

Title: Phase II Clinical Trial Evaluating Neoadjuvant Endocrine Therapy Combined With Chemotherapy in the Treatment of ER-positive, HER2-negative Stage I – III Breast Cancer

Date:2024.02.22

1 Methods

1.1 Study design and participants

This investigator-initiated, prospective, single-centre, non-randomised trial took place at the Second Affiliated Hospital of Anhui Medical University. We enrolled patients from 20 March 2024 to 13 June 2025 and followed them to the prespecified data cut-off of 15 July 2026. The institutional ethics committee approved the protocol (approval no. YX2023-203(F1)); all participants gave written informed consent before enrolment. Eligible patients received the prespecified treatment. We recorded clinical, laboratory, imaging, treatment, and pathological data prospectively on standardised case report forms, then checked the records against the hospital's electronic medical record, laboratory, and pathology systems.

Eligible patients were aged ≥ 18 years and had histologically confirmed ER-positive, human epidermal growth factor receptor 2 (HER2)-negative breast cancer with a Ki67 labelling index $>20\%$. ER positivity required nuclear staining in $\geq 1\%$ of tumour cells by immunohistochemistry (IHC). HER2-negative disease was defined as an IHC score of 0 or 1+, or a score of 2+ with negative fluorescence in situ hybridization (FISH). Patients had stage I-III operable or locally advanced breast cancer according to the eighth edition of the American Joint Committee on Cancer (AJCC) staging system. Those with clinically node-negative disease (cN0) also needed high-risk features, including a primary tumour classified as cT2 or higher and/or a markedly elevated Ki67 labelling index. Other requirements were measurable disease according to the Response Evaluation Criteria in Solid Tumours (RECIST) version 1.1, an Eastern Cooperative Oncology Group performance status (ECOG PS) of 0-2, and adequate bone marrow and major organ function. Previous systemic chemotherapy or endocrine therapy for breast cancer was not permitted.

We excluded patients who were pregnant or lactating, had a severe or uncontrolled infection, or had clinically significant cardiovascular disease. Cardiovascular exclusions included New York Heart Association class III or IV heart failure, unstable angina, and myocardial infarction within 6 months before enrolment. A left ventricular ejection fraction <50%, previous organ transplantation requiring immunosuppression, and known human immunodeficiency virus infection were also exclusion criteria. The investigator could exclude another malignancy if it could interfere with study participation or outcome assessment. Patients were also ineligible if psychiatric, cognitive, or other medical conditions could compromise adherence to the protocol.

1.2 Procedures

Patients received standard anthracycline- and taxane-based chemotherapy with letrozole. The chemotherapy regimen was either six cycles of docetaxel, epirubicin, and cyclophosphamide (TEC) every 3 weeks or four cycles of epirubicin plus cyclophosphamide followed by four cycles of a taxane (EC→T). The investigator could replace docetaxel or paclitaxel with nanoparticle albumin-bound paclitaxel (nab-paclitaxel), depending on the patient's clinical features and treatment tolerance.

Patients took letrozole 2.5 mg orally once daily throughout neoadjuvant systemic therapy. Premenopausal patients also received ovarian function suppression with leuporelin acetate (3.75 mg) subcutaneously every 28 days, from the start of neoadjuvant treatment until definitive surgery.

We assessed radiological response with contrast-enhanced breast magnetic resonance imaging (MRI) at baseline, after four cycles of therapy, and after neoadjuvant treatment but before surgery. Best radiological response was classified according to the Response Evaluation Criteria in Solid Tumours (RECIST) version 1.1. Best percentage tumour reduction was the largest reduction in target-lesion size from baseline across serial MRI examinations. Positive values denoted shrinkage; negative values denoted growth.

We scheduled definitive surgery about 3–6 weeks after neoadjuvant therapy, once treatment-related toxicities had resolved adequately and the patient was medically fit.

A multidisciplinary team chose the type and extent of surgery after considering pretreatment disease characteristics, radiological response, patient preference, and surgical feasibility.

Dedicated breast pathologists reviewed the pretreatment core-needle biopsies and postoperative surgical specimens according to institutional practice. They graded pathological response with the Miller-Payne system and recorded residual invasive tumour size and postoperative pathological stage (ypT and ypN). They assessed ER, progesterone receptor (PR), HER2, and Ki67 by immunohistochemistry using standardised institutional procedures.

