

# Clinical and Pathological Characteristics of Young Female Breast Cancer

**Dan Zheng**

West China Hospital of Medicine: Sichuan University West China Hospital

**Ting Luo**

West China Hospital of Medicine: Sichuan University West China Hospital

**Xiaorong Zhong**

West China Hospital of Medicine: Sichuan University West China Hospital

**Chengshi Wang**

West China Hospital of Medicine: Sichuan University West China Hospital

**Ping He**

West China Hospital of Medicine: Sichuan University West China Hospital

**Tinglun Tian**

West China Hospital of Medicine: Sichuan University West China Hospital

**Yanping Wang**

West China Hospital of Medicine: Sichuan University West China Hospital

**Hong Zheng (✉ [hzheng@scu.edu.cn](mailto:hzheng@scu.edu.cn))**

West China Hospital, Sichuan University <https://orcid.org/0000-0001-5802-330X>

---

## Research article

**Keywords:** Young women, Breast cancer, Epidemiology, Prognosis, Propensity score matching

**Posted Date:** February 1st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-184597/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

With the increase in socioeconomic status and development of early screening technologies, the proportion of young breast cancer has gradually increased. However, epidemiological research on breast cancer in young women is lagging. There is a lack of diagnosis and treatment guidelines specifically for young breast cancer patients.

## Methods

This is a single-center, retrospective cohort study which adopted 2,142 women  $\leq 41$  years who were diagnosed with stage I-III invasive breast cancer. Patients were grouped into hormone receptor-positive and -negative groups. Variance of common characteristics between the two groups were compared using Chi-square test, Fisher-exact test and Wilcoxon rank sum test. Cox proportional hazards regression was employed for survival estimation and Kaplan-Meier curves were used to graphically present the survival data. Propensity score matching was used to balance covariates between patients who received or not received the same treatment.

## Results

The median age of the whole cohort was 37 (16-40), and 75.0% suffered from hormone receptor (HR) positive tumors. Modified radical mastectomy was the most frequent surgery (77.7%), and 78.7% women received adjuvant chemotherapy. Adjuvant radiotherapy was implemented in 39.0% of patients, and 58.3% women did not receive radiotherapy. The HR-positive status independently predicted unfavorable overall survival (OS, HR = 1.50, 95% CI 1.03-2.21,  $P = 0.04$ ) and invasive disease-free survival (iDFS, HR = 1.47, 95% CI 1.05-2.05,  $P = 0.02$ ). After propensity score matching (PSM), adjuvant chemotherapy (HR = 0.47, 95% CI 0.26-0.87,  $P = 0.02$ ), and adjuvant radiotherapy (HR = 0.54, 95% CI 0.37-0.78,  $P = 0.001$ ) improved OS significantly. Adjuvant chemotherapy predicted favorable iDFS (HR = 0.60, 95% CI 0.38-0.94,  $P = 0.03$ ). Endocrine therapy improved both OS and iDFS in patients with HR-positive disease.

## Conclusion

The number of young women with breast cancer is gradually increasing, and these women have worse survival outcomes than their elder counterparts. HR-positive disease predicted worse long-term survival outcomes. Adjuvant chemotherapy and adjuvant radiotherapy were required for all of the young patients. Young women with HR-positive disease can benefit from endocrine therapy. No clear benefit was seen from neoadjuvant chemotherapy in young women with early-stage breast cancer.

# Introduction

Breast and cervical cancer have become the two most malignant tumors with the highest incidence in young adult women [1]. According to the latest breast cancer (BC) statistics data from America, the

estimated BC incidence in females < 40 years old is approximately 4%, while the mortality rate is 3% [2]. The proportion of BC patients younger than 35 years was reported even up to 9.5–12% in Asian and has continually increased[3]. The occurrence of BC in patients younger than 40 years in China is 15.2%, much higher than in America, which is 5.6%, according to 2012 data of GLOBOCAN ([http://globocan.iarc.fr/old/age-specific\\_table\\_r.asp](http://globocan.iarc.fr/old/age-specific_table_r.asp)). As the optimization of diagnosis and treatment strategies increases, overall mortality from BC has decreased rapidly. However, the rate of decline of mortality in patients younger than 50 years old has decreased compared with female patients older than 50 since 2007 [4]. Young BC patients have a longer life expectancy after diagnosis, and their requirement for childbearing, work, and social intercourse is higher.

Previous studies reported that BC in young patients demonstrates a higher tumor histological grade and a higher proportion of triple negative and Human epidermal growth factor receptor 2 (HER2) positive subtype, resulting in a worse prognosis [5]. Differences in survival outcomes between young and older patients are more significant in the luminal A subtype, which may be caused by lower compliance to endocrine therapy in young women [6]. In addition, fewer young patients undergo menopause after chemotherapy, leading to a higher recurrence risk [7]. HR-positive expression had been reported as an independent risk factor for BC-specific survival (BCSS) in premenopausal females with advanced BC [8]. Additionally, a higher proportion of endocrine therapy resistance has been reported in young BC patients [9]. The incidence of *de novo* metastatic breast cancer (dnMBC) is approximately 2.6%, and the metastasis rate of primary early-stage or locally advanced BC is 27.1% in young patients [10]. Research validating the Oncotype DX Recurrence Score enrolled only 4.4% patients younger than 41 years [11], and a study on the 70-gene signature enrolled 1.8% patients under age 35 [12]. Resulting relatively insufficient evidence to extend the outcomes to young patients.

We enrolled young females (< 41 years) diagnosed with BC from a breast cancer registration database. We hypothesized that HR status plays a key role in impacting the prognosis of young women with BC. We aimed to identify epidemiologic and pathological characteristics of BC occurring in young women who were diagnosed during different time periods. Survival outcomes and effectiveness of different treatment patterns were analyzed to identify prognostic factors and suitable treatments for young BC women.

## Materials And Methods

### Database

The Breast Cancer Information Management System (BCIMS) is a patient registration database of the West China Hospital, Sichuan University[13]. BC patients treated in the West China Hospital of Sichuan University signed informed consents and were included in this database for registration. Data adopted in this study were from BCIMS and approved by the Clinical Test and Biomedical Ethics Committee of the West China Hospital, Sichuan University. Researchers regularly monitored the quality of data to ensure the completeness and accuracy of data collection. Regular follow-up was initiated according to the NCCN and CSCO BC patient follow-up guidelines. Follow up with patients generally occurred at least every 4

months during the first 3 years after diagnosis, every 6–12 months during the third to fifth year, and every year after 5 years.

## **Patients**

Female patients under age 41 who were diagnosed with primary anatomic-stage I–III invasive BC via pathological examination at the West China Hospital of Sichuan University, received surgical treatment and systematic therapy, and had complete follow-up data were included in this study. We referred to the definition of the third international consensus guidelines of the European Society for Medical Oncology (ESMO) and chose female breast patients under age 41 as young patients in this study [14]. Records missing HR status, anatomic stage, or follow-up information were excluded. Pathological diagnosis was determined by a pathologist via puncture biopsy samples or surgically resected samples. Estrogen receptor (ER) and progestin receptor (PR) status was determined according to the American Society of Clinical Oncology ER and PR immunohistochemical testing guidelines [15]. HER2-positive cancer is defined as pathological immunohistochemical staining, showing HER2 “3+” or fluorescence *in situ* hybridization (FISH) suggesting HER2 gene amplification [16]. Based on the HR expression status as determined by immunohistochemistry, patients were divided into an HR-positive group and an HR-negative group. Common demographic and baseline characteristics of patients enrolled in this research are displayed in Table 1.

**Table 1**  
**Baseline characteristics of the whole cohort**

| <b>Characteristics</b>                          | <b>N = 2142</b> |
|---|-----------------|
| Median age (years)                              | 37(16–40)       |
| Menopausal status at diagnosis                  |                 |
| Post-menopausal                                 | 40(1.9%)        |
| Pre-menopausal                                  | 2084(97.3%)     |
| NA  | 18(0.8%)        |
| Education level                                 |                 |
| Elementary school and lower                     | 145(6.8%)       |
| Junior/senior high school                       | 1305(60.9%)     |
| College and higher                              | 294(13.7%)      |
| Other/NA *                                      | 398(18.6%)      |
| Living region                                   |                 |
| Rural   | 592(27.6%)      |
| Urban   | 1528(71.3%)     |
| NA  | 22(1.0%)        |
| BMI   |                 |
| BMI < 18.5                                      | 117(5.5%)       |
| BMI ≥ 18.5 < 24                                 | 1247(58.2%)     |
| BMI > 24  | 416(19.4%)      |
| NA  | 362(16.9%)      |
| <b>Family history of breast cancer</b>          |                 |
| Yes   | 23(1.1%)        |
| No  | 2119(98.9%)     |
| <b>Family history of other malignant tumors</b> |                 |

# pT = pathological tumor stage, pN = pathological N stage

\* Contain patients whose educational level was not available and unspecified

Abbreviations: NA = not available, BMI = body mass index

| <b>Characteristics</b>    | <b>N = 2142</b> |
|---------------------------|-----------------|
| Yes                       | 176(8.2%)       |
| No                        | 1966(91.8%)     |
| pT stage #                |                 |
| 0                         | 56(2.6%)        |
| 1                         | 638(29.8%)      |
| 2                         | 959(44.8%)      |
| 3                         | 134(6.3%)       |
| 4                         | 101(4.7%)       |
| NA                        | 254(11.8%)      |
| pN stage #                |                 |
| 0                         | 1007(47.0%)     |
| 1                         | 666(31.1%)      |
| 2                         | 247(11.5%)      |
| 3                         | 221(11.3%)      |
| NA#                       | 1(0.05%)        |
| Anatomical Clinical stage |                 |
| 0                         | 309(14.1%)      |
| I                         | 381(17.8%)      |
| II                        | 923(43.1%)      |
| III                       | 529(24.7%)      |
| Hormone receptor          |                 |
| positive                  | 1606(75.0%)     |
| negative                  | 536(25.0%)      |
| Ki67                      |                 |

# pT = pathological tumor stage, pN = pathological N stage

\* Contain patients whose educational level was not available and unspecified

Abbreviations: NA = not available, BMI = body mass index

| <b>Characteristics</b>      | <b>N = 2142</b> |
|-----------------------------|-----------------|
| < 14%                       | 458(21.4%)      |
| ≥ 14%                       | 14414(67.3%)    |
| NA                          | 243(11.3%)      |
| HER2                        |                 |
| negative                    | 1491(69.6%)     |
| positive                    | 404(18.9%)      |
| equivocal                   | 26(1.2%)        |
| NA                          | 206(9.6%)       |
| Surgery                     |                 |
| Radical mastectomy          | 99(4.6%)        |
| Modified radical mastectomy | 1683(78.6%)     |
| Breast-conserving surgery   | 259(12.1%)      |
| Other/NA                    | 101(4.7%)       |
| Chemotherapy                |                 |
| Neoadjuvant                 | 314(14.7%)      |
| Adjuvant                    | 1714(80.0%)     |
| Salvage                     | 16(0.7%)        |
| Never                       | 98(4.6%)        |
| Radiation                   |                 |
| Pre-operational             | 25(1.2%)        |
| Post-operative              | 840(39.2%)      |
| Salvage                     | 29(1.3%)        |
| Never                       | 1248(58.3%)     |
| Surgical castration         |                 |

