Official Title: A Phase III Randomized Study to Investigate the Efficacy and Safety of Atezolizumab (Anti-PD-L1 Antibody) in Combination With Neoadjuvant Anthracycline/Nab-Paclitaxel-Based Chemotherapy Compared With Placebo and Chemotherapy in Patients With Primary Invasive Triple-Negative Breast Cancer

NCT Number: NCT03197935

Document Date: SAP Version 2: 06-November-2018

STATISTICAL ANALYSIS PLAN

TITLE: **A PHASE III RANDOMIZED STUDY TO INVESTIGATE THE EFFICACY AND SAFETY OF ATEZOLIZUMAB (ANTI–PD-L1 ANTIBODY) IN COMBINATION WITH NEOADJUVANT ANTHRACYCLINE /NAB-PACLITAXEL-BASED CHEMOTHERAPY COMPARED WITH PLACEBO AND CHEMOTHERAPY IN PATIENTS WITH PRIMARY INVASIVE TRIPLE-NEGATIVE BREAST CANCER**

PROTOCOL NUMBER: WO39392

STUDY DRUG: Atezolizumab (MPDL3280A; RO5541267)

VERSION NUMBER: 2

IND NUMBER: 123277

EUDRACT NUMBER: 2016-004734-22

SPONSOR: F. Hoffmann-La Roche Ltd

PLAN PREPARED BY: [redacted], Ph.D.

DATE FINAL: Version 1: 6 December 2017

DATE AMENDED: Version 2: See electronic date stamp below

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

Name	Reason for Signing	Date and Time (UTC)
[redacted]	Company Signatory	06-Nov-2018 13:15:32

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

This Statistical Analysis Plan (SAP) Version 2 for Study WO39392 (IMpassion031) has been amended to incorporate an adaptive, two-stage design that uses accumulating data to modify aspects of the study as it continues, without undermining the validity and integrity of the trial.

Study WO39392 is currently designed (protocol version 4) to randomize approximately 204 patients in total (approximately 102 patients per arm) to control the type I error for the primary endpoint of pathological complete response (pCR) at the 5% level of significance (two-sided) in the intent-to-treat (ITT) population. However, we have observed strong evidence of difference in efficacy effect across the programmed death–ligand 1 (PD-L1) subgroups of atezolizumab plus nab-paclitaxel arm in metastatic triple negative breast cancer (TNBC) study (WO29522) that could impact the WO39392 study. To address the uncertainty surrounding design choices based on the effect of PD-L1 and to allow review of accumulating information during the ongoing WO39392 study, we plan to change the current fixed design to an adaptive design. The proposed adaptive design consists of two trial stages. After an interim analysis at the end of stage 1 (i.e., approximately 204 patients), recommendations will be made to either continue the study unchanged or to expand the target population for a subsequent stage 2 of the trial (i.e., approximately 120 additional patients).

To maintain the integrity of the trial, an independent Data Coordinating Center will perform the interim analysis and an independent Data Monitoring Committee (iDMC) will evaluate the interim analysis results. The iDMC will provide a recommendation as to whether to continue the study unchanged or expand to stage 2. The decision rule to be applied at the interim analysis will be clearly expressed to the iDMC and documented in the iDMC charter so that the study can be conducted with the Sponsor remaining completely blinded to all results unless iDMC recommendation is not to expand into stage 2. The iDMC will also monitor safety and study conduct for stage 2 of Study WO39392.

To reflect these changes, this SAP has been amended alongside the study protocol. Additional minor changes have been made to improve clarity and consistency.

TABLE OF CONTENTS

1. BACKGROUND 5

2. STUDY DESIGN 5

2.1 Protocol Synopsis 5

2.2 Outcome Measures 5

2.2.1 Primary Efficacy Outcomes 5

2.2.2 Secondary Efficacy Outcomes 6

2.2.3 Safety Outcomes 6

2.2.4 Pharmacokinetic Outcomes 6

2.2.5 Immunogenicity Outcomes 6

2.3 Determination of Sample Size 6

3. STUDY CONDUCT 7

3.1 Randomization 7

3.2 Data Monitoring 8

4. STATISTICAL METHODS 8

4.1 Definition of Analysis Populations 8

4.2 Analyses of Study Conduct 8

4.3 Analyses of Treatment Group Comparability 9

4.4 Analyses of Efficacy 9

4.4.1 Primary Efficacy Endpoint 9

4.4.1.1 Primary Analyses 9

4.4.2 Secondary Efficacy Endpoints 12

4.4.2.1 Event-Free Survival 12

4.4.2.2 Disease-Free Survival 12

4.4.2.3 Overall Survival 13

4.4.2.4 Patient-Reported Outcomes of Role and Physical Function and Global Health Status/ Health-Related Quality of Life–EORTC Data 13

4.4.3 Exploratory Efficacy Analyses 14

4.4.3.1 Patient-Reported Outcomes of Disease/Treatment Symptoms, Emotional and Social Function – EORTC Data 14

4.4.3.2 Patient-Reported Outcome of Treatment Bother – FACT-G, GP5 Data 14
4.4.3.3 Health Economic EQ-5D-5L Data 14
4.4.3.4 Proportion of Patients Undergoing Breast-Conserving Surgery 14
4.4.3.5 Residual Cancer Burden Index 14
4.5 Analyses of Safety 15
4.5.1 Exposure of Study Medication 15
4.5.2 Adverse Events 15
4.5.3 Laboratory Data 16
4.5.4 Vital Signs 16
4.6 Pharmacokinetic Analyses 16
4.7 Immunogenicity Analyses 16
4.8 Biomarker Analyses 17
4.9 Exploratory Analyses 17
4.9.1 Exploratory Biomarker Analyses 17
4.9.2 Subgroup Analyses 17
4.10 Interim Analyses 18
4.10.1 Planned Interim Analysis 18
4.10.2 Optional Interim Analysis 19
5. REFERENCES 20

LIST OF FIGURES

Figure 1 Flowchart of the Decision Rules 19

LIST OF APPENDICES

Appendix 1 Protocol Synopsis 21
Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2) 32
Appendix 3 Adaptive Multiple Test Design 39

1. BACKGROUND

This Statistical Analysis Plan (SAP) describes the analyses that are planned to be performed for Study WO39392 (also known as IMpassion031).

2. STUDY DESIGN

2.1 PROTOCOL SYNOPSIS

Study WO39392 (IMpassion031) is a global Phase III, double-blind, randomized, placebo-controlled study designed to evaluate the efficacy and safety of neoadjuvant treatment with TECENTRIQ®(atezolizumab) and nab-paclitaxel followed by doxorubicin and cyclophosphamide with atezolizumab compared with neoadjuvant treatment with placebo and nab-paclitaxel followed by doxorubicin and cyclophosphamide with placebo in patients with primary invasive early triple-negative breast cancer (TNBC), that are eligible for surgery. The study will enroll patients with initial clinically assessed T2-4d TNBC.

The study design consists of two trial stages, 1 and 2, with an a priori fixed sample size for both stages. In stage 1, approximately 204 patients will be randomized in a 1:1 ratio to the atezolizumab arm or the placebo arm as described in Section 3.1.

After an interim analysis conducted at the completion of stage 1 (i.e., once all patients enrolled in stage 1 have had surgery and pCR assessment), a decision will be made to either continue the study unchanged or to expand patient enrolment for a subsequent stage 2 of the trial (i.e., approximately 120 additional patients). If after the first stage, the trial is not read out for efficacy or lack of efficacy, the sponsor will proceed into the second stage with the selected population. Approximately 120 additional patients will be randomized in a 1:1 ratio to the atezolizumab arm or the placebo arm. At the conclusion of stage 2, the primary analysis of pCR is conducted by combining the p-values of both stage 1 and 2 into a single test statistic using a predefined combination function.

Note that the primary analysis of pCR could occur at the end of stage 1 (early stop for efficacy) projected to take place approximately 21 months after the first patient has been randomized or at the end of stage 2 projected to take place approximately 35–44 months after the first patient has been randomized depending on the targeted population selected at the end of stage 1.

The protocol synopsis for Study WO39392 including a description of the study design is in Appendix 1.

2.2 OUTCOME MEASURES

2.2.1 Primary Efficacy Outcomes

The co-primary efficacy endpoint for this study is pCR defined as eradication of invasive tumor from both breast and lymph nodes (ypT0/is ypN0) in the intent-to-treat (ITT)

population (full population) and in the subpopulation with programmed death-ligand 1 (PD-L1)–positive tumor status (tumor-infiltrating immune cell [IC] IC1/2/3) (see Section 4.1). Patients whose pCR assessment was missing will be counted as not achieving a pCR.

2.2.2 Secondary Efficacy Outcomes

- Event-free survival (EFS) defined as the time from randomization until documented disease recurrence, progression, or death from any cause in all patients and in the subpopulation with PD-L1–positive tumor status. More details for EFS definition are given in Section 4.4.2.1.
- Disease-free survival (DFS) defined as the time from surgery until documented disease recurrence or death from any cause in all patients (ITT population) who undergo surgery and in the subpopulation of patients with PD-L1–positive tumor status who undergo surgery. More details for DFS definition are given in Section 4.4.2.1.
- Overall survival (OS) defined as the time from randomization to the date of death from any cause in all patients and in the subpopulation with PD-L1–positive tumor status.
- Mean and mean changes from baseline score in function (role, physical) and global health status (GHS)/ health-related quality of life (HRQoL) by cycle and between treatment arms as assessed by the functional and HRQoL scales of the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire Core 30 (QLQ C30).

2.2.3 Safety Outcomes

- Incidence, nature, and severity of adverse events graded according to National Cancer Institute Common Terminology Criteria in Adverse Events, Version 4.0 (NCI CTCAE v4.0).
- Changes in vital signs and clinical laboratory results.

2.2.4 Pharmacokinetic Outcomes

Serum concentration of atezolizumab at specified timepoints (see Appendix 2).

2.2.5 Immunogenicity Outcomes

Incidence of anti-drug antibodies (ADAs) during the study and the prevalence of ADAs at baseline.

2.3 DETERMINATION OF SAMPLE SIZE

The study will randomize approximately 204 patients in stage 1 (approximately 102 patients per arm) and approximately 120 patients in stage 2 (approximately 60 patients per arm).

The performance of the study design depends on several factors including the underlying proportion of (pCR) responders in each treatment arms for the subgroup S

(PD-L1 IC1/2/3) and the complement group C (PD-L1 IC0). Simulations were carried out to evaluate the operation characteristics of the current design under a variety of conditions including the overall power defined as the probability to reject any null hypothesis at any analysis time point in the study, which is observed to be at least 68% across the simulation scenarios under investigation. More details are provided in Appendix 1.

3. STUDY CONDUCT

3.1 RANDOMIZATION

Randomization will occur in a 1:1 ratio using a permuted-block randomization method. Patients will be randomized to one of two treatment arms:

- **Arm A**: atezolizumab (840 mg fixed dose) administered via intravenous (IV) infusion every two weeks (q2w) in combination with nab-pac (125 mg/m^2) administered via IV infusion every week (qw) for 12 weeks followed by atezolizumab (840 mg fixed dose) administered q2w in combination with doxorubicin (60 mg/m^2)+cyclophosphamide (600 mg/m^2) administered q2w via IV infusions with filgrastim/ pegfilgrastim support for 4 cycles. Patients randomized to the atezolizumab arm will continue to receive unblinded atezolizumab post-surgery at a fixed dose of 1200 mg by IV infusion every 3 weeks (q3w) for 11 cycles, for a total of approximately 12 months of atezolizumab therapy.
- **Arm B**: placebo administered q2w via IV infusion in combination with nab-paclitaxel (125 mg/m^2) administered qw via IV infusion for 12 weeks followed by placebo administered q2w in combination with doxorubicin (60 mg/m^2)+cyclophosphamide (600 mg/m^2) administered q2w via IV infusions with filgrastim/ pegfilgrastim support for 4 cycles. Patients randomized to the placebo arm will be unblinded post-surgery and will continue to be followed up as per the schedule of activities outlined in the study protocol.

Postoperative patient management for those in either treatment arm may include radiotherapy as clinically indicated, and management of patients who do not achieve a pCR should follow current standard-of-care guidelines. For patients in Arm A, chemotherapy and/or radiotherapy may be administered concurrently with atezolizumab after discussion with Medical Monitor.

