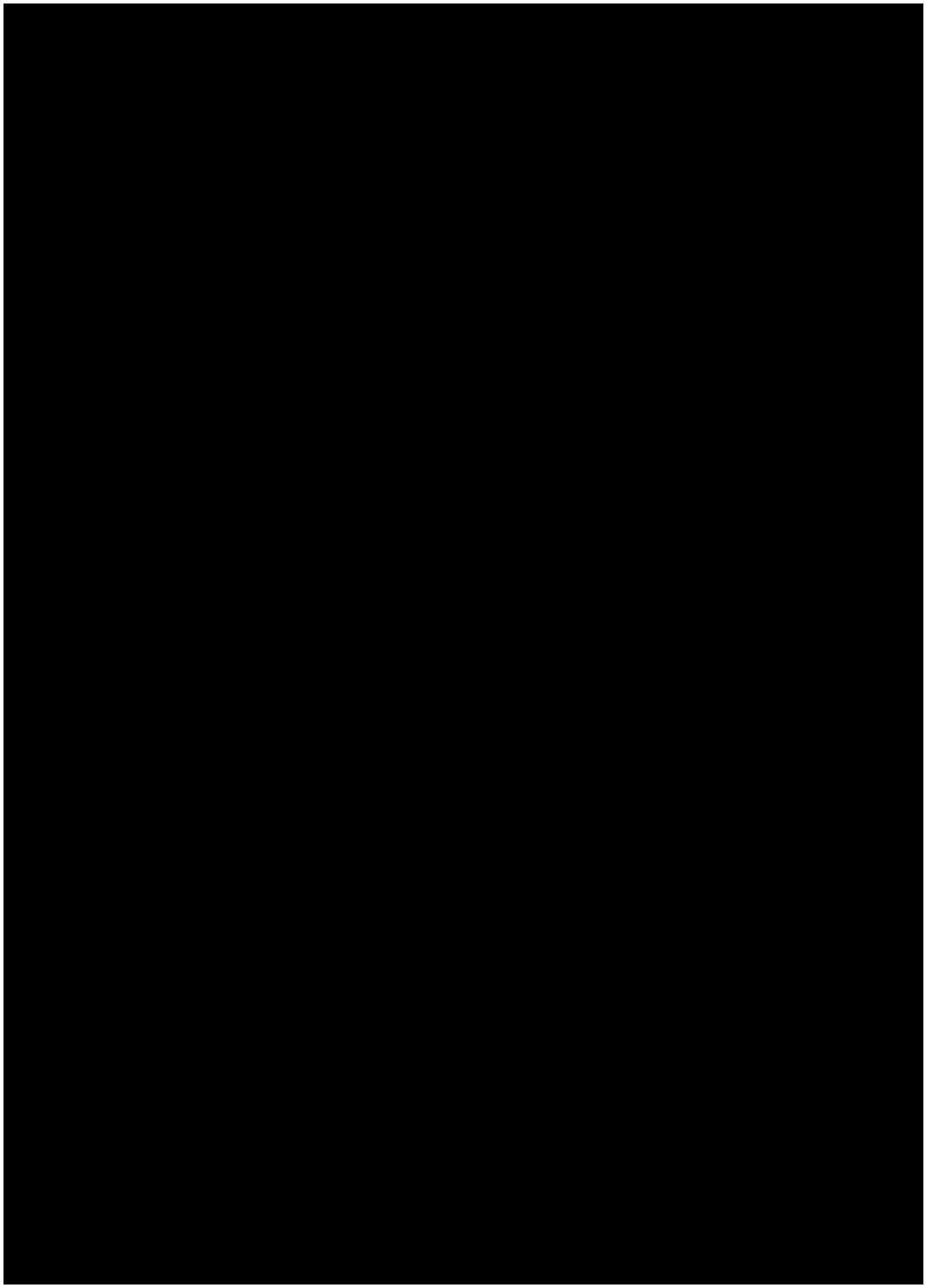
2.8.1.8 Other safety data

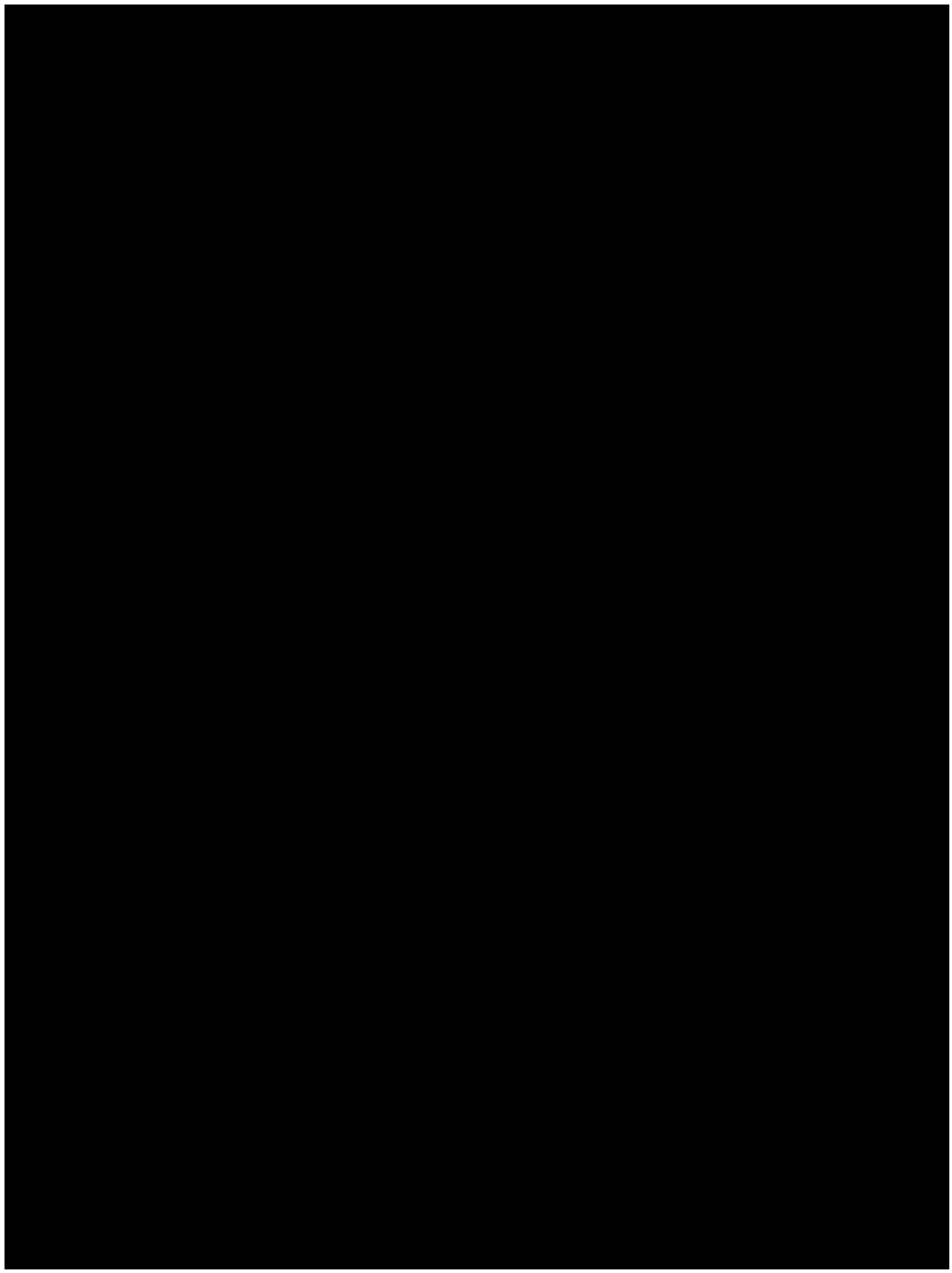
Data from other tests will be listed, notable values will be flagged, and any other information collected will be listed as appropriate.

