Early and late outcomes of bevacizumab plus chemotherapy versus chemotherapy alone as a neoadjuvant treatment in HER2-negative nonmetastatic breast cancer: a meta-analysis of randomized controlled trials

Jinli Wei*
Yulin Luo*
Deyuan Fu
Department of Thyroid and Breast Surgery, Northern Jiangsu People’s Hospital, Yangzhou 225001, Jiangsu, People’s Republic of China
*These authors contributed equally to this work

Purpose: To better clarify the efficacy of neoadjuvant bevacizumab plus chemotherapy (BEV + CT) vs chemotherapy (CT) alone in the treatment of HER2-negative nonmetastatic breast cancer.

Methods: PubMed, Embase, Web of Science, and Cochrane Library databases were searched for relevant articles published from January 1, 2000 to July 31, 2018. Review Manager software version 5.3 was used to perform this meta-analysis.

Results: Six randomized controlled trials matched the selection criteria, yielding a total of 4,354 patients with early outcomes and 3,777 patients with late outcomes. Pooled pathological complete response (pCR) and 5-year disease-free survival (DFS) rates were higher for the neoadjuvant BEV + CT group (OR =1.37 [1.19, 1.58]; P<0.001 and HR =0.84 [0.72, 0.98]; P=0.020, respectively), but 5-year overall survival (OS) rate showed no significant difference (HR =0.79 [0.55, 1.11]; P=0.180). Subgroup analysis showed that the pCR rate was significantly higher in both patients with hormone receptor (HR)-positive breast cancer (OR =1.30 [1.01, 1.66]; P=0.040) and those with HR-negative breast cancer (OR =1.52 [1.25, 1.83]; P<0.001) in BEV + CT group.

Conclusion: Compared with CT alone, neoadjuvant BEV + CT significantly improved the 5-year DFS rate of HER2-negative breast cancer patients, but showed no benefit in terms of 5-year OS rate.

Keywords: bevacizumab, neoadjuvant chemotherapy, breast cancer, HER2-negative, randomized controlled trial

Introduction

Bevacizumab (BEV), approved by the US Food and Drug Administration in February 26, 2004, is a humanized monoclonal antibody directed against vascular endothelial growth factor, which promotes tumor angiogenesis via proliferation and migration of vascular endothelial cells.1,2 According to a review published in 2014, BEV improves overall survival (OS) and/or progression-free survival (PFS) when used in the first- and second-line treatment of metastatic cancers, such as colorectal cancer and nonsquamous non-small-cell lung cancer, and improves PFS when used in the first-line treatment of advanced renal cell carcinoma, breast cancer, and recurrent platinum-sensitive or platinum-resistant epithelial ovarian, fallopian tube, or...
BEV and paclitaxel were also found to be effective in treating disseminated bone marrow carcinoma arising from breast cancer. In the case of HER2-negative breast cancer, BEV plus chemotherapy (BEV + CT) regimen resulted in improved PFS in the metastatic setting and improved pathological complete response (pCR) rate in the neoadjuvant setting. Similarly, a randomized, Phase III trial demonstrated that BEV, administered in combination with paclitaxel or capecitabine, was effective and may be used as an option for the first-line treatment of patients with HER2-negative metastatic breast cancer if other chemotherapy (CT) regimens, including taxanes or anthracyclines, are not considered appropriate. However, no benefit was found by adding BEV for 1 year to standard CT regimen in treating both colon and breast cancers.

There are many factors that cannot be controlled in the studies of BEV as a treatment for advanced breast cancer; thus, some trials found only modest gain in PFS and OS. On the contrary, comparing the efficacy of BEV + CT regimen to CT-alone regimen for early, operable, or nonmetastatic breast cancer and then evaluating its long-term efficacy is more convincing. In this study, we performed a meta-analysis aimed at comparing the early and late outcomes of these two neoadjuvant chemotherapy (NAC) regimens in HER2-negative nonmetastatic breast cancer.

