CONFIDENTIAL INFORMATION

Statistical Analysis Plan (SAP)

Version 1
7th September 2021

EudraCT Number	2018-003367-58
Protocol Number	SOLTI-1716
Protocol Version Date	10-May-2021. Version number 2.0.
Title	TATEN trial Targeting non-Luminal disease by PAM50 with pembrolizumab + paclitaxel in Hormone Receptor-positive/HER2-negative advanced/metastatic breast cancer, who have progressed on or after CDK 4/6 inhibitor treatment.
Sponsor	PPD

PPD

STATISTICAL ANALYSIS PLAN (SAP)

Signature Page

EudraCT Number: **2018-003367-58**
Protocol Number: **SOLTI-1716**

Written by:

Signature: PPD

Date: 2021/09/09

PPD

SAIL S.L.

Barcelona, Spain

Sponsor Signature:

The undersigned hereby declare that they have examined the Statistical Analysis Plan document and agree to its form and content.

Represented by:

Signature: PPD

Date: 2021/09/07

Name: PPD

Title: PPD

Date:

Name: PPD

Title: PPD

SAP Revision History:

Version Number	Date	Changes
1.0		New

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse event
AESI	Adverse Event of Special Interest
ATC	Anatomical Therapeutic Chemical
CBR	Clinical Benefit Rate
CI	Confidence Interval
CR	Complete Response
CTCAE	Common Terminology Criteria for Adverse Events
ER	Estrogen Receptor
HER2	Human Epidermal Growth Factor Receptor 2
ICH	International Conference on Harmonization
IHC	Immunohistochemistry
IQR	Interquartile Range
ISH	In Situ Hybridization
ITT	Intention-To-Treat
IV	Intravenous
BC	Breast Cancer
MEDDRA	Medical Dictionary for Regulatory Activities
MG	Milligrams
NYHA	New York Heart Association
ORR	Objective Response Rate
PD	Progressive Disease
PGR	Progesterone Receptor
PP	Protocol Population
PR	Partial Response
PT	Preferred Term
RDI	Relative Dose Intensity
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Stable Disease
SD	Standard Deviation
SI	International System of Units
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TLF	Tables, Listings, Figures

TABLE OF CONTENTS

1 INTRODUCTION

1.1 General

The purpose of this statistical analysis plan (SAP) is to provide a protocol specific description of the statistical analysis that will be performed to produce an integrated clinical/statistical report.

This SAP is based upon the following study documents:

- Protocol version: 10-May-2021, Version 2.0
- electronic Case Report Form (eCRF) release version: 01-July-2020

1.2 Type of Study

This is an open-label, single arm, multicenter phase II study.

Pembrolizumab in combination with paclitaxel. Route of administration I.V.

1.3 Study Population

Pre- and post-menopausal women with locally advanced or metastatic non-luminal HR+/HER2-negative endocrine resistant breast cancer who had recurrence or progression while receiving previous therapy with a CDK inhibitor in the adjuvant setting or to treat advanced disease.

1.4 Study Design

The study will utilize the optimal Simon's two-stage design with one interim and a final analysis. The interim analysis will be conducted when 15 patients are evaluable for ORR as determined locally by the investigator using RECIST v.1.1. If 5 or fewer responses are observed in up to 15 patients of evaluable population (EP), the trial will be terminated in favor of the null for futility. Otherwise, up to a further 31 patients may be evaluated, for a maximum total of 46 evaluable patients. If a total of 19 or more responses are seen at the end of the second stage, then the null will have been rejected in favor of the alternative; and further investigation of the combination is warranted.

Recruitment will not be halted during the interim analysis period. Therefore, no interruption in the accrual will be done during the interim analysis to maintain the dynamic of accrual in the trial.

After confirmation of all eligibility criteria, eligible patients will receive pembrolizumab 200 mg every 3 weeks (on D1 of each 21-day cycle, beginning in Cycle 1) in combination with paclitaxel 80 mg/m2 administered at days 1, 8, 15 of each 21-day cycle beginning at cycle 2. Treatment will continue until disease progression, the development of unacceptable toxicity, withdrawal of consent, 24 months from the date of the first dose of pembrolizumab or end of study, whichever occurs first.

All patients will be followed for survival from screening until the last patient recruited has been followed for 12 months, has progressed, or has died, whichever occurs first. The patient will be followed for survival approximately every 3 months (± 21 days) until death, withdrawal of consent, loss to follow-up, or study termination by SOLTI. In addition, information regarding use of subsequent anti-cancer agents for metastatic HR+/HER2- during the survival follow-up period will be collected.

1.5 Study Schedule

Trial Period:	Screening Phase		Treatment Cycles[b]									End of Treatment	Post-Treatment		
Treatment Cycle/Title:	Pre-screening (Visit 1)	Main Study Screening[a] (Visit 2)	1			2			All Cycles			Discon	Safety Follow-up[c]	Follow-Up	Survival Follow-Up
Scheduling Window (Days):		-28 to -1	1+ 3	8± 1	15± 1	1+3	8± 1	15± 1	1+ 3	8± 1	15± 1	At time of Discon	30 days post discon	Every 9 weeks post discon	Every 12 weeks[t] (±21 days)
Administrative Procedures															
Pre-screening Consent	X														
Informed Consent		X													
Inclusion/Exclusion Criteria	X	X													
Demographics and Medical History[d]		X													
Prior and Concomitant Medication Review[e]		X	X			X			X			X	X		
Pembrolizumab Administration			X			X			X						
Paclitaxel[f]						X	X	X	X	X	X				
Post-study anticancer therapy status														X	X
Survival Status															X
Clinical Procedures/Assessments															
Review Adverse Events			X			X			X			X	X	X	
Full Physical Examination[g]		X													
Directed Physical Examination[h]			X			X			X			X			
Height		X													