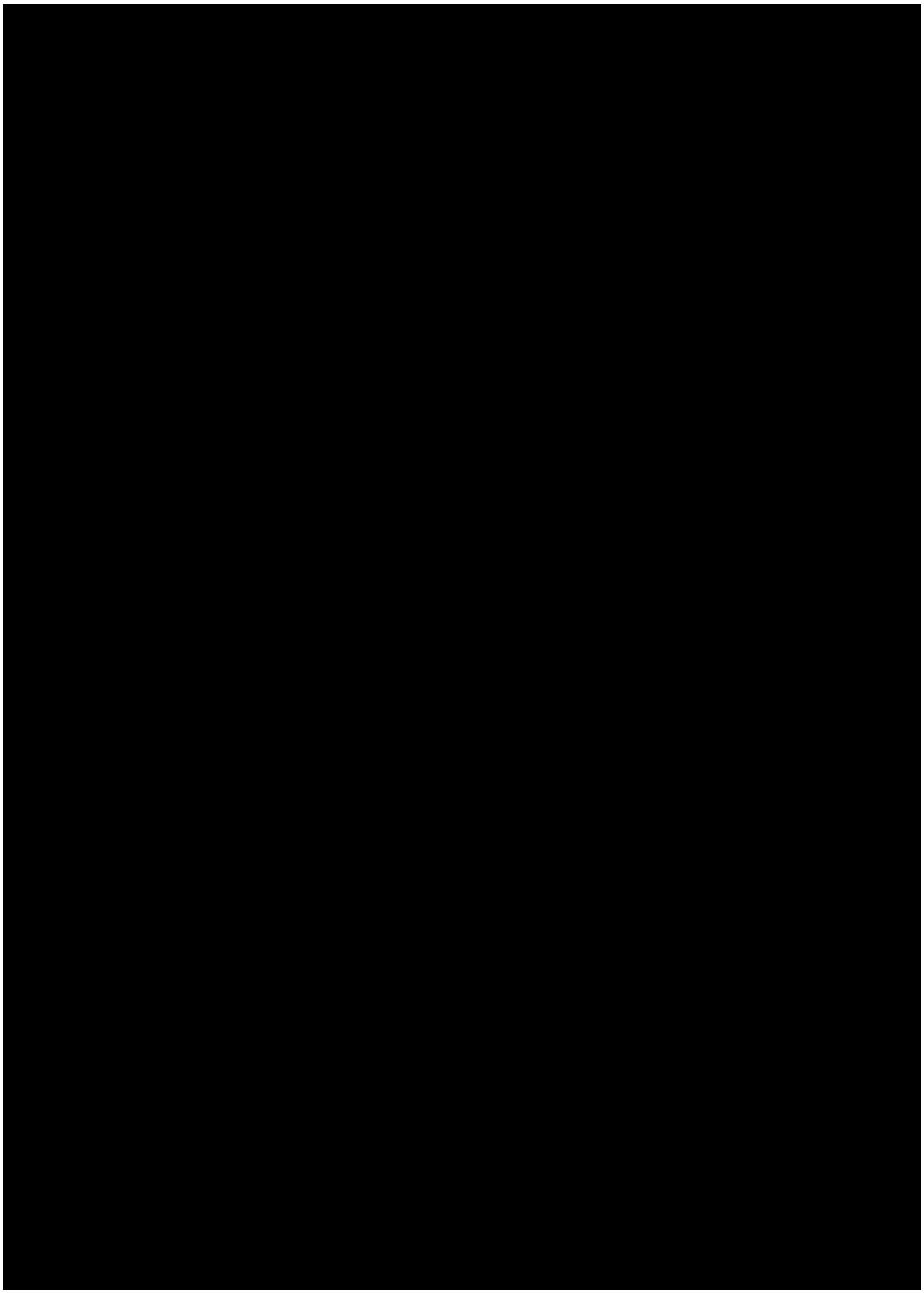
All assessments collected later than 30 days after the last treatment date will be flagged in the listings.