Methods

Search strategy

Eligible articles published between January 1, 2000 and July 31, 2018 were identified by searching the PubMed, Embase, Web of Science, and Cochrane Library databases. The following medical subject heading terms were used for the search: (breast OR mammary) AND (cancer OR tumo* OR carcinoma* OR neoplasm*) AND (bevacizumab OR avastin) AND (randomized controlled trial [Publication Type] OR randomized [Title/Abstract] OR placebo [Title/Abstract]). There was no language restriction placed on the search. References in the eligible articles were also reviewed to identify relevant citations.

Study selection

Only studies that fulfilled the following criteria were included: 1) randomized controlled trials (RCTs) that compared CT with BEV as a neoadjuvant treatment for nonmetastatic breast cancer patients without any prior cancer-related CT; 2) studies focusing on HER2-negative breast cancer; 3) studies with sufficient data on post-therapy pCR, disease-free survival (DFS), or OS to assess the short- or long-term efficacy of adding BEV to the CT treatment; and 4) articles published in English. There were no limitations on CT regimens.

The following studies were excluded: 1) studies on metastatic breast cancer, 2) non-randomized or single-arm clinical trials, 3) studies focusing on adjuvant CT alone, and 4) studies on HER2-positive breast cancer.

Data extraction and quality assessment

Two reviewers independently assessed all the potentially relevant studies. In the case of any disagreements, consensus was reached by discussion between the two reviewers or decided by senior investigators. Additionally, all authors independently assessed the risk of bias in each eligible trial. The Cochrane tool for assessing the risk of bias in randomized clinical trials was used to complete the risk assessment picture using Review Manager software version 5.3. Each study was rated based on the assessment criteria in the Cochrane Handbook for Systematic Reviews of Interventions version 5.3 (Figure 1 and Table 1).

Statistical analysis

The Review Manager 5.3 statistical package was used to analyze the data on study variables; four of them were
dichotomous variables expressed as ORs, and two of them were generic inverse variance variables expressed as HRs. To measure overall heterogeneity, we used the $I^2$ statistic. $I^2$ values greater than 50% indicated high heterogeneity. First, random-effect model was used for meta-analysis, and then heterogeneity was judged. If $I^2$ value was less than 50%, then fixed-effect model was used instead of the random-effect model. Results were reported with 95% CIs, and a P-value of less than 0.05 was considered statistically significant; that is, the 95% CI did not include the value “1”.

Our study adhered to the PRISMA guidelines.

**Results**

**Study quality and characteristics**

A total of 1,041 records related to BEV and CT were retrieved by searching the PubMed, Embase, Web of Science, and Cochrane Library databases. After removing duplicate articles, 695 studies were found eligible, of which 15 full-text articles met the inclusion criteria. Of these 15 articles, two non-RCTs, two single-arm clinical trials, four HER2-positive breast cancer studies, and one study focusing on adjuvant CT alone were excluded. Finally, a total of six studies were included in the meta-analysis. The selection process is depicted in the flow diagram (Figure S1).

Table 2 presents the baseline characteristics of the six studies used for meta-analysis. A total of four RCTs assessed the early outcomes of 4,354 patients given CT alone (50.1%) and BEV + CT (49.9%), and two RCTs assessed the late outcomes of 3,777 patients given CT alone (49.9%) and BEV + CT (50.1%). In three of the six RCTs, patients were divided into several subgroups based on their CT regimens. Four RCTs, including seven subgroups, provided data on pCR, and two RCTs, including four subgroups, estimated DFS and OS. All the trials focused on patients without metastatic disease diagnosed by core needle biopsy of breast and lymph nodes (T1-4d, N0-3). Details about clinical nodal (CN) status and HR status were available in all six trials for all the subgroups. For approximately 1% of patients, CN status was unknown. All CT regimens included anthracycline and taxane, while some also included cyclophosphamide, capecitabine, gemcitabine, carboplatin, and fluorouracil (Table 2).