The randomization scheme is designed to ensure that an approximately equal number of patients will be enrolled in each treatment arm within the categories defined for the following stratification factors at baseline:

- American Joint Committee on Cancer (AJCC) Stage at diagnosis (II vs. III) for patients enrolled in both stages 1 and 2.
- Tumor infiltrating immune cells PD-L1 status (IC0 vs. IC1/2/3) for patients enrolled in stage 1 as well as stage 2 IF the study continues to enroll all-comer patients in stage 2.

3.2 DATA MONITORING

An iDMC will evaluate the primary efficacy endpoint of pCR in the ITT and PD-L1–positive populations at the prespecified interim analysis based on data of approximately 204 patients enrolled in stage 1 and will make a recommendation to either complete the study unchanged or to expand the target population for the subsequent stage 2 of the trial (i.e., approximately 120 additional patients). The decision rules to be applied at the interim analysis will be clearly expressed to the iDMC and documented in the iDMC charter so that the study can be conducted with the Sponsor remaining completely blinded to all results at this stage.

The iDMC will also evaluate safety data and study conduct on a regular basis during the study until the primary analysis of pCR, which is performed at stage 1 (approximately 204 patients) or stage 2 (approximately 324 patients), if the iDMC recommends extending the target population after which, iDMC review of the study data will be discontinued.

Sponsor affiliates will be excluded from iDMC membership. All summaries and analyses for the iDMC review will be prepared by an independent Data Coordinating Center. The iDMC will follow a charter that outlines the iDMC's roles and responsibilities.

4. STATISTICAL METHODS

The analyses outlined in this SAP supersede those specified in the protocol for regulatory filing purposes.

4.1 DEFINITION OF ANALYSIS POPULATIONS

The primary analysis population for efficacy is the ITT population (full population), defined as all randomized patients, and the PD-L1–positive subpopulation, defined as patients in the ITT population whose PD-L1 status is IC1/2/3 at the time of randomization. Patients will be assigned to the treatment group to which they were randomized.

The primary analysis population for safety is the safety evaluable population defined as all patients who received at least one dose of study medication. Patients will be assigned to treatment groups as treated, and all patients who received any dose of atezolizumab will be included in the atezolizumab treatment arm.

4.2 ANALYSES OF STUDY CONDUCT

For all randomized patients (i.e., ITT population), a participant flowchart for depicting the progress of subjects through the phases of the trial will be provided by treatment arm and study stage (i.e., 1 and 2) of enrollment, including a complete description of patient disposition specifying the number of randomized patients and completed and discontinued patients from trial treatment and study with reasons for premature discontinuation. Documented major protocol deviations including those related to study

inclusion/ exclusion criteria, conduct of the study, patient management, or patient assessment will be also tabulated by treatment arm and study stage of enrollment.

4.3 ANALYSES OF TREATMENT GROUP COMPARABILITY

Demographic variables such as age, sex, race/ethnicity, and baseline characteristics (in particular, stratification variables) will be summarized by treatment arm and study stage (i.e., 1 and 2) of enrollment for the ITT as well as PD-L1−positive population. Only descriptive analyses are planned; no formal statistical tests will be applied. Continuous variables will be reported and summarized by use of standard measures of central tendency and dispersion (mean, SD, median, and range including minimum and maximum), and categorical (i.e., discrete) data will be reported and summarized by frequencies and percentages.

The baseline value of any variable will be defined as the last available value prior to the first administration of study treatment.

4.4 ANALYSES OF EFFICACY

4.4.1 Primary Efficacy Endpoint

4.4.1.1 Primary Analyses

The ITT as well as PD-L1−positive population will be used for the primary analysis of the co-primary endpoints of pCR in these populations as follows. The **overall one-sided type I error** for testing the co-primary endpoints of pCR is $\boldsymbol{\alpha = 0.025}$.

Two main sources for multiplicity in statistical testing from the current design are:

- Multiple target populations, namely ITT and PD-L1−positive populations.
- Study having 2 stages with target population at stage 2 depending on results from stage 1.

Simes' closed testing procedure will be used to address multiplicity in target populations and weighted inverse p-value combination method is applied to account for the adaptive choice of target population at stage 2 as follows.

At Stage 1

For the interim analysis at end of stage 1, the co-primary endpoints of pCR in the ITT and PD-L1−positive population will be tested at the pre-defined one-sided type I error of 0.0125 (i.e. 50% of the total type I error) following Simes' procedure. This involves the following p-values:

- $\boldsymbol{p_1^{\{F\}}}$ is the p-value from the **one-sided** Cochran-Mantel-Haenszel (CMH) χ^2 test for difference in pCR rates between treatment groups in the **ITT population** (Δ_F) stratified according to tumor PD-L1 status (IC0 vs. IC1/2/3) and clinical stage at presentation (Stage II vs. III). The associated elementary null hypothesis is denoted as $H_0^{\{F\}}: \Delta_F \leq 0$.

- $\boldsymbol{p_1^{\{S\}}}$ is the p-value from the CMH χ^2 test for difference in pCR rates between treatment groups in the **PD-L1–positive population** (Δ_S) stratified according to clinical stage at presentation (Stage II vs. III). The associated elementary null hypothesis is denoted as $H_0^{\{S\}}$:$\Delta_S \leq 0$.
- $\boldsymbol{p_1^{\{F,S\}}} = min\left\{2min\left(p_1^{\{S\}}, p_1^{\{F\}}\right), max\left(p_1^{\{S\}}, p_1^{\{F\}}\right)\right\}$, p-value based on Simes' method for the intersection hypothesis $H_0^{\{F,S\}}$, which specifies that there is no proportion difference in either F or S.

In addition, the observed difference in pCR rates between arms in S, denoted as $\widehat{\Delta}_1^S$ and its complementary sub-population of patients with PD-L1 IC0 (C), denoted as $\widehat{\Delta}_1^C$ will also be computed.

Based on $p_1^{\{F\}}$, $p_1^{\{S\}}$, $p_1^{\{F,S\}}$, $\widehat{\Delta}_1^S$ and $\widehat{\Delta}_1^C$, the following decision could be made at stage 1:

1. If $p_1^{\{F,S\}} \leq \alpha_1$, the analysis at stage 1 is considered the primary efficacy analysis for pCR with evidence for efficacy declared in either F (if $p_1^{\{F\}} \leq \alpha_1$) or S (if $p_1^{\{S\}} \leq \alpha_1$) or both, the study will therefore not enroll extra patients in stage 2.
2. Otherwise, if both $\widehat{\Delta}_1^S$ and $\widehat{\Delta}_1^C$ are below the pre-specified thresholds, $d_S = 0.12$ and $d_C = 0.1$, respectively, the study will also not enroll extra patients in stage 2 for lack of efficacy.
3. Otherwise, if $\widehat{\Delta}_1^S \geq d_S$ and $\widehat{\Delta}_1^C < d_C$, the study will enroll only patients in S at stage 2, with S being the only target population in the primary efficacy analysis for pCR.
4. Otherwise, if $\widehat{\Delta}_1^S < d_S$ and $\widehat{\Delta}_1^C \geq d_C$, the study will enroll all-comer patients (F) in stage 2; however, only F will be the target population in the primary efficacy analysis for pCR.
5. If $\widehat{\Delta}_1^S \geq d_S$ and $\widehat{\Delta}_1^C \geq d_C$, the study will enroll all-comer patients in stage 2, and the co-primary efficacy endpoints of pCR in F and S will be considered in the primary efficacy analysis for pCR.

<u>At Stage 2 (if needed)</u>

If the study advances to stage 2, at the primary efficacy analysis for pCR to protect the overall α the corresponding boundary for statistical significance is α_2=0.0184, accounting for the fact that and overall $\alpha_1 = 0.0125$ having been spent at stage 1 and the information fraction at stage 1, namely $w_1 = N_1/(N_1 + N_2)$. Indeed, the p-value p_1 based on stage 1 data only and the combination p-value p_2 have the same joint distribution under the null hypothesis as the p-values from a group-sequential test with two stages at information times $t_1 = w_1$ and $t_2 = 1$. Thus standard statistical software for group sequential designs can be used for the determination of critical values for the adaptive design (Wassmer and Brannath 2016, Section 6.2.4).

Final test(s) involving the target population(s) selected at stage 1 will be based on weighted inverse normal p-value combination method (Wassmer and Brannath 2016) as follows:

Let $w_1 = \sqrt{N_1/(N_1+N_2)}$ and $w_2 = \sqrt{N_2/(N_1+N_2)}$ be the a priori chosen weights associated with stage 1 and stage 2, respectively.

If both F and S are selected as the co-primary target populations, the following quantities are computed:

- $\boldsymbol{p_2^{\{S\}}}$ and $\boldsymbol{p_2^{\{F\}}}$ are defined and derived similarly as $\boldsymbol{p_1^{\{S\}}}$ and $\boldsymbol{p_2^{\{F\}}}$, respectively, with derivation based solely on patients enrolled in stage 2.
- $\boldsymbol{p_2^{\{F,S\}}} = min\left\{2min\left(p_2^{\{S\}}, p_2^{\{F\}}\right), max\left(p_2^{\{S\}}, p_2^{\{F\}}\right)\right\}$, p-value based on Simes' method for the intersection hypothesis $H_0^{\{F,S\}}$ based only on patients enrolled in stage 2.
- $\boldsymbol{p_{12}^{\{S\}}} = 1 - \Phi\left\{w_1 \cdot \Phi^{-1}\left(1 - p_1^{\{S\}}\right) + w_2 \cdot \Phi^{-1}\left(1 - p_2^{\{S\}}\right)\right\}$, the combined p-value for $\mathrm{H}_0^{\{\mathrm{S}\}}$ from stage 1 and 2 following the weighted inverse normal combination method, where Φ denotes the cumulative distribution function of a standard normal variate.
- $\boldsymbol{p_{12}^{\{F\}}} = 1 - \Phi\left\{w_1 \cdot \Phi^{-1}\left(1 - p_1^{\{F\}}\right) + w_2 \cdot \Phi^{-1}\left(1 - p_2^{\{F\}}\right)\right\}$, the combined p-value for $\mathrm{H}_0^{\{\mathrm{F}\}}$ from stage 1 and 2 following the weighted inverse normal combination method.
- $\boldsymbol{p_{12}^{\{F,S\}}} = 1 - \Phi\left\{w_1 \cdot \Phi^{-1}\left(1 - p_1^{\{F,S\}}\right) + w_2 \cdot \Phi^{-1}\left(1 - p_2^{\{F,S\}}\right)\right\}$, the combined p-value for $\mathrm{H}_0^{\{\mathrm{F,S}\}}$ from stage 1 and 2 following the weighted inverse normal combination method.

If either F or S, but not both of them, is selected as the only target population, there remains only one elementary null hypothesis, which is also the intersection hypothesis. In general, let Q be the selected target population, that is, Q being either S or F the following quantities are computed:

- $\boldsymbol{p_2^{\{Q\}}}$ is derived as above and $\boldsymbol{p_2^{\{F,S\}}} = p_2^{\{Q\}}$.
- $\boldsymbol{p_{12}^{\{Q\}}}$ and $\boldsymbol{p_{12}^{\{F,S\}}}$ are then computed based on $(p_1^{\{Q\}}, p_2^{\{Q\}})$ and $(p_1^{\{F,S\}}, p_2^{\{F,S\}})$, respectively, as mentioned above. Please note that $p_1^{\{F,S\}}$ is the same as when both F and S are selected as the target populations. In general $p_1^{\{F,S\}}$ is invariant to selection of the target population(s) for the primary efficacy analysis for pCR.

In all cases, at the primary efficacy analysis for pCR, the null hypothesis associated with a selected target population (Q) is rejected if both $p_{12}^{\{Q\}} \leq \alpha_2$ and $p_{12}^{\{F,S\}} \leq \alpha_2$ following Simes' procedure.

Operating characteristics of the current design across several simulation scenarios are provided in Appendix 1.

A summary table that presents the number and proportion of pCR in each treatment arm, together with the 2-sided 95% CIs with use of the Clopper–Pearson method (Clopper and Pearson 1934) will be produced by stage and overall for both the ITT and PD-L1–positive populations. Confidence intervals for the difference in pCR rate between the two arms will be determined using the normal approximation to the binomial distribution.