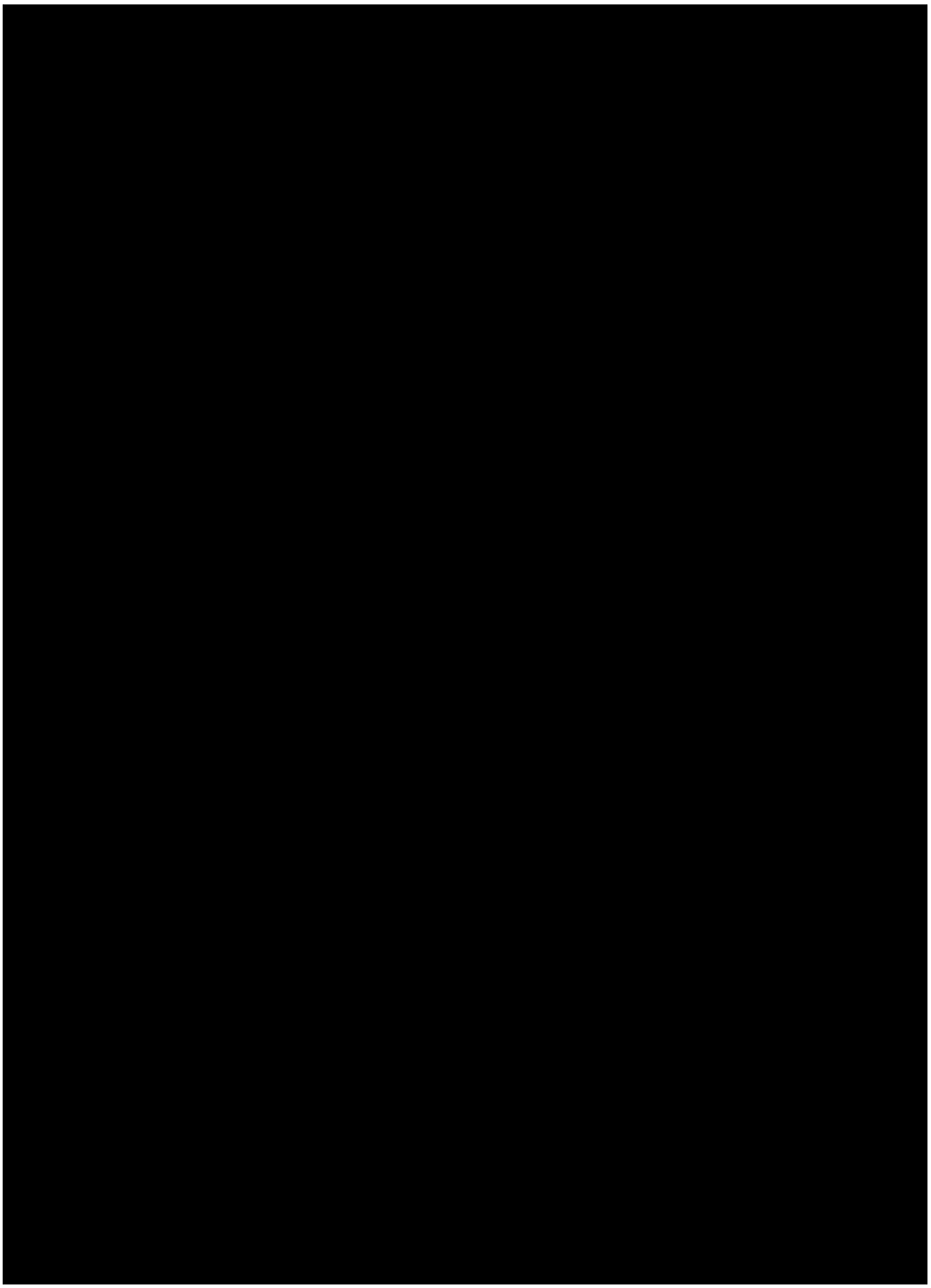
Subgroup analyses will be explored as described in Section 2.2.3.

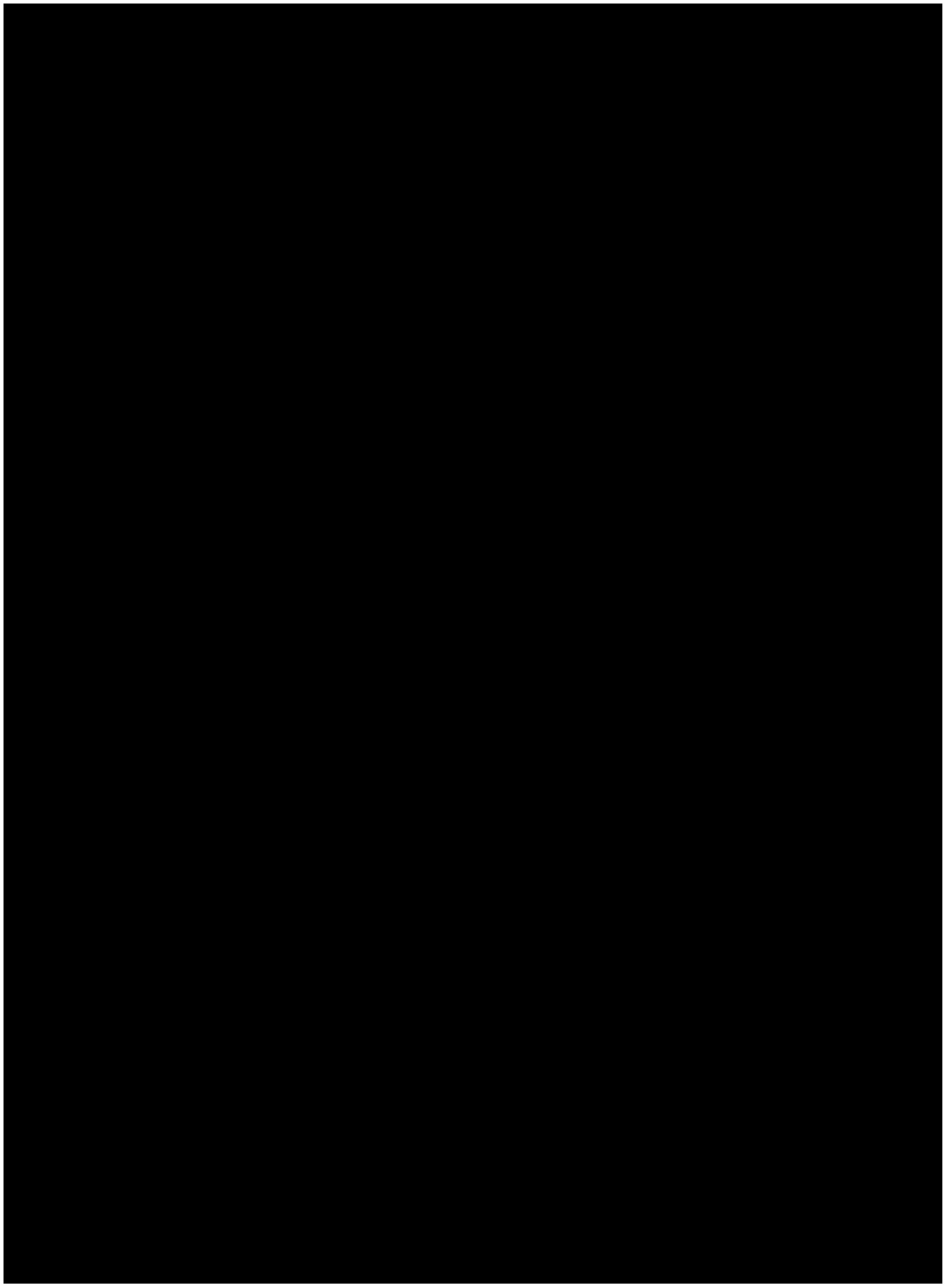
[redacted]