**Outcomes of interest**

A total of 4,354 patients were included in pCR analysis. pCR was classified into two types after NAC: 1) pCR breast, which was defined as the absence of residual invasive disease with or without ductal carcinoma in situ regardless of nodes (ypT0/isN0/+), and 2) pCR breast/axillary which was defined as pCR breast and the absence of any tumor deposit $\geq 0.2$ mm in sampled axillary nodes (ypT0/isN0). Invasive disease-free survival (IDFS) in the BEATRICE trial was classified as DFS.

The pooled event rates of pCR breast (ypT0/isN0/+), for the BEV + CT and CT-alone groups were 29% and 23%, respectively. Based on our analysis, the pooled estimate of OR was 1.37, and the 95% CI was 1.19–1.58 ($P<0.001$) (Figure 2A). Similar results were obtained when further subgroup analysis was performed (Figure 2B), suggesting that BEV + CT group showed distinctly improved rate of pCR in breast regardless of the presence or absence of axillary tumors (ypT0/isN0/+). With regard to pCR breast/axillary (ypT0/isN0), a higher incidence of pCR was also observed in the BEV + CT group (25%). The pooled estimate of OR was 1.32, and the 95% CI was 1.14–1.53 with or without subgroup analysis ($P<0.001$) (Figure 3A and B). This result suggests that BEV + CT group showed distinctly improved rate of pCR not only in breast but also in axillary nodes. In terms of HR status, pooled pCR was higher in triple-negative breast cancer (TNBC) (OR =1.52 [1.25, 1.83]; $P<0.001$) (Figure 4A) and HR-positive breast cancer (OR =1.30 [1.01, 1.66]; $P=0.040$) (Figure 4B) treated with BEV + CT.

<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (attrition bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Other bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bear et al (2012)11</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>von Minckwitz et al (2012)12</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>CALGB 40603 (2014)14</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>ARTEMIS (2015)13</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>NSABP B-40 (2015)27</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>BEATRICE (2017)28</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>
### Table 2 Baseline characteristics of the six studies with 11 cohort groups for meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>No of patients (total, Std vs Tes)</th>
<th>HR(+) (Std vs Tes)</th>
<th>HR(−) (Std vs Tes)</th>
<th>CN(+) (Std vs Tes)</th>
<th>CN(−) (Std vs Tes)</th>
<th>CN(u) (Std vs Tes)</th>
<th>TS (Std vs Tes)</th>
<th>Regimens (Std)</th>
<th>Regimens (Tes)</th>
<th>Bev schedule (cycles)</th>
<th>End points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bear et al sub-1</td>
<td>2012</td>
<td>395, 200 vs 195</td>
<td>120 vs 117</td>
<td>80 vs 78</td>
<td>96 vs 91</td>
<td>104 vs 104</td>
<td>0 vs 0</td>
<td>T → AC</td>
<td>T + Bev → AC + Bev</td>
<td></td>
<td>15 mg/kg iv, q3w (×6)</td>
<td>pCR</td>
</tr>
<tr>
<td>Bear et al sub-2</td>
<td>2012</td>
<td>397, 200 vs 197</td>
<td>116 vs 116</td>
<td>84 vs 81</td>
<td>92 vs 91</td>
<td>108 vs 106</td>
<td>0 vs 0</td>
<td>TX → AC</td>
<td>TX + Bev → AC + Bev</td>
<td></td>
<td>15 mg/kg iv, q3w (×6)</td>
<td>pCR</td>
</tr>
<tr>
<td>Bear et al sub-3</td>
<td>2012</td>
<td>394, 192 vs 202</td>
<td>113 vs 121</td>
<td>79 vs 81</td>
<td>88 vs 93</td>
<td>104 vs 109</td>
<td>0 vs 0</td>
<td>TG → AC</td>
<td>TG + Bev → AC + Bev</td>
<td></td>
<td>15 mg/kg iv, q3w (×6)</td>
<td>pCR</td>
</tr>
<tr>
<td>von Minckwitz et al</td>
<td>2012</td>
<td>1,925, 969 vs 956</td>
<td>629 vs 633</td>
<td>340 vs 323</td>
<td>542 vs 554</td>
<td>391 vs 376</td>
<td>36 vs 26</td>
<td>T1–T4d</td>
<td>EC → T</td>
<td></td>
<td>15 mg/kg iv, q3w (×8)</td>
<td>pCR</td>
</tr>
<tr>
<td>CALGB 40603 sub-1</td>
<td>2014</td>
<td>218, 108 vs 110</td>
<td>0 vs 0</td>
<td>108 vs 110</td>
<td>55 vs 61</td>
<td>49 vs 42</td>
<td>4 vs 7</td>
<td>II–III</td>
<td>wP → ddAC</td>
<td></td>
<td>10 mg/kg iv, q2w (×9)</td>
<td>pCR</td>
</tr>
<tr>
<td>CALGB 40603 sub-2</td>
<td>2014</td>
<td>225, 113 vs 112</td>
<td>0 vs 0</td>
<td>113 vs 112</td>
<td>55 vs 57</td>
<td>48 vs 47</td>
<td>10 vs 8</td>
<td>II–III</td>
<td>wP + Cb → ddAC</td>
<td></td>
<td>10 mg/kg iv, q2w (×9)</td>
<td>pCR</td>
</tr>
<tr>
<td>ARTemis</td>
<td>2015</td>
<td>800, 401 vs 399</td>
<td>275 vs 277</td>
<td>126 vs 122</td>
<td>208 vs 209</td>
<td>193 vs 190</td>
<td>0 vs 0</td>
<td>T &gt; 2 cm</td>
<td>T → FEC</td>
<td></td>
<td>15 mg/kg iv, q2w (×4)</td>
<td>pCR</td>
</tr>
<tr>
<td>NSABP B-40 sub-1</td>
<td>2015</td>
<td>394, 199 vs 195</td>
<td>118 vs 117</td>
<td>81 vs 78</td>
<td>96 vs 91</td>
<td>103 vs 104</td>
<td>0 vs 0</td>
<td>T → AC</td>
<td>T + Bev → AC + Bev</td>
<td></td>
<td>15 mg/kg iv, q3w (×6+10 PS)</td>
<td>DFS, OS</td>
</tr>
<tr>
<td>NSABP B-40 sub-2</td>
<td>2015</td>
<td>400, 204 vs 196</td>
<td>117 vs 117</td>
<td>87 vs 79</td>
<td>93 vs 90</td>
<td>111 vs 106</td>
<td>0 vs 0</td>
<td>TX → AC</td>
<td>TX + Bev → AC + Bev</td>
<td></td>
<td>15 mg/kg iv, q3w (×6+10 PS)</td>
<td>DFS, OS</td>
</tr>
<tr>
<td>NSABP B-40 sub-3</td>
<td>2015</td>
<td>392, 191 vs 201</td>
<td>116 vs 122</td>
<td>75 vs 79</td>
<td>88 vs 91</td>
<td>103 vs 110</td>
<td>0 vs 0</td>
<td>TG → AC</td>
<td>TG + Bev → AC + Bev</td>
<td></td>
<td>15 mg/kg iv, q3w (×6+10 PS)</td>
<td>DFS, OS</td>
</tr>
<tr>
<td>BEATRICE</td>
<td>2017</td>
<td>2,591, 1,290 vs 1,301</td>
<td>66 vs 72</td>
<td>1,224 vs 1,229</td>
<td>476 vs 477</td>
<td>814 vs 824</td>
<td>0 vs 0</td>
<td>T1–T3</td>
<td>T or A or TA</td>
<td>(T or A or TA) + Bev</td>
<td>5 mg/kg iv, q1w (×52)</td>
<td>DFS, OS</td>
</tr>
</tbody>
</table>

**Abbreviations:** Bev, bevacizumab; Cb, carboplatin; CN(+), clinical nodal positive; CN(−), clinical nodal negative; CN(u), clinical nodal status unknown; C, cyclophosphamide; DFS, disease-free survival; dd, dose-dense; A, doxorubicin; T, docetaxel; E, epirubicin; F, fluorouracil; G, gemcitabine; HR(+), hormone receptor positive; HR(−), hormone receptor negative; iv, intravenous injection; OS, overall survival; pCR, pathological complete response; P, paclitaxel; PS, postsurgery; Sub, subgroup; Std, standard; Tes, testing; TS, tumor stage; w, week; X, capecitabine.
Figure 2 (A) Forest plot of ORs for pCR in breast regardless of lymph node status (ypT0/isN0+) without subgroup analysis. (B) Subgroup analysis of pCR in breast regardless of lymph node status (ypT0/isN0+).

Abbreviations: BEV, bevacizumab; CT, chemotherapy; pCR, pathological complete response.

Figure 3 (A) Subgroup analysis of pCR in breast and lymph nodes (ypT0/isN0). (B) Forest plot of ORs for pCR in breast and lymph nodes (ypT0/isN0) without subgroup analysis.

Abbreviations: BEV, bevacizumab; CT, chemotherapy; pCR, pathological complete response.
Two trials including 3,777 patients reported 5-year late-outcome data for OS and DFS. The final efficacy results of the BEATRICE trial published in April 2017 made it possible to evaluate the efficacy of neoadjuvant BEV in HER2-negative nonmetastatic breast cancer within 5 years. In order to make the index consistent, IDFS in the BEATRICE trial was defined as DFS. The percentage of the total number of patients with no follow-up information was 7.11% in the 5-year OS analysis (Table 3A) and 7.16% in the 5-year DFS analysis (Table 3B). There was no significant difference in 5-year OS between treatment groups (HR = 0.79, 95% CI 0.55–1.11; P = 0.180) (Figure 5A). However, BEV + CT was associated with better DFS (HR = 0.84, 95% CI 0.72–0.98; P = 0.020) (Figure 5B).

**Publication bias**
In the process of making forest plots, we conducted a sensitivity and heterogeneity analysis of the inclusion and elimination of each study. In essence, each study had no effect on outcome, and heterogeneity was less than
50%. We assessed publication bias using funnel plots (Figure 6).

**Discussion**

Randomized studies have shown that BEV + CT improves DFS and overall response rate compared to CT alone in metastatic breast cancer, but the relative efficacy of BEV + CT in early breast cancer is still controversial, especially for TNBC, which has a poor clinical outcome owing to limited treatment. Thus, we performed this meta-analysis aimed at comparing the early and late outcomes of these two NAC regimens in HER2-negative nonmetastatic breast cancer.

Of the four trials consisting of seven subgroups which reported pCR in breast while ignoring axillary lymph nodes status (ypT0/isN0/+), only Bear et al sub-1 reported a decreased pCR rate (1.5%) for BEV + CT group. Both Bear et al sub-3 and CALGB 40603 sub-1 reported a nonsignificant increase in pCR rate for the BEV + CT group. However, both the study of Bear et al and CALGB 40603 trials (without subgroup analysis) showed increased pCR rates for the BEV + CT group, which contradicts the results of the subgroup analysis. These conflicting results suggest that BEV + TX (docetaxel + capecitabine) → AC (doxorubicin + cyclophosphamide) was more effective than BEV + docetaxel (T) → AC and BEV + TG (docetaxel + gemcitabine) → AC, whereas BEV + wP (weekly paclitaxel) + carboplatin (Cb) → ddAC (dose-dense doxorubicin + cyclophosphamide) was more effective than BEV + wP → ddAC (Table 2).

In four trials which included five subgroups, pCR in breast and axillary lymph nodes (ypT0/isN0/) was determined (Figure 3A and B). In CALGB 40603 sub-1, CALGB 40603 sub-2, CALGB 40603, and Bear et al study, a higher pCR rate in the BEV + CT group was not observed. However, ARTemis and von Minckwitz et al showed increased pCR rates for BEV + CT group, which were in accordance with the results of the total analysis shown in Figure 3A and B. These conflicting results suggest that the addition of neoadjuvant BEV to EC (epirubicin + cyclophosphamide) → T or T → FEC (fluorouracil + epirubicin + cyclophosphamide) was more effective than the addition of BEV to T → AC, T → AC (with capecitabine or gemcitabine), and wP → ddAC (with or without Cb) (Table 2).

We tried to determine which group of patients, HR positive or HR negative, benefited more from the addition of BEV to NAC. As shown in Figure 4A and B, CALGB 40603 reported pCR in HR-negative patients only; ARTemis, Bear et al, and von Minckwitz et al reported pCR in both HR-positive and HR-negative patients. Subgroup analysis in the Bear et al study (Figure 4B) showed that the effect of BEV on pCR was greater in HR-positive patients (OR = 1.68 [1.15, 2.47]; P = 0.008), and there was no statistically significant difference in TNBC patients (OR = 1.19 [0.83, 1.70]; P = 0.340) (Figure 4A). However, in the von Minckwitz et al trial, the opposite trend was found in the effect of BEV on pCR rate when HR levels were considered. Bear et al demonstrated that adding BEV to T → AC (with or without capecitabine or gemcitabine) was more effective in HER2-negative/HR-positive patients. In addition, von Minckwitz et al demonstrated...
that adding BEV to EC → T was more effective in TNBC patients. In conclusion, regardless of the HR levels or CT regimens, the rate of pCR was higher in the BEV + CT group than the CT-alone group.

The NSABP B-40\(^{27}\) and BEATRICE\(^{28}\) trials reported data on OS (Table 3A). Though the pooled 5-year OS did not differ significantly between the BEV + CT group and CT-alone group, the two trials did yield contrary results (Figure 5A). In the BEATRICE trial, there was no statistically significant difference in OS between the two treatment arms in either the total population (HR =0.93 [0.74, 1.17]; \(P=0.520\)) or the pre-specified subgroups.\(^{28}\) In addition, high heterogeneity could not be ignored (\(I^2=73\%\)). The different results and high heterogeneity in the two trials may be due to differences in HR status, CN status, regimens, and BEV dose. First and most importantly, there was a distinct difference in HR status: the proportion of HR-positive patients was 59.6% in NSABP B-40 vs 5.3% in BEATRICE. Second, the proportion of patients with positive axillary nodes was 50% in NSABP B-40 vs 36% in BEATRICE. Third, the additional

---

**Figure 6** (A) Funnel plot of Figure 2A. (B) Funnel plot of Figure 2B. (C) Funnel plot of Figure 3A. (D) Funnel plot of Figure 3B. (E) Funnel plot of Figure 4A. (F) Funnel plot of Figure 4B.
drugs (cyclophosphamide or capecitabine or gemcitabine) in the NSABP B-40 trial (Table 2) may have played an important role. Fourth, in BEATRICE, patients were randomized 1:1 and received either BEV for 52 weeks (5 mg/kg/week equivalent) or four of more cycles of CT (T or A or TA) alone after definitive surgery, whereas in NSABP B-40 patients received BEV for 48 weeks (15 mg/kg every 3 weeks for six cycles of NAC plus 10 doses given postoperatively).

There are several limitations of this meta-analysis that need to be addressed. First, this is a retrospective study, resulting in selective reporting and failure to exclude incomplete outcome data. Second, the number of RCTs included was relatively small, especially in the analysis of OS and DFS. Third, the definition of pCR and DFS varied from trial to trial. Fourth, there were no specific causes of death reported in the BEATRICE and NSABP B-40 trials. Fifth, for the subgroup of HR-positive patients, a follow-up of only 5 years was short to find differences in OS in stages II–III. Despite these limitations, a meta-analysis is needed to determine which subgroup could benefit most from NAC with the addition of BEV. To obtain more convincing results, we suggest that more large trials with long-term follow-up to be conducted.

Disclosure

The authors report no conflicts of interest in this work.

References


Supplementary material

Figure S1 Prisma 2009 flow diagram.


Reference