STATISTICAL ANALYSIS PLAN APPROVAL PAGE

To compare the effectiveness and safety of different adjuvant chemotherapy regimens for high-risk triple-negative breast cancer patients predicted by mRNA-lncRNA model (an open-label, multi-center, randomized controlled study)

Study institute: Fudan University Shanghai Cancer Center

Major study investigator: Zhi-Ming Shao, MD. PhD.

ClinicalTrials.gov Identifier: NCT02641847

Version: 1.1

TABLE OF CONTENTS

1. Introduction

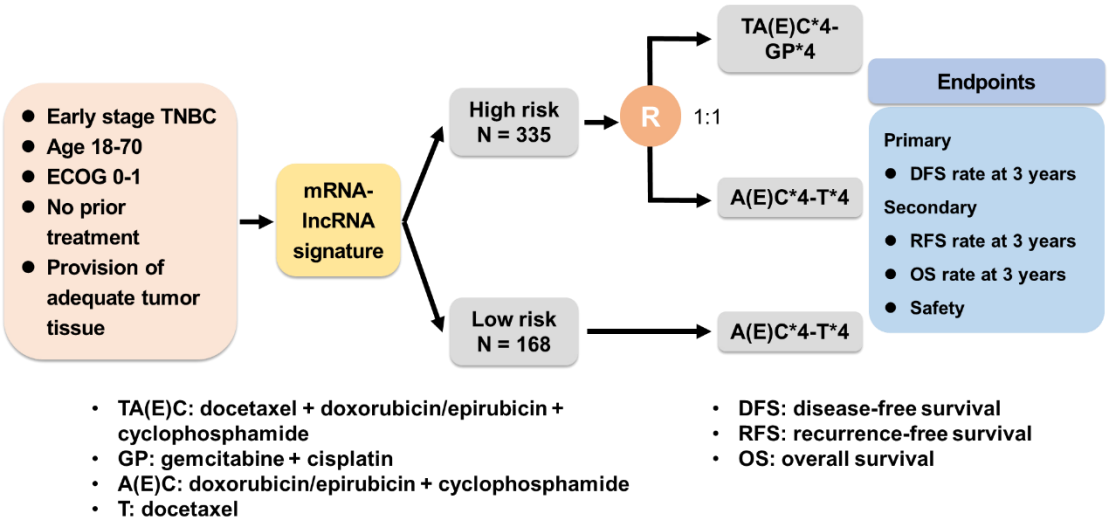
This SAP is based on the following documents: Study Protocol, Version 1.2 (16-Aug-2023).

This statistical analysis plan (SAP) details the methods of statistical analysis that will be used in the analysis of this study. This SAP does not include tables, listings, and figures (TLFs). TLFs will be shown in another article. Changes to the methods of statistical analysis will require amendment to the SAP.

1.1 Study Design

This study is a prospective, open-label, multicenter, randomized, phase III clinical trial. It is planned to screen patients with newly diagnosed, operable, unilateral invasive triple-negative breast cancer. Based on the integrated mRNA-lncRNA signature score, patients were assigned to one of three treatment groups. Women with a low-risk score were assigned to receive doxorubicin/epirubicin and cyclophosphamide followed by docetaxel [A(E)C-T]. Women with a high-risk score were randomly assigned in a 1:1 ratio to receive either A(E)C-T or docetaxel, doxorubicin/epirubicin, and cyclophosphamide followed by gemcitabine and cisplatin [TA(E)C-GP].

The study protocol provides a comprehensive depiction of the flow chart, elucidating the sequential steps and procedures involved. The study design diagram is presented as followed.



1.2 Sample Size

The primary endpoint was DFS rate at 3 years between patients in the high-risk

cohort receiving TA(E)C-GP versus A(E)C-T treatment. We expected that the 3-year DFS rate would rise from 79% to 90%, with a corresponding hazard ratio (HR) of 0.51. Employing a two-sided significance level of 0.05, a power of 80%, a 3-year enrollment period and 2-year follow-up period, and accounting for a 9% lost to follow-up rate, a total of 335 high-risk TNBC patients with 50 DFS events are required (equally distributed between the control and experimental groups). To compare the efficacy of standard treatment between high and low-risk groups receiving standard treatment, with a 1:1 ratio and a prospective observational trial design, it is planned to include 168 low-risk patients receiving standard treatment. Hence, the overall required sample size for this study is 503.

The primary analysis for high-risk patients will be conducted after collecting 50 DFS events between arm A and arm B or at least 80 months after first patient enrollment.