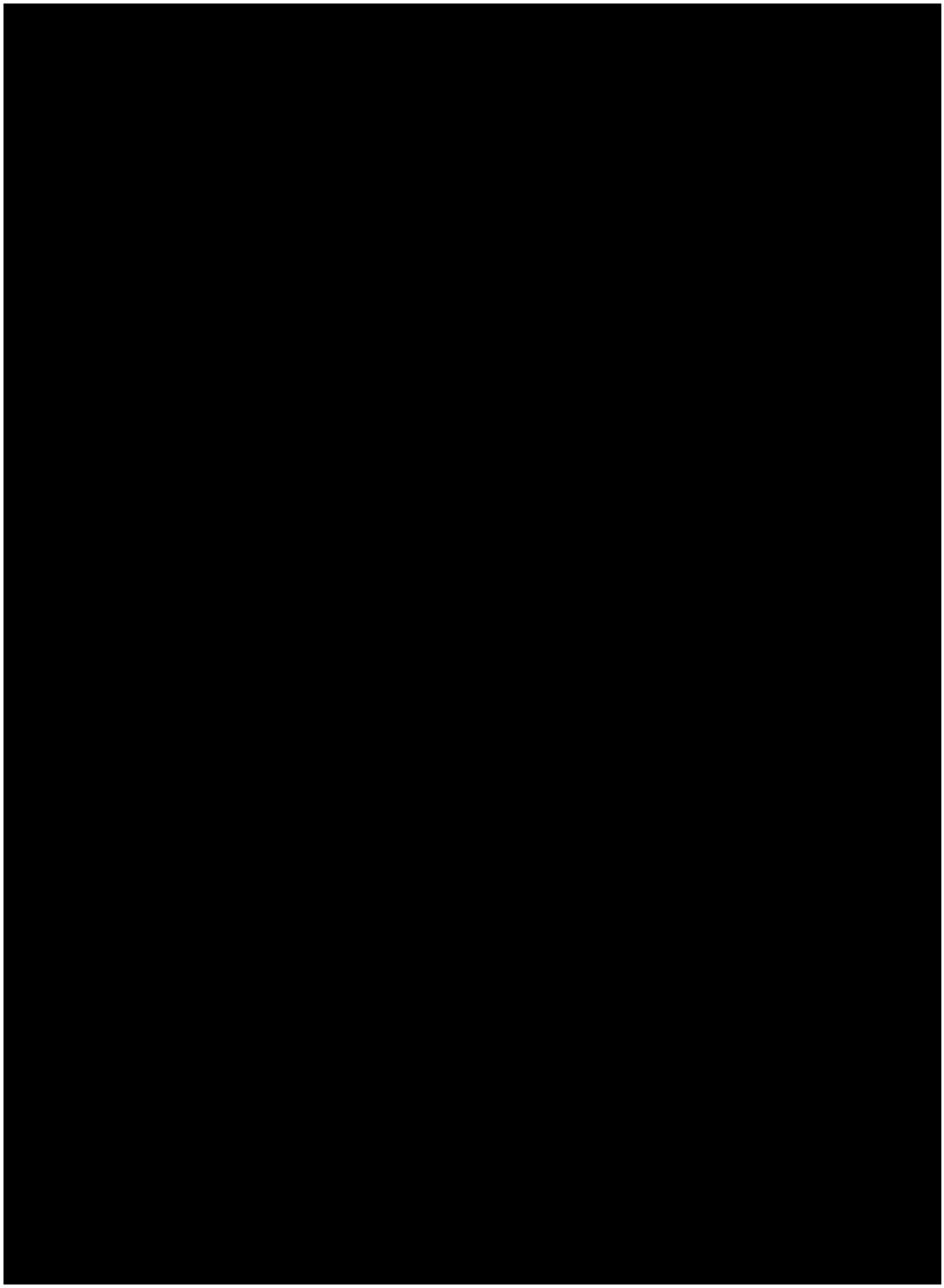












2.10 Interim analysis

Interim analyses are planned for the monitoring of safety data by DMC, and will be performed approximately every 6 months during the course of the study.

Part B1

No formal interim efficacy analysis is planned for this part B1. The primary analysis has been performed after all participants had completed 6 months treatment or had discontinued treatment prior to 6 months.