# pT = pathological tumor stage, pN = pathological N stage

\* Contain patients whose educational level was not available and unspecified

Abbreviations: NA = not available, BMI = body mass index

| Characteristics  | N = 2142    |
|--|-------------|
| Yes  | 204(9.5%)   |
| No   | 1938(90.5%) |
| # pT = pathological tumor stage, pN = pathological N stage                   |             |
| * Contain patients whose educational level was not available and unspecified |             |
| Abbreviations: NA = not available, BMI = body mass index                     |             |

## Propensity score matching (PSM)

Due to the significant difference in baseline clinicopathological characteristics between patients receiving or not receiving chemotherapy and radiotherapy, propensity score (PS) was estimated by running a logistic model in which the allocated treatment was the dependent variable. Covariates were selected based on univariable Cox proportional regression and preexisting clinical knowledge of the most important confounders in BC. Patients receiving neoadjuvant chemotherapy, adjuvant chemotherapy, and adjuvant radiotherapy were matched with patients not receiving corresponding treatments who had the closest PS (in 1:1 matching). Caliper matching was used, with caliper values of 0.2, 0.001, and 0.005 for no replacement in neoadjuvant chemotherapy, adjuvant chemotherapy, and adjuvant radiotherapy matching, respectively. Then, absolute standardized differences (ASD) was used to assess covariate balance before and after PSM. ASD below 0.1 was defined as an acceptable covariate balance degree. To improve the robustness of our outcomes, sensitivity analysis was performed separately for the three treatments. Treatment regimens, outcome evaluation and statistical analysis were described in supplementary materials.

## Results

### Baseline characteristics

A total of 2,142 female patients aged  $\leq 40$  who were diagnosed with anatomical stage I-III in the West China hospital, Sichuan University were enrolled. Specifically, we collected data from 13,399 females with primary invasive BC, 10,778 women over 40 years old were excluded, 6 patients were excluded for recurrent disease, 257 cases were deleted for missing HR information, 27 cases were eliminated for missing clinical stage information, and 135 cases were eliminated for having a follow-up period less than 0.5 months. Additionally, 9 patients were also excluded due to being diagnosed before January 2000 and another 45 patients were deleted for having stage IV disease (Fig. 1). The median follow-up period was 7.1 years. Common characteristics of the whole cohort are listed in Table 1.

To explore the change of BC baseline characteristics in young women from 2000 to 2017, we analyzed the age of patients at diagnosis, HR status, HER2 status, anatomic stage, and surgery pattern according to diagnosis year. During the three time periods before 2008, from 2008 to 2012, and since 2013, the number of women younger than 30 when diagnosed with early breast cancer (EBC) has been increasing. The proportion of women aged 30–35 diagnosed with EBC was slightly decreased, while this percentage fluctuated in the three phases in women equal to or older than 35 years. In all of the three time phases, the percentage of the HER2 negative subtype was the highest, while the percentage of the HER2 positive subtype continuously rose from 2000 to 2017. More patients were diagnosed with HR-positive tumors than HR-negative tumors since 2013 (36.6% vs. 28.0%), while a lower proportion of women were diagnosed with HR-positive BC than HR-negative BC between 2008 and 2012 (37.4% vs. 46.1%,  $P < 0.001$ ). Interestingly, the proportion of clinical stage I-II increased from 2000 to 2017, and fewer patients were diagnosed with stage III disease. Regrettably, women with clinical stage 0 disease significantly decreased overall. Further, the number of breast conserving surgeries was remarkably increased in the last ten years and the percentage of radical/extensive radical surgeries decreased (Fig. 2). In the whole cohort, most women suffered from HR-positive disease in the three diagnosis time periods.

## Intergroup differences in clinicopathological traits between HR-positive and HR-negative breast cancer

To explore the role of HR in young BC patients, we first analyzed intergroup differences in clinical and pathological characteristics between the HR-positive and HR-negative groups. Patients were divided into an HR-positive group ( $N = 1,606$ ) and an HR-negative group ( $N = 536$ ) based on HR expression status. The median age of both groups was 37 years old ( $P = 0.99$ ). In addition, 30.5% patients suffered pT1 and 5.1% had pT4 disease in the HR-positive group, while only 27.6% patients suffered pT1 and 3.5% had pT4 stage tumors in the HR-negative group ( $P < 0.001$ ). Additionally, there were fewer patients with pT3 disease in the HR-positive group than the HR-negative group (5.5% vs. 8.4%,  $P < 0.001$ ). There were more patients without axillary lymph node metastasis in the HR-negative group than in the HR-positive group (49.4% vs. 46.2%,  $P < 0.001$ ). The proportion of pN2 disease in the HR-positive group was 12.2%, and in the HR-negative group it was 9.5%. Furthermore, 68.6% of HR-positive patients suffered clinical stage II-III disease and in the HR-negative group, 65.5% had stage II-III disease. Fewer patients with stage 0-I disease (31.5%) were in the HR-positive group compared with the HR-negative group (34.5%,  $P < 0.001$ ). Ki67 expression and HER2 status were also clearly different between the two groups. A higher percentage of HR-negative women had tumors with Ki67  $\geq 14\%$  (76.5%) and HER2-positive expression (27.0%), while these proportions in the HR-positive group were only 64.2% and 16.1%, respectively ( $P < 0.001$ , Table 2).

Table 2  
Intergroup difference of clinicopathological factors between HR-positive group and HR-negative group.

| Characteristics                        | HR-positive<br>(n = 1606) | HR-negative<br>(n = 536) | P*      |
|--|---------------------------|--------------------------|---------|
| <b>Median age (years)</b>              | 37                        | 37                       | 0.99    |
| <b>Diagnosis period</b>                |                           |                          | < 0.001 |
| Before 2008                            | 417(26.0)                 | 139(25.9%)               |         |
| 2008–2012                              | 601(37.4%)                | 247(46.1%)               |         |
| Since 2013                             | 588(36.6%)                | 150(28.0%)               |         |
| <b>Menopausal status at diagnosis</b>  |                           |                          | 0.19    |
| Post-menopausal                        | 34(2.1%)                  | 6(1.1%)                  |         |
| Pre-menopausal                         | 1557(96.9%)               | 527(98.3%)               |         |
| NA #                                   | 15(0.9%)                  | 3(0.6%)                  |         |
| <b>Education level</b>                 |                           |                          | 0.24    |
| Elementary school and lower            | 111(6.9%)                 | 34(6.3%)                 |         |
| Junior/senior high school <sup>δ</sup> | 972(60.5%)                | 333(62.1%)               |         |
| College and higher                     | 233(14.5%)                | 61(11.4%)                |         |
| Other/NA*                              | 290(18.1%)                | 108(20.1%)               |         |
| <b>Living region</b>                   |                           |                          | 0.14    |
| Rural                                  | 431(26.8%)                | 161(30.0%)               |         |
| Urban                                  | 1161(72.3%)               | 367(68.5%)               |         |
| NA #                                   | 14 (0.9%)                 | 8(1.5%)                  |         |
| <b>BMI</b>                             |                           |                          | 0.40    |
| BMI < 18.5                             | 85(5.3%)                  | 32(6.0%)                 |         |

# Excluded by P-value calculation

\* Contained patients whose educational level was not available and unspecified

<sup>δ</sup> Wilcoxon rank sum test

Abbreviations: NA = not available, BMI = body mass index, PTX = paclitaxel

| Characteristics  | HR-positive<br>(n = 1606) | HR-negative<br>(n = 536) | P*      |
|--|---------------------------|--------------------------|---------|
| BMI ≥ 18.5 < 24  | 956(59.5%)                | 291(54.3%)               |         |
| BMI > 24   | 308(19.2%)                | 108(20.1%)               |         |
| NA #   | 257(16.0%)                | 105(19.6%)               |         |
| <b>Family history of breast cancer</b>   |                           |                          | 0.48    |
| Yes  | 19(1.2%)                  | 4(0.7%)                  |         |
| No   | 1587(98.8%)               | 532(99.3%)               |         |
| <b>Family history of other malignant tumor</b>                                 |                           |                          | 0.32    |
| Yes  | 126(7.8%)                 | 50(9.3%)                 |         |
| No   | 1480(92.2%)               | 486(90.7%)               |         |
| <b>pT stage<sup>δ</sup></b>  |                           |                          | < 0.001 |
| 0  | 38(2.4%)                  | 18(3.4%)                 |         |
| 1  | 490(30.5%)                | 148(27.6%)               |         |
| 2  | 720(44.8%)                | 239(44.6%)               |         |
| 3  | 89(5.5%)                  | 45(8.4%)                 |         |
| 4  | 82(5.1%)                  | 19(3.5%)                 |         |
| NA#  | 187(11.6%)                | 67(12.5%)                |         |
| <b>pN stage<sup>δ</sup></b>  |                           |                          | < 0.001 |
| 0  | 742(46.2%)                | 265(49.4%)               |         |
| 1  | 501(31.2%)                | 165(30.8%)               |         |
| 2  | 196(12.2%)                | 51(9.5%)                 |         |
| 3  | 166(10.3%)                | 55(10.3%)                |         |
| NA#  | 1(0.1%)                   | 0(0%)                    |         |
| # Excluded by P-value calculation  |                           |                          |         |
| * Contained patients whose educational level was not available and unspecified |                           |                          |         |
| δ Wilcoxon rank sum test   |                           |                          |         |
| Abbreviations: NA = not available, BMI = body mass index, PTX = paclitaxel     |                           |                          |         |

| Characteristics                   | HR-positive<br>(n = 1606) | HR-negative<br>(n = 536) | P*      |
|-----------------------------------|---------------------------|--------------------------|---------|
| <b>Clinical stage<sup>δ</sup></b> |                           |                          | < 0.001 |
| 0                                 | 226(14.1%)                | 83(15.5%)                |         |
| I                                 | 279(17.4%)                | 102(19.0%)               |         |
| II                                | 700(43.6%)                | 223(41.6%)               |         |
| III                               | 401(25.0%)                | 128(23.9%)               |         |
| <b>Ki67</b>                       |                           |                          | < 0.001 |
| < 14%                             | 408(25.4%)                | 50(9.3%)                 |         |
| ≥ 14%                             | 1031(64.2%)               | 410(76.5%)               |         |
| NA <sup>#</sup>                   | 167(10.4%)                | 76(14.2%)                |         |
| <b>HER2</b>                       |                           |                          | < 0.001 |
| negative                          | 1161(72.3%)               | 345(64.4%)               |         |
| positive                          | 259(16.1%)                | 145(27.0%)               |         |
| equivocal                         | 25(1.6%)                  | 1(0.2%)                  |         |
| NA <sup>#</sup>                   | 161(10.0%)                | 45(8.4%)                 |         |
| <b>Surgery</b>                    |                           |                          | 0.16    |
| Radical mastectomy                | 69(4.3%)                  | 30(5.6%)                 |         |
| Modified radical mastectomy       | 1252(78.0%)               | 431(80.4%)               |         |
| Breast-conserving surgery         | 206(12.8%)                | 53(9.9%)                 |         |
| No/other                          | 79(4.9%)                  | 22(4.1%)                 |         |
| <b>Chemotherapy pattern</b>       |                           |                          | 0.87    |
| Neoadjuvant                       | 232(14.4%)                | 82(15.6%)                |         |
| Adjuvant                          | 1285(80.0%)               | 429(79.8%)               |         |