2. Study Endpoints

2.1 Efficacy Endpoints

2.1.1 Disease-Free Survival (DFS)

Refers to the duration from the moment of random assignment to the earliest occurrence of disease recurrence (either locoregional or distant), invasive contralateral cancer, a second primary cancer, or death from any cause.

2.1.2 Recurrence-Free Survival (RFS)

Refers to the duration from the moment of random assignment to the date of diagnosis of invasive breast cancer recurrence or death.

2.1.3 Overall Survival (OS)

Refers to the period from the moment of random assignment to the date of death for any reason.

2.2 Safety Endpoints

Refers to the incidence and grade of adverse events (AE) and severe adverse events (SAE); AE is assessed according to the National Cancer Institute Common

Toxicity Grading Criteria (NCI-CTCAE) version 4.0.

AE refers to any untoward medical occurrence in a study subject administered an investigational product which does not necessarily have a causal relationship with the treatment.

Adverse events that occurred during this period between the start of randomization and the last visit, whether related to the study drug or not, should be entered into the case report form. In addition, any adverse events that occur after the adverse event reporting period, which the investigators estimate may be related to the study drug, should also be reported as adverse events.

Adverse events are defined as serious adverse events when they meet one or more of the following criteria:

- Death;
- Life-threatening;
- Events that require hospitalization or prolonged hospitalization;
- Events that can cause permanent or severe disability/dysfunction/disruption of work ability;
- Events that can cause congenital anomaly or birth defects, Other important medical events.

Treating investigators must report any SAEs occurring during the course of the study within 24 hours of discovery of the event. This includes events both related and unrelated to the product.

3.Statistical Analysis

After the completion of the trial scheme and case report form, the analysis plan shall be formulated, and necessary modifications shall be made during the trial process. The plan shall be completed before data locking, and the statistical analysis report will be provided after data analysis.

3.1 Analysis Data Set

1) Intention-to-treat population (ITT) :All randomized patients were included in ITT.

Efficacy analyses were done on an ITT population in this study.

2) Safety Analysis Set (SS): It refers to subjects who have taken the study drug at least once;

3.2 Method of Statistics

1) Whether subject to normal distribution: if not, modify statistical methods or perform data transformation.
2) Whether there is outlier: make statistical and professional analysis and decide whether to include or not.
3) Whether there is missing value: when a primary therapeutic index of individual subject fails to be measured, the last observation data should be transferred.
4) The percentage of dropouts should not exceed 20%, otherwise it requires analysis and explanation.
5) Descriptive statistical analysis: e.g., mean, standard deviation, maximum, minimum, confidence interval, rate, etc.

3.3 Method of Analysis

1) Measurement data: Use t test, paired t test, rank sum test, paired rank sum test, etc.
2) Enumeration data: Use Fisher's exact test, etc., rank sum test is adopted to ranked data.
3) Analysis of efficacy indicators: CMH test, chi square test or logistic regression will be used for enumeration data. Analysis of variance or rank sum test will be used for measurement data according to the feature of the data. Kaplan Meier method or Cox regression will be used for survival data.
4) ITT analysis: ITT analysis will be conducted for the main efficacy indicators.

3.4 Statistic Expression

1) The report is mainly represented by tables with title, annotation and number of cases, which are self-evident.

2) Two-sided P values will be calculated for all statistical tests. A value of $P < 0.05$ is considered significant.

4.References

1. Eiermann, W. et al. Phase III study of doxorubicin/cyclophosphamide with concomitant versus sequential docetaxel as adjuvant treatment in patients with human epidermal growth fA(E)Ctor receptor 2-normal, node-positive breast cancer: BCIRG-005 trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 29, 3877-3884, doi:10.1200/jco.2010.28.5437 (2011).
2. Martin, M. et al. Adjuvant docetaxel for node-positive breast cancer. The New England journal of medicine 352, 2302-2313, doi:10.1056/NEJMoa043681 (2005).
3. Martin, M. et al. Adjuvant docetaxel for high-risk, node-negative breast cancer. The New England journal of medicine 363, 2200-2210, doi:10.1056/NEJMoa0910320 (2010).
4. Mamounas, E. P. et al. PA(E)Clitaxel after doxorubicin plus cyclophosphamide as adjuvant chemotherapy for node-positive breast cancer: results from NSABP B-28. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 23, 3686-3696, doi:10.1200/jco.2005.10.517 (2005).