Progression-free survival (PFS)

Part A

No interim efficacy analysis will be performed for Part A. After the halt of recruitment, the primary PFS analysis for Part A will be performed when all participants have completed at least 6 months of study treatment or have discontinued from study treatment.

Part B2

Not applicable. Part B2 was not initiated based on the results from Part B1.

3 Sample size calculation

Note the below subsections illustrate how the original sample size was considered. Based on the decision of recruitment halt, Part A will not wait for the required number of PFS/OS events to perform the originally planned interim/final analysis. Instead, similar to Part B1, the primary analysis for Part A will be descriptive only and will be undertaken when all participants have completed 6 months of study treatment or have been discontinued from study treatment, whichever occurs earlier.

3.1 Primary endpoint(s)

The sample size calculation in Parts A and B2 is based on the primary variable of PFS. The hypotheses to be tested and details of the testing strategy are described in Section 2.5.2.

The median PFS in the control group (placebo + nab-paclitaxel) of this study is estimated to be approximately 5 months, based on data from first and second line mTNBC patients treated with either nab-paclitaxel or paclitaxel in the PAKT study (Schmid et al 2018), the LOTUS study (Kim et al 2017), and the Impassion130 study (Schmid et al 2018).

It is expected that treatment with alpelisib plus nab-paclitaxel will result in a 40% reduction in the hazard rate for PFS, i.e. an expected hazard ratio of 0.6 (which corresponds to an increase in median PFS to 8.33 months under the exponential model assumption).

Part A

In order to ensure 90% power to test the null hypothesis: PFS hazard ratio = 1, versus the specific alternative hypothesis: PFS hazard ratio = 0.6, it is calculated that a total of 192 PFS events need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, participants randomized to the two treatment groups in a 1:1 ratio, and a 3-look group sequential design with a gamma spending function to define a futility rule at the 1st interim analysis and a Haybittle-Peto alpha spending function, using information fractions of (0.40, 0.85, 1).

Assuming that enrolment will continue for 29 months enrolling 3 participants/month up to month 6, 8 participants/month up to month 12 and 11 participants/month thereafter, assuming losses to follow-up for PFS of 10%, a total of 252 participants will need to be randomized to observe the targeted 192 PFS events at about 6 months after the randomization date of the last participant, i.e. 35 months after the randomization date of the first participant.

Part B2

In order to ensure 90% power to test the null hypothesis: PFS hazard ratio = 1, versus the specific alternative hypothesis: PFS hazard ratio = 0.6, it is calculated that a total of 192 PFS events need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, participants randomized to the two treatment groups in a 1:1 ratio, and a 3-look group sequential design with a gamma spending function to define a futility rule at the 1st interim analysis and a Haybittle-Peto alpha spending function, using information fractions of (0.40, 0.75, 1).

Assuming that enrolment will continue for 16.5 months enrolling 12 participants/month up to month 6 and 20 participants/month thereafter, assuming losses to follow-up for PFS of 10%, a total of 282 participants will need to be randomized to observe the targeted 192 PFS events at about 5.5 months after the randomization date of the last participant, i.e. 22 months after the randomization date of the first participant.

These calculations were made using the software package East 6.4.

Part B1

The sample size calculation is based on the primary variable ORR considering the statistical model, hypothesis and method of analysis detailed in Section 2.5.2. Proof of preliminary efficacy (PPE) will be declared if both of the following conditions are met:

- the mean of the posterior distribution of ORR is at least 35%

and

- the posterior probability that the ORR is ≥ 25% is at least 0.9

Approximately 32 participants will be enrolled. With 32 participants, the probability of declaring proof of preliminary efficacy (PPE) is at most 8% when ORR ≤25%. The probability of declaring PPE is at least 80% for ORR ≥ 45% (Table 3-1).

Table 3-1 Part B1 PPE Operating characteristics

True ORR	Probability of declaring PPE (≥12 responders)	Probability of missing PPE (≤11 responders)
25%	0.079	0.921
30%	0.228	0.772
35%	0.448	0.552
40%	0.675	0.325
45%	0.847	0.153

3.2 Power for analysis of key secondary variables

Part A

Per original design, OS, as the key secondary variable, was to be formally statistically tested, provided that the primary variable of PFS is statistically significant. The hypotheses to be tested and details of the testing strategy are provided in original Section 2.6.2. The median OS in the control group (placebo + nab-paclitaxel) of this study is estimated to be approximately 12

months, based on data from first and second line mTNBC participants treated with either nab-paclitaxel or paclitaxel in the PAKT study (Schmid et al 2018), the LOTUS study (Kim et al 2017), and the IMpassion130 study (Schmid et al 2018). It is hypothesized that treatment with alpelisib plus nab-paclitaxel will result in a 33.3% reduction in the hazard rate for OS, i.e., an expected hazard ratio of 0.667 (which corresponds to an increase in median OS to 18 months under the exponential model assumption). Then in order to ensure 80% power to test the null hypothesis: OS hazard ratio = 1, versus the specific alternative hypothesis: OS hazard ratio = 0.667, it is calculated that a total of 198 deaths need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, participants randomized to the two treatment groups in a 1:1 ratio, and a 3-look group sequential design with a Lan-DeMets (O'Brien and Fleming 1979) alpha spending function using information fractions of (0.66 [if at final PFS analysis], 0.85, 1).

Based on the same number of participants that are planned to be enrolled in this study to provide sufficient power for the primary endpoint (i.e. 252 participants), and assuming losses to follow-up for OS of 10%, it is estimated that these 198 deaths will be observed approximately 66 months after the randomization date of the first participant. Therefore the cut-off date for the final analysis of OS will be approximately 31 months after the cut-off date for the final analysis of PFS. These calculations were made using the software package East 6.4.

Part B1

Not applicable.