4.4.2 Secondary Efficacy Endpoints

4.4.2.1 Event-Free Survival

Event-free survival (EFS) is defined as the time from randomization to the first documented occurrence of disease recurrence, disease progression, or death from any cause. EFS events covered under "disease recurrence" will include local, regional, or distant recurrence and contralateral breast cancer. Ipsilateral or contralateral in situ disease and second primary non-breast cancers will not be counted as EFS events, even though they will be recorded in the electronic Case Report Form accordingly.

Patients without an event at the time of the analysis will be censored on the date on which they are last known to be alive and event free, on or before the clinical data cutoff date for the respective analysis. Patients with no post-baseline information will be censored at the date of randomization.

EFS will be compared between treatment arms with use of the stratified log-rank test. The hazard ratio (HR) for EFS will be estimated using a stratified Cox proportional hazards model. The 95% CI for the HR will be provided. Kaplan–Meier methodology will be used to estimate the median EFS (if reached) for each treatment arm, and Kaplan-Meier curves will be produced. The Brookmeyer–Crowley methodology will be used to construct the 95% CI for the median EFS for each treatment arm (Brookmeyer and Crowley 1982). The Kaplan–Meier approach will be used to estimate 3-year EFS rates and corresponding 95% CIs for each treatment arm.

The stratification factors for all analyses based on the ITT population will be: tumor stage (AJCC) at baseline (Stage II vs. III) and PD-L1 status (IC0 vs. IC1/2/3).

The stratification factors for EFS analysis in the PD-L1–positive subpopulation will be AJCC at baseline (Stage II vs. III).

Results from unstratified analyses will also be provided.

4.4.2.2 Disease-Free Survival

DFS is defined as the time from surgery (i.e., the first date of no disease) to the first documented disease recurrence or death from any cause, whichever occurs first. DFS will be analyzed with the use of the same methodology as specified for EFS for both the ITT population and the PD-L1 positive subpopulation.

The DFS analysis will be performed approximately 36 months after the randomization of the last patient.

Patients who do not undergo surgery at the end of neoadjuvant treatment will be excluded from the analysis of DFS. Patients without a DFS event at the time of analysis will be censored at the date when they were last known to be alive and event free. Patients who do not have information after surgery will be censored at the date of surgery.

4.4.2.3 Overall Survival

Overall Survival (OS) is defined as the time from the date of randomization to the date of death due to any cause. Patients who are not reported as having died at the time of analyses will be censored at the date when they were last known to be alive. Patients who do not have information after baseline will be censored at the date of randomization. OS will be analyzed with the use of the same methodology as specified for EFS for both the ITT population and the PD-L1–positive subpopulation.

4.4.2.4 Patient-Reported Outcomes of Role and Physical Function and Global Health Status/ Health-Related Quality of Life–EORTC Data

The EORTC QLQ-C30 (version 3) data will be scored according to the EORTC scoring manual (Fayers et al. 2001). Missing data will be assessed and reported by cycle. In the event of incomplete data, if the scale has more than 50% of the constituent items completed, a pro-rated score will be computed consistent with the scoring manual and validation papers of the measure. For subscales with less than 50% of the items completed, the subscale will be considered missing. Patient-reported outcomes (PRO) completion, compliance rates, and reasons for missing data will be summarized at each timepoint by treatment arm.

The primary patient-reported endpoints are mean and mean changes from baseline score in function (role, physical) and GHS/ HRQoL. Summary statistics (mean, SD, median, and range) of linearly transformed absolute scores and mean changes from baseline will be calculated for the functional (role [Question {Q}6, Q7], physical [Q1–Q5]) and the GHS/ HRQoL (Q29, Q30) scales of the EORTC QLQ-C30 at each assessment time point for each arm. The mean change from baseline (and 95% CI) will be assessed on patients with at least one post-baseline measurement. Previously published meaningful thresholds of change important differences will be used to identify clinically meaningful change from baseline within each treatment group on the functional and GHS/ HRQoL scales (Osoba et al. 1998, Cocks et al. 2011).

Longitudinal analysis will be conducted to estimate the effect difference on PRO repeated responses over a selected period of time and between the treatment arms. Mixed effect models on a set of covariates (baseline domain score, patient demographic, and clinical variables) will be conducted. Change from baseline at subsequent cycles

will be presented by treatment arm and will include least squares mean (LS Mean), difference in LS Mean between two treatment arms, and 95% CIs for the differences. The standard error will also be calculated for each LS Mean.

4.4.3 Exploratory Efficacy Analyses

4.4.3.1 Patient-Reported Outcomes of Disease/Treatment Symptoms, Emotional and Social Function – EORTC Data

Summary statistics (mean, SD, median, and range) of linearly transformed absolute scores and mean changes from baseline will be calculated for all disease/treatment-related symptom items and scales, and the emotional, social function scales of the EORTC QLQ-C30 at each assessment timepoint for each arm. The analyses of this exploratory endpoint will be analogous to the secondary PRO endpoints (see Section 4.4.2.4 for further details).

4.4.3.2 Patient-Reported Outcome of Treatment Bother – FACT-G, GP5 Data

A descriptive analysis of absolute scores and the proportion of patients selecting each response option at each assessment time-point by treatment arm will be reported for item GP5 ("I am bothered by side effects of treatment") from the Functional Assessment of Cancer Therapy-General (FACT-G) physical well-being subscale. Item GP5 from version 4 of the FACT-G questionnaire will be scored according to the FACIT scoring manual (Cella 1997).

4.4.3.3 Health Economic EQ-5D-5L Data

Health utility data from the EuroQoL 5 Dimension, 5 Level (EQ-5D-5L), will be evaluated in pharmacoeconomic models. The results from the health economic data analyses will be reported separately from the clinical study report.

4.4.3.4 Proportion of Patients Undergoing Breast-Conserving Surgery

The proportion of patients undergoing breast-conserving surgery will be compared between the treatment arms using the same methodology for pCR analysis in Section 4.4.1.1, for the subpopulation of patients having surgery.

4.4.3.5 Residual Cancer Burden Index

Pathological review of the primary tumor site and sampled lymph nodes by the local pathologists at surgery will also include the calculation of the Residual Cancer Burden (RCB) index. Such index is computed by taking into account the primary tumor bed area, overall cancer cellularity, the percentage of in situ disease, number of positive lymph nodes and the diameter of the largest lymph node metastasis (Symmans et al. 2007). Pathological complete response would correspond to RCB=0, whereas minimal, moderate, and extensive residual disease would be indicated by RCB-I, RCB-II, and RCB-III, respectively.

The following analyses will be carried out for the population of patient with evaluable RCB index.

A summary table that presents the number and proportion of RCB-0, -I, -II and -III in each treatment arm, together with the respective 2-sided 95% CIs with use of the Clopper-Pearson method (Clopper and Pearson 1934) will be produced.

Comparison between treatment arms will be done using ordered logit (proportional-odds) regression model (Peter McCullagh 1980) stratified for the stratification factors as for the primary pCR analysis.

4.4.3.5.1 Association between RCB Index and EFS, OS

The analyses for EFS, DFS and OS, described in Section 4.4.2.1, Section 4.4.2.2, and Section 4.4.2.3 respectively, will be repeated for the subgroups of patients with RCB-0, RCB-I, RCB-II, and RCB-III.

4.5 ANALYSES OF SAFETY

Safety analyses will be performed on the safety population defined as all patients who received any dose of study medication.

Safety will be assessed through summaries of AEs, changes in laboratory test results, changes in vital signs, study treatment exposures, and immunogenicity as measured by ADAs and will be presented by treatment arm.

4.5.1 Exposure of Study Medication

Study drug exposure, including but not limited to treatment duration, number of cycles, and dose intensity, will be summarized with descriptive statistics for each study treatment on each treatment arm if deemed appropriate.

4.5.2 Adverse Events

Verbatim descriptions of AEs will be mapped to Medical Dictionary for Regulatory Activities (MedDRA) thesaurus terms and graded according to NCI CTCAE v4.0.

Treatment-emergent adverse events (TEAEs), defined as events occurring on or after the first dose of study treatment will be summarized by MedDRA term, appropriate MedDRA levels, and NCI CTCAE v4.0 grade, regardless of relationship to study drug as assessed by the investigator. For each patient, if multiple incidences of the same AEs occur, the maximum severity reported will be used in the summaries.

The following TEAEs will be summarized separately:

- AEs leading to withdrawal of study drug,
- AEs leading to dose reduction or interruption,
- Grade $\geq$ 3 AEs, Grade 5 AEs, serious adverse events (SAEs), and
- Adverse events of special interest (AESIs).

All deaths and causes of death will be summarized.

4.5.3 Laboratory Data

Laboratory data with values outside of the normal ranges will be identified. Relevant laboratory values will be summarized by treatment arm over time, with NCI CTCAE v4.0 Grade 3 and Grade 4 values identified, where appropriate. Changes from baseline in NCI CTCAE v4.0 grade (i.e., shift tables) will be also provided by treatment arm. Of note, abnormal laboratory data that are clinically significant will be reported as adverse events and summarized in the adverse event tables.

A Hy's law analysis will be provided.

The finding of an elevated ALT or AST (>3×**baseline value**) in combination with either an elevated total bilirubin (>2×ULN) **or clinical jaundice in the absence of cholestasis or other causes of hyperbilirubinemia** is considered to be an indicator of severe liver injury (as defined by Hy's Law). Therefore, investigators must report as an adverse event the occurrence of either of the following:

- Treatment-emergent ALT or AST >3 × ULN (or baseline value if baseline value was above the ULN) in combination with total bilirubin >2 × ULN (of which ≥35% is direct bilirubin).
- Treatment-emergent ALT or AST >3 × ULN (or baseline value if baseline value was above the ULN) in combination with clinical jaundice.

4.5.4 Vital Signs

Changes in selected vital signs will be summarized by treatment arm and by change over time including change from baseline.

4.6 PHARMACOKINETIC ANALYSES

Atezolizumab serum concentration data (C_{min} and C_{max}) will be tabulated and summarized. Descriptive statistics will include means, medians, ranges, and SDs, as appropriate.

Additional pharmacokinetic and pharmacodynamic analyses will be conducted as appropriate.

4.7 Immunogenicity Analyses

The immunogenicity analyses will include patients with at least one post-baseline ADA assessment, with patients grouped according to treatment received.

The numbers and proportions of ADA-positive patients and ADA-negative patients at baseline (baseline prevalence) and after baseline (post-baseline incidence) will be summarized by treatment group. When determining post-baseline incidence, patients are considered to be ADA-positive if they are ADA-negative or are missing data at baseline but develop an ADA response following study drug exposure (treatment induced ADA response), or if they are ADA-positive at baseline and the titer of one or

more post-baseline samples is at least 0.60 titer unit greater than the titer of the baseline sample (treatment-enhanced ADA response). Patients are considered to be ADA-negative if they are ADA-negative or are missing data at baseline and all post baseline samples are negative, or if they are ADA-positive at baseline but do not have any post-baseline samples with a titer that is at least 0.60 titer unit greater than the titer of the baseline sample (treatment unaffected).

The relationship between ADA status and safety, efficacy, and pharmacokinetics may be investigated.

4.8 BIOMARKER ANALYSES

EFS, DFS, and OS will be assessed in the PD-L1 IC1/2/3 population as secondary endpoints as described in Sections 2.2.2, 4.4.2.1, and 4.4.2.3 respectively.

4.9 EXPLORATORY ANALYSES

4.9.1 Exploratory Biomarker Analyses

Exploratory biomarker analyses will be performed in baseline pretreatment, on-treatment, residual-disease, and post recurrence samples in an effort to understand the association of these markers with study treatment outcome, including efficacy and/or adverse events. The biomarkers may include but will not be limited to PD-L1-expressing tumor cells, tumor infiltrating immune cells, RNA-based T-effector signature and other biomarkers in tumor and blood, as defined by immunohistochemistry, quantitative reverse transcription polymerase chain reaction, next-generation DNA and RNA sequencing, or other methods. Pharmacodynamic changes of these biomarkers at baseline, on-treatment samples and at surgical resection.

Circulating tumor DNA (ctDNA) will be evaluated in post-surgery plasma samples and compared between control and experimental arms.

Whole genome sequencing data will be analyzed in the context of this study and explored in aggregate with data from other studies to increase researchers' understanding of disease pathobiology and guide the development of new therapeutic approaches.

Results from these analyses will be presented in a separate report.