# Excluded by P-value calculation

\* Contained patients whose educational level was not available and unspecified

δ Wilcoxon rank sum test

Abbreviations: NA = not available, BMI = body mass index, PTX = paclitaxel

| Characteristics                         | HR-positive<br>(n = 1606) | HR-negative<br>(n = 536) | P*      |
|---|---------------------------|--------------------------|---------|
| Salvage                                 | 13(0.8%)                  | 3(0.8%)                  |         |
| Never                                   | 76(4.7%)                  | 22(3.8%)                 |         |
| <b>Radiation</b>                        |                           |                          | 0.06    |
| Pre-operational                         | 20(1.2%)                  | 5(0.9%)                  |         |
| Post-operative                          | 655(40.8%)                | 185(34.5%)               |         |
| Salvage                                 | 22(1.4%)                  | 7(1.3%)                  |         |
| Never                                   | 909(56.6%)                | 339(63.2%)               |         |
| <b>Adjuvant endocrine therapy</b>       |                           |                          | < 0.001 |
| AI                                      | 107(6.7%)                 | 2(0.4%)                  |         |
| SERM                                    | 1210(75.3%)               | 43(8.0%)                 |         |
| Never                                   | 289(18.0%)                | 491(91.6%)               |         |
| <b>Neoadjuvant chemotherapy regimen</b> |                           |                          | < 0.001 |
| Anthracycline                           | 49(3.1%)                  | 15(2.8%)                 |         |
| Anthracycline combined with PTX         | 156(9.7%)                 | 42(7.8%)                 |         |
| PTX based                               | 11(0.7%)                  | 16(3.0%)                 |         |
| Others                                  | 16(1.0%)                  | 9(1.7%)                  |         |
| Never                                   | 1374(85.5)                | 454(84.7%)               |         |
| <b>Adjuvant chemotherapy regimen</b>    |                           |                          | < 0.001 |
| Anthracycline                           | 402(25.0%)                | 67(12.5%)                |         |
| Anthracycline combined with PTX         | 677(42.2%)                | 301(56.2%)               |         |
| PTX based                               | 113(7.0%)                 | 28(5.2%)                 |         |
| others                                  | 93(5.8%)                  | 33(6.1%)                 |         |

# Excluded by P-value calculation

\* Contained patients whose educational level was not available and unspecified

δ Wilcoxon rank sum test

Abbreviations: NA = not available, BMI = body mass index, PTX = paclitaxel

| Characteristics  | HR-positive<br>(n = 1606) | HR-negative<br>(n = 536) | P*      |
|--|---------------------------|--------------------------|---------|
| Never  | 321(20.0)                 | 107(20.0)                |         |
| <b>Surgical castration</b>   |                           |                          | < 0.001 |
| Yes  | 191(11.9%)                | 13(2.4%)                 |         |
| No   | 1415(88.1%)               | 523(97.6%)               |         |
| # Excluded by P-value calculation  |                           |                          |         |
| * Contained patients whose educational level was not available and unspecified |                           |                          |         |
| δ Wilcoxon rank sum test   |                           |                          |         |
| Abbreviations: NA = not available, BMI = body mass index, PTX = paclitaxel     |                           |                          |         |

Table 3  
Survival outcomes in HR-negative group and HR-positive group

|                   | HR-negative (n = 536) | HR-positive (n = 1606) | Hazards ratio <sub>HR positive</sub> (95% CI) |
|-------------------|-----------------------|------------------------|---|
|                   | N (%)                 | N (%)                  |   |
| Overall mortality | 60 (11.2%)            | 143 (8.9%)             | 1.50(1.03–2.21) #                             |
| Invasive disease  | 77(14.4%)             | 242(15.1%)             | 1.47(1.05–2.05) *                             |

# HR was additionally adjusted for diagnosis period(before 2008, 2008–2012, since 2013), education level (primary school or lower, high school, undergraduate or higher), pT (0–1, 2, 3–4, NA), pN (0,1, 2, 3), histological grade (I-II, III, NA), Ki67(< 14%, ≥ 14%, NA), family history of non-breast cancer malignant disease (no, yes), chemotherapy pattern (never/salvage, adjuvant, neoadjuvant), radiotherapy mode (never/salvage, adjuvant/neoadjuvant), surgery (never, breast conserving, modified radical, radical/extensive radical) and endocrine therapy (no, AI, ,SERM).

\* HR was additionally adjusted for age, diagnosis period(before 2008, 2008–2012, since 2013), pT (0–1, 2, 3–4, NA), pN (0,1, 2, 3), Ki67(< 14%, ≥ 14%, NA), histological grade (I-II, III, NA), surgical castration (never, yes), chemotherapy patterns, radiotherapy mode (never/salvage, adjuvant/neoadjuvant) and endocrine therapy (no, AI, ,SERM).

First stratum of every factor was the reference.

## Treatment patterns and chemotherapy regimens

The main surgery types in this study included breast conserving surgery, modified radical surgery, and radical or extensive radical surgery. There was no significant difference in the surgery pattern between the HR-positive group and the HR-negative group. Differences in chemotherapy between the two groups were not obvious either. There were marginally more women free from radiotherapy during the whole disease process in the HR-negative group compared with their counterparts (63.2% vs. 56.6%,  $P= 0.06$ ). In both

groups, the percentage of patients receiving anthracycline combined with taxanes as neoadjuvant or adjuvant chemotherapy was the highest. Anthracycline was the second highest frequently used chemotherapy regimen. Taxanes were used more frequently in the adjuvant phase than in the neoadjuvant phase. Over 80% of women never received neoadjuvant chemotherapy, while approximately 20% of women were free from adjuvant chemotherapy in both the HR-positive group and the HR-negative group. In addition, 11.9% of women received surgical castration after a diagnosis of HR-positive BC, and 2.4% of women underwent surgical castration during the treatment process in the HR-negative group (Table 2).

## Survival outcomes

To estimate survival outcomes in the HR-positive and HR-negative groups, we drew Kaplan-Meier survival curves of OS and iDFS in the two groups. There was no significant difference in OS (Log-rank  $P=0.075$ ) between the two groups. Three-year and five-year OS were 92.7% and 89.8% in the HR-negative group, while three-/five-year OS in the HR-positive group were 97.3% and 94.7%, respectively (Fig. 3A). Univariate Cox hazards proportional regression demonstrated that HR status was not significantly associated with OS (HR 0.76, 95% CI: 0.56–1.03,  $P=0.076$ ). Factors including time-period diagnosed, patient educational level, other malignant diseases, family history (except for BC), tumor histologic grade, pT and pN, anatomic stage, Ki67 expression intensity in tumor tissue, surgery, chemotherapy, radiotherapy, and endocrine therapy were all OS relative factors ( $P<0.05$ , Fig. 4A). After additionally adjusting independent prognosis relative factors and HR status, the latter demonstrated remarkable relativity with OS (HR = 1.51, 95% CI: 1.03–2.21,  $P=0.04$ ). Besides, more advanced pT and pN, Ki67  $\geq 14\%$ , adjuvant radiotherapy, and endocrine therapy also independently prognosed worse OS in multivariate cox regression. The three-year and five-year iDFS were 90.0% and 86.6% in the HR-negative group, and the three-/five-year iDFS in the HR-positive group were 92.3% and 87.7%, respectively (Log-rank  $P=0.85$ , Fig. 3B). Univariate Cox hazards proportional regression showed that age, time-period diagnosed, surgical castration, tumor histologic grade, pT and pN, anatomic clinical stage, Ki67 expression intensity in tumor tissue, chemotherapy and radiotherapy patterns, as well as endocrine therapy were all iDFS relative factors ( $P<0.05$ , Fig. 4B). Factors that were significant in the univariate analysis were included in the multivariate Cox hazards proportional regression combined with HR status. The HR-positive status predicted worse iDFS in the multivariate analysis (HR = 1.47, 95% CI: 1.05–2.05,  $P=0.02$ ). Diagnosis after 2008 and receiving endocrine treatment with SERMs were independent favorable outcome predictors, while higher pT and pN stage, surgical castration, and Ki67  $\geq 14\%$  were independent risk factors in multivariate cox regression analysis.

Considering the important role of HR status in young female BC, we further analyzed factors related to OS and iDFS in HR-positive and HR-negative subgroups. Factors which were associated with OS and iDFS ( $P < 0.1$ , Supplementary Fig. 1/2) in univariate cox regression were further adopted in multivariate analysis.

Higher gravidity, more advanced N stage, Ki67  $\geq$  14% and radical mastectomy were independent risk factors while endocrine therapy was independent protective factor in HR-positive subgroup for OS (Table 4). In HR-negative subgroup, higher T stage and N stage independently predicted worse OS while modified mastectomy prognosed better OS. Multivariate analysis also showed that later diagnosis period and endocrine therapy independently prognosed better iDFS and more advanced T stage, N stage and histological grade related to worse iDFS outcomes in HR-positive subgroup (Table 4). Only women who diagnosed with BC at older ages prognosed better iDFS in HR-negative subgroup (Table 4). Endocrine therapy significantly improved OS ( $HR_{AI} = 0.25$ , 95%CI 0.09–0.65,  $P = 0.005$ ,  $HR_{SERM} = 0.27$ , 95%CI 0.17–0.44,  $P < 0.001$ ) and iDFS ( $HR_{AI} = 0.58$ , 95%CI 0.32–1.09,  $P = 0.09$ ,  $HR_{SERM} = 0.60$ , 95%CI 0.42–0.86,  $P = 0.006$ ) outcomes in HR-positive subgroup patients after adjusting all the prognosis associated factors (Supplementary table 1/2).