Part B2

Similarly, OS, as the key secondary variable per original design, will be formally statistically tested, provided that the primary variable PFS is statistically significant. The hypotheses to be tested and details of the testing strategy are provided in the original section 2.6.2. The median OS in the control group (placebo + nab-paclitaxel) of this study is estimated to be approximately 12 months, based on data from first and second line mTNBC patients treated with either nab-paclitaxel or paclitaxel in the PAKT study (Schmid et al 2018), the LOTUS study (Kim et al 2017), and the IMpassion130 study (Schmid et al 2019). It is hypothesized that treatment with alpelisib plus nab-paclitaxel will result in a 33.3% reduction in the hazard rate for OS, i.e., an expected hazard ratio of 0.667 (which corresponds to an increase in median OS to 18 months under the exponential model assumption). Then in order to ensure 80% power to test the null hypothesis: OS hazard ratio = 1, versus the specific alternative hypothesis: OS hazard ratio = 0.667, it is calculated that a total of 197 deaths need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, participants randomized to the two treatment groups in a 1:1 ratio, and a 3-look group sequential design with a Lan-DeMets (O'Brien and Fleming 1979) alpha spending function using information fractions of (0.60 [if at final PFS analysis], 0.85, 1).

Based on the same number of participants that are planned to be enrolled in this study to provide sufficient power for the primary endpoint (i.e. 282 participants), and assuming losses to follow-up for OS of 10%, it is estimated that these 197 deaths will be observed approximately 41 months after the randomization date of the first participant. Therefore the cut-off date for the

final analysis of OS will be approximately 19 months after the cut-off date for the final analysis of PFS. These calculations were made using the software package East 6.4.

4 Change to protocol specified analyses

No change from protocol specified analysis has been specified.

5 Appendix

5.1 Imputation rules

The missing or partial date imputation rules will be described in the programming datasets specification document.

5.2 AEs coding/grading

Adverse events are coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

AEs will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.3.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death.

If CTCAE grading does not exist for an adverse event, grades 1 – 5 corresponding to the severity of mild, moderate, severe, life-threatening, and death will be used. Information on deaths will also be collected on the 'Death' eCRF.

5.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version v4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters shown below.

For laboratory tests where grades are not defined by CTCAE v4.03, results will be graded by the low/normal/high classifications based on laboratory normal ranges.

A severity grade of 0 will be assigned for all non-missing lab values not graded as 1 or higher. Grade 5 is not applicable for the lab data. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

If laboratory values are provided as '<X' (i.e. below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, these numeric values are set to X-0.0001 and X+0.0001, respectively.

5.4 Statistical models

5.4.1 Baseline comparability

Appropriate descriptive summary statistics of baseline variables (see Section 2.3.2) will be provided by Study Part. No p-values will be provided.

5.4.2 Analysis of time to event Data

The following sections present a general methodology to be used to analyze time-to-event variables. Inferential testing however will only be conducted for PFS and OS, i.e. the primary and key secondary endpoints as detailed in Section 2.5 and Section 2.6. The following parameters are considered:

- Progression-free survival
- Overall survival
- Time to definitive deterioration of the ECOG score by at least one category of the score from baseline
- Time to response: defined as the time between date of randomization until first documented response (CR or PR) according to RECIST 1.1
- Duration of response
- Time to definitive deterioration of PRO scores

Kaplan-Meier estimates

The survival function in each treatment group will be estimated using the Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG.

Median survival for each treatment group will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the method of (Brookmeyer and Crowley 1982). Kaplan-Meier estimates of the survival function with 95% confidence intervals at specific time points will be summarized. The standard error of the Kaplan-Meier estimate will be calculated using Greenwood's formula (Collett 1994).

Hazard ratio

Hazard ratio will be estimated by fitting the Cox proportional hazards model using SAS procedure PHREG (with TIES=EXACT option in the MODEL statement).

For the stratified unadjusted Cox model (where the baseline hazard function is allowed to vary across strata), the MODEL statement will include the treatment group variable as the only covariate and the STRATA statement will include stratification variable(s) per IRT.

Hazard ratio with two-sided 95% confidence interval will be based on Wald test.

Treatment of ties

The STRATA statement in LIFETEST procedure will be used to analyze time to event data with ties. The PHREG procedure in SAS with option TIES=EXACT will be used to fit the Cox proportional hazards model.

Checking proportionality of hazard assumption

Plots (SURVIVAL LOGSURV LOGLOGS) generated by LIFETEST procedure in SAS will be used to provide visual checks of the proportional hazard assumption.

- SURVIVAL plots estimated survivor functions. The shape of the curves should be basically the same if hazards are proportional.
- LOGSURV plots the cumulative hazard functions The larger cumulative hazard should be a multiple of smaller if hazards are proportional
- LOGLOGS plots log (cumulative hazard). The LOGLOG plot will show parallel curves if hazards are proportional.

5.4.3 Analysis of binary data

The two criteria to assess the proof of preliminary efficacy (PPE) in part B1 are:

- Clinical relevance: posterior mean ≥ 35%
- Bayesian statistical significance: pr(ORR ≥ 25% | data) ≥ 0.90

Let p_i denote the ORR for treatment group i and which follows a beta prior distribution Beta [a,b], where a>0, b>0. Let y_i out of n_i participants be responders. Therefore, the posterior distribution of p_i is Beta($a+ y_i$, $b+ n_i-y_i$) (Spiegelhalter et al. 2004).

A minimally informative unimodal Beta prior (Neuenschwander et al. 2008) Beta [0.35/(1-0.35),1] will be used. The parameters were chosen so that the mean of the prior distribution is equal to 0.35, which ensures that the clinical relevance criterion is met, if the observed ORR is exactly equal to 35%.

The efficacy criteria will be assessed based on the actual number of participants enrolled in the study. For example, if the total number of participants is 32 in a given treatment group, the first efficacy criterion requires that at least 12 participants are responders. In that case, the posterior distribution is Beta (0.35/(1-0.35)+12 , 1+32-12) and the posterior probability pr(ORR ≥ 25% | data) = 0.939.

For further details, see Appendix 4 of the study protocol.

Confidence interval for response rate and clinical benefit rate

Responses will be summarized in terms of percentage rates with 95% confidence intervals using exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement for one-way table (Clopper and Pearson 1934).