Trial Period:	Screening Phase		Treatment Cycles[b]									End of Treatment	Post-Treatment		
Treatment Cycle/Title:	Pre-screening (Visit 1)	Main Study Screening[a] (Visit 2)	1			2			All Cycles			Discon	Safety Follow-up[c]	Follow-Up	Survival Follow-Up
Scheduling Window (Days):		-28 to -1	1+ 3	8± 1	15± 1	1+3	8± 1	15± 1	1+ 3	8± 1	15± 1	At time of Discon	30 days post discon	Every 9 weeks post discon	Every 12 weeks[t] (±21 days)
Vital Signs and Weight[i]		X	X			X			X						
ECOG Performance Status[j]		X	X			X			X			X	X	X	
12- Lead Electrocardiogram (Locally performed)		X													
Laboratory Procedures/Assessments: analysis performed by LOCAL laboratory															
Pregnancy Test – Urine or Serum β-HCG[k]		X	X			X			X						
PT/INR and aPTT		X													
CBC with Differential[l]		X	X	X	X	X	X	X	X	X	X		X		
Comprehensive Serum Chemistry Panel[m]		X	X			X			X				X		
Urinalysis[n]		X	X			X			X				X		
T3 (free or total), Free T4 and TSH			Every 2 cycles										X		
Efficacy Measurements															
Tumor Imaging[o]			Every 9 weeks (63 ±5 days)									X[p]		X[q]	
Tumor Biopsies/Archival Tissue Collection/Correlative Studies Blood															
Archival or Newly Obtained Tissue Collection[r]	X														
Blood Collection[s]			X			X						X			
Confirmation of *PAM50* eligibility & inclusion criteria															
Medical monitor eligibity confirmation		X													

a) Written informed consent is required before performing any study-specific tests or procedures. Signing of the Informed Consent Form cannot occur outside the 28-day screening period. All screening evaluations must be completed and reviewed to confirm that patients meet all eligibility criteria before randomization. Results of standard-of-care tests or examinations performed prior to obtaining informed consent and within 28 days prior to study entry (except where otherwise specified) may be used for screening assessments rather than repeating such tests. The investigator will maintain a screening log to record details of all patients screened and to confirm eligibility or record reasons for screening failure, as applicable.

b) Assessment window of + 3 days for Cycles ≥2, Day 1. If scheduled dosing is precluded because of a holiday, then dosing may be postponed to the soonest following date, with subsequent dosing continuing the specified schedule. If treatment was postponed for fewer than 2 days, the patient can resume the original schedule. If scheduled study assessments cannot be obtained because of a holiday, these assessments should then be obtained at the soonest following date, provided that the soonest following date is not within 2 days of other regularly scheduled study assessments. Assessments scheduled on the day of study treatment administration (Day 1) of each cycle should be performed prior to study treatment infusion unless otherwise noted.

c) Patients will be asked to return to the clinic not more than 30 days after the decision to discontinue treatment for a treatment discontinuation visit. The visit at which the decision is made to discontinue treatment (e.g., disease progression is determined or confirmed) may be used as the treatment discontinuation visit. If a participant initiates a new anti-cancer therapy within 30 days after the last dose of trial treatment, the 30-day Safety Follow-up visit must occur before the first dose of the new therapy.

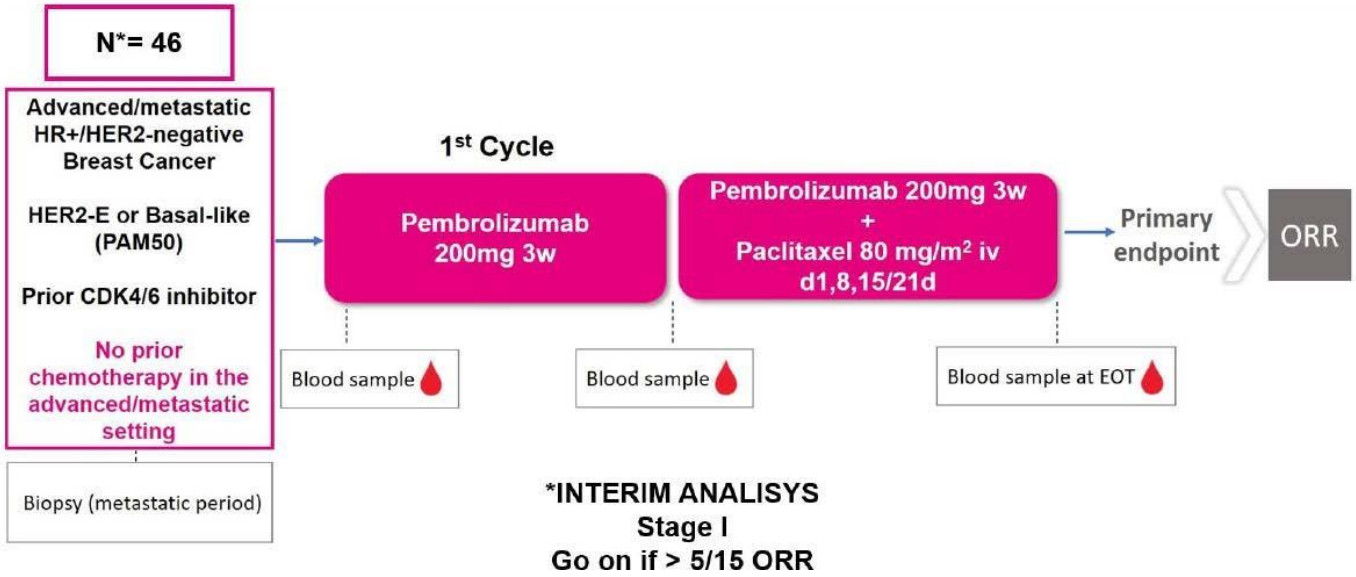
d) Cancer history includes stage, date of diagnosis, and prior anti-tumor treatment. Demographic information includes age and self-reported race/ethnicity. Reproductive status and smoking history should also be captured.