4.9.2 Subgroup Analyses

To assess the consistency of study results in subgroups defined by demographic and baseline characteristics (including subtypes of triple-negative breast cancer [TNBC]), pCR, EFS, DFS, and OS in these subgroups will be examined. Summaries of EFS, DFS and OS, including unstratified HRs estimated from Cox proportional hazards models and Kaplan-Meier estimates of the median (if reached), will be produced separately for each level of the categorical variables.

4.10 INTERIM ANALYSES

4.10.1 Planned Interim Analysis

An interim analysis will be performed and reviewed by the iDMC at the end of stage 1 (i.e., once all of the approximately 204 patients enrolled in stage 1 have had surgery and pCR assessment). At this interim analysis, a decision concerning the target population for the subsequent stage of the trial has to be made if efficacy or lack thereof cannot be declared after stage 1 (as described in Section 4.4.1.1). This decision is based on the observed pCR proportions (in the ITT and PD-L1–positive populations) in the first stage of the trial and entails either of the following outcomes:

1. The analysis at stage 1 is considered the primary efficacy analysis for pCR with statistical evidence for efficacy in either the ITT (F) or PD-L1–positive population (S) or both. The study will therefore not enroll extra patients in stage 2.
2. If no statistical evidence for efficacy can be found in both F and S, and both the observed difference in pCR proportion between treatment arms in PD-L1-positive population (S) (denoted $\widehat{\Delta}_1^S$) and its complementary (C, IC0 patients) (denoted $\widehat{\Delta}_1^C$) are below the pre-specified threshold, d_S and d_C as specified in Section 4.4.1, respectively, the study will not enroll additional patients in stage 2. Hence, the analysis at stage 1 is considered the primary efficacy analysis for pCR.
3. The study will enroll only patients in S at stage 2, with S being the only target population in the primary efficacy analysis of pCR.

   If $\widehat{\Delta}_1^S \geq d_S$ and $\widehat{\Delta}_1^C < d_C$, the study will enroll only PD-L1-positive patients in stage 2.
4. The study will enroll all-comer patients (F) at stage 2 with F being the only target population in the primary efficacy analysis of pCR.

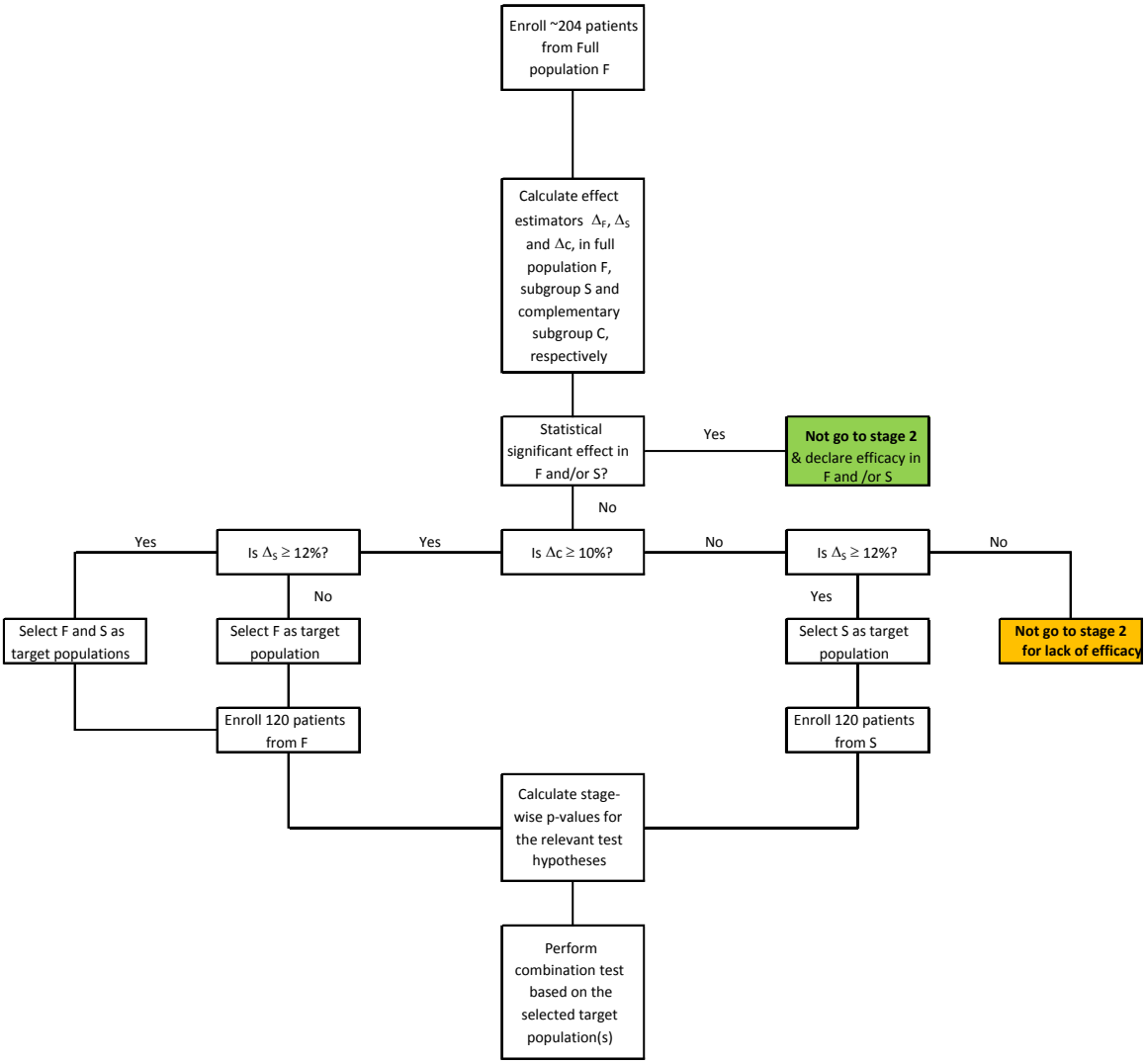
   If $\widehat{\Delta}_1^S < d_S$ and $\widehat{\Delta}_1^C \geq d_C$, the study will enroll all-comer patients in stage 2; however, only the pCR endpoint for F will be considered in the primary efficacy analysis of pCR.
5. The study will enroll all-comer patients at stage 2, and the co-primary efficacy endpoints of pCR in F and S will be considered in the primary efficacy analysis of pCR.

   If $\hat{\Delta}_1^S \geq d_S$ and $\hat{\Delta}_1^C \geq d_C$, the study will enroll all-comer patients in stage 2, and the co-primary efficacy endpoint of pCR in F and S will be considered in the primary efficacy analysis of pCR.

However, the integrity of the trial is not dependent on strict adherence to these thresholds, i.e. the proposed p-value combination test protects type I error regardless of how the decision regarding stage 2 is made (Brannath, Gutjahr, and Bauer 2012).

A graphical presentation of the decision rules is given in Figure 1 below with more details including selection and determination of the stage-wise boundaries for statistical significance provided in Section 4.4.1.1.

Figure 1 Flowchart of the Decision Rules



Although these guidelines are in place, the iDMC has authority to deviate from these guidelines if safety or additional efficacy analyses indicate that a different recommendation is more appropriate.

4.10.2 Optional Interim Analysis

To adapt to information that may emerge during the course of this study, the Sponsor may choose to conduct optional interim efficacy analyses after the primary analysis for pCR and before the final analysis for EFS, DFS and OS, if needed (e.g., for regulatory or publication purposes). The decision to conduct an optional interim analysis, along with the rationale, timing, and statistical details for the analysis, will be documented in the Sponsor's trial master file prior to the conduct of the interim analysis.

5. REFERENCES

Brannath W, Gutjahr G, and Bauer P. Probabilistic foundation of confirmatory adaptive designs. J Am Stat Assoc, 2012;107:824–832.

Brookmeyer R and Crowley J. A confidence interval for the median survival time. Biometrics 1982;38:29–4.

Cella D. Manual of the Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System. Center on Outcomes, Research and Education (CORE), Evanston Northwestern Healthcare and Northwestern University, Evanston IL, Version 4, 1997.

Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 1934;26:404–13.

Cocks K., King MT, Velikova G, et al. Evidence-based guidelines for determination of sample size and interpretation of the European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. J Clin Oncol 2011;29:89–96.

Fayers PM, Aaronson NK, Bjordal K, et al. On behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 scoring manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer, Brussels, 2001.

Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. Chapman and Hall/CRC, 1999.

Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality of life score. J Clin Oncol 1998;16:139–44.

Peter McCullagh. Regression Models for Ordinal Data. J R Stat Soc Series B Stat Methodol. 1980 42 (2):109-42.

Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol. 2007 Oct 1;25(28):4414–22.

Wassmer G, Brannath W. Group sequential and confirmatory adaptive designs in clinical trials. Heidelberg: Springer, 2016.

Appendix 1 Protocol Synopsis

TITLE:	**A PHASE III RANDOMIZED STUDY TO INVESTIGATE THE EFFICACY AND SAFETY OF ATEZOLIZUMAB (ANTI–PD-L1 ANTIBODY) IN COMBINATION WITH NEOADJUVANT ANTHRACYCLINE/NAB-PACLITAXEL-BASED CHEMOTHERAPY COMPARED WITH PLACEBO AND CHEMOTHERAPY IN PATIENTS WITH PRIMARY INVASIVE TRIPLE-NEGATIVE BREAST CANCER**
PROTOCOL NUMBER:	WO39392
VERSION NUMBER:	5
EUDRACT NUMBER:	2016-004734-22
IND NUMBER:	123277
TEST PRODUCT:	Atezolizumab (RO5541267)
PHASE:	III
INDICATION:	Triple-negative breast cancer
SPONSOR:	F. Hoffmann-La Roche Ltd

Objectives and Endpoints

Study WO39392 (also known as IMpassion031) will evaluate the efficacy, safety, and pharmacokinetics of neoadjuvant nab-paclitaxel and atezolizumab followed by doxorubicin and cyclophosphamide with atezolizumab (referred to as atezolizumab + nab-pac–AC) or neoadjuvant nab-paclitaxel and placebo followed by doxorubicin and cyclophosphamide with placebo (referred to as placebo + nab-pac–AC) in patients with T2–4d triple-negative breast cancer (TNBC). Specific objectives and corresponding endpoints for the study are as follows:

Objectives	Corresponding Endpoints
Primary Efficacy Objective:	
• To evaluate the efficacy of atezolizumab + nab-pac–AC compared with placebo + nab-pac–AC in the neoadjuvant setting	• pCR defined as eradication of tumor from both breast and lymph nodes (ypT0/is ypN0) *in the following:* – *All patients (ITT population)* – *Subpopulation of patients with PD-L1–positive tumor status (IC1/2/3)*

Objectives	Corresponding Endpoints
Secondary Efficacy Objectives:	
• To evaluate the efficacy of atezolizumab + nab-pac–AC compared with placebo + nab-pac–AC in the neoadjuvant setting	• EFS defined as the time from randomization until documented disease recurrence, progression, or death from any cause in all patients *(ITT population)* and in the subpopulation with PD-L1–*positive* tumor status • *DFS defined as the time from surgery until documented disease recurrence or death from any cause in all patients (ITT population) who undergo surgery and in the subpopulation of patients with PD-L1–positive tumor status who undergo surgery* • OS defined as the time from randomization to the date of death from any cause in all patient*s (ITT population) and in t*he subpopulation with PD-L1–*positive* tumor status
• To evaluate PROs of function and HRQoL associated with atezolizumab + nab-pac–AC compared with placebo + nab-pac–AC, measured by the functional and HRQoL scales of the EORTC QLQ-C30	• Mean and mean changes from baseline score in function (role, physical) and GHS/HRQoL by cycle and between treatment arms as assessed by the functional and HRQoL scales of the EORTC QLQ-C30
Exploratory Efficacy Objectives:	
• To evaluate the efficacy of atezolizumab + nab-pac–AC compared with placebo + nab-pac–AC in the neoadjuvant setting	• Proportion of patients undergoing breast-conserving surgery • RCB index • Correlation of RCB with other clinical endpoints (if deemed appropriate)
• To evaluate PROs of disease/treatment-related symptoms associated with atezolizumab + nab-pac–AC compared with placebo + nab-pac–AC, as measured by the EORTC QLQ-C30	• Mean and mean changes from baseline score in disease/treatment-related symptoms by cycle and between treatment arms as assessed by all symptom items/scales of the EORTC QLQ-C30
• To evaluate any treatment burden patients may experience associated with the addition of atezolizumab to nab-pac–AC compared with placebo + nab-pac–AC, as measured by a single item (GP5: “I am bothered by side effects of treatment”) from the physical well-being subscale of the FACT-G Quality of Life instrument	• Proportion of patients reporting each response option at each assessment timepoint by treatment arm for item GP5 from the FACT-G
• To evaluate and compare between treatment arms patient’s health utility as measured by the EQ-5D-5L questionnaire to generate utility scores for use in economic models	• Utility scores of the EQ-5D-5L questionnaire
Safety Objective:	
• To evaluate the safety and tolerability of atezolizumab + nab-pac–AC compared with placebo + nab-pac–AC	• Occurrence and severity of adverse events as defined by NCI CTCAE v4.0