Table 4  
Hormonal receptor subgroup multivariate cox regression

| HR-positive subgroup                  | OS                        |         |                                       | DFS                       |         |
|---------------------------------------|---------------------------|---------|---------------------------------------|---------------------------|---------|
| Factors                               | Hazards ratio<br>(95% CI) | P       | Factors                               | Hazards ratio<br>(95% CI) | P       |
| <b>Gravity</b>                        | 1.14(1.01–1.28)           | 0.03    | <b>Diagnosis period</b>               |                           |         |
| <b>pN stage</b>                       |                           |         | <b>2008–2012 vs.<br/>Before 2008</b>  | 0.6(0.43–0.83)            | 0.002   |
| <b>1 vs. 0</b>                        | 1.58(0.87–2.89)           | 0.14    | <b>Since 2013 vs.<br/>Before 2008</b> | 0.69(0.47–1.01)           | 0.06    |
| <b>2 vs. 0</b>                        | 2.44(1.26–4.74)           | 0.008   | <b>Castrate vs. Never</b>             | 1.77(1.28–2.45)           | 0.001   |
| <b>3 vs. 0</b>                        | 4.63(2.37–9.04)           | < 0.001 | <b>pT stage</b>                       |                           |         |
| <b>Ki67</b>                           |                           |         | <b>2 vs. 0–1</b>                      | 1.34(0.93–1.92)           | 0.12    |
| <b>≥ 14% vs. &lt;14%</b>              | 2.11(1.17–3.82)           | 0.01    | <b>3–4 vs. 0–1</b>                    | 2.57(1.66–3.98)           | < 0.001 |
| <b>Surgery pattern</b>                |                           |         | <b>pN stage</b>                       |                           |         |
| <b>Conserve vs.<br/>Never/salvage</b> | 0.4(0.07–2.25)            | 0.3     | <b>1 vs. 0</b>                        | 1.43(1.01–2.03)           | 0.04    |
| <b>Modified vs.<br/>Never/salvage</b> | 1.35(0.5–3.67)            | 0.55    | <b>2 vs. 0</b>                        | 1.88(1.26–2.81)           | 0.002   |
| <b>Radical vs.<br/>Never/salvage</b>  | 3.28(1.08–9.91)           | 0.04    | <b>3 vs. 0</b>                        | 2.94(1.96–4.42)           | < 0.001 |
| <b>Endocrine therapy</b>              |                           |         | <b>Ki67</b>                           |                           |         |
| <b>AI vs. Never</b>                   | 0.20(0.08–0.55)           | 0.002   | <b>≥ 14% vs. &lt;14%</b>              | 1.70(1.18–2.45)           | 0.005   |
| <b>SERM vs. Never</b>                 | 0.32(0.20–0.51)           | < 0.001 | <b>Histology grade</b>                |                           |         |
|                                       |                           |         | <b>III vs. II</b>                     | 1.38(2.01–0.09)           | 0.09    |
|                                       |                           |         | <b>Endocrine therapy</b>              |                           |         |
|                                       |                           |         | <b>AI vs. No</b>                      | 0.54(0.31–0.97)           | 0.04    |
|                                       |                           |         | <b>SERM vs. No</b>                    | 0.55(0.39–0.76)           | < 0.001 |
| <b>HR-negative<br/>subgroup</b>       | OS                        |         |                                       | DFS                       |         |
| Factors                               | Hazards ratio<br>(95% CI) | P       | Factors                               | Hazards ratio<br>(95% CI) | P       |

| HR-positive subgroup       | OS               |       | DFS                            |                 |      |
|----------------------------|------------------|-------|--------------------------------|-----------------|------|
| <b>T stage</b>             | <b>Age</b>       |       |                                |                 |      |
| <b>2 vs. 0–1</b>           | 1.7(0.77–3.76)   | 0.19  | <b>≥ 30 &lt; 35 vs. &lt;30</b> | 0.49(0.22–1.07) | 0.07 |
| <b>3–4 vs. 0–1</b>         | 4.19(1.73–10.17) | 0.001 | <b>≥ 35 vs. &lt;30</b>         | 0.51(0.26–0.98) | 0.04 |
| <b>N stage</b>             | <b>T stage</b>   |       |                                |                 |      |
| <b>1 vs. 0</b>             | 1.36(0.66–2.78)  | 0.41  | <b>2 vs. 0–1</b>               | 1.75(0.95–3.24) | 0.07 |
| <b>2 vs. 0</b>             | 3.76(1.71–8.30)  | 0.001 | <b>3–4 vs. 0–1</b>             | 2.18(1.00–4.75) | 0.05 |
| <b>3 vs. 0</b>             | 3.36(1.53–7.39)  | 0.003 | <b>N stage</b>                 |                 |      |
| <b>Surgery pattern</b>     | <b>1 vs. 0</b>   |       |                                | 0.97(0.54–1.71) | 0.91 |
| <b>Conserved vs. Never</b> | 0.59(0.16–2.18)  | 0.43  | <b>2 vs. 0</b>                 | 1.78(0.85–3.70) | 0.13 |
| <b>Modified vs. Never</b>  | 0.37(0.14–0.98)  | 0.04  | <b>3 vs. 0</b>                 | 2.32(1.19–4.55) | 0.01 |
| <b>Radical vs. Never</b>   | 0.46(0.14–1.47)  | 0.19  |                                |                 |      |

## Propensity score matching (PSM)

There were significant differences in baseline clinicopathological characteristics between patients receiving or not receiving chemoradiotherapy. To investigate the effects of these treatments, PSM was first conducted to achieve post-randomization. Before matching, variables including diagnosis time-period, pT, pN, Ki67, histological grade, and HER2 status were imbalanced between patients receiving or not receiving all three treatments. HR status was also significantly different between patients receiving or not receiving adjuvant radiotherapy (Table 5). We matched patients based on their propensity to receive neoadjuvant chemotherapy, adjuvant chemotherapy, and adjuvant radiotherapy. The propensity score (PS) was the conditional probability of receiving an exposure given a group of covariates that can be used to adjust intergroup bias to estimate the effects of different treatment patterns. The intergroup difference of all variables had been eliminated after PSM (Table 5), and the balance of covariates of the three treatments in Table 5 was substantially improved after matching (Fig. 5). We did not conduct PSM for surgery because most of the patients received modified radical surgery (77.7%, Table 1).

Table 5

Clinical and pathological factors by treatment pattern before and after propensity score matching (PSM)

| Covariates                | Before PMS               |                       | P | After PMS                |                       | P*   |      |
|---------------------------|--------------------------|-----------------------|---|--------------------------|-----------------------|------|------|
|                           | No adjuvant chemotherapy | Adjuvant chemotherapy |   | No adjuvant chemotherapy | Adjuvant chemotherapy |      |      |
| N                         | 427                      | 1714                  |   | 249                      | 249                   |      |      |
| Diagnosis time period (%) |                          |                       |   | 0.008                    |                       |      | 0.56 |
| before 2008               | 86 (20.1)                | 469 (27.4)            |   | 71 (28.5)                | 82 (32.9)             |      |      |
| 2008–2012                 | 187 (43.8)               | 661 (38.6)            |   | 103 (41.4)               | 95 (38.2)             |      |      |
| since 2013                | 154 (36.1)               | 584 (34.1)            |   | 75 (30.1)                | 72 (28.9)             |      |      |
| pT (%)                    |                          |                       |   | < 0.001                  |                       |      | 0.97 |
| 0–1                       | 118 (27.6)               | 576 (33.6)            |   | 84 (33.7)                | 84 (33.7)             |      |      |
| 2                         | 159 (37.2)               | 800 (46.7)            |   | 104 (41.8)               | 107 (43.0)            |      |      |
| 3–4                       | 117 (27.4)               | 118 (6.9)             |   | 34 (13.7)                | 34 (13.7)             |      |      |
| NA                        | 33 (7.7)                 | 220 (12.8)            |   | 27 (10.8)                | 24 (9.6)              |      |      |
| pN (%)                    |                          |                       |   | < 0.001                  |                       |      | 0.89 |
| 0                         | 134 (31.4)               | 873 (50.9)            |   | 105 (42.2)               | 105 (42.2)            |      |      |
| 1                         | 139 (32.6)               | 527 (30.7)            |   | 82 (32.9)                | 77 (30.9)             |      |      |
| 2                         | 70 (16.4)                | 177 (10.3)            |   | 37 (14.9)                | 43 (17.3)             |      |      |
| 3                         | 84 (19.7)                | 137 (8.0)             |   | 25 (10.0)                | 24 (9.6)              |      |      |
| HR-positive (%)           | 320 (74.9)               | 1285 (75.0)           | 1 | 195 (78.3)               | 196 (78.7)            | 1.00 |      |
| Ki67 (%)                  |                          |                       |   | < 0.001                  |                       |      | 0.63 |
| < 14%                     | 138 (32.3)               | 319 (18.6)            |   | 63 (25.3)                | 62 (24.9)             |      |      |
| ≥ 14%                     | 240 (56.2)               | 1201 (70.1)           |   | 159 (63.9)               | 153 (61.4)            |      |      |
| NA                        | 49 (11.5)                | 194 (11.3)            |   | 27 (10.8)                | 34 (13.7)             |      |      |
| HER2 (%)                  |                          |                       |   | 0.001                    |                       |      | 0.97 |
| negative                  | 275 (64.4)               | 1230 (71.8)           |   | 192 (77.1)               | 194 (77.9)            |      |      |

|                                  | Before PMS                      |                              | After PMS  |                                 |                              |
|----------------------------------|---------------------------------|------------------------------|------------|---------------------------------|------------------------------|
| positive                         | 85 (19.9)                       | 319 (18.6)                   | 38 (15.3)  | 36 (14.5)                       |                              |
| equivocal/NA                     | 67 (15.7)                       | 165 (9.6)                    | 19 (7.6)   | 19 (7.6)                        |                              |
| <b>WHO grade (%)</b>             |                                 |                              | < 0.001    |                                 | 0.86                         |
| I-II                             | 58 (13.6)                       | 435 (25.4)                   | 39 (15.7)  | 39 (15.7)                       |                              |
| III                              | 80 (18.7)                       | 648 (37.8)                   | 59 (23.7)  | 54 (21.7)                       |                              |
| NA                               | 289 (67.7)                      | 631 (36.8)                   | 151 (60.6) | 156 (62.7)                      |                              |
| <b>Covariates</b>                | <b>No adjuvant radiotherapy</b> | <b>Adjuvant radiotherapy</b> | P          | <b>No adjuvant radiotherapy</b> | <b>Adjuvant radiotherapy</b> |
| <b>N</b>                         | 1301                            | 840                          |            | 626                             | 626                          |
| <b>Diagnosis time period (%)</b> |                                 |                              | < 0.001    |                                 | 0.28                         |
| before 2008                      | 303 (23.3)                      | 252 (30.0)                   |            | 156 (24.9)                      | 180 (28.8)                   |
| 2008–2012                        | 502 (38.6)                      | 346 (41.2)                   |            | 259 (41.4)                      | 239 (38.2)                   |
| since 2013                       | 496 (38.1)                      | 242 (28.8)                   |            | 211 (33.7)                      | 207 (33.1)                   |
| <b>pT (%)</b>                    |                                 |                              | < 0.001    |                                 | 0.84                         |
| 0–1                              | 473 (36.4)                      | 221 (26.3)                   |            | 177 (28.3)                      | 191 (30.5)                   |
| 2                                | 565 (43.4)                      | 394 (46.9)                   |            | 294 (47.0)                      | 289 (46.2)                   |
| 3–4                              | 111 (8.5)                       | 124 (14.8)                   |            | 81 (12.9)                       | 76 (12.1)                    |
| NA                               | 152 (11.7)                      | 101 (12.0)                   |            | 74 (11.8)                       | 70 (11.2)                    |
| <b>pN (%)</b>                    |                                 |                              | < 0.001    |                                 | 0.87                         |
| 0                                | 822 (63.2)                      | 185 (22.0)                   |            | 185 (29.6)                      | 185 (29.6)                   |
| 1                                | 306 (23.5)                      | 360 (42.9)                   |            | 284 (45.4)                      | 295 (47.1)                   |
| 2                                | 89 (6.8)                        | 158 (18.8)                   |            | 76 (12.1)                       | 68 (10.9)                    |
| 3                                | 84 (6.5)                        | 137 (16.3)                   |            | 81 (12.9)                       | 78 (12.5)                    |
| <b>HR-positive (%)</b>           | 950 (73.0)                      | 655 (78.0)                   | 0.011      | 490 (78.3)                      | 484 (77.3)                   |
|                                  |                                 |                              |            |                                 | 0.73                         |