e) Concomitant medications include any prescription medications or over-the-counter medications. At screening, any medications the patient has used within the

28 days prior to Cycle 1, Day 1. At subsequent visits, changes to current medications or medications used since the last documentation of medications will be recorded. Concomitant medications administered after 30 days after the last dose of trial treatment should be recorded.
f) Weekly paclitaxel beginning in Cycle 2.
g) A complete physical examination should include an evaluation of the head, eyes, ears, nose, and throat, and the cardiovascular, dermatological, musculoskeletal, respiratory, gastrointestinal, genitourinary, and neurological systems as well as weight (in kilograms) and height (in centimeters; height is measured at the screening visit only). Symptom-directed physical exam after baseline assessment.
h) Directed physical examination should include symptom-directed examination or as clinically indicated. Limited physical examination will be performed monthly every treatment cycle, with additional examinations as clinically indicated at baseline.
i) Vital signs include heart rate, blood pressure, and temperature.
j) Have an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1, as assessed within 10 days prior to the start of study treatment
k) Serum/urine pregnancy test within 72 hours before Cycle 1, Day 1. Afterward, perform every cycle. A positive urine test must be confirmed with a serum test
l) Hematology consists of CBC, including RBC count, hemoglobin, hematocrit, WBC count with differential (neutrophils, eosinophils, lymphocytes, monocytes, basophils, and other cells), and platelet count. A manual differential can be done if clinically indicated. Local laboratory assessments from each cycle must be reviewed prior to study treatment administration for each cycle. Laboratory tests for screening should be performed within 10 days prior to the first dose of treatment. After Cycle 1, pre-dose laboratory procedures can be conducted up to 72 hours prior to dosing.
m) Serum chemistry includes creatinine, sodium, potassium, magnesium, calcium, glucose, total bilirubin, ALT, AST, alkaline phosphatase, Lactate dehydrogenase (LDH). Local laboratory assessments from each cycle must be reviewed prior to study treatment administration for each cycle. Screening results may be valid for Cycle 1, Day 1 if performed within 10 days prior to Week 1, Day 1.
n) Urinalysis (specific gravity, pH, glucose, protein, ketones, and blood). Screening results may be valid for Week 1, Day 1 if performed within 10 days prior to Week 1, Day 1.
o) Tumor assessments performed as standard of care prior to obtaining informed consent and within 28 days prior to Cycle 1, Day 1 may be used rather than repeating tests. All measurable and evaluable lesions should be assessed and documented at the screening visit. Radiologic imaging performed during the screening period should consist of 1) CT and/or MRI of the chest/abdomen/pelvis 2) Bone scan or PET scan should be performed to evaluate for bone metastases and 3) any other imaging studies (CT neck, plain films, bone scan etc.) as clinically indicated by the treating physician. The same radiographic procedures and technique must be used throughout the study for each patient (e.g., if the patient had CT chest/abdomen/pelvis performed during screening, then she should subsequently undergo CT performed using the same radiologic protocol throughout the remainder of the study). Results must be reviewed by the investigator before dosing at the next cycle. Tumor assessments will be performed at baseline, every 9 weeks (63 days ± 5 days). All known sites of disease documented at screening should be re-assessed at each subsequent tumor evaluation. Tumor assessments performed after the screening period should consist of 1) CT and/or MRI of the chest/abdomen/pelvis, 2) bone scan or PET scan if there were osseous sites of disease identified or if these studies are felt to be clinically indicated by the treating physician, and 3) any other imaging studies felt to be clinically indicated by the treating physician. Tumor response will be evaluated using both RECIST v1.1 and immune-modified RECIST criteria. In the absence of disease progression, tumor assessments should continue regardless of whether patients discontinue study treatment, unless they withdraw consent or the study is terminated by SOLTI, whichever occurs first.
p) If previous imaging was obtained within 4 weeks prior to the date of discontinuation, then imaging at treatment discontinuation is not mandatory. In participants who discontinue study treatment due to documented disease progression and the Investigator elects not to implement iRECIST, this is the final required tumor imaging.
q) In subject who discontinue trial treatment without documented disease progression, every effort should be made to continue monitoring their disease status by radiologic imaging every 9 weeks (+ 5 days) until (1) the star of new anti-cancer treatment, (2) disease progression, (3) death, or (4) the end of the study, whichever occurs the first.
r) Tumor tissue should be of good quality based on total and viable tumor. An FFPE block should be provided. Fine-needle aspiration, brushing, cell pellets from pleural effusion, and lavage samples are not acceptable. Retrieval of archival tumor sample can occur outside the 28-day screening period (centrally perform PAM50 test). Tumor sample from metastatic disease period is mandatory.
s) Plasma from 30 ml of blood will be collected and banked at C1D1 (samples may be extracted within the ten previous days to week 1 day 1), C2D1 and at progression.
t) Survival follow-up information will be collected via telephone calls, patient medical records, and/or clinic visits approximately every 3 months (± 21 days) until death, loss to follow-up, withdrawal of consent or until study termination by SOLTI. All patients will be followed for survival and new anti-cancer therapy (including targeted therapy and immunotherapy) information unless the patient requests to be withdrawn from follow-up; this request must be documented in the source documents and signed by the investigator. If the patient withdraws from study, the study staff may use a public information source (e.g., county records) to obtain information about survival status only.

1.6 Trial Diagram

Figure 1. Trial diagram



1.7 Sample Size

This study will utilize the optimal two-stage Simon's design (1989). The null hypothesis that the true response rate is not more than 30% ($p \leq 0.3$) versus the alternative hypothesis that the true response rate is at least 50% ($p \geq 0.50$) will be tested. The first stage will be conducted with 15 evaluable patients. If there are 5 or fewer responses in these 15 patients, the study will be stopped. Otherwise, 31 additional patients will be accrued for a total of 46. The null hypothesis will be rejected if 19 or more responses are observed in 46 patients of EP. This design yields a one-sided type I error rate of 0.05 (alpha) and power of 80%.