5.5 Implementation of RECIST

Response and progression evaluation will be performed according to Novartis RECIST guideline (as described in detail in Section 16.1 of the study protocol), which is based on RECIST version 1.1 (Eisenhauer et al 2009). The text below gives instructions and rules to provide details needed for programming.

5.5.1 Overall lesions response for participants with only non-measurable lesions at baseline

Participants without measurable disease per RECIST 1.1 are eligible to enter the study if they have at least one predominantly lytic bone lesion. For participants with non-measurable lesions only at baseline, the overall lesion response will be based solely on non-target lesion response or an occurrence of a new lesion. Non-measurable lesions will be entered as non-target lesions. Therefore, the best overall response is determined from non-target lesion response and presence of new lesions (refer to Table 16-4 in Section 16.1.3.2.9 of the study protocol).

5.5.2 Best overall response (BOR)

The best overall tumor response will be assessed as per RECIST 1.1 criteria. The definitions and the details on the derivation are given in Section 16.1 of the study protocol.

Only tumor assessments performed before the start of any new antineoplastic therapies (i.e. any additional antineoplastic therapy or surgery) and within 30 days after the last administration of study treatment will be included in the assessment of best overall response.

- New antineoplastic therapies will be identified from the data collected on 'Antineoplastic medications since discontinuation of study treatment' eCRF.
- Palliative radiotherapy is the only setting of radiotherapy allowed during the study. Therefore, palliative radiotherapy will not be considered as an anti-neoplastic therapy for assessment of BOR unless reported on the 'Anti-neoplastic Radiotherapies since discontinuation of study treatment' eCRF. As per RECIST 1.1, it should not be delivered to a target lesion.
- Continuation of study combination partner therapy alone after end of study treatment without confirmed progression, will also not be considered as a new antineoplastic therapy.

The standard definition of a best overall response evaluation of 'stable disease', 'disease progression' or 'unknown' given in the Section 16.1.3.1 of the study protocol will be used for this study. Best overall response with confirmation of response for each participant is determined from the sequence of overall (lesion) responses (as reported by the investigator for local BOR) according to the following rules:

- CR = at least two determinations of CR at least 4 weeks apart before progression.
- PR = at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR).
- SD = at least one SD assessment (or better) > 7 weeks after randomization (and not qualifying for CR or PR).
- Non-CR/non-PD = at least one non-CR/non-PD assessment (or better) > 7 weeks after randomization date (and not qualifying for CR). This applies only for participants with non-measurable disease alone at baseline.
- PD = progression ≤ 17 weeks after randomization (and not qualifying for CR, PR, SD and non-CR/non-PD)

- UNK = all other cases (i.e. not qualifying for confirmed CR or PR and without SD or non-CR/non-PD after more than 7 weeks or without progression within the first 17 weeks).

Participants with best overall response "unknown" will be summarized by reason for having unknown status. The following reasons will be used:

- No valid post-baseline assessment
- All post-baseline assessments have overall response UNK
- New anti-neoplastic therapy started before first post-baseline assessment
- SD too early (≤7 weeks after randomization)
- PD too late (>17 weeks after randomization and not qualifying for CR, PR, SD or non-CR/non-PD)

Special (and rare) cases where BOR is unknown due to both early SD and late PD will be classified as "SD too early".

5.5.3 Disease progression

Progressive disease should only be assigned if it is confirmed by an assessment method as per RECIST 1.1 guidelines (e.g. radiologic assessment, photos for skin lesions, etc.). If a new lesion is detected using an objective assessment method other than radiologic assessment, then it should also be entered as a new lesion in the eCRF with the appropriate method. Discontinuation due to disease progression or death due to study indication, without corresponding supportive data in the RECIST CRF (as defined above), will not be considered as progressive disease in the calculation of best overall response and in the analysis of PFS.

5.5.4 Change in imaging modality

Per RECIST 1.1, a change in methodology can be defined as either a change in contrast use (e.g. keeping the same technique, like CT, but switching from 'with' to 'without' contrast use or vice-versa, regardless of the justification for the change), a major change in technique (e.g. from CT to MRI, or vice-versa), or a change in any other imaging modality. A change from conventional to spiral CT or vice versa will not constitute a major "change in method" for the purposes of response assessment. A change in methodology will result by default in an UNK (unknown) overall lesion response based on the Novartis calculation. However, a response from the investigator or the central blinded reviewer that differs from the Novartis calculated UNK is acceptable, if a definitive response assessment can be justified based on the available information.

Potential discrepancies between the modality used and overall lesion response (e.g. change in modality but response is different from 'Unknown') will be queried during the data validation process.

5.5.5 Determination of missing adequate assessments

The term 'missing adequate assessment' refers to assessments that are not done or for which the overall lesion response is 'Unknown'. 'Missing adequate assessment' will also be referred to as 'missing assessment'.

As detailed in Section 16.1.3.2.10 of the study protocol, PFS censoring and event date options depend on the presence and the number of missing tumor assessments.

An exact rule to determine whether there are no, one or two missing TAs is therefore needed. This rule is based on the interval between the last adequate tumor assessment (LATA) date and the event date. The scheduled date of tumor assessments (in weeks from randomization), protocol specified window for tumor assessments, and the thresholds for LATA that belong to a visit can be found in the following table.

Table 5-1 Schedule for tumor assessment and time windows

Assessment schedule		Scheduled date – 1 week	Scheduled date (weeks from randomization)	Scheduled date +1 week	Threshold (weeks)*
Every 8 weeks for the first 18 months	Baseline	0	0^	1	n/a
	C3D1	7	8	9	12
	C5D1	15	16	17	20
	C7D1	23	24	25	28
	C9D1	31	32	33	36
	C11D1	39	40	41	44
	C13D1	47	48	49	52
	C15D1	55	56	57	60
	C17D1	63	64	65	68
	C19D1	71	72	73	78
Every 12 weeks after 18 months	C22D1	83	84	85	90
	C25D1	95	96	97	102
	C28D1	107	108	109	114
	C31D1	119	120	121	126

* The mid-point between current and next visit (except for baseline) and the upper limit for LATA to be matched to a certain scheduled assessment, e.g. if LATA is at week 13, this is after threshold for C3D1 and before that for C5D1, so the matching scheduled assessment is C5D1.
^ Day of randomization is taken as 0.