Objectives	Corresponding Endpoints
Pharmacokinetic Objectives:	
• To characterize the pharmacokinetics of atezolizumab when administered in combination with nab-pac–AC chemotherapy	• Serum concentration of atezolizumab at specified timepoints
Immunogenicity Objective:	
• To evaluate the immune response to atezolizumab	• Incidence of ADAs during the study and the prevalence of ADAs at baseline
Exploratory Immunogenicity Objective:	
• To evaluate potential effects of ADAs	• Relationship between ADA status and efficacy, safety, and PK endpoints
Exploratory Biomarker Objective:	
• To assess predictive, prognostic, and pharmacodynamic exploratory biomarkers in archival and/or fresh tumor tissue and blood and their association with efficacy endpoints including but not limited to pCR	• Relationship between PD-L1 IHC and efficacy endpoints *other than pCR* • Relationship between tumor derived RNA-based immune gene signatures and efficacy endpoints, including but not limited to pCR • Relationship between tumor-based tumor infiltrating lymphocytes and/or CD8 IHC and efficacy endpoints, including but not limited to pCR • Pharmacodynamic changes in cancer-related immune, stroma and tumor immune biology parameters on-treatment by, but not limited to, gene expression and IHC in baseline, on-treatment, and residual disease tumor tissues • Relationship of baseline and on-treatment plasma biomarkers and efficacy endpoints, including but not limited to pCR

ADA = anti-drug antibody; atezolizumab + nab-pac–AC = nab-paclitaxel and atezolizumab followed by doxorubicin and cyclophosphamide with atezolizumab; *DFS = disease-free survival;* EFS = event-free survival; EORTC = European Organisation for Research and Treatment of Cancer; EQ-5D-5L = EuroQoL 5-Dimension, 5-Level; FACT–G = Functional Assessment of Cancer Therapy–General; GHS = global health status; HRQoL = health-related quality of life; IC = tumor-infiltrating immune cell; IHC = immunohistochemistry; ITT = intent to treat; NCI CTCAE = National Cancer Institute Common Terminology Criteria for Adverse Events; OS = overall survival; pCR = pathologic complete response; PD–L1 = programmed death–ligand 1; PK = pharmacokinetics; placebo + nab-pac–AC = neoadjuvant nab-paclitaxel and placebo followed by doxorubicin and cyclophosphamide with placebo; PRO = patient-reported outcome; QLQ-C30 = Quality of Life Questionnaire Core 30; RCB = residual cancer burden.

Study Design

Description of Study

This is a global Phase III, double-blind, randomized, placebo-controlled study designed to evaluate the efficacy and safety of neoadjuvant treatment with atezolizumab + nab-pac-AC, or placebo + nab-pac–AC in patients eligible for surgery with initial clinically assessed T2-4d TNBC.

Female and male patients aged ≥ 18 years with an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1 who have histologically confirmed invasive TNBC with a primary tumor size of > 2 cm are eligible.

Human epidermal growth factor receptor 2 (HER2) and estrogen/progesterone receptor (ER/PR) status will be used to define TNBC. HER2 negativity will be defined by central laboratory assessment using in situ hybridization (ISH) or immunohistochemistry (IHC) assays per American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) criteria, and ER/PgR negativity will be defined by using IHC per ASCO/CAP criteria. Central laboratory assessment will occur prior to randomization. Patients whose tumors are not confirmed to be triple negative will not be eligible. Patients whose tumor tissue is not evaluable for PD-L1 will not be eligible.

Patients who do not initially meet all eligibility criteria, other than TNBC status, may be rescreened once.

This study has an adaptive design consisting of two stages. Stage 1 is the randomization and treatment of approximately 204 patients. At the end of Stage 1, an interim analysis for efficacy will be done by an iDMC. Then, a recommendation will be made to either continue the study unchanged or to expand the patient population (Stage 2; approximately 120 additional patients). If the recommendation is to expand the patient population into Stage 2, the Sponsor will remain blinded to the results of the interim analysis performed at the end of Stage 1.

Patients who have consented and are eligible will be randomized in a 1:1 ratio to either of the following treatment groups:

- **Arm A**: atezolizumab (840 mg) administered via intravenous (IV) infusion Q2W in combination with nab-paclitaxel (125 mg/m^2) administered via IV infusion QW for 12 weeks followed by atezolizumab (840 mg) administered Q2W in combination with doxorubicin (60 mg/m^2) + cyclophosphamide (600 mg/m^2) administered Q2W via IV infusions with filgrastim/pegfilgrastim support for 4 cycles. Patients randomized to the atezolizumab arm will continue to receive unblinded atezolizumab post-surgery at a fixed dose of 1200 mg by IV infusion every 3 weeks (Q3W) for 11 cycles, for a total of approximately 12 months of atezolizumab therapy.
- **Arm B**: placebo administered Q2W via IV infusion in combination with nab-paclitaxel (125 mg/m^2) administered QW via IV infusion for 12 weeks followed by placebo administered Q2W in combination with doxorubicin (60 mg/m^2) + cyclophosphamide (600 mg/m^2) administered Q2W via IV infusions with filgrastim/pegfilgrastim support for 4 cycles. Patients randomized to the placebo arm will be unblinded post-surgery and will continue to be followed.

For patients in Stage 1 of the the study, randomization will be stratified by the following factors:

- *American Joint Committee on Cancer (AJCC) stage at diagnosis (II vs. III; see below for evaluation and classification of lymph nodes)*
- *Tumor PD-L1 status (tumor-infiltrating immune cell [IC] IC0 [<1% PD-L1 expressing IC per tumor area] vs. IC1/2/3 [≥1% PD-L1 expressing IC per tumor area])*

Depending on the iDMC recommendation, Stage 2 could follow either an all-comer or PD-L1 enrichment design. If an all-comer design is recommended, the patients enrolled in Stage 2 will be stratified the same as patients enrolled in Stage 1. If an enrichment design is recommended, only PD-L1–positive patients will be enrolled, and the AJCC stage at diagnosis will be used as the sole randomization stratification factor.

Patients who discontinue neoadjuvant therapy early as a result of disease progression must be discontinued from all study treatment, will be managed as per local practice, and will be followed for survival only. Patients who discontinue prematurely from the study will not be replaced.

Any patient who receives non-protocol therapy prior to surgery will be discontinued from study treatment and will be managed as per local practice; these patients will remain on study for survival follow-up.

The primary efficacy endpoint (pathologic complete response [pCR]; ypT0/is ypN0) will be established via local review following completion of neoadjuvant therapy and surgery. Pathologists who review study specimens must utilize the evaluations and assessments outlined in the Pathology Manual. Investigator/individual patient unblinding will occur after pCR assessment. Surgery should be performed at least 14 days after the last dose of neoadjuvant therapy but no later than 6 weeks after the last infusion. Platelet counts should be checked prior to surgery and should be $\geq$75,000 cells/μL.

Patients with clinically positive axillary nodes by physical examination or by any radiographic imaging at baseline should undergo fine-needle aspiration or core-needle biopsy prior to randomization followed by axillary lymph node dissection (ALND) at time of definitive surgery. The results of the baseline fine-needle aspiration or core-needle biopsy will determine the nodal staging, so that patients with a positive biopsy result should be staged as lymph-node-positive (N1-N3c) whereas patients with a negative or equivocal biopsy result should be staged as lymph-node-negative (N0) regardless of any other clinical measurements.

In patients with clinically or fine-needle biopsy (FNA)/core needle biopsy-proven negative axillary nodes at baseline, axillary surgical management after completion of neoadjuvant therapy should include sentinel lymph node biopsy (SLNB) or ALND. If SLNB is conducted, it is strongly recommended that more than one lymph node (two to three minimum) be removed and all patients with positive macrometastases in sentinel nodes should undergo ALND regardless of the number of positive nodes. All patients with T4 tumors should undergo ALND or current standard of care as described in international or national guidelines.

Postoperative patient management for those in either treatment arm should include radiotherapy as clinically indicated, and management of patients who do not achieve a pCR should follow current standard-of-care guidelines. For those randomized to receive atezolizumab, patients may receive this therapy simultaneously with atezolizumab.

For those randomized to receive atezolizumab, the first dose of postoperative atezolizumab should be administered within 45 days of surgery.

Efficacy, safety, laboratory measurements, patient-reported outcomes (PROs), and pharmacokinetics will be assessed throughout the study. The first 26 patients enrolled (approximately 13 patients in the control arm and approximately 13 patients in the atezolizumab containing arm) will undergo additional cardiac safety monitoring as part of a cardiac safety cohort. Following completion of study treatment and surgery, all patients will continue to be followed for efficacy, safety, and PRO objectives until the end of the study. No interim efficacy analyses for early stopping are planned.

Safety assessments will include the occurrence and severity of adverse events and laboratory abnormalities graded per National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) Version 4.0. Laboratory safety assessments will include the regular monitoring of hematology and blood chemistry. Serum samples will be collected to monitor chemotherapy and atezolizumab pharmacokinetics and to detect the presence of antibodies to atezolizumab. Patient samples, including tumor tissues, as well as plasma and blood, will be collected for exploratory biomarker assessments.

An iDMC will evaluate *the primary efficacy endpoint of pCR in the intent-to-treat (ITT) population (defined as all randomized patients) and in the PD-L1–positive subpopulation based on an interim analysis of efficacy data from the Stage 1 patients (approximately 204 enrolled patients). The iDMC will make a recommendation either to continue the study unchanged or to enroll an additional 120 patients (Stage 2). The decision rules to be applied at the interim analysis will be clearly expressed to the iDMC and documented in the iDMC charter so that the study can be conducted with the Sponsor remaining completely blinded to all results at this stage.*

The iDMC will evaluate safety data and study conduct on a regular basis during the study until the primary analysis of pCR, which is performed at Stage 1 (approximately 204 patients) or Stage 2 (approximately 324 patients), if the iDMC recommends extending the target population. Afterwhich, iDMC review of the study data will be discontinued.

Sponsor affiliates will be excluded from iDMC membership. The iDMC will follow a charter that outlines the iDMC roles and responsibilities.

Number of Patients

Approximately 204 patients will be enrolled in this study during Stage 1 and approximately 120 additional patients will be enrolled if the iDMC recommends to expand to Stage 2. Enrollment will take place at approximately 75 global sites.

Target Population

Inclusion Criteria

Patients must meet the following criteria for study entry:

- Signed Informed Consent Form
- Ability to comply with protocol, in the investigator's judgment
- Women or men aged ≥ 18 years
- ECOG performance status of 0 or 1
- Histologically documented TNBC (negative HER2, ER, and PgR status); HER2 negativity will be defined by central laboratory assessment using ISH or IHC assays per ASCO/CAP criteria and ER/PgR negativity will be defined by central laboratory assessment using IHC per ASCO/CAP criteria. Central laboratory assessment will occur prior to randomization.

  Patients with multifocal tumors (more than one tumor confined to the same quadrant as the primary tumor) are eligible provided all discrete lesions are sampled and centrally confirmed as TNBC.
- Confirmed tumor PD-L1 evaluation as documented through central testing of a representative tumor tissue specimen

  In Stage 2, if the iDMC recommendation is to expand to a PD-L1–positive population, only patients with confirmed tumor PD-L1 positive (IC1/2/3) will be considered eligible.
- Primary breast tumor size of > 2 cm by at least one radiographic or clinical measurement
- Stage at presentation: cT2–cT4, cN0–cN3, cM0
- Patient agreement to undergo appropriate surgical management including axillary lymph node surgery and partial or total mastectomy after completion of neoadjuvant treatment
- Baseline LVEF ≥ 53% measured by echocardiogram (ECHO) or multiple-gated acquisition (MUGA) scans
- Adequate hematologic and end-organ function, as defined by the following laboratory results obtained within 14 days prior to the first study treatment:

  ANC ≥ 1500 cells/μL (without granulocyte colony–stimulating factor [G-CSF] support within 2 weeks prior to Cycle 1, Day 1)

  Lymphocyte count ≥ 500 cells/μL

  Platelet count ≥ 100,000 cells/μL (without transfusion within 2 weeks prior to Cycle 1, Day 1)

  Hemoglobin ≥ 9.0 g/dL

  AST, ALT, and alkaline phosphatase ≤ 2.5 × the upper limit of normal (ULN)

  Serum bilirubin ≤ 1.0 × ULN

  Patients with known Gilbert syndrome who have serum bilirubin level ≤ 3 × ULN may be enrolled.