Up to 46 evaluable patients of EP may therefore be enrolled to this clinical trial. It is expected that approximately 403 patients will be screened to identify 55 patients who meet all the inclusion and exclusion criteria and who are evaluable for response. Assuming a dropout rate of 10%, the study will aim to recruit 55 patients in total to have at least 46 evaluable patients of EP to attain 80% power at nominal level of one-sided alpha of 0.05.

2 STUDY OBJECTIVES

2.1 Primary objective

The primary objective of the study is to assess the proportion of patients with best overall response of CR or PR, as per local investigator's assessment and according to RECIST v1.1.

2.2 Secondary Objectives

The secondary objectives in this study are to evaluate the progression free survival (PFS), clinical benefit rate (CBR) and duration of response (DOR), Time to response (TtR), Pre-PFS, overall survival (OS) and safety.

3 DEFINITION OF ENDPOINTS

3.1 Primary Endpoint

- Overall Response rate (ORR) defined as the proportion of patients with best overall response of complete response (CR) or partial response (PR), as per local investigator´s assessment and according to Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 criteria.

3.2 Secondary Endpoints

- Clinical Benefit Rate (CBR) defined as the proportion of patients with a best overall response of CR, PR or an overall lesion response of Stable Disease (SD) or Non-PR/Non-progression disease (PD) lasting ≥ 24 weeks, based on local investigator´s assessment according to RECIST v1.1.
- Progression free survival (PFS) defined as the time from allocation to the first occurrence of disease progression, as determined locally by the investigator according to RECIST v.1.1, or death from any cause, whichever occurs first.
- Duration of response (DoR) defined as the time from the first occurrence of a documented objective response to disease progression, as determined locally by the investigator through use of RECIST v.1.1, or death from any cause, whichever occurs first
- Time to response (TtR) defined as the time from allocation to the first objective tumor response (tumor shrinkage of ≥30%) observed for patients who achieved a CR or PR.
- Overall survival (OS) defined as the time from allocation to death from any cause (OS will be determined at the end of the study).
- PFS on study treatment compared to PFS on prior line of therapy (pre-PFS).
- ORR according to PD1 mRNA expression.

3.3 Safety Endpoints

Analysis of safety-related data will be:

- The degree of exposure (dose, duration, number of patients) will be assessed to determine the degree to which study safety can be assessed.
- Incidence, duration, and severity of Adverse Events (AEs) assessed by the NCI Common Terminology for Classification of Adverse Events (CTCAE) version 5, including dose reductions, delays, and treatment discontinuations.

3.4 Exploratory Endpoints

- According to iRECIST 1.1:
    - ORR
    - PFS
    - DoR
    - TtR
- To determine ORR according to PD-L1 protein expression by IHC.
- To evaluate ORR according to stromal TILs.
- To identify a new gene signature predictive of pembrolizumab and paclitaxel therapy benefit.
- To identify new biomarkers of response to the combination treatment, we aim to further evaluate the expression of 752 genes that encompass important genomic signatures and individual genes. The nCounter® Breast Cancer 360 Panel includes 752 genes that cover established breast cancer diagnostic and research signatures as well as key pathways at the interface of the tumor, tumor microenvironment and immune response. Special attention will be given to innate immune response genes as well as markers of antigen presentation, which are expected to be determinant for this combination treatment. The following genes/signatures will be evaluated among others: PAM50 genes and signatures, risk of recurrence (ROR) score, ER signaling biology, immune cell marker (e.g., BCL2; CD163; CD68; CD84; CD8A; CD8B; CHIT1), chemo-endocrine score (CES), immune infiltration (e.g. CCL5; CD27; CD274; CD276; CD8A; CMKLR1; CXCL9) or proliferation.

4 ANALYSIS POPULATIONS

The following populations will be analyzed:

- **Screening Population**: All patients who were present at the screening visit will be included in the screening population.
- **Intent-to-treat (ITT) population**: All patients that are enrolled in the study. This population will be used for the efficacy analysis, except for overall response rate and clinical benefit rate, which used the evaluable population
- **Evaluable population (EP):** includes all patients who have received at least one combination dose of pembrolizumab and paclitaxel who had measurable disease according to the Investigator site assessment at baseline and who had at least one postbaseline tumor assessment . This population will be used for overall response rate and cinical benefit rate analysis.
- **Safety population (SP):** All patients who received at least one (even incomplete) part of the study treatments. The SP population will be analyzed for the safety endpoints.
- **Exploratory Analysis Set:** Exploratory analyses will be performed on those patients who were evaluable for exploratory endpoints including enough quality of the sample.

5 STATISTICAL METHODS

5.1 General Methodology

The statistical analysis will be conducted following the principles as specified in International Conference on Harmonization (ICH) Topic E9 (CPMP/ICH/363/96). The significance level will be α=0.05 for all tests. As an exploratory study, multiple testing without adjustment of the significance level is considered acceptable.

Definition of baseline: For each safety or efficacy parameter, the last valid assessment made before first study drug administration will be used as the baseline for all analyses of that safety or efficacy parameter unless otherwise specified.

Continuous data will be summarized in terms of the number of observations, mean, standard deviation (SD), median, minimum, maximum, and first and third quartiles, unless otherwise stated. Where data are collected over time, both the observed data and change from baseline will be summarized at each time point.

The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, median, and first and third quartiles will be reported to one more decimal place than the raw data recorded in the database. The SD will be reported to two more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be four for any summary statistic.

Categorical data will be summarized in terms of the number of subjects providing data at the relevant time point (n), frequency counts and percentages. Changes from baseline in categorical data will be summarized using shift tables where appropriate.