|                        | Before PMS                  |                       | After PMS |                             |                          |                |
|------------------------|-----------------------------|-----------------------|-----------|-----------------------------|--------------------------|----------------|
| Covariates             | No neoadjuvant chemotherapy | Adjuvant chemotherapy | P         | No neoadjuvant chemotherapy | Neoadjuvant chemotherapy | P <sup>z</sup> |
| <b>N</b>               | 1827                        | 314                   |           | 278                         | 278                      |                |
| <b>Mean age</b>        | 35.77 (4.05)                | 34.89 (4.46)          | < 0.001   | 35.03 (4.28)                | 34.94 (4.29)             | 0.80           |
| <b>pT (%)</b>          |                             |                       | < 0.001   |                             |                          | 0.92           |
| 0–1                    | 638 (34.9)                  | 56 (17.8)             |           | 53 (19.1)                   | 56 (20.1)                |                |
| 2                      | 829 (45.4)                  | 130 (41.4)            |           | 119 (42.8)                  | 123 (44.2)               |                |
| 3–4                    | 122 (6.7)                   | 113 (36.0)            |           | 88 (31.7)                   | 84 (30.2)                |                |
| NA                     | 238 (13.0)                  | 15 (4.8)              |           | 18 (6.5)                    | 15 (5.4)                 |                |
| <b>pN (%)</b>          |                             |                       | < 0.001   |                             |                          | 0.81           |
| 0                      | 951 (52.1)                  | 56 (17.8)             |           | 58 (20.9)                   | 55 (19.8)                |                |
| 1                      | 552 (30.2)                  | 114 (36.3)            |           | 97 (34.9)                   | 103 (37.1)               |                |
| 2                      | 181 (9.9)                   | 66 (21.0)             |           | 53 (19.1)                   | 58 (20.9)                |                |
| 3                      | 143 (7.8)                   | 78 (24.8)             |           | 70 (25.2)                   | 62 (22.3)                |                |
| <b>HR-positive (%)</b> | 1373 (75.2)                 | 232 (73.9)            | 0.684     | 204 (73.4)                  | 203 (73.0)               | 1.00           |
| <b>Ki67 (%)</b>        |                             |                       | < 0.001   |                             |                          | 0.78           |
| < 14%                  | 350 (19.2)                  | 107 (34.1)            |           | 78 (28.1)                   | 84 (30.2)                |                |
| ≥ 14%                  | 1263 (69.1)                 | 178 (56.7)            |           | 168 (60.4)                  | 166 (59.7)               |                |
| NA                     | 214 (11.7)                  | 29 (9.2)              |           | 32 (11.5)                   | 28 (10.1)                |                |
| <b>HER2 (%)</b>        |                             |                       | 0.001     |                             |                          | 0.91           |
| negative               | 1312 (71.8)                 | 193 (61.5)            |           | 180 (64.7)                  | 184 (66.2)               |                |
| positive               | 330 (18.1)                  | 74 (23.6)             |           | 64 (23.0)                   | 63 (22.7)                |                |
| equivocal/NA           | 185 (10.1)                  | 47 (15.0)             |           | 34 (12.2)                   | 31 (11.2)                |                |
| <b>WHO grade (%)</b>   |                             |                       | < 0.001   |                             |                          | 0.99           |
| I-II                   | 447 (24.5)                  | 46 (14.6)             |           | 44 (15.8)                   | 45 (16.2)                |                |

|   | Before PMS |            | After PMS  |            |
|---|------------|------------|------------|------------|
| III   | 667 (36.5) | 61 (19.4)  | 62 (22.3)  | 61 (21.9)  |
| NA  | 713 (39.0) | 207 (65.9) | 172 (61.9) | 172 (61.9) |
| *Age was excluded from PSM for patients receiving or nor receiving adjuvant chemotherapy because an ideal balance could not achieve when age was included.  |            |            |            |            |
| #Only diagnosis time period pT, pN, and HR status were included in PSM for patients receiving or nor receiving adjuvant radiotherapy. Covariates included age, Ki67, HER2 status, WHO grade were balanced between the two groups and ideal balance could not be acquired when add any of these four covariates. |            |            |            |            |
| ?Covariate of diagnosis time period was excluded from PSM for patients receiving or nor receiving neoadjuvant chemotherapy because an ideal balance could not achieve when this factor was included.  |            |            |            |            |

Multivariable Cox hazards proportional regression was done to explore the effectiveness of neoadjuvant chemotherapy, adjuvant chemotherapy and adjuvant radiotherapy in pre-matched and post-matched cohorts. The prognosis-related variables of OS and iDFS were adjusted in the regression models together with treatment status (Table 6). Before matching, neoadjuvant chemotherapy was an independent risk factor for OS ( $HR = 1.57$ , 95% CI 1.07–2.33,  $P = 0.02$ ), adjuvant radiotherapy prognosed better OS ( $HR = 0.61$ , 95% CI 0.44–0.84,  $P = 0.003$ ), while adjuvant chemotherapy showed no significant effect on OS ( $HR = 0.75$ , 95% CI 0.54–1.06,  $P = 0.10$ ). After matching, neoadjuvant chemotherapy was no longer a risk predictor for OS ( $HR = 1.48$ , 95% CI 0.93–2.37,  $P = 0.10$ ). Adjuvant chemotherapy ( $HR = 0.47$ , 95% CI 0.26–0.87,  $P = 0.02$ ) and adjuvant radiotherapy ( $HR = 0.54$ , 95% CI 0.37–0.78,  $P = 0.001$ ) significantly reduced the risk of death by 53% and 46%, respectively. Neoadjuvant chemotherapy ( $HR = 1.50$ , 95% CI 1.10–2.05,  $P = 0.01$ ) predicted unfavorable iDFS, while adjuvant chemotherapy ( $HR = 0.71$ , 95% CI 0.54–0.93,  $P = 0.01$ ) improved iDFS outcomes in the pre-matched cohorts. In the post-matched cohorts, adjuvant chemotherapy also predicted better iDFS ( $HR = 0.60$ , 95% CI 0.38–0.94,  $P = 0.03$ ), however, the impact of neoadjuvant chemotherapy on iDFS was reduced ( $HR = 1.42$ , 95% CI 0.98–2.06,  $P = 0.07$ ). No association was seen between adjuvant radiotherapy and iDFS both before and after matching (Table 6).

Table 6  
Cox proportional hazards regression analysis

|                          | OS                                   |       | iDFS                                 |       |                                      |      |                                      |      |
|--------------------------|--------------------------------------|-------|--------------------------------------|-------|--------------------------------------|------|--------------------------------------|------|
|                          | Before PSM                           |       | After PSM                            |       | Before PSM                           |      | After PSM                            |      |
|                          | HR                                   | P     | HR                                   | P     | HR                                   | P    | HR                                   | P    |
| Neoadjuvant chemotherapy | 1.57<br>(1.07–<br>2.33) <sup>a</sup> | 0.02  | 1.48<br>(0.93–<br>2.37) <sup>a</sup> | 0.10  | 1.50<br>(1.10–<br>2.05) <sup>b</sup> | 0.01 | 1.42<br>(0.98–<br>2.06) <sup>b</sup> | 0.07 |
| Adjuvant chemotherapy    | 0.75<br>(0.54–<br>1.06) <sup>a</sup> | 0.10  | 0.47<br>(0.26–<br>0.87) <sup>a</sup> | 0.02  | 0.71<br>(0.54–<br>0.93) <sup>b</sup> | 0.01 | 0.60<br>(0.38–<br>0.94) <sup>b</sup> | 0.03 |
| Adjuvant radiotherapy    | 0.61<br>(0.44–<br>0.84) <sup>c</sup> | 0.003 | 0.54<br>(0.37–<br>0.78) <sup>c</sup> | 0.001 | 0.90<br>(0.70–<br>1.16) <sup>d</sup> | 0.43 | 0.80<br>(0.61–<br>1.07) <sup>d</sup> | 0.13 |

<sup>a</sup> HR was additionally adjusted for HR (negative, positive), diagnosis time period (before 2008, 2008–2012, since 2013), age, education level (primary school or lower, high school, undergraduate or higher, NA), family history of non-breast cancer malignant disease (no, yes), pT (0–1, 2, 3–4), pN (0, 1, 2, 3), Ki67(< 14%, ≥ 14%, NA), HER2 status(negative, positive, equivocal/NA), histological grade (I-II, III, NA), adjuvant radiotherapy (no, yes) and endocrine therapy (no, AI, SERM)

<sup>b</sup> HR was additionally adjusted for HR (negative, positive), diagnosis time period (before 2008, 2008–2012, since 2013), age, pT (0–1, 2, 3–4), pN (0, 1, 2, 3), Ki67(< 14%, ≥ 14%, NA), HER2 status(negative, positive, equivocal/NA), histological grade (I-II, III, NA), surgery castration (no, yes), adjuvant radiotherapy (no, yes) and endocrine therapy (no, AI, SERM)

<sup>c</sup> HR was additionally adjusted for HR (negative, positive), diagnosis time period (before 2008, 2008–2012, since 2013), age, education level (primary school or lower, high school, undergraduate or higher, NA), family history of non-breast cancer malignant disease (no, yes), pT (0–1, 2, 3–4), pN (0, 1, 2, 3), Ki67(< 14%, ≥ 14%, NA), HER2 status(negative, positive, equivocal/NA), histological grade (I-II, III, NA), adjuvant chemotherapy (no/salvage, neoadjuvant, adjuvant) and endocrine therapy (no, AI, SERM)

<sup>d</sup> HR was additionally adjusted for HR (negative, positive), diagnosis time period (before 2008, 2008–2012, since 2013), age, pT (0–1, 2, 3–4), pN (0, 1, 2, 3), Ki67(< 14%, ≥ 14%, NA), HER2 status(negative, positive, equivocal/NA), histological grade (I-II, III, NA), surgery castration (no, yes), adjuvant chemotherapy (no/salvage, neoadjuvant, adjuvant) and endocrine therapy (no, AI, SERM).

To identify potential hidden biases which may have confounded our conclusions, we carried out sensitivity analyses. Patients with pT3-4 tumor or lymph node metastasis or HER2-positive BC or triple negative disease were included in the sensitivity analysis for neoadjuvant chemotherapy. Neoadjuvant chemotherapy increased the risk of death and invasive disease events by 81% (HR = 1.81, 95% CI 1.07–3.04,  $P=0.03$ ) and 70% (HR = 1.70, 95% CI 1.12–2.60,  $P=0.01$ ), respectively in the neoadjuvant chemotherapy cohort. Patients with pT3-4 tumor or lymph node metastasis ( $N = 1179$ ) were grouped to conduct sensitivity analysis for adjuvant radiotherapy and adjuvant chemotherapy. After matching, adjuvant radiotherapy clearly improved OS by reducing death risk 51% (HR = 0.49, 95% CI 0.33–0.73,  $P<$

0.001). No iDFS benefit was found for adjuvant radiotherapy ( $HR = 0.84$ , 95% CI 0.61–1.14,  $P = 0.25$ ). Furthermore, a significant iDFS benefit was attributed to adjuvant chemotherapy ( $HR = 0.47$ , 95% CI 0.32–0.70,  $P < 0.001$ ), while death risk was marginally decreased by adjuvant chemotherapy in this group of patients ( $HR = 0.51$ , 95% CI 0.25–1.02,  $P = 0.06$ ).