  For patients not receiving therapeutic anticoagulation: INR or aPTT ≤ 1.5 × ULN within 14 days prior to initiation of study treatment

  For patients receiving therapeutic anticoagulation: stable anticoagulant regimen and stable INR during the 14 days immediately preceding initiation of study treatment

  Creatinine clearance ≥ 30 mL/min (calculated using the Cockcroft-Gault formula)

  Serum albumin ≥ 25 g/L (≥ 2.5 g/dL)

- Representative formalin-fixed, paraffin-embedded (FFPE) tumor specimen in paraffin blocks (preferred) or at least 20 unstained slides, with an associated pathology report documenting ER, PgR, and HER2 negativity

  Tumor tissue should be of good quality based on total and viable tumor content and must be evaluated for PD-L1 expression prior to enrollment. Patients whose tumor tissue is not evaluable for PD-L1 expression are not eligible.

  If multiple tumor specimens are submitted, patients may be eligible if at least one specimen is evaluable for PD-L1. For the purpose of stratification, the PD-L1 score of the patient will be the maximum PD-L1 score among the samples.

  In Stage 2, if the recommendation from the iDMC is to expand to a PD-L1–positive population, no further stratification based on PD-L1 status will be conducted.

  Acceptable samples include core-needle biopsies for deep tumor tissue (minimum three cores) or excisional, incisional, punch, or forceps biopsies for cutaneous, subcutaneous, or mucosal lesions.

  Fine-needle aspiration, brushing, and cell pellet from cytology specimens are not acceptable.
- For women of childbearing potential: agreement to remain abstinent (refrain from heterosexual intercourse) or use contraceptive methods, and agreement to refrain from donating eggs, as defined below:

  Women must remain abstinent or use contraceptive methods that result in a failure rate of $<$ 1% per year during the treatment period and for at least 5 months after the last dose of atezolizumab, or 1 month after the last dose of nab-paclitaxel, or 6 months after the last dose of doxorubicin, or 12 months after the last dose of cyclophosphamide, whichever is later. Women must refrain from donating eggs during this same period.

  A woman is considered to be of childbearing potential if she is postmenarcheal, has not reached a postmenopausal state ($\geq$ 12 continuous months of amenorrhea with no identified cause other than menopause), and has not undergone surgical sterilization (removal of ovaries and/or uterus).

  Examples of contraceptive methods with a failure rate of $<$ 1% per year when used consistently and correctly, include combined (estrogen and progestogen containing) hormonal contraception associated with inhibition of ovulation, progestogen-only hormonal contraception associated with inhibition of ovulation, bilateral tubal occlusion; male sterilization; intrauterine devices; intrauterine hormone-releasing system; and sexual abstinence.

  The reliability of sexual abstinence should be evaluated in relation to the duration of the clinical study and the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not acceptable methods of contraception.
- For men: agreement to remain abstinent (refrain from heterosexual intercourse) or use contraceptive measures and agreement to refrain from donating sperm, as defined below:

  With female partners of childbearing potential, men must remain abstinent or use a condom plus an additional contraceptive method that together result in a failure rate of $<$ 1% per year during the treatment period and for 6 months after the last dose of nab-paclitaxel, cyclophosphamide, or doxorubicin. Men must refrain from donating sperm during this same period.

  With pregnant female partners, men must remain abstinent or use a condom during the treatment period and for 6 months after the last dose of nab-paclitaxel, cyclophosphamide, or doxorubicin to avoid exposing the embryo.

  The reliability of sexual abstinence should be evaluated in relation to the duration of the clinical trial and the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not acceptable methods of contraception.

- Women who are not postmenopausal (≥ 12 months of non–therapy-induced amenorrhea) or have undergone a sterilization procedure must have a negative serum pregnancy test result within 14 days prior to initiation of study drug
- Willingness and ability to comply with scheduled visits, treatment plans, laboratory tests, and other study procedures, including the completion of PRO questionnaires

Exclusion Criteria

Patients who meet any of the following criteria will be excluded from study entry:

- Prior history of invasive breast cancer
- Stage IV (metastatic) breast cancer
- Prior systemic therapy for treatment and prevention of breast cancer
- Previous therapy with anthracyclines or taxanes for any malignancy
- History of ductal carcinoma in situ (DCIS), except for patients treated exclusively with mastectomy > 5 years prior to diagnosis of current breast cancer
- History of pleomorphic lobular carcinoma in situ (LCIS), except for patients surgically managed > 5 years prior to diagnosis of current breast cancer (note that patients with non-pleomorphic LCIS [either untreated or treated with surgery] are allowed)
- Bilateral breast cancer
- Undergone incisional and/or excisional biopsy of primary tumor and/or axillary lymph nodes. Patients who have undergone SLNB at the baseline may be eligible only if the SLNB was free of invasive carcinoma. Any patient with a positive SLN (involved with invasive carcinoma) is ineligible to participate in this study.
- Axillary lymph node dissection prior to initiation of neoadjuvant therapy
- History of other malignancy within 5 years prior to screening, with the exception of those with a negligible risk of metastasis or death (e.g., 5-year OS of > 90%), such as adequately treated carcinoma in situ of the cervix, non-melanoma skin carcinoma, localized prostate cancer, or Stage I uterine cancer
- History of cerebrovascular accident within 12 months prior to randomization
- Cardiopulmonary dysfunction as defined by any of the following prior to randomization:

    History of NCI CTCAE v4.0 Grade ≥ 3 symptomatic congestive heart failure or New York Heart Association (NYHA) criteria Class ≥ II

    Angina pectoris requiring anti-anginal medication, serious cardiac arrhythmia not controlled by adequate medication, severe conduction abnormality, or clinically significant valvular disease

    High-risk uncontrolled arrhythmias (i.e., atrial tachycardia with a heart rate > 100/min at rest, significant ventricular arrhythmia [ventricular tachycardia], or higher-grade atrioventricular [AV]-block [second-degree AV-block Type 2 [Mobitz 2] or third degree AV-block])

    Significant symptoms (Grade ≥ 2) relating to left ventricular dysfunction, cardiac arrhythmia, or cardiac ischemia

    Myocardial infarction within 12 months prior to randomization

    Uncontrolled hypertension (systolic blood pressure > 180 mmHg and/or diastolic blood pressure > 100 mmHg)

    Evidence of transmural infarction on ECG

    Requirement for oxygen therapy
- History of severe allergic, anaphylactic, or other hypersensitivity reactions to chimeric or humanized antibodies or fusion proteins
- Known hypersensitivity to biopharmaceuticals produced in Chinese hamster ovary cells
- Known allergy or hypersensitivity to the components of the atezolizumab formulation

- Known allergy or hypersensitivity to the components of the nab-paclitaxel, cyclophosphamide, or doxorubicin formulations
- Known allergy or hypersensitivity to filgrastim or pegfilgrastim formulations
- Active or history of autoimmune disease or immune deficiency, including but not limited to myasthenia gravis, myositis, autoimmune hepatitis, systemic lupus erythematosus, rheumatoid arthritis, inflammatory bowel disease, antiphospholipid syndrome, Wegener granulomatosis, Sjögren syndrome, Guillain-Barré syndrome, or multiple sclerosis with the following exceptions:

  Patients with a history of autoimmune-related hypothyroidism on a stable dose of thyroid replacement hormone may be eligible for this study.

  Patients with controlled Type I diabetes mellitus on a stable dose of insulin regimen may be eligible for this study.

  Patients with eczema, psoriasis, lichen simplex chronicus, or vitiligo with dermatologic manifestations only (e.g., patients with psoriatic arthritis are excluded) are permitted provided all of following conditions are met:

  Rash must cover $<10\%$ of body surface area

  Disease is well controlled at baseline and requires only low-potency topical corticosteroids

  No occurrence of acute exacerbations of the underlying condition requiring psoralen plus ultraviolet A radiation, methotrexate, retinoids, biologic agents, oral calcineurin inhibitors, or high-potency or oral corticosteroids within the previous 12 months
- History of idiopathic pulmonary fibrosis, organizing pneumonia (e.g., bronchiolitis obliterans), drug-induced pneumonitis, idiopathic pneumonitis, or evidence of active pneumonitis on screening chest computed tomography (CT) scan

  History of radiation pneumonitis in the radiation field (fibrosis) is permitted.
- Positive HIV test at screening
- Active hepatitis B virus (HBV) infection, defined as having a positive hepatitis B surface antigen (HBsAg) test at screening

  Patients with a past or resolved HBV infection, defined as having a negative HBsAg test and a positive total hepatitis B core antibody (HBcAb) test at screening, are eligible for the study if active HBV infection is ruled out on the basis of HBV DNA viral load per local guidelines.
- Active hepatitis C virus (HCV) infection, defined as having a positive HCV antibody test at screening

  Patients who have a positive HCV antibody test are eligible for the study if a polymerase chain reaction (PCR) assay is negative for HCV RNA.
- Active tuberculosis
- Severe infections within 4 weeks prior to initiation of study treatment, including but not limited to hospitalization for complications of infection, bacteremia, or severe pneumonia
- Treatment with therapeutic oral or IV antibiotics within 2 weeks prior to initiation of study treatment

  Patients receiving prophylactic antibiotics (e.g., for prevention of a urinary tract infection or to prevent chronic obstructive pulmonary disease exacerbation) are eligible for the study.
- Major surgical procedure within 4 weeks prior to initiation of study treatment or anticipation of need for a major surgical procedure (other than anticipated breast surgery) during the course of the study
- Prior allogeneic stem cell or solid organ transplantation

- Administration of a live attenuated vaccine within 4 weeks prior to initiation of study treatment or anticipation of need for such a vaccine during the atezolizumab/placebo treatment or within 5 months after the last dose of atezolizumab/placebo

  Patients must agree not to receive live, attenuated influenza vaccine (e.g., FluMist®) within 4 weeks prior to randomization, during treatment or within 5 months following the last dose of atezolizumab (for patients randomized to atezolizumab).
- Any other disease, metabolic dysfunction, physical examination finding, or clinical laboratory finding giving reasonable suspicion of a disease or condition that contraindicates the use of an investigational drug or that may affect the interpretation of the results or render the patient at high risk from treatment complications
- Prior treatment with CD137 agonists or immune checkpoint–blockade therapies, including anti-CD40, anti–CTLA-4, anti–PD-1, and anti–PD-L1 therapeutic antibodies
- Treatment with systemic immunostimulatory agents (including but not limited to interferons, IL-2) within 4 weeks or 5 half-lives of the drug, whichever is longer, prior to initiation of study treatment
- Treatment with systemic immunosuppressive medications (including but not limited to prednisone, cyclophosphamide, azathioprine, methotrexate, thalidomide, and anti–tumor necrosis [anti-TNF] factor agents) within 2 weeks prior to initiation of study treatment or anticipation of need for systemic immunosuppressive medications during the study

  Patients who have received acute, low-dose, systemic immunosuppressant medications (e.g., a one-time dose of dexamethasone for nausea) may be enrolled in the study after discussion with and approval by the Medical Monitor.

  The use of inhaled corticosteroids and mineralocorticoids (e.g., fludrocortisone for adrenal insufficiency) is allowed.
- Pregnant or lactating, or intending to become pregnant during the study

  Women of childbearing potential must have a negative serum pregnancy test result within 14 days prior to initiation of study treatment.

End of Study

The end of the study is defined as the date when the last patient, last visit (LPLV) occurs for evaluation of secondary endpoints, or the date of Sponsor decision to end the study, whichever is earlier.

Length of Study

The total duration of the study *for Stage 1* is expected to be approximately 51 months. *If the iDMC recommends to expand enrollment (Stage 2), the duration of the study may increase to approximately 74 months.*

Investigational Medicinal Products

The investigational medicinal products (IMPs) for this study are atezolizumab, its placebo, and nab-paclitaxel.