The Kaplan-Meier method will be used to estimate the time to event function the median of time to event and the 95% confidence interval of the same will be calculated. These confidence intervals will be calculated based on the Greenwood method. Number and proportion of events, median survival time and survival rates, with corresponding 95%Cl will be calculated.

Percentages will be presented to one decimal place. A percentage of 100% will be reported as 100%. Percentages will not be presented for zero counts. Unless otherwise stated, percentages will be calculated using n as the denominator, for frequency tables not assessed by time point the set will be used as denominator. If sample sizes are small, the data displays will show the percentages, but any textual report (e.g., clinical study report) will describe frequencies only.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. However, if a p-value is only presented to four decimal places (by SAS) it will not be rounded again but will be presented to four decimal places. P-values less than 0.0001 will be presented as "<0.0001".

Confidence intervals will be presented to one more decimal place than the raw data. A two-sided significance level of 5% will be used for confidence intervals.

For binary endpoints, the 95% confidence intervals (CIs) will be constructed based on an exact binary distribution.

All scores and change from baseline will be summarized in terms of the number of observations, mean, standard deviation, 95%CI of mean, median, range and interquartile range. We will examine the residuals to assess model assumptions.

All report outputs will be produced using SAS® version 9.4 version in a secure and validated environment. All report outputs will be provided to the Sponsor in a single Microsoft Word document.

5.2 Subject Disposition

Descriptive statistics will be provided for the following:

- Overall number of subjects in the screening population, the number of patients eligible to participate in the study, and number of screening failures.
- Number and percent of subjects in each of the analysis populations.

- Number and percent of subjects excluded from each of the analysis populations along with reason for exclusion.
- Listing of subjects excluded from each of the analysis populations along with reason for exclusion.
- Listing of protocol deviations.
- Study termination:
    - Number and percent of subjects who completed the study.
    - Frequency of premature termination reasons.
    - Listing of all dropouts along with reason for termination, treatment group and time of termination.

No statistical tests are planned for these data.

5.3 Baseline Characteristics

Baseline characteristics will be provided overall for the Safety, Efficacy and EP populations.

Descriptive statistics, including number of subjects, mean, standard deviation (SD), median and range for continuous variables and frequency and percent for categorical variables will be provided.

Baseline Characteristics:

- Demographics
- Toxic background
- Oncological history
- Prior significant diseases
- Prior concomitant medication
- History of Breast Cancer
- Oncological history
- Previous treatments for breast cancer
- Previous surgeries for breast cancer
- Physical examination
- Tumor clinical evaluation
- ECOG
- TNM classification
- Histologic grade
- Estrogen receptor (ER)
- Progesterone receptor (PgR)
- HR status
- Ki67 value (%)
- HER2 status

No statistical tests are planned for these data.

A by-subject listing of all demographic and other baseline characteristics will be provided.

5.4 Efficacy

All efficacy analysis will be based on the ITT and Evaluable population.

5.4.1 Response Efficacy Definitions

- Best Overall Response, defined as the best overall response recorded from the start of the study treatment.
    - When CR or PR is the best response across all time points, then best overall response will be CR or PR respectively.
    - When SD is the best response for ≥ 24 weeks the best overall response will be SD ≥ 24 w.
    - When SD is the best response for < 24 weeks the best overall response will be SD < 24 w.
    - When non-target disease only and Non-CR/Non-PD is the best response for ≥ 24 weeks the best overall response will be SD ≥ 24 w.
    - When non-target disease only and Non-CR/Non-PD is the best response for < 24 weeks the best overall response will be SD < 24 w.
    - When PD is the best response across all time points, best overall response will be PD.
    - When there is no evaluable tumor assessments best overall response will be NE.
- Objective Response Rate (ORR) is defined as the proportion of patients with best overall response of CR or PR.
- Clinical Benefit Rate (CBR) is defined as the proportion of patients with best overall response of CR or PR or SD ≥ 24 w.
- PFS is defined as the time from start dose until death by any cause or objective tumor progression or clinical disease progression. Patients with no progression or death will be censored at the date of their last evaluable imaging. Censoring rules are specified below:

Situation	**Date of progression or censoring**	**Outcome**
Progression documented between scheduled visits	Earliest of: • Date of assessment by investigator (if progression is based on clinical criteria); or • Date of assessment showing new lesion (if progression is based on new lesion); or • Date of last radiological assessment of measured lesions (if progression is based on increase in sum of measured lesions).	Progressed
Death before first progression disease (PD) assessment	Date of death.	Progressed
Death between adequate assessment visits	Date of death.	Progressed
No progression	Date of last radiological assessment of measured lesions.	Censored
Treatment discontinuation for undocumented progression	Date of last radiological assessment of measured lesions.	Censored
Treatment discontinuation for toxicity or other reason	Date of last radiological assessment of measured lesions.	Censored
Death or progression after more than one missed visit	Date of last radiological assessment of measured lesions.	Censored

- The Overall Survival (OS) is defined as the time from start dose until death from any cause. Patients with no death will be censored on the last available follow-up date.