## Discussion

In this research, we described a large retrospective cohort of young women with BC women from west China with a median follow-up of 7.1 years. We assessed a series of clinical and pathological characteristics based on different diagnosis time periods and HR status to display epidemiological and clinical features of BC in young females. Our data demonstrated a trend of increasing incidence of BC in young women in western China.

In our study, data showed that there was a higher proportion of HR-positive disease (75.0%) in young BC women, which was 17.7% luminal A and 50.1% luminal B. This rate was approximate to results from other research reports [17–20]. The proportions of the triple negative subtype (15.9%) and HER2-enriched subtype (HR-negative, 6.8%) were also similar to previous studies [21, 22]. The luminal A subtype in a group of elder patients (median age 55.5) was 34.4%, and the proportion of luminal B, HER2-enriched, and basal-like subtypes were 17.5%, 20.8%, and 24.7%, respectively [23]. Combined with our data, this suggests that there is a higher proportion of HR-positive tumors in young patients. Notably, previous data showed that BC in young patients is more likely to be of the triple-negative and HER2-enriched subtypes [17, 24, 25]. However, the proportions of the two subtypes in our cohort were lower than in elder patients in previous studies [23].

HR-positive disease was an independent risk factor of OS and iDFS in young patients, which was validated in our study and others [20, 21]. It is worthy of noting that endocrine therapy was definitely protective factor in HR-positive subgroup for OS and iDFS (Supplementary table 1/2), suggesting that endocrine therapy is particularly relevant to the long-term prognosis to HR-positive BC. Tumor and lymph node stage were the main factors impacting the survival outcomes of young patients with both HR-positive and HR-negative BC. It had been reported HR positive was an independent factor for more later period BC-specific death rate, especially in the subgroup of 20–40 years old patients[8]. Another research reported that HR positive only prognoses worse survival outcome in young women (< 60 or premenopausal women) rather in old women[26]. In view of a significant impact of HER2 on the prognosis of BC, we had adopted HER2 expression status into multivariate cox regression analysis. After combined adjusted HR and other prognosis related factors with HER2 status, HR-positive expression was still prognosed worse survival outcomes (OS:  $HR = 1.46$ , 95% CI 0.99–2.16,  $P = 0.057$ , DFS:  $HR = 1.41$ , 95% CI 1.01–1.98,  $P = 0.04$ ). We also reviewed study which reported that HER2 did not have an prognostic influence on HR-positive young BC[27]. Age has been defined as an independent risk factor for survival in BC patients [9, 28, 29]. Xu et al. reported a 95.2% 5-year OS in patients under age 60 (mainly composed of post-menopausal women )[30]. In the present research, the median iDFS of the whole cohort was 6.3 years and the 5-year OS was 67.7%, clearly inferior to that of the elder patients. Furthermore, studies had

reported that BC in young patients was more prone to lymph node metastasis [31–33]. Our data displayed that 53% of patients had lymph node metastasis, while this rate in a study of patients aged 40–50 was 49.2% [22]. Few studies had explored the relationship between the HR status and node metastasis. There were slightly more patients (53.7%) suffering from lymph node metastasis in the HR-positive group than in the HR-negative group (50.6%, Table 2). Although logistic regression in this study did not show a definite correlation between the HR status and the node metastasis ( $OR = 1.14$ ,  $P = 0.20$ ). We found that breast conserving surgery and modified radical surgery marginally improved OS in young BC patients, probably due to the majority patients received modified radical surgery. In addition, adjuvant chemotherapy and adjuvant radiation improved survival outcomes significantly, emphasizing the importance of adjuvant systematic cytotoxic treatment and adjuvant radiotherapy in young patients.

## Conclusion

Although we did not find that age act as independent prognostic factor, survival outcomes in this young cohort were clearly worse than in elder patients from previous studies. HR-positive status predicted worse survival outcomes. Adjuvant chemotherapy and adjuvant radiotherapy were required for all young patients. In addition, endocrine therapy was necessary for young women with HR-positive disease. No clear benefit was seen from neoadjuvant chemotherapy, and the role of neoadjuvant chemotherapy in young women with early-stage BC requires more exploration.

## Declarations

### Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Consent for publication

Not applicable.

### Ethical approval and consent to participate

The present study was approved by the Clinical Test and Biomedical Ethics Committee of West China Hospital, Sichuan University (reference number 2012-130), and all procedures performed in this study involving human participants were with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in this study.

## Author Contributions

DZ, XZ and HZ conceived the study concept and design. TL, TT, PH and XZ collected data. DZ performed data clean and statistical analysis.

## Funding

This work was supported by the National Key Development Plan for Precision Medicine Research (grant number: 2017YFC0910004).

## Data Availability Statement

The original datasets generated during analyzed are not publicly available due individual privacy could be compromised but are available from the corresponding author on reasonable request.

## Acknowledgments

We thank LetPub ([www.letpub.com](http://www.letpub.com)) for its linguistic assistance during the preparation of this manuscript.

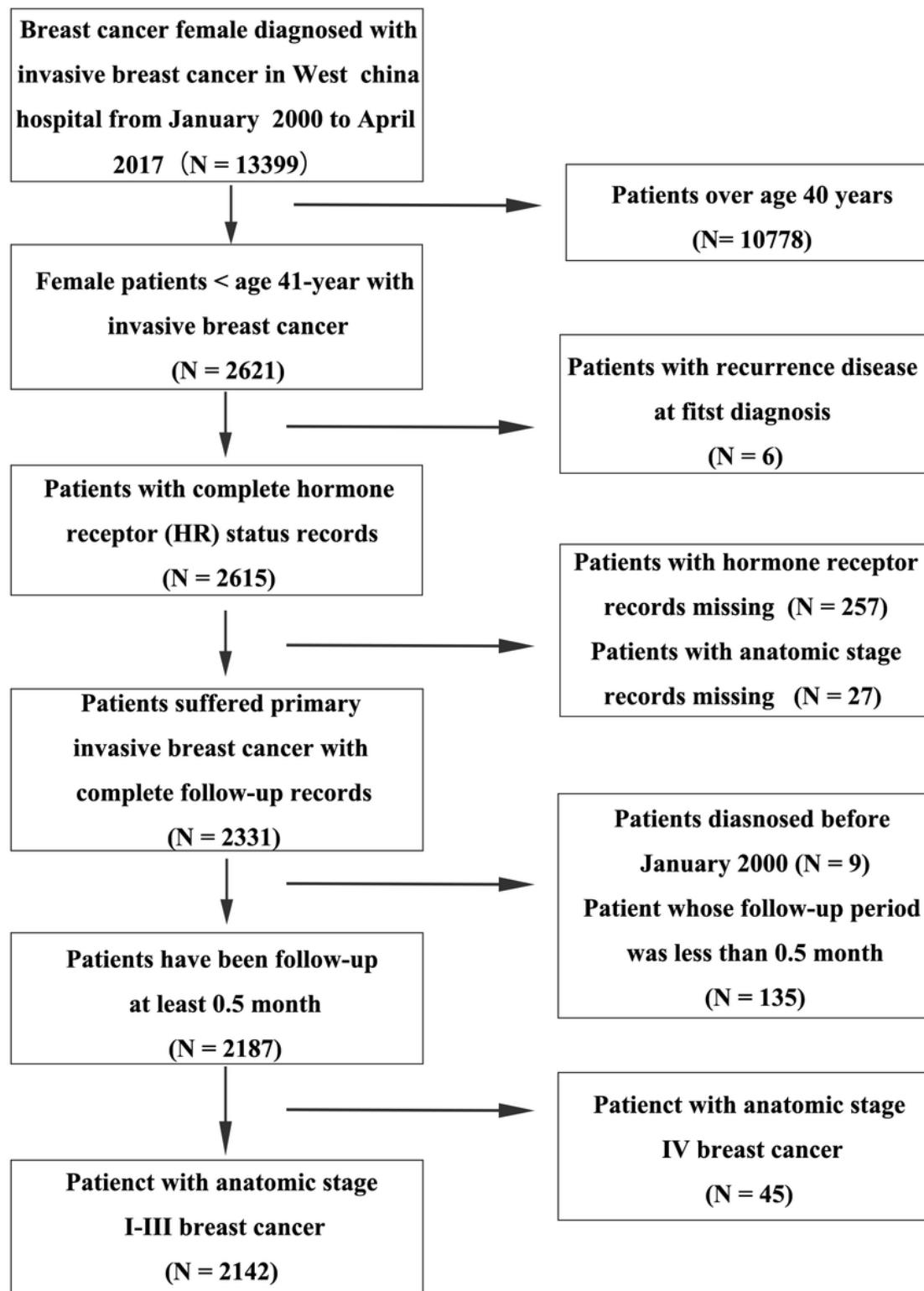
## References

1. Fidler MM, Gupta S, Soerjomataram I, et al. Cancer incidence and mortality among young adults aged 20–39 years worldwide in 2012: a population-based study. *The Lancet Oncology*. 2017;18[12]:1579–89.
2. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin*. 2019;69[6]:438–51.
3. Ahn SH, Son BH, Kim SW, et al. Poor outcome of hormone receptor-positive breast cancer at very young age is due to tamoxifen resistance: nationwide survival data in Korea—a report from the Korean Breast Cancer Society. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2007;25[17]:2360–8.
4. DeSantis CE, Ma J, Goding Sauer A, et al. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin*. 2017;67[6]:439–48.
5. Han W, Kim SW, Park IA, et al Young age: an independent risk factor for disease-free survival in women with operable breast cancer. *BMC Cancer* 2004. 4[82].
6. Cluze C, Rey D, Huiart L, et al Adjuvant endocrine therapy with tamoxifen in young women with breast cancer: determinants of interruptions vary over time. *Ann Oncol* 2012. 23[4]: 882–890.
7. Azim HA, Jr.andPartridge AH. Biology of breast cancer in young women. *Breast Cancer Res* 2014. 16[4]: 427.