Atezolizumab and Placebo

Patients will receive 840 mg atezolizumab or placebo administered by IV infusion Q2W (14 [±3] days) for 20 weeks (i.e., 10 doses) in combination with nab-paclitaxel, doxorubicin, and cyclophosphamide chemotherapy (see below). Postoperatively, patients randomized to the atezolizumab arm will continue to receive unblinded atezolizumab post-surgery at a fixed dose of 1200 mg by IV infusion every 3 weeks (Q3W) for 11 cycles, for a total of approximately 12 months of atezolizumab therapy; patients randomized to the placebo arm will stop receiving placebo.

Nab-Paclitaxel

Nab-paclitaxel will be administered as background treatment along with the non-IMPs doxorubicin, cyclophosphamide, and filgrastim/pegfilgrastim as specified below.

Non-Investigational Medicinal Products

Patients will receive nab-paclitaxel (125 mg/m^2) administered via IV infusion given over 30 minutes weekly for 12 weeks followed by doxorubicin (60 mg/m^2) + cyclophosphamide (600 mg/m^2) administered via IV infusion Q2W with filgrastim/pegfilgrastim support for 4 cycles (i.e., a total of 4 doses of doxorubicin and cyclophosphamide). The dose of cyclophosphamide should be capped at 1200 mg.

Statistical Methods

Primary Analysis

The primary efficacy objective for this study is to evaluate the efficacy of neoadjuvant atezolizumab + nab-pac–AC compared with placebo + nab-pac–AC in patients with T2-4d TNBC, as measured by pCR defined as eradication of tumor from both breast and lymph nodes (ypT0/is ypN0). The primary efficacy endpoint will be established following completion of neoadjuvant therapy and surgery.

The ITT *as well as PD-L1–positive* population*s* will be used for the primary analysis of pCR. In the primary analysis, patients whose pCR assessment was missing will be counted as not achieving a pCR. An estimate of the pCR rate and its 95% CI (Clopper and Pearson 1934) will be calculated for each treatment arm. The difference in pCR rates will be provided with 95% CI, using the normal approximation to the binomial distribution. *For the ITT population, t*he Cochran-Mantel-Haenszel χ^2 test stratified according to tumor PD-L1 status (IC0 vs. IC1/2/3) and clinical stage at presentation (Stage II vs. III) will be used to test pCR rates between treatment groups at a two-sided significance level of 5%. *For the PD-L1–positive population, similar* test will *be used with stratification only for clinical stage at presentation. An unstratified χ^2 versions of these tests will* also be provided as a sensitivity analysis.

Determination of Sample Size

The study will first randomize approximately N_1=204 patients in Stage 1 (1:1 randomization ratio). Based on information from these patients and a pre-specified adaptive rule, the decision will be made regarding whether or not to randomize approximately N_2=120 patients in Stage 2 (1:1 randomization ratio).

The pre-defined one-sided type I error for the interim analysis at Stage 1 is α_1=0.0125 (i.e., 50% of the total type I error). The co-primary endpoints in the ITT and PD-L1-positive populations at Stage 1 will be tested using a closed testing procedure using Simes' test for the intersection hypothesis. Test statistics in ITT and PD-L1-positive populations are always positively correlated, hence type I error rate control of the Simes test can be guaranteed in general.

Importantly, the p-value p_1 based on Stage 1 data only and the combination p-value p_{comb} have the same joint distribution under the null hypothesis as the p-values from a group-sequential test with two stages at information times $t_1 = w_1$ and $t_2 = 1$. Thus standard statistical software for group sequential designs can be used for the determination of critical values for the adaptive design. As the study design uses a critical value of α_1=0.0125 (i.e., 50% of the total type I error) for Stage 1, this implies that a critical value of α_{comb}=0.0184 can be applied to the combination p-values.

Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2)

	Screening [a]	Treatment			Completion of Study Therapy/Early Term. Visit [c]	Survival Follow-Up [d]
		Neoadjuvant Treatment (28-Day Cycles) (Cycles 1–5; Weeks 1–20)	Pre-Surgery Visit/ Surgery [b]	Arm A: Adjuvant Treatment Arm B: Monitoring (21-Day Cycles) (Cycles 6–16)		
	Days –28 to –1	Day 1 (±3 days)		Day 1 (±3 days)	≤30 Days after Last Dose (Arm A) or Last Monitoring Visit (Arm B)	
Informed consent	x [e]					
Baseline tumor tissue sample submission for HER2 and ER/PgR determination and exploratory biomarkers (mandatory)	x [f]					
Demographic data	x					
Medical history and baseline conditions	x					
Disease status assessments [g]	x	x	x [g]	x [g]	x [g]	x [g]
Tumor Staging [h]	x					
Ultrasound [g]	x		x			
EORTC QLQ-C30, EQ-5D-5L [i]		x		x	x	x [j]
FACT-G, Single Item GP5 [i]		x [k]		x	x	x [j]
Vital signs [l]	x	On each infusion day	x	x [m]	x	
Weight	x	x		x [m]		
Height	x					

Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2) (cont.)

Complete physical examination [n]	x					
Limited physical examination [o]		x [g]	x	x [m]	x	
ECOG Performance Status [p]	x	x	x	x	x	
ECG (12-lead) [q]	x	As clinically indicated				
ECHO or MUGA scan [r]	x	As described in footnote "r"			x	x
Spirometry (FVC, FEV_1, FEV_1:FVC, FEF_{25-75})	x	As clinically indicated				
Hematology [s]	x [t]	On each infusion day	x	x [m]	x	
Chemistry [u]	x [t]	On each infusion day	x	x [m]	x	
Pregnancy test [v]	x [t]	x [v]		x [v]	x [v]	x [v]
Coagulation (INR, aPTT)	x [t]		x		x	
TSH, free T3 (or total T3 [w]), free T4	x [t]	x [w]			x	
Viral serology [x]	x [t]					
Urinalysis [y]	x [t]	As clinically indicated				
Serum PK sample for atezolizumab		See Appendix 2 for detailed schedule				
Serum ADA sample for atezolizumab		See Appendix 2 for detailed schedule				
Blood and plasma samples for biomarkers		See Appendix 2 for detailed schedule				
Blood sample for RBR (optional) [z]		x				

Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2) (cont.)

Tumor tissue (fresh sample preferred) at screening, on-study, and time of disease recurrence [aa]	See Appendix 3 for detailed schedule					
Radiographic assessments (e.g., CT scan, MRI, PET scan)	As clinically indicated					
Bilateral mammogram	x [bb]				x	x [cc]
Concomitant medications	x [dd]	x		x [ee]	x [ee]	x [g, ff]
Adverse events [gg]	x [gg]	x [gg]		x	x	x [gg]
Study treatment administration [hh]		x		x [ii, m]		
Survival follow-up and anti-cancer treatment						x [d, jj]

AC = doxorubicin + cyclophosphamide; ADA = anti-drug antibody; CT = computed tomography; ECHO = echocardiogram; ECOG = Eastern Cooperative Oncology Group; eCRF = electronic Case Report Form; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30; EQ-5D-5L = EuroQoL 5-Dimension, 5-Level; ER = estrogen receptor; FACT-G = Functional Assessment of Cancer Therapy-General; FCV = Forced Vital Capacity; FEV_1 = Forced Expiratory Volume 1; FEF_{25-75} = Forced Expiratory Flow 25%–75%; FFPE = formalin-fixed, paraffin-embedded; HBcAb = hepatitis B core antibody; HBsAb = hepatitis B surface antibody; HBsAg = hepatitis B surface antigen; HBV = hepatitis B virus; HCV = hepatitis C virus; HER2 = human epidermal growth factor receptor 2; MRI = magnetic resonance imaging; MUGA = multiple-gated acquisition; PD-L1 = programmed death–ligand-1; PET = positron emission tomography; PgR = progesterone receptor; PK = pharmacokinetic; PRO = patient-reported outcome; RBR = Research Biosample Repository; T3 = triiodothyronine; T4 = thyroxine; Term. = termination; TSH = thyroid-stimulating hormone.

Notes: On treatment days, all assessments should be performed prior to dosing, unless otherwise specified.

Assessments shaded in gray should be performed as scheduled, but the associated data do not need to be recorded on the eCRF (except in the case of an adverse event).

If treatment is withheld (e.g., by adverse events or delays in initiating post-surgical therapy), the schedule of assessments should be held accordingly (e.g., Day 1 of Week 21 = first administration of atezolizumab post-surgery or Day 1 of Cycle 6).

[b] Results of standard-of-care tests or examinations performed prior to obtaining informed consent and within 30 days prior to Day 1 may be used; such tests do not need to be repeated for screening.

[b] Pre-surgical visit and associated assessments should occur within 14 days of surgery. Surgery should be conducted no earlier than 14 days and no later than 6 weeks after last dose of neoadjuvant therapy. Platelet counts should be checked prior to surgery and should be ≥ 75,000 cells/μL.

Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2) (cont.)

[c] Patients who discontinue study treatment will return to the clinic for a treatment discontinuation visit not more than 30 days after the last dose of study treatment.

[d] The survival follow-up period begins from the date of treatment completion/early termination visit, *and has a* duration of up to *approximately* 51 months from the date of randomization of the first patient *in Stage 1). This may increase to approximately 74 months if the study includes Stage 2 patients.* Visit windows are ±28 days for quarterly and semiannual assessments.

[e] Informed consent must be documented before any study-specific screening procedure is performed and may be obtained more than 28 days before initiation of study treatment.

[f] After signing of the Informed Consent Form, retrieval and submission of tumor tissue sample can occur outside the 28-day screening period. Tumor tissue should be of good quality based on total and viable tumor content (sites will be informed if the quality of the submitted specimen is inadequate to determine tumor PD-L1 status). An FFPE block or at least 20 unstained slides should be provided. Fine-needle aspiration, brushing, cell pellets from pleural effusion, and lavage samples are not acceptable. For core-needle biopsy specimens, at least three cores should be submitted for evaluation. Retrieval of tumor sample can occur outside the 28-day screening period.

[g] Assessment of primary tumor and regional lymph nodes should be done by physical examination at baseline and prior to administration of each cycle of study treatment during neoadjuvant therapy. Additionally, standard breast imaging modality should include an ultrasound of breast and axilla disease. (Physical examination and ultrasound is mandatory within 28 days prior to randomization and within 14 days pre-surgery.) At baseline, if there is evidence of suspicious axillary lymph nodes, then fine-needle aspiration is required. Ultrasound-detected axillary lymph nodes suspicious of malignancy include those with cortical thickness >2 mm. Disease status based on all available clinical assessments should be documented every 3 months during adjuvant study treatment and follow-up up to 3 years after surgery and every 6 months thereafter. In addition, liver function tests, bone scans, chest X-ray/diagnostic CT scan, liver imaging, and/or other radiographic modalities may be considered when clinically indicated to exclude metastatic disease; these assessments should be performed within a timeline as per current local standard of practice. Whenever possible, disease recurrence should be confirmed pathologically. If disease recurrence is diagnosed at any time during the study, patients will discontinue scheduled study assessments and will be followed for survival, anti-cancer medications, and new relapse events.

[h] See Section 4.5.6 in protocol.

[i] All PRO assessments (EORTC QLQ-C30, followed by the FACT-G single item GP5, and then the EQ-5D-5L questionnaires) must be completed by the patient at the investigational site at the start of the clinic visit before discussion of the patient's health state, lab results or health record, before administration of study treatment, and/or prior to the performance of any other study assessments that could bias the patient's responses. Interview assessment by a member of the clinical staff will be allowed if the patient is not able to complete the measure on her or his own. Study personnel should review all questionnaires for completeness before the patient leaves the investigational site.

The EORTC QLQ-C30 and EQ-5D-5L questionnaires will be completed by patients at baseline (Cycle 1, Day 1) and on Day 1 of every cycle thereafter. The FACT-G, single item GP5 will not be completed by patients at the baseline visit (Cycle 1, Day 1).

Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2) (cont.)

[j] Patients who discontinue study treatment for any reason will continue to complete the EORTC QLQ-C30, FACT-G single item GP5, and EQ-5D-5L questionnaires in-clinic during the survival follow-up period at the following timepoints: every 3 months (±28 days) for Year 1, every 6 months (± 28 days) for years 2-3, and then annually (±28 days) thereafter.