- The duration of response (DoR) is defined as the time from start dose of first tumor response (either CR or PR) to disease progression or death due to any cause. The DoR will be calculated for the patients with CR or PR.
- The duration of Clinical Benefit (DoCB) is defined as the time from documentation of first clinical benefit (either CR or PR or SD ≥ 24 w) to disease progression or death due to any cause. The DoCB will be only calculated for the participants with clinical benefit.
- The Time to Progression (TTP) is defined as the time from start dose to objective tumor progression or clinical disease progression (TTP does not include deaths). Patients with no progression will be censored at the date of their last evaluable imaging.
- The Time to Response (TTR) is defined as the time from start dose to ORR date. Patients without ORR will be censored at the date of their last evaluable imaging.
- Maximum Tumor Shrinkage (best percentage of change from baseline in the size of target tumor lesions) is defined as the biggest percentage of tumor shrinkage from baseline (obtained from the sum of the largest diameters of the target lesions).
- PFS on study treatment compared to PFS on prior line of therapy (pre-PFS). Pre-PFS is defined as the time from the date of first dose of the previous medication to the date of the first documented progression in the previously metastatic line. Pre-PFS only applies to patients whose received have previous treatment for metastatic disease
- ORR according to PD1 mRNA expression. The PD1 mRNA will be pre-established cutoff with median or tertiles. Median: Cutoff points will be calculated according to the median value for the mRNA expression. Samples with mRNA expression above or equal to the median were considered as samples with high expression, while those with value below the median as samples with low expression. Tertiles: Cutoff points will be calculated according to the tertiles value for the mRNA expression. Samples with mRNA expression above or equal to the tertile-1 (PD1-high) were considered as samples with high expression, samples with mRNA expression above or equal to the tertile-2 (PD1-high) were considered as samples with intermediate expression, while those with value below the tertile-2 as samples with low expression.

5.4.2 Primary Efficacy Analysis

The primary analysis is the ORR as per local investigator's assessment and according RECIST v1.1 in the Evaluable Population.

- ORR, as per local investigator's assessment and according RECIST v1.1, is defined as the proportion of patients with best overall response of CR or PR .

To be consistent with the sample size calculation, the primary hypothesis will be tested at 0.05 level of significance (one-sided) and 90% confidence intervals will be reported. Additionally, the ORR will be reported along with associated 95% confidence interval.

The ORR with its 95% Wilson Score confidence interval, will be calculated at stage II. One-sided p-value will be obtained with exact binomial test at a nominal alpha level of 0.05. The stage II will be declared positive if the number of responding patients is ≥19. In this case the null hypothesis (H0: ORR<=30%) will be rejected.

5.4.3 Secondary Efficacy Analysis

For binary endpoints, the number and proportion of patients with response with the 95% Pearson-Clopper CI will be calculated.

The Kaplan-Meier estimate of the PFS survival function will be estimated and displayed. The resulting median PFS time will be given for each cohort with 95% confidence intervals, as well as 25th and 75th percentiles will be reported.

The OS curve will be estimated by the Kaplan-Meier methodology, and the 95% CI will be estimated by the Cox proportional-hazards models.

Maximum tumor shrinkage will be described with the median, range, mean, standard deviation, and interquartile range. Waterfall plots describing the percentage of change in target tumor lesions will be provided.

ORR according to PD1 mRNA expression analysis will be performed by SOLTI. To compare distribution of variables between two groups, we used Fisher's exact test. Proportions and 95% CI were also provided. Univariate and multivariable logistic regression analyses are done to investigate the association of each variable with ORR. Odds ratios (ORs) and 95% CIs are calculated for each variable. All the laboratory analyses are performed blinding to the clinical data.

For analyses of PFS, Kaplan–Meier curves will be plotted, and groups compared using the log-rank test. Hazard ratios and associated 95% CIs will be obtained from Cox proportional hazards regression models. In order to determine the optimal cut-point for CDR21 in relation to PFS, separate models were fitted with data cut at each observed value of CDR21 with the cut-point giving the highest Harrell's c-index considered optimal.

For all tests, we will use two-sided p-values with alpha ≤ 0.05 level of significance. P-values emerging from these analyses will not be interpreted in a confirmative sense; they will be considered of descriptive nature only.

5.4.4 Handling of Missing Data

Study variables could be missing for patients who withdrawn from the trial or for specific visits. We will report reasons for withdrawal.

Patient with missing in response outcomes (ORR and CBR) will considered as no responders. Patients without any post-baseline assessment will be considered as non-responders or without clinical benefit.

For the analysis of maximum tumor shrinkage only observed cases will be used.

The analysis of timed endpoints is based on a Kaplan-Meier method (PFS, DoR and OS), therefore, not affected by patient withdrawals (as they are censored) provided that dropping out is unrelated to prognosis.

For PFS, patients without a date of disease progression will be analyzed as censored observations on the date of last tumor assessment. If no post-baseline tumor assessment is available, patients will be censored at the date of treatment initiation + 1 day. Data for patients with an event who missed two or more scheduled assessments immediately prior to the event will be censored at the last tumor assessment prior to the missed visits.

For OS, patients who are not reported as having died will be analyzed as censored observations on the date they were last known to be alive. If no post-baseline data are available, OS will be censored at the date of treatment initiation + 1 day.

The other variables will be managed with simple imputations methods (last observation carried forward). The effect that any missing data might have on results will be assessed via sensitivity analysis.

5.4.5 Exploratory Analysis

These statistical analyses will be exploratory. Exploratory analyses will be performed on the exploratory analysis set.

iRECIST:

- ORR iRECIST is defined as the proportion of patients with best overall response of iRECIST Complete Response (iCR) or iRECIST Partial Response (iPR)
- PFS iRECIST is defined as the time from start dose until death by any cause or iRECIST Progressive Disease (iCPD) or clinical disease progression.
- DoR iRECIST is defined as the time from start dose of first tumor response (either iCR or iPR) date to iCPD or death due to any cause. The DoR will be calculated for the patients with iCR or iPR.
- TTR iRECIST is defined as the time from start dose to first tumor response (either iCR or iPR) date. Patients without response will be censored at the date of their last evaluable.

The MTS will be described (N, mean, median, standard deviation, minimum, and maximum) and analyzed based on Wilcoxon test. The use of multiple regression adjust by baseline factors will be deliberated.