The treating team selected postoperative radiotherapy, endocrine therapy, and other systemic treatment according to pathological findings, recurrence risk, and current institutional and national guidelines. Follow-up then continued according to institutional practice until the predefined data cut-off of 15 July 2026.

1.3 Endpoints

The primary endpoint was objective response rate (ORR), defined as the proportion of patients whose best radiological response during neoadjuvant treatment was complete response (CR) or partial response (PR) according to RECIST version 1.1.

Secondary endpoints were pCR, defined as no residual invasive disease in the breast or axillary lymph nodes (ypT0/is ypN0); pathological response, defined as Miller-Payne grade 3-5; paired change in Ki67 between biopsy and surgery; disease-free survival (DFS); and treatment-related adverse events (TRAEs).

Exploratory endpoints included changes in serum tumour markers, tumour and nodal downstaging, and radiological-pathological concordance. We also examined how well radiological complete response predicted breast pathological complete response. An exploratory analysis compared DFS between Miller-Payne grades ≥ 3 and ≤ 2 .

DFS was measured from definitive breast surgery to the first qualifying event: local, regional, or distant invasive recurrence; contralateral invasive breast cancer; a second primary malignancy; or death from any cause. Patients without an event were censored at their most recent disease assessment.

We monitored safety throughout neoadjuvant treatment. Treatment-related adverse events were recorded by incidence, type, and severity, then graded with the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE), version 5.0.

1.4 Statistical analysis

The sample-size calculation used the primary endpoint of objective response rate. We assumed rates of 65% with conventional neoadjuvant chemotherapy and 90% with concurrent neoadjuvant chemo-endocrine therapy. A one-sample proportion test required 28 evaluable patients for 90% power at a two-sided significance level of 0.05. We increased this to 30 because the expected response rate was uncertain. After allowing about 10% for dropout or nonevaluable patients, planned enrolment was 33.

Efficacy analyses used the intention-to-treat (ITT) population, which included every enrolled patient. The safety population included anyone who received at least one dose of study treatment. Continuous variables are reported as the mean $\pm$ standard deviation or the median with interquartile range (IQR) and range, as appropriate. Categorical variables are reported as counts and percentages.

We reported objective response, pCR, and pathological response rates as proportions. Two-sided 95% confidence intervals (CIs) were calculated with the exact Clopper-Pearson method. We used the Wilcoxon signed-rank test to compare paired pretreatment and post-treatment measurements of ER, PR, Ki67, carcinoembryonic antigen (CEA), and cancer antigen 15-3 (CA15-3). Patients without paired measurements were omitted from the relevant analysis.

We used the Stuart-Maxwell test for marginal homogeneity to compare paired distributions of clinical and postoperative pathological T category, N category, and overall AJCC stage.

We used Spearman's rank correlation coefficient to examine radiological and pathological measurements. Agreement between RECIST response categories and Miller-Payne grades was assessed with quadratic-weighted Cohen's κ . We described the ability of radiological complete response to predict breast pathological complete

response (ypT0/is) using sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Exact two-sided 95% confidence intervals came from the Clopper-Pearson method. Agreement between MRI residual tumour size and residual invasive tumour size in the surgical specimen was assessed with a two-way mixed-effects, single-measure intraclass correlation coefficient (ICC) for absolute agreement. Bland-Altman analysis provided the mean difference and 95% limits of agreement.

We estimated disease-free survival (DFS) with the Kaplan-Meier method and calculated 95% confidence intervals using Greenwood's formula. Patients without a DFS event were censored at their most recent disease assessment. The reverse Kaplan-Meier method provided median follow-up. For the exploratory analysis, we used the log-rank test to compare DFS between Miller-Payne grade ≥ 3 and grade ≤ 2 . With few DFS events, this comparison was treated as hypothesis-generating and was not adjusted for covariates or multiple testing.

Safety analyses were descriptive and covered the incidence, severity, and maximum grade of treatment-related adverse events. We classified events with the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE), version 5.0.

All tests were two-sided, and $P < 0.05$ was considered statistically significant. Analyses used R version 4.6.0 and Python version 3.11.