[k] While on study treatment, all patients will complete the FACT-G, single item GP5 beginning on Cycle 2, Day 1 and at Day 1 of every cycle thereafter.

[l] Includes respiratory rate, pulse rate, systolic and diastolic blood pressure while the patient is in a seated position, and temperature. Record abnormalities observed at baseline on the General Medical History and Baseline Conditions eCRF. At subsequent visits, record new or worsened clinically significant abnormalities on the Adverse Event eCRF. For the first infusion, vital signs should be measured within 60 minutes prior to the infusion and, if clinically indicated, every 15 (±5) minutes during and 30 (±10) minutes after the infusion. For subsequent infusions, vital signs should be measured within 60 minutes prior to the infusion and, if clinically indicated or if symptoms occurred during the previous infusion, during and 30 (±10) minutes after the infusion.

[m] For patients in Arm A only.

[n] Includes evaluation of the head, eyes, ears, nose, throat, and the cardiovascular, dermatological, musculoskeletal, respiratory, gastrointestinal, and neurological systems. Record abnormalities observed at baseline on the General Medical History and Baseline Conditions eCRF. At subsequent visits, record new or worsened clinically significant abnormalities on the Adverse Event eCRF.

[o] Perform a limited, symptom-directed examination at specified timepoints and as clinically indicated at other timepoints. Record new or worsened clinically significant abnormalities on the Adverse Event eCRF.

[p] See Appendix 10.

[q] ECG recordings will be obtained during screening and as clinically indicated. Patients should be resting in a supine position for at least 10 minutes prior to ECG recording.

[r] Cardiac monitoring (ECHO or MUGA scan) will be performed on all patients enrolled in the study. ECHO is the preferred method. The same method used for a given patient *at* screening should be used throughout the study. ECHO or MUGA scan should be obtained at **baseline** *and* **after the second dose of AC** (*which would correspond to* Week 16 ±1 week *if no study treatment interruptions or discontinuations have occurred) during neoadjuvant study treatment. During the adjuvant (Arm A) or monitoring (Arm B) phase of the study, ECHO or MUGA scan should be obtainedat* ***Cycle 6, Day 1*** *(which would approximately correspond to* Week *21* ±1 week *if no study treatment interruptions or discontinuations have occurred), and* ***every 3 months afterwards*** *while on the Adjuvant/Monitoring Phase (which would correspond to Cycle 10, Day 1 and Cycle 14, Day 1, or approximately* Week *33* ±1 week*,* and Week *45* ±1 week *respectively, if no treatment interruptions or discontinuations have occurred*). ECHO or MUGA scan should be obtained at the **early termination visit** if not performed within the previous 6 weeks. During the survival follow-up period, ECHO or MUGA scan should be obtained **annually until the end of study**. For additional cardiac screening tests for patients in the cardiac safety cohort, please see Appendix 4.

[s] Hematology includes WBC count, RBC count, hemoglobin, hematocrit, platelet count, and differential count (neutrophils, eosinophils, basophils, monocytes, lymphocytes, other cells).

[t] Screening laboratory test results must be obtained within 14 days prior to initiation of study treatment.

Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2) (cont.)

[u] Chemistry panel (serum or plasma) includes sodium, potassium, chloride, bicarbonate or total CO2, glucose, BUN or urea, creatinine, total protein, albumin, calcium, total bilirubin, alkaline phosphatase, ALT, AST, and LDH. Magnesium and phosphorus should be included at screening and as clinically indicated during study treatment.

[v] All women of childbearing potential will have a serum pregnancy test at screening. Urine pregnancy tests will be performed at the following specified subsequent visits for women of child-bearing potential (including premenopausal women who have had tubal ligation) and women not meeting the definition of postmenopausal: Day 1 of Cycles 1 – 5; Day 1 of Cycles 6, 8, 10, 12, 14, and 16; at treatment discontinuation (unless administered within 30 days); and at 3 months and 6 months after treatment discontinuation. For all other women, documentation must be present in medical history confirming that the patient is not of childbearing potential. If a urine pregnancy test is positive, it must be confirmed by a serum pregnancy test.

[w] TSH, free T3 (or total T3 for sites where free T3 is not performed), and free T4 will be assessed on Day 1 of Cycle 1 and every fourth cycle thereafter.

[x] At screening, patients will be tested for HIV, HBsAg, HBsAb, total HBcAb, and HCV antibody. If a patient has a negative HBsAg test and a positive total HBcAb test at screening, an HBV DNA test should be performed to rule out active HBV infection prior to initiation of study treatment. If a patient has a positive HCV antibody test at screening, an HCV RNA test should be performed to rule out active HCV infection prior to initiation of study treatment.

[y] Includes pH, specific gravity, glucose, protein, ketones, and blood; dipstick permitted.

[z] Not applicable for a site that has not been granted approval for RBR sampling. Performed only for patients at participating sites who have provided written informed consent to participate. Whole blood for DNA isolation will be collected from patients who have consented to optional RBR sampling at Week 1, Day 1. If, however, the RBR genetic blood sample is not collected during the scheduled visit, it may be collected as soon as possible (after randomization) during the conduct of the clinical study.

[aa] Tumor tissue should be of good quality based on total and viable tumor content (sites will be informed if the quality of the submitted specimen is inadequate to determine tumor PD-L1 status). For tissue sample provided at screening, an FFPE block or at least 20 unstained slides should be provided. Retrieval of tumor screening sample can occur outside the 28-day screening period. Fine-needle aspiration, brushing, cell pellets from pleural effusion, and lavage samples are not acceptable. For core-needle biopsy specimens, at least three cores should be submitted for screening and on-study evaluation. At least two cores should be submitted for disease recurrence specimens. See Section 4.5.8 and Appendix 3 for specific tissue sample requirements for each time point.

[bb] The unaffected breast should have been imaged within 60 days prior to randomization. The affected breast should be imaged within 28 days prior to randomization.

[cc] Mammograms of any remaining breast tissue should be performed at least annually.

[dd] Includes any medication (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by a patient in addition to protocol-mandated study treatment from 30 days prior to initiation of study treatment (for the purposes of screening) until the treatment discontinuation visit. Record all prior anti-cancer therapies.

Appendix 2 Schedule of Activities (for Both Stage 1 and Stage 2) (cont.)

[ee] To be collected for both study arms. For patients in Arm B (Monitoring), only medications given for reportable adverse events as per protocol (see Section 5.3.1) as well as new anti-cancer treatments should be collected.

[ff] Medications related to the treatment of serious adverse events are to be reported during the follow-up period, as well as new anti-cancer treatments.

[gg] After informed consent has been obtained but prior to initiation of study treatment, only serious adverse events caused by a protocol-mandated intervention should be reported. After initiation of study treatment, all adverse events will be reported until 30 days after the last dose of study treatment or until initiation of new anti-cancer therapy, whichever occurs first, and serious adverse events and adverse events of special interest will continue to be reported until 90 days after the last dose of study treatment or until initiation of new anti-cancer therapy, whichever occurs first. After this period, all deaths, regardless of cause, should be reported. In addition, the Sponsor should be notified if the investigator becomes aware of any serious adverse event that is believed to be related to prior exposure to study treatment (see Section 5.6). The investigator should follow each adverse event until the event has resolved to baseline grade or better, the event is assessed as stable by the investigator, the patient is lost to follow-up, or the patient withdraws consent. Every effort should be made to follow all serious adverse events considered to be related to study treatment or trial-related procedures until a final outcome can be reported.

[hh] The initial dose of atezolizumab will be delivered over 60 (±15) minutes. Subsequent infusions will be delivered over 30 (±10) minutes if the previous infusion was tolerated without infusion-associated adverse events, or 60 (±15) minutes if the patient experienced an infusion-associated adverse event with the previous infusion.

[ii] Study drug administration during the maintenance phase for the atezolizumab-containing arm only.

jj After treatment discontinuation, information on survival follow-up and new anti-cancer therapy (including targeted therapy and immunotherapy) will be collected via telephone calls, patient medical records, and/or clinic visits (unless the patient withdraws consent or the Sponsor terminates the study). If a 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, where allowable by local regulation.

Appendix 3 Adaptive Multiple Test Design

The performance of the proposed design depends on several factors including the underlying proportion of (pCR) responders in each treatment arms for the subgroup S (PD-L1 IC1/2/3) and the complement group C (PD-L1 IC0). Simulations were carried out to evaluate the operation characteristics of the current design under a variety of conditions as follows.

Three simulation scenarios were considered with numerical details provided in Table 2.

- **Subgroup S with very strong credentials**

  Subgroup S has very strong credentials when convincing evidence indicates that the benefits of the treatment are limited to the subgroup S.

- **Subgroup S with strong credentials**

  Subgroup S has strong credentials when evidence for the predictive ability of the subgroup is convincing enough to assume that the treatment is more likely to be effective (and is probably more effective) in the subgroup S than in the complement subgroup C, but the evidence is not sufficiently compelling to rule out a clinically meaningful effect in complement subgroup C.

- **Subgroup S with weak credentials**

  If the subgroup S has weak credentials when convincing evidence for predictive value of the subgroup S is lacking and the treatment is expected to be broadly effective.

Table 1 Definition of Weak, Strong, and Very Strong Credentials on Subgroup S

Subgroup S Credentials	pCR proportion under treatment in Full population F (π_F)[a]	pCR Proportion under treatment in Subgroup S (π_S)	pCR Proportion under treatment in complement Subgroup C (π_C)	pCR Proportion response under Control ($\pi_{control}$)
Weak	0.68 (Δ_F=0.2)	0.68 (Δ_S=0.20)	0.68 (Δ_C=0.20)	0.48
Strong	0.6376 (Δ_F=0.1576)	0.68 (Δ_S=0.20)	0.60 (Δ_C=0.12)	0.48
Very Strong	0.5952 (Δ_F=0.1152)	0.68 (Δ_S=0.20)	0.52 (Δ_C=0.04)	0.48

[a]Assuming a prevalence of 0.47 of Subgroup S in the full population.

Appendix 3 Adaptive Multiple Test Design (cont.)

Table 2 Relative Frequencies of Decisions at Stage 1under Weak, Strong, and Very Strong Credentials on Subgroup S

Subgroup S Credentials	Δ_S	Δ_C	Relative Frequency[a,b,c]				
			Stop for efficacy	Stop for lack of efficacy	Continue only S	Continue only F	Continue S and F
Weak	0.20	0.20	0.64	0.04	0.08	0.14	0.10
Strong	0.20	0.12	0.46	0.10	0.22	0.11	0.11
Very Strong	0.20	0.04	0.33	0.17	0.38	0.05	0.06

[a]Assuming a prevalence of 0.47 of Subgroup S in the full population.
[b]Assuming probability of drop-out of 0.05 for each treatment arm.
[c]Frequency based on 100000 simulations.

Table 3 Rejection Probabilities under Weak, Strong, and Very Strong Credentials on Subgroup S

Subgroup S Credentials	Δ_S	Δ_C	Probability[a,b,c]		
			Reject $H_0^{\{F\}}$	Reject $H_0^{\{S\}}$	Reject $H_0^{\{F\}}$ or $H_0^{\{S\}}$ i.e. Overall Power
Weak	0.20	0.20	0.80	0.49	0.88
Strong	0.20	0.12	0.55	0.57	0.76
Very Strong	0.20	0.04	0.28	0.62	0.67

[a]Assuming a prevalence of 0.47 of Subgroup S in the full population.
[b]Assuming probability of drop-out of 0.05 for each treatment arm.
[c]Probability based on 100000 simulations.

Table 4 Conditional Rejection Probabilities if Stage 2 Activated under Weak, Strong, and Very Strong Credentials on Subgroup S

Subgroup S Credentials	Δ_S	Δ_C	Conditional Probability[a,b,c,d] (if stage 2 activated)		
			Reject $H_0^{\{F\}}$	Reject $H_0^{\{S\}}$	Reject $H_0^{\{F\}}$ or $H_0^{\{S\}}$ i.e. Overall "Go" Power
Weak	0.20	0.20	0.54	0.36	0.74
Strong	0.20	0.12	0.29	0.51	0.68
Very Strong	0.20	0.04	0.10	0.63	0.68

[a]Assuming a prevalence of 0.47 of Subgroup S in the full population.
[b]Assuming probability of drop-out of 0.05 for each treatment arm.
[c]Probability based on 100000 simulations.
[d]Conditional probability of event happening once the trial continues to enroll patients in stage 2 ("go").