The statistical analysis plan did not include a provision for correcting for multiplicity in tests for secondary and exploratory analyses. For all tests, we will use two-sided P-values with alpha ≤ 0.05 level of significance. The P-values emerging from these analyses will not be interpreted in a confirmative sense but will be considered of descriptive nature only. Changes in biomarkers will be evaluated on a univariate level regarding their association with treatment. Baseline and post-treatment values of quantitative biomarkers will be described with median, standard deviations, median and range in quantitative measures (copy numbers, immunoscore…). Baseline and post-treatment values of categorical biomarkers (mutations, molecular profiling, and breast cancer subtypes…) will be described with frequency and percentage.

Change from baseline between values of quantitative biomarkers will be analyzed with mean and median differences. The corresponding 95% confidence intervals, applicable test statistics and p-values will be presented. P-values and 95%CI for mean differences will be based on paired T-test. P-values for ranks will be based on Wilcoxon paired test. The 95%CI for median difference will be based on bootstrap percentile method. Change from baseline between percentage of categorical biomarkers will be analyzed with McNemar's test.

Markers will be evaluated on a univariate level regarding their change over potential for prediction of the clinical endpoints (ORR, PFS and OS). Biomarker and response correlations with clinical covariates will be investigated. It will be checked whether covariates can improve the prediction and whether there is an interaction with the biomarkers. Further multivariate techniques (e.g., Multiple Logistic Regression, Cox regression and penalized regressions models) will be evaluated to study combinations of markers. Techniques to control false discovery rate and overfitting (cross-validation) will be also considered. Analysis will be performed on exploratory analysis set.

ORR according to PD1 mRNA expression analysis will be performed by SOLTI. Analysis involving gene expression data will be performed by SOLTI.

5.5 Safety

All safety analyses will be performed on the Safety Population. These data will include report of AEs, laboratory analysis, vital signs, ECOG PS, changes in ECG, time of trial dosing, and proportion of interruptions. Terms and severity of AE will be obtained and reported according to the NCI CTCAE v.5. Statistical summaries of characteristics, frequency, grading, duration, and relationship with treatment will be presented.

All safety tables will list or summarize subjects on the safety population. These safety assessments will be subjected to clinical review and summarized by appropriate descriptive statistics.

5.5.1 Duration and Extent of Exposure of Study Treatment

The study treatment period is defined as the time between the study entry and the last dose of study combination therapy.

Duration and extent of exposure will be based on the safety set.

The following parameters will be calculated:

- b: "Actual Cycle Duration" is the difference between start and stop date dose.
- c: "Actual Cycle Dose Days" is the number of days with dose administration in the cycle, considering the interruptions.
- d: "Actual Total Dose" is the total dose a patient took in a cycle, considering interruptions and reductions.
- e: "Intended Daily Dose per Cycle".
- f: "Intended Cycle Duration".
- g: "Intended Cycle Dose Days".
- A: "Total number of cycles" = 1
- B: "Treatment Duration" = Sum over all cycles of (b).

- C: "Days on drug" = Sum over all cycles of (c).
- D: "Total Actual Total Dose" = Sum over all cycles of (d).
- E: "Mean Intended Daily Dose" = Mean over all cycles of (e).
- F: "Total Intended Duration" = Sum over all cycles of (f).
- G: "Total Intended Dose Days" = Sum over all cycles of (g).
- H: "Intended Total Dose" = G*E
- I: "Actual Average Daily Dose on Dose Days" = D/C
- J: "Ratio for Dose Interruption" = C/G
- K: "Ratio for Cycle Duration" = F/B
- L: "Actual Average Daily Dose Intensity" = I*J*K
- M: "Relative Dose Intensity (RDI)" = L/E*100

The treatment duration (days), days on Drug and Treatment compliance (%) will be summarized in terms of the number of observations, mean, standard deviation (STD), median, minimum and maximum.

Extent of exposure measured as RDI (dose level) will be described with median, interquartile range (IQR) and range. The RDI will be dichotomized in different cutoffs (≥50%, ≥70%, ≥80%, ≥90%, ≥100%) and described with frequencies and percentages.

5.5.2 Concomitant Medications

The number and percent of unique patients taking concomitant medications will be summarized by therapeutic classification, coded term, and dose level. Elective surgeries/procedures performed during the study will be presented in a listing.

The following are conventions that will be used to classify individual medications as prior and/or concomitant:

- Medications with stop dates prior to screening visit date will be considered prior.
- Medications with missing stop dates or stop dates the day of or after screening visit date will be considered concomitant, regardless of start date. Additionally, if the start date is prior to screening visit date or missing, the medication will also be considered prior.

Frequencies and by-subject listing of all prior and concomitant medications will be provided, containing variables listed on Prior/Concomitant Assessment eCRF, their corresponding categories (Prior or Concomitant), and WHO Anatomical Therapeutic Chemical (ATC) level 2 and ATC Name.

5.5.3 Adverse Events

All AEs will be recorded on the eCRF "Adverse Events" page and will be coded using the current version of MedDRA® to give a system organ class (SOC) and preferred term (PT) for each event. All adverse event safety data will be updated to the version of MedDRA that is current at the time of the database lock and statistical analyses. Adverse events will be coded with grades defined according to CTCAE V5.0 criteria.

Treatment-emergent AEs (i.e., those events occur after the first study medication administration and were not present at baseline or worsened in severity following the start of treatment) will be tabulated. The TEAE will be tabulated according to intensity and causality. If intensity of an AE or causality of an AE to the study medication is missing, a worst-case scenario will prevail (severe in intensity or probably related will be assumed). In the summary tables the number of subjects with events and the number of events will be presented.

The onset date of an AE will be compared to the date of first dose of study drug to determine whether the AE is treatment emergent. Adverse events with an onset date on or after the date of first dose of study drug will be classified as treatment emergent.