8. Fu J, Zhong C, Wu L, et al. Young Patients with Hormone Receptor-Positive Breast Cancer Have a Higher Long-Term Risk of Breast Cancer Specific Death. *J Breast Cancer*. 2019;22[1]:96–108.
9. Shah AN, Metzger O, Bartlett CH, et al. Hormone Receptor-Positive/Human Epidermal Growth Receptor 2-Negative Metastatic Breast Cancer in Young Women: Emerging Data in the Era of Molecularly Targeted Agents. *The oncologist* 2020.
10. McKenzie HS, Maishman T, Simmonds P, et al. Survival and disease characteristics of de novo versus recurrent metastatic breast cancer in a cohort of young patients. *British journal of cancer* 2020.
11. Sparano JA, Gray RJ, Makower DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. *The New England journal of medicine* 2015. 373[21]: 2005–2014.
12. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *The New England journal of medicine* 2016. 375[8]: 717–729.
13. Zhong X, Luo T, Deng L, et al. Multidimensional Machine Learning Personalized Prognostic Model in an Early Invasive Breast Cancer Population-Based Cohort in China: Algorithm Validation Study. *JMIR Med Inform* 2020. 8[11]: e19069.
14. Paluch-Shimon S, Pagani O, Partridge AH, et al. ESO-ESMO 3rd international consensus guidelines for breast cancer in young women (BCY3). *Breast (Edinburgh, Scotland)* 2017. 35[203–217].
15. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2010;28[16]:2784–95.
16. Recommended by Breast Cancer Expert P. [Guideline for HER2 detection in breast cancer, the 2019 version]. *Zhonghua Bing Li Xue Za Zhi*. 2019;48[3]:169–75.
17. Collins LC, Marotti JD, Gelber S, et al. Pathologic features and molecular phenotype by patient age in a large cohort of young women with breast cancer. *Breast Cancer Res Treat*. 2012;131[3]:1061–6.
18. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *Jama*. 2006;295[21]:2492–502.
19. Yang XR, Sherman ME, Rimm DL, et al. Differences in risk factors for breast cancer molecular subtypes in a population-based study. *Cancer Epidemiol Biomarkers Prev*. 2007;16[3]:439–43.
20. Keegan TH, DeRouen MC, Press DJ, et al. Occurrence of breast cancer subtypes in adolescent and young adult women. *Breast Cancer Res* 2012. 14[2]: R55.
21. Partridge AH, Hughes ME, Warner ET, et al. Subtype-Dependent Relationship Between Young Age at Diagnosis and Breast Cancer Survival. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2016;34[27]:3308–14.
22. Tang LC, Jin X, Yang HY, et al. Luminal B subtype: a key factor for the worse prognosis of young breast cancer patients in China. *BMC Cancer* 2015. 15[201].
23. Bastien RR, Rodriguez-Lescure A, Ebbert MT, et al. PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers. *BMC Med Genomics* 2012. 5[44].

24. Anders CK, Hsu DS, Broadwater G, et al. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2008;26[20]:3324–30.
25. Ogiya R, Sagara Y, Niikura N, et al Impact of Subtype on Survival of Young Patients With Stage IV Breast Cancer. *Clin Breast Cancer* 2019. 19[3]: 200–7 e201.
26. Fu J, Wu L, Jiang M, et al. Real-world impact of non-breast cancer-specific death on overall survival in resectable breast cancer. *Cancer*. 2017;123[13]:2432–43.
27. Cvetanovic A, Popovic L, Filipovic S, et al. Young age and pathological features predict breast cancer outcome - report from a dual Institution experience in Serbia. *J BUON*. 2015;20[6]:1407–13.
28. Anastasiadi Z, Lianos GD, Ignatiadou E, et al. Breast cancer in young women: an overview. *Updates Surg*. 2017;69[3]:313–7.
29. Mustillo A, Ayoub JP, Charpentier D, et al. Prognosis in young women less than 40 years of age with brain metastasis from breast cancer. *Curr Oncol*. 2020;27[1]:39–45.
30. Xu YB, Liu H, Cao QH, et al Evaluating overall survival and competing risks of survival in patients with early-stage breast cancer using a comprehensive nomogram. *Cancer Med* 2020.
31. Kroman N, Jensen MB, Wohlfahrt J, et al. Factors influencing the effect of age on prognosis in breast cancer: population based study. *BMJ*. 2000;320[7233]:474–8.
32. Walker RA, Lees E, Webb MB, et al Breast carcinomas occurring in young women (< 35 years) are different. *British journal of cancer* 1996. 74[11]: 1796–1800.
33. Assi HA, Khoury KE, Dbouk H, et al. Epidemiology and prognosis of breast cancer in young women. *J Thorac Dis*. 2013;5(Suppl 1):2–8].

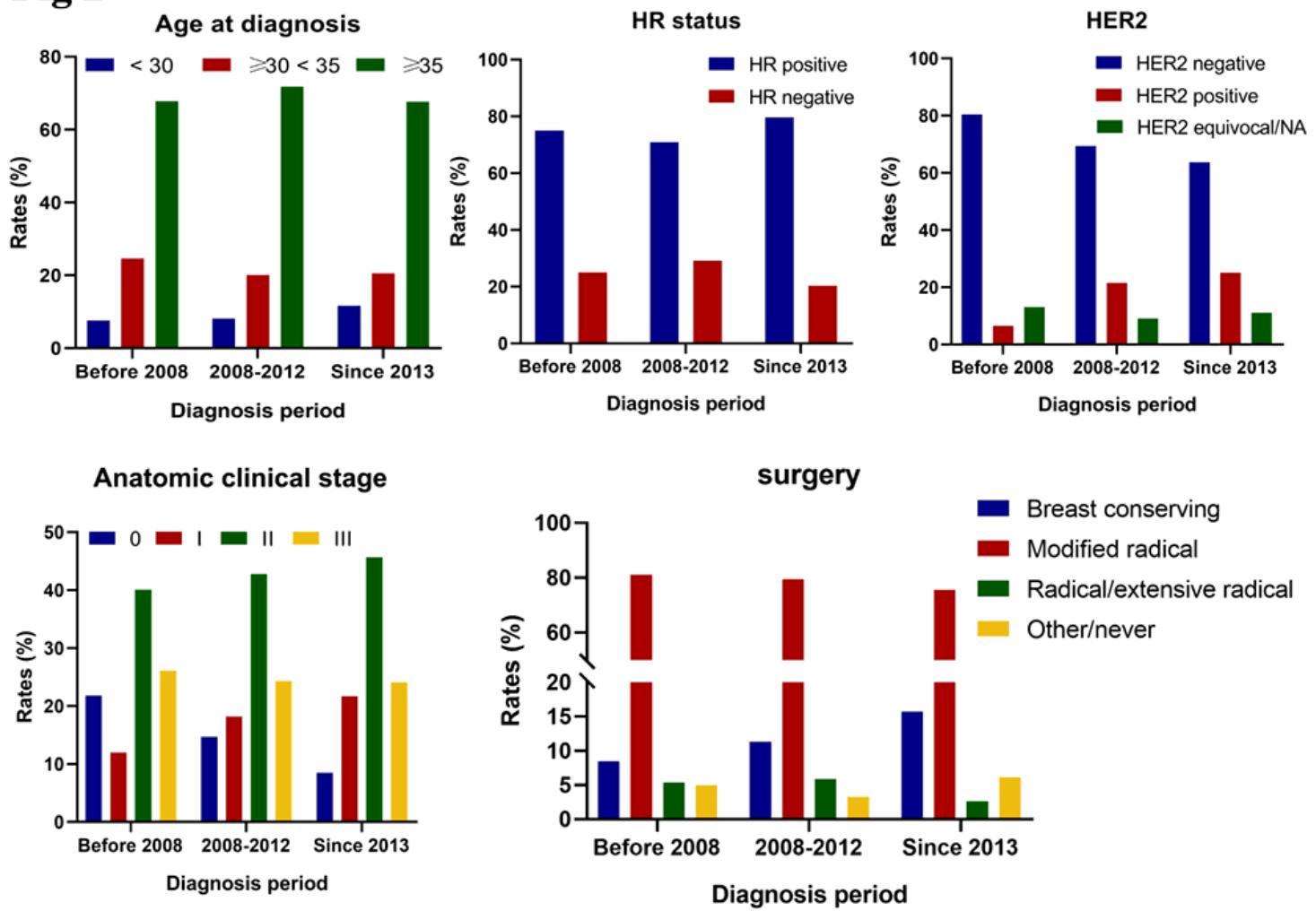
## Figures



**Figure 1**

Study design flowchart A total of 2,621 females aged  $\leq 40$  diagnosed with anatomical clinical stage I-IV breast cancer in the West China hospital, Sichuan University were enrolled. Six patients were excluded for recurrent disease, 257 for missing hormone receptor information, 26 for missing clinical stage information, 9 due to diagnosis before January 2000, and 135 for having a follow-up period less than 0.5 months.

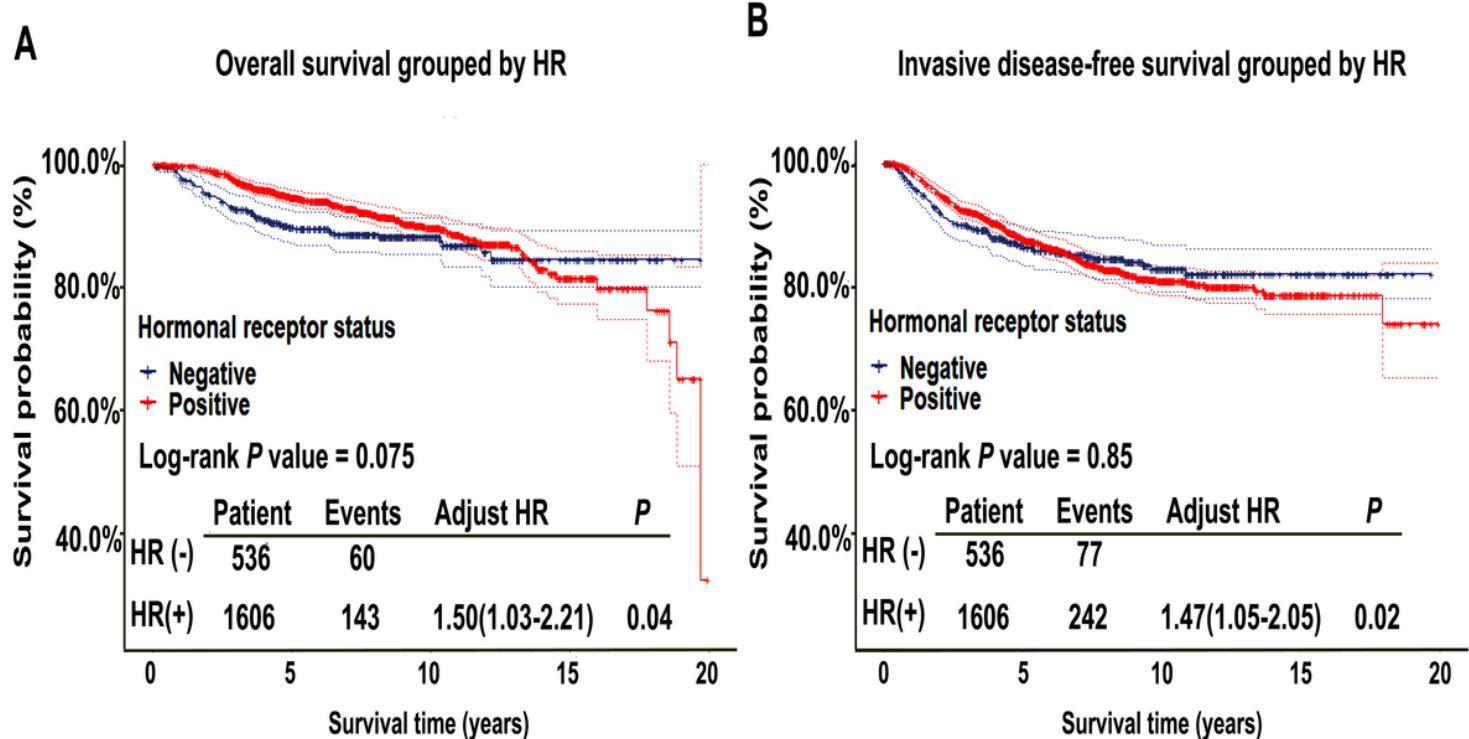
## Fig 2



## Figure 2

Distribution of clinicopathological characteristics in different diagnosis time periods Distribution of baseline age, HR status, HER2 status, anatomic stages, and surgery patterns in different diagnosis time periods (before 2008, 2008-2012, and since 2013) were changed over time. The overall trend of the ratio of age and HR status remained stable. There were clearly more patients with HER2-positive tumors. Anatomic stage I-II breast cancer increased continuously, while stage 0 disease persistently significantly decreased over time. In all of the three time periods, modified radical surgery accounted for the highest proportion and the rate of breast conserving surgery climbed gradually. The proportion of radical surgery declined slightly since 2013.