To calculate the number of missing tumor assessments, the LATA before an event is matched with a scheduled tumor assessment using the time window in Table 5-1 (essentially whichever scheduled assessment it is closest to). Two thresholds, D1 and D2 are calculated for that scheduled assessment based on the protocol-specified schedule and windows

- An event after LATA+D1 will be considered as having >=1 missing assessment
- An event after LATA+D2 will be considered as having >=2 missing assessments

Since there is a change of schedule for tumor assessments at 18 months, D1 and D2 are defined differently depending on when LATA happens.

Rule 1: if LATA happens within 60 weeks from randomization (the matched scheduled tumor assessment is C15D1 or before)

- D1=8+2=10 weeks
- D2=2*8+2=18 weeks

Rule 2: if LATA happens after 60 weeks but within 68 weeks from randomization (the matched scheduled tumor assessment is C17D1)

- D1=8+2=10 weeks
- D2=8+12+2=22 weeks

Rule 3: if LATA happens after 68 weeks from randomization (the matched scheduled tumor assessment is C19D1 or later)

- D1=12+2=14 weeks
- D2=2*12+2=26 weeks.

Therefore, using the D2 definition above, the censoring of an event occurring after ≥2 missing TAs (in the PFS supplementary analysis in Parts A and B2) can be refined as follows: if the distance between the last adequate TA date and the PFS event date is larger than D2, then the participant will be censored and the censoring reason will be 'Event documented after two or more missing tumor assessments.

The same definition of D2 will be used to determine the PFS censoring reason. If the distance between the last adequate tumor assessment date and the earliest of the following dates (analysis cut off, consent withdrawal etc.) is less than or equal to D2:

- Analysis cut-off date
- Date of consent withdrawal
- Date of loss to follow-up

then the censoring reason will be 1. 'Ongoing without event', 2. 'Withdrew consent' or 3. 'Lost to follow-up', respectively. However, if this distance is larger than D2 with no event observed, then the censoring reason will be 'Adequate assessment no longer available'.

5.5.6 No baseline tumor assessments

For the PFS analysis, as specified in Table 16-5 in Section 16.1.3.2.10 of the study protocol, since the timing of disease progression cannot be determined for participants with missing baseline tumor assessment, these participants are censored in the PFS analysis at the date of randomization. This rule however only applies to the disease progression component of the PFS assessment, and not to the survival component. Participants without baseline tumor assessments who die within D2 distance (see Section 5.5.5 for definition) of randomization will be counted as having an event in the derivation of PFS at the date of death (Note: all deaths will be counted in the overall survival analysis regardless of presence or absence of the baseline tumor assessment).

6 References

- Amit O, Mannino F, Stone AM, Bushnell W, Denne J, Helterbrand J and Burger HU. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. European Journal of Cancer 2011; 47: 1772-1778.

- Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.
- Committee for Medicinal Products for Human Use (CHMP). Guideline on adjustment for baseline covariates in clinical trials, EMA/CHMP/295050/2013.
- Cleeland CS (2009) The Brief Pain Inventory. User Guide. Houston, Texas
- Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrical, 26, 404-413.
- Collett D (1994). Modelling survival data in medical research. London, Chapman & Hall.
- Dodd LE, Korn EL, Freidlin B, Gray R, Bhattacharya S. An audit strategy for progression-free survival. Biometrics, 2011, 67 (3): 1092-99.
- Eisenhauer E, et al (2009) New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). European Journal of Cancer, Vol.45: 228-47.
- Fayers PM, Aaronson NK, Bjordal K, et al, on behalf of the EORTC Quality of Life Group (2001) The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.
- FDA briefing Document Oncologic Drugs Advisory Committee Meeting, July 24, 2012, Evaluation of radiologic review of progression-free survival in Non-hematologic malignancies.
- Fleming TR, Harrington DP and O'Brien PC (1984). Designs for group sequential tests. Controlled Clinical Trials, 5, 348-361.
- Glimm E, Maurer W and Bretz F (2010). Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine, 29, 219-228.
- Haybittle JL (1971) Repeated assessment of results in clinical trials of cancer treatment. Br J Radiol; 44(526):793-7.
- Kim SB, Dent R, Im SA, et al (2017) Ipatasertib plus paclitaxel versus placebo plus paclitaxel as first-line therapy for metastatic triple-negative breast cancer (LOTUS): a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Oncol.; 18(10):1360-72.
- Lan KKG, DeMets DL (1983). Discrete sequential boundaries for clinical trials. Biometrika, 70, 659-63.
- Neuenschwander B, Branson M and Gsponer T (2008). Critical aspects of the Bayesian approach to Phase I cancer trials. Statistics in Medicine, 27, 2420-2439.
- O'Brien PC and Fleming TR (1979). A multiple testing procedure for clinical trials. Biometrics, 35, 549-53.
- Oemar M, Janssen B. (2013) EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. Version 2.0, October 2013. © EuroQol Group 2013
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J and Smith PG (1976). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. British Journal of Cancer, 34, 585-607.

- Schmid P, Abraham J, Turner N (2018) AZD5363 (Capivasertip) plus paclitaxel versus placebo plus paclitaxel as first-line therapy for metastatic triple-negative breast cancer (PAKT): a randomised, double-blind, placebo-controlled, phase II trial; 1007.
- Schmid P, Adams S, Rugo HS (2019) IMpassion130: updated overall survival (OS) from a global, randomized, double-blind, placebo-controlled, Phase III study of atezolizumab (atezo) + nab-paclitaxel (nP) in previously untreated locally advanced or metastatic triple-negative breast cancer (mTNBC); Hall D1.
- Schmid P, Adams S, Rugo HS, et al (2018) Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. N. Engl. J. Med.; 379(22):2108-21.
- Spiegelhalter DJ, Abrams KR and Myles JP (2004). Bayesian Approaches to Clinical Trials and Health-care Evaluation. Chichester, Wiley.