All deaths and SAEs, regardless of cause, from treatment start until 90 days after final dose of treatment or 30 days following cessation of study treatment if the participant initiates new anticancer therapy. Non-fatal AEs occurring after treatment start regardless of cause, up until 30 days after final dose of treatment or until start of new anti-cancer treatment, whichever is first. Disease progression is not considered a treatment emergent adverse event unless the patient dies of disease prior to 28 days after discontinuation of treatment. Events that are continuations of baseline abnormalities are considered treatment emergent adverse events only if there is an increase in grade over baseline, or if there is an increase following a decrease during the study.

Treatment emergent adverse events with cause possibly, probably, or definitely related to treatment as judged by the investigator. Events that are continuation of baseline abnormalities are not considered treatment related unless there is an increase in grade, or if there is an increase following a decrease, and the increase is judged by the investigator to be due to treatment.

Descriptive statistics will be used to characterize the profiles of treatment-related AEs, treatment-related deaths, SAEs, treatment-related delays, dose reductions, and/or treatment discontinuations. All AEs will be graded according to the current version of the NCI-CTCAE v.5 and the Medical Dictionary for Regulatory Activities (MedDRA).

The following summaries will be provided:

- Summary of Adverse Events
    - Subjects with at least one Adverse Event (AE)
    - Subjects with at least one Treatment Emergent AE (TEAE)
    - Subjects with at least one related study drug TEAE
    - Subjects with at least one grade 3 or 4 or 5 TEAE
    - Subjects with at least one related grade 3 or 4 or 5 TEAE
    - Subjects with at least one serious TEAE
    - Subjects with at least one related serious TEAE
    - Subjects with at least one non-serious TEAE
    - Subjects with at least one Adverse Event of Special Interest (AESI)
    - Deaths due to TEAE
    - Subjects with TEAE leading to Discontinuation of Study Treatment
    - Subjects dropped out due to AE
- Summary by SOC and PT of the number and percentage of subjects reporting each:
    - Treatment Emergent Adverse Events
    - Treatment Emergent Adverse Events by Treatment and Haematologic
    - Treatment Emergent Adverse Event Related to Study Drug
    - Related Treatment Emergent Adverse Event by Treatment and Haematologic
    - AESI Treatment Emergent Adverse Events
    - Treatment Emergent Adverse Events with Grade 3 or 4 or 5
    - Related Treatment-Emergent Adverse Event with Grade 3 or 4 or 5
    - Serious Treatment Emergent Adverse Event
    - Serious and Related Treatment Emergent Adverse Event
    - Treatment Emergent Adverse Event by Maximum Severity
    - Treatment Emergent Adverse Event Leading to Discontinuation of Study Drug

5.5.4 Clinical Laboratory Parameters

The following summaries will be produced for all hematology and biochemistry laboratory parameters:

- Shift tables of low, normal, high distribution (n; %) with respect normal ranges of center, at each post-baseline cycle by baseline (cycle 1 day 1) distribution.
- Shift tables of low, normal, high distribution (n; %) of clinically significant at each post-baseline cycle comparing with baseline (cycle 1 day 1) distribution.
- The value distribution will be displayed using a serial box plot at each cycle, for each treatment arm.

5.5.5 Vital Signs

The following summaries will be produced for all vital sign's parameters:

- Shift tables of low, normal, high distribution (n; %) at each post-baseline cycle by baseline (cycle 1 day 1) distribution.

5.5.6 Physical Examination

The following summaries will be produced for all physical examination's parameters:

- Shift tables of low, normal, high distribution (n; %) at each post-baseline cycle by baseline (cycle 1 day 1) distribution.

5.6 Interim Analysis

The Stage I interim analysis will be conducted when 15 patients are evaluable for ORR as determined locally by the investigator using RECIST v.1.1. If 5 or fewer responses are observed in up to 15 patients of evaluable population (EP), the trial will be terminated in favor of the null for futility. Otherwise, up to a further 31 patients may be evaluated, for a maximum total of 46 evaluable patients. Recruitment will not be halted during the interim analysis period. Therefore, no interruption in the accrual will be done during the interim analysis to maintain the dynamic of accrual in the trial.

Stage I interim analysis report will summarize descriptively the following results:

- Baseline Characteristics
- ORR (Overall Response Rate), as per local investigator´s assessment and according RECIST v1.1
- Median PFS (Progression Free Survival), as determined locally by the investigator according RECIST v1.1.
- DOR (Duration of response), as determined locally by the investigator according RECIST v1.1
- CBR (Clinical Benefit Rate), as determined locally by the investigator according RECIST v1.1
- OS (Overall Survival)
- Maximum Tumor shrinkage, according RECIST v1.1
- Extent of Exposure
- Concomitant Medication
- Adverse Events

No statistical tests are planned for these data.

5.7 Deviations from SAP

Any deviations from the original statistical plan will be described and justified in the final clinical study report.

6 APPENDICES

6.1 SAS Codes for Primary Endpoint

All report outputs will be produced using SAS® version 9.4 (TS1M5) version in a secure and validated environment.

ORR

The number and proportion of patients with ORR, with its 95% Wilson confidence intervals will be calculated.

```
proc freq data=EP;
  tables ORR / binomial(wilson p=.30) alpha=.05;
  title 'ORR (EP)';
run;
```

6.2 List of Tables, Listings, Figures

A complete list of tables, listings and figures (TLFs) will be given in a separate document which can be updated without updating the SAP. The list will serve as a reference for both the Sponsor, the trial statistician and the statistical programmer and includes the totality of statistical output to be produced.

All output will be headed with an appropriate heading specifying the study ID and abbreviated study title.

All output will be dated and have page numbers in the form 'Page [x / y]' where x denotes the current page within an output and y the total number of pages of that output.

All statistical output will identify the underlying analysis populations and indicate the number of patients/events in this population (N) and the number of patient/events actuals contributing to the output (n). All statistical output will be presented per treatment (if applicable).

All patient listings will contain additionally to the patient identification the analysis population and the treatment.