**Fig 3.**



**Figure 3**

Kaplan-Meier curve of whole EBC cohort grouped by HR (A) OS Kaplan-Meier curve of all EBC patients grouped by HR. There was a marginally significant difference of OS between the HR-positive and HR-negative groups (Log-rank *P* = 0.075). After adjusting for factors including diagnosis period, education level, pT, pN, histological grade, Ki67, family history of non-breast cancer malignant disease, chemotherapy pattern, adjuvant radiotherapy, surgery, and endocrine therapy, the hazards ratio of positive HR status was 1.50 (95% CI: 1.03-2.21, *P* = 0.04). (B) iDFS Kaplan-Meier curve of all EBC patients grouped by HR. No obvious difference was seen in iDFS between the HR-positive and HR-negative groups (Log-rank *P* = 0.85). Hazards ratio of positive HR status was 1.47 (95% CI: 1.05-2.05, *P* = 0.02) after adjusting for age, diagnosis period, pT, pN, Ki67, histological grade, surgical castration, chemotherapy patterns, adjuvant radiotherapy, and endocrine therapy.

Fig 4

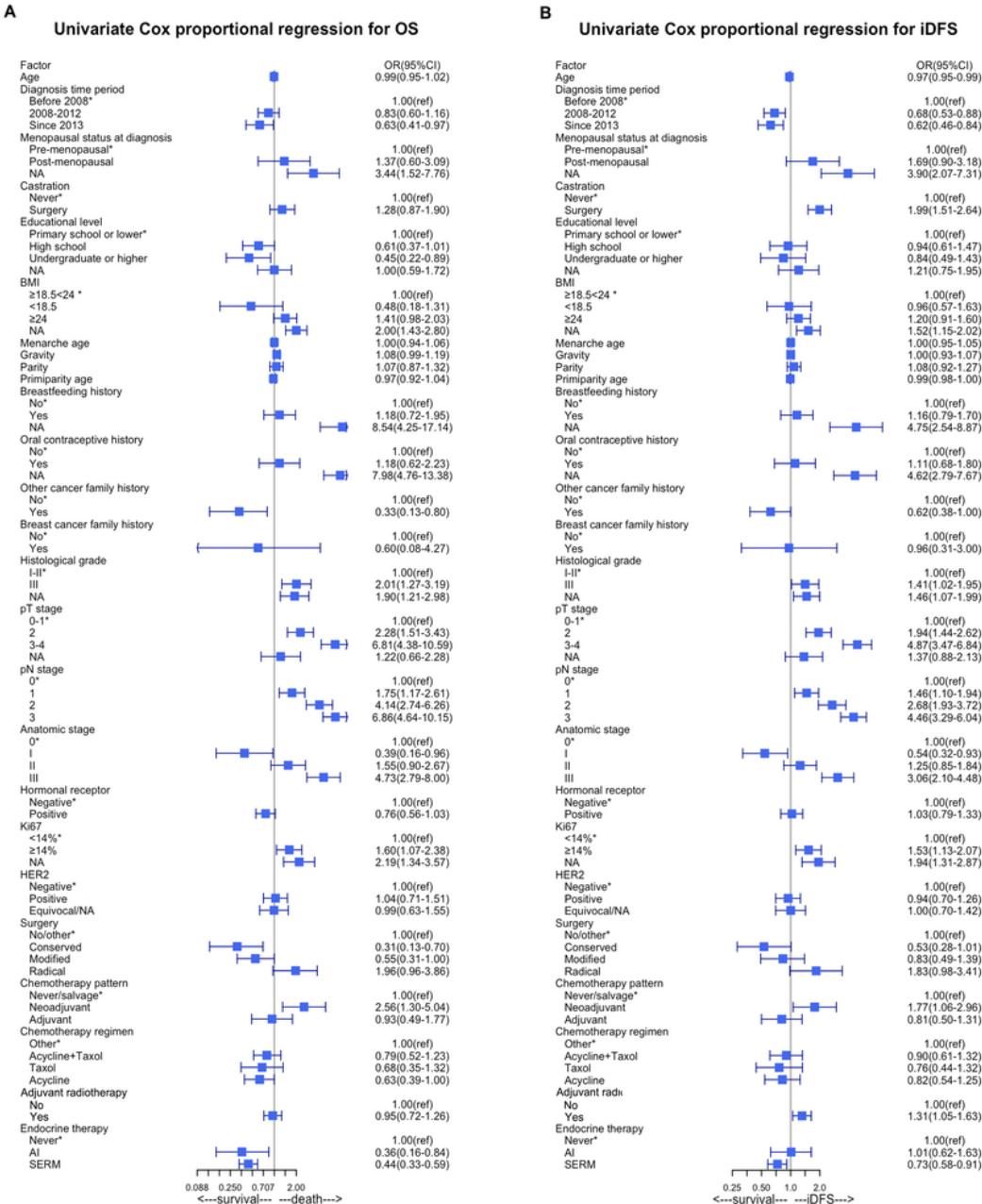
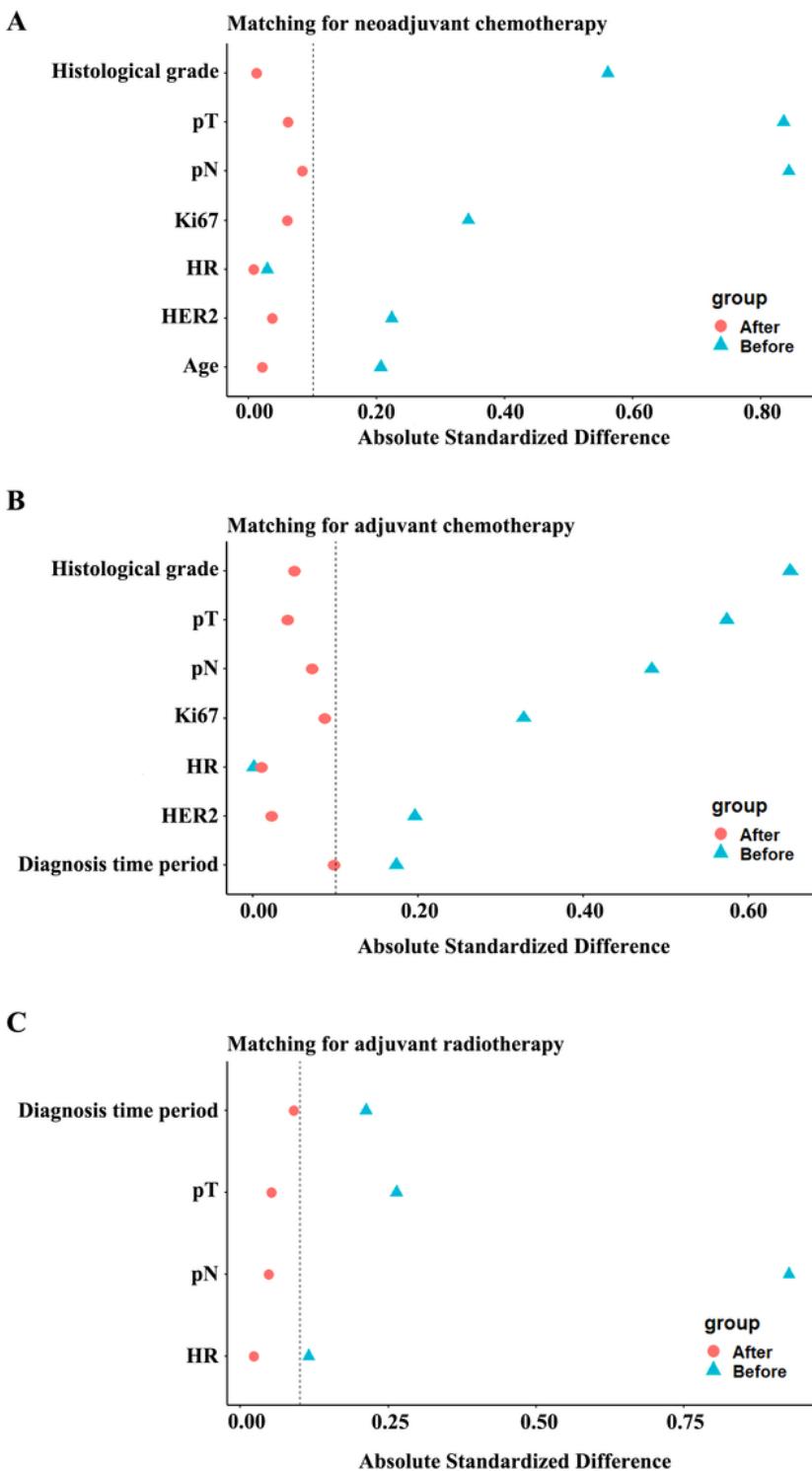


Figure 4

Univariate Cox hazards proportional regression analysis for OS and iDFS (A) Variables including diagnosis time period, education level, pT, pN, anatomic stage, histological grade, Ki67, family history of non-breast cancer malignant disease, chemotherapy pattern, surgery, and endocrine therapy were significantly associated with OS. (B) Variables, including age, diagnosis period, pT, pN, anatomic stage,

Ki67, histological grade, surgical castration, chemotherapy patterns, adjuvant radiotherapy, and endocrine therapy, were significantly associated with iDFS.

**Fig 5**



**Figure 5**

Absolute standardized difference in prognosis-related features before and after propensity score matching. (A) ASD between groups receiving and not receiving neoadjuvant chemotherapy before and after PMS. After PMS, the balance of all seven covariates was substantially improved (<0.1) compared

with before PMS. (B) ASD between groups receiving and not receiving adjuvant chemotherapy before and after PMS. Age was not included in PMS for adjuvant chemotherapy because the ideal matching could not be obtained after including it. ASD of five out of six covariates was over 0.1 before PSM and reduced to less than 0.1 after PSM. (C) ASD between groups receiving and not receiving adjuvant radiotherapy before and after PMS. pT, pN, and HR were imbalanced before PMS and were balanced after PMS.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfig1.tif](#)
- [Supplementaryfig2.tif](#)
- [Supplementarymetarials.docx](#)