The Breast 73 (2024) 103672

Contents lists available at ScienceDirect

The Breast

journal homepage: www.journals.elsevier.com/the-breast

Original article

"Heterogeneity of treatment effect on patients' long-term outcome according to pathological response type in neoadjuvant RCTs for breast cancer."

Laura Pala [a,1], Isabella Sala [b,c,1], Eleonora Pagan [b], Tommaso De Pas [a], Emma Zattarin [a], Chiara Catania [a], Emilia Cocorocchio [a], Giovanna Rossi [a], Daniele Laszlo [a], Giovanni Ceresoli [a], Jacopo Canzian [a], Elena Valenzi [a], Vincenzo Bagnardi [b,2], Fabio Conforti [a,d,*,2]

[a] Department of Medical Oncology, Humanitas Gavazzeni, Bergamo, Italy
[b] Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy
[c] Department of Medicine and Surgery, University of Milan-Bicocca, Milan, Italy
[d] Department of Oncology and Hemato-oncology, University of Milan, Milan, Italy

A B S T R A C T

Introduction: To provide evidence explaining the poor association between pCR and patients' long-term outcome at trial-level in neoadjuvant RCTs for breast cancer (BC), we performed a systematic-review and meta-analysis of all RCTs testing neoadjuvant treatments for early-BC and reporting the hazard ratio of DFS (HR_{DFS}) for the intervention versus control arm stratified by pathological response type (i.e., pCR yes versus no).

Methods: The objective was to explore differences of treatment effects on DFS across patients with and without pCR.

We calculated the pooled HR_{DFS} in the two strata of pathological response (i.e., pCR yes versus no) using a random-effects model, and assessed the difference between these two estimates using an interaction test.

Results: Ten RCTs and 8496 patients were included in the analysis.

Patients obtaining pCR in the intervention-arm had a higher, although not statistically significant, risk of DFS-event as compared with patients obtaining pCR in the control-arm: the pooled HR_{DFS} for the experimental versus control arm was 1.23 (95%CI, 0.91–1.65). On the opposite, the risk of DFS-event was higher for control as compared with the intervention-arm in the stratum of patients without pCR: the pooled HR_{DFS} was 0.86 (95%CI, 0.78–0.95).

Treatment effect on DFS was significantly different according to pathological response type (interaction test p: 0.014).

Conclusion: We reported new evidence that contributes to explaining the poor surrogacy value of pCR at trial-level in neoadjuvant RCTs for early-BC.

1. Introduction

The pathological complete response (pCR) is supported by regulatory agencies as surrogate endpoint for long-term patients' clinical outcome in neoadjuvant randomized clinical trials (RCTs) for early breast cancer (BC) [1]. However, a meaningful association between pCR and patients' survival has been proven only at patient-level (i.e., significantly better survival of patients who achieved pCR as compared with those who did not), but not at trial-level (i.e., poor association between degree of improvement in pCR-rate and survival reported across trials) [2,3].

Several hypotheses have been put forward to explain the weak association between pCR and both DFS and OS at trial-level [1].

One explanation is that pCR measures the effects of a therapy only on the primary tumor and not on micrometastases, and they could be meaningfully different [1].

The strong association between pCR and long-term outcome observed at patient-level in early neoadjuvant trials testing older chemotherapy regimens seemed to disprove such hypothesis [3]. However, since new neoadjuvant treatment regimens achieve substantially higher pCR rates than older chemotherapy regimens, it is possible that the degree of dissociation between the response obtained on primary tumors versus micrometastases increases and becomes clinically relevant in recent neoadjuvant RCTs [1].

To test this hypothesis, we performed a systematic review and meta-analysis of all randomized clinical trials (RCTs) testing neoadjuvant treatments in patients with early BC and reporting the long-term

* Corresponding author. Department of Medical Oncology, Cliniche Humanitas Gavazzeni, Bergamo, 24125, Italy.
E-mail address: fabio.conforti@gavazzeni.it (F. Conforti).
[1] These authors shared first authorships.
[2] These authors shared last authorships.

https://doi.org/10.1016/j.breast.2024.103672
Received 31 August 2023; Received in revised form 2 December 2023; Accepted 8 January 2024
Available online 19 January 2024

patients' outcome (i.e., HR_{DFS}) stratified by pathological response type (i.e., pCR yes versus no).

2. Methods

We followed the Preferred-Reporting-Items-for-Systematic-Reviews-and-Meta-analyses (PRISMA) and guidelines in this study [4].

We systematically searched PubMed, MEDLINE, Embase, and Scopus, up to December 1, 2022, for all RCTs testing neoadjuvant chemotherapy given alone or combined with other treatments.

The following search terms were used: "breast cancer", "neoadjuvant therapy", "preoperative therapy" and "pathologic complete response".

We used the following inclusion criteria:

i) RCTs performed in neoadjuvant setting in patients with early-BC, independently of the included disease molecular subtypes;
ii) RCTs comparing different types and/or schedules and/or duration of chemotherapy, and/or chemotherapy combined with targeted therapy and/or antivascular agents and/or immunotherapy;
iii) RCTs with available data on HR-DFS for intervention versus control arm stratified by pathological response type.

We deemed eligible any trial in which additional post-surgical adjuvant treatments were delivered if it was the same for all patients.

We excluded neoadjuvant RCTs testing endocrine therapy, due to the very low rate of pCR achieved with endocrine therapy.

Two investigators (FC and LP) independently reviewed the list of retrieved articles to choose relevant articles, and disagreements were resolved by consensus with all investigators. Two investigators (IS and EP) independently extracted data from the studies and discrepancies were resolved by consensus with all investigators. The following data were extracted: study design, number of patients enrolled, type of treatments administered, pCR-rate, definition of pCR used, number of disease-free-survival events, duration of follow-up.

The primary objective of the analysis was to explore potential difference between treatment effects, measured in terms of log (HR-DFS), across patients with and without pCR.

For each trial, we used the classification reported in the original paper to define treatments as either intervention or control arms.

2.1. Statistical analysis

From each trial, we extracted treatment effects (i.e., HR-DFS with corresponding 95 % confidence intervals (95 % CI)) according to strata of pathological response type (i.e., pCR yes or not). When data were reported only in Kaplan-Meier (KM) survival curves, we used a web based validated tool (WebPlotDigitizer) to extract data coordinates from published KM curves [5]. Then, pseudo IPD were reconstructed using the validated algorithm proposed by Guyot et al. [6] In brief, this algorithm converts digitized curves to KM data by finding numerical solutions to inverted KM equations, using available information on number of events and numbers at risk. For each trial, we calculated the pCR relative risk (RR-pCR), and its corresponding 95 % CI, as the ratio of the risk of event in the intervention group versus the risk of event in the control group. For trials for which information was available, we also extracted (or reconstructed) the effect of pCR (yes vs no) in terms of HR-DFS with the corresponding 95 % CI. HRs and RRs were considered on a log scale in the models.

Using random-effects models, we calculated the pooled RR-pCR of the effect of treatment (intervention vs control arm) on pCR and the pooled HR-DFS of the effect of pCR (yes vs no) on DFS. The DerSimonian-Laird method was used to fit the models.

Then, we calculated the pooled HR-DFS in the two strata of the pathological response using a random-effects model. The heterogeneity between the two estimates was assessed with an interaction test to give the p for interaction. The I^2 statistics, which express the percentage of the total observed variability due to heterogeneity, was also calculated.

Finally, a χ^2 test with 2 degrees of freedom was used to test the heterogeneity of results, according to the three types of treatment administered in the experimental arm (i.e., i) different schedule or new chemotherapy agents; ii) anti-HER2 drugs; iii) antivascular agents).

All reported p-values are two-sided. Analyses were performed using the SAS software v. 9.4 (SAS Institute, Cary, NC) and the R software v. 4.0.2.

3. Results

Ten RCTs, for a total of 8496 patients, fulfilled the inclusion criteria and were included in the analysis (Table 1) [7–16].

Two trials evaluated an intense dose-dense chemotherapy regimen *versus* standard-dose or different intensified/dose-dense regimens, 3 trials an anti-HER2 agent, 3 trials the antivascular agent bevacizumab, one trial the addition of a new chemotherapy agent (nab-paclitaxel) versus a standard-dose anthracycline and taxane-based chemotherapy, and one trial an anti-PD1 drug combined with chemotherapy.

Three trials only enrolled HER2-positive tumors, one trial only triple-negative tumors, while all other trials enrolled mixed molecular subtypes.

All trials but one applied a pCR definition to breast and lymph nodes (i.e., ypT0/is-ypN0).

Notably, none of the trials included was published before 2010 (year of publication range: 2010 to 2020).

The pooled pCR rate in the intervention arm was 37.5 % (95 % CI: 28.1%–47.8 %). The pCR rate was higher in the intervention arm in 8 out of 10 trials, and the pooled pCR relative risk (RR-pCR) was 1.22 (95 % CI: 1.12–1.33), significantly favoring the intervention arm (Fig. S1).

Patients who achieved a pCR had a significantly longer DFS than those without pCR (pooled-HR-DFS: 0.29 (95 % CI: 0.21–0.42); Fig. S2, three RCTs were not included in this analysis due to data unavailability).

The analysis of DFS by pathological response type showed that patients who achieved pCR had higher, although not statistically significant, risk of a DFS event when treated in the intervention arm: the pooled HR-DFS for intervention versus control arm in patients with pCR (n = 2826) was 1.23 (95 % CI, 0.91–1.65; I^2: 26.46 %; Fig. 1A). On the contrary, in patients without pCR (n = 5426), the risk of a DFS event was higher for control as compared with the intervention arm: the pooled HR-DFS was 0.86 (95 % CI, 0.78–0.95; I^2: 0 %; Fig. 1A).

Subgroup analysis revealed no significant heterogeneity of results according to type of treatment administered in the intervention arm (p-heterogeneity = 0.17).

4. Discussion

Our results showed significantly different treatments effects on patients' long-term outcome in neoadjuvant RCTs for early-BC, according to the type of pathological response obtained. Indeed, while patients without pCR had longer DFS when treated in the intervention arm, on the contrary, those who achieved a pCR had longer DFS when treated in the control arm.

Several biological and methodological hypotheses have been put forward to explain the weak association between pCR and both DFS and OS at trial-level in neoadjuvant RCTs for early-BC [17,18]. For example, in many neoadjuvant RCTs, patients receive additional adjuvant systemic treatments after surgery. Such post-surgical therapies can improve the prognosis of patients, particularly those who do not reach pCR, diluting the association at trial-level between pCR and long-term patients' outcome [17,18].

Another potential explanation is that patients who do not achieve a pCR may not be disadvantaged, as shown by patients with endocrine-responsive BCs, who derive impressive survival benefit from endocrine treatments but rarely obtain a pCR [17,18].

Table 1
Characteristics of RCTs included in the analysis.

Study	Experimental arm	Control Arm	Treatment Group	N patients	pCR definition	Median FUP	RR-pCR (95 % CI)	HR-DFS (95 % CI)	HR-OS (95 % CI)
PREPARE	ddE-ddP-CMF	EC-P	Intensified/dose-dense *vs* standard-dose or different	733	Breast and lymph nodes	43.2	1.41 (1.01–1.98)	0.88 (0.66–1.18)	0.79 (0.54–1.16)
GeparOcto	wPMy (Cb) (H + Per)	ddE-ddP-ddC (H + Per)	intensified/dose-dense regimens	945	Breast and lymph nodes	47.0	0.99 (0.86–1.13)	1.16 (0.85–1.59)	0.90 (0.58–1.40)
NOAH	H + AP-P-CMF	AP-P-CMF	Chemotherapy plus antiHER2-targeted	235	Breast	38.4	1.94 (1.30–2.89)	0.59 (0.38–0.90)	0.62 (n.a.)
KRISTINE	T-DM1+Pert	TCb + H + Pert	therapies	444	Breast and lymph nodes	36.8	0.84 (0.70–1.00)	1.11 (0.52–2.40)	1.21 (0.37–3.96)
HannaH	T-FEC + H sc	T-FEC + H iv		591	Breast and lymph nodes	40.3	1.16 (0.93–1.45)	0.95 (0.69–1.30)	0.76 (0.44–1.32)
ARTemis	Bev + T-FEC	T-FEC	Chemoterapy plus Bevacizumab	800	Breast and lymph nodes	42.0	1.34 (1.00–1.78)	1.18 (0.89–1.57)	1.26 (0.90–1.76)
GeparQuinto group 1	EC-T + Bev	EC-T		1925	Breast and lymph nodes	45.6	1.24 (1.01–1.52)	1.03 (0.84–1.25)	0.97 (0.75–1.26)
CALGB 40603	AC-P + Bev or AC-P + Bev + Cb	AC-P or AC-P + Cb		443	Breast and lymph nodes	39.0	1.15 (0.94–1.41)	0.80 (0.55–1.17)	0.76 (0.49–1.19)
GeparSepto	EC-nabP	EC-P	Others	1206	Breast and lymph nodes	49.6	1.33 (1.13–1.56)	0.66 (0.51–0.86)	0.82 (0.59–1.16)
KEYNOTE-522	PCb-A(E)C + pembrolizumab	PCb-A(E)C		1174	Breast and lymph nodes	39.1	1.13 (1.02–1.26)	0.63 (0.48–0.82)	n.a.

Abbreviations: A: adriamicin; Bev: bevacizumab; CddP; Cb: carboplatin; E: epirubicin; F:fluoruracil; H: trastuzumab; My: myocet; NabP: Nab-paclitaxel; P: paclitaxel; Pem: pemetrexed; Pert: pertuzumab; T: docetaxel; dd: dose dense; w: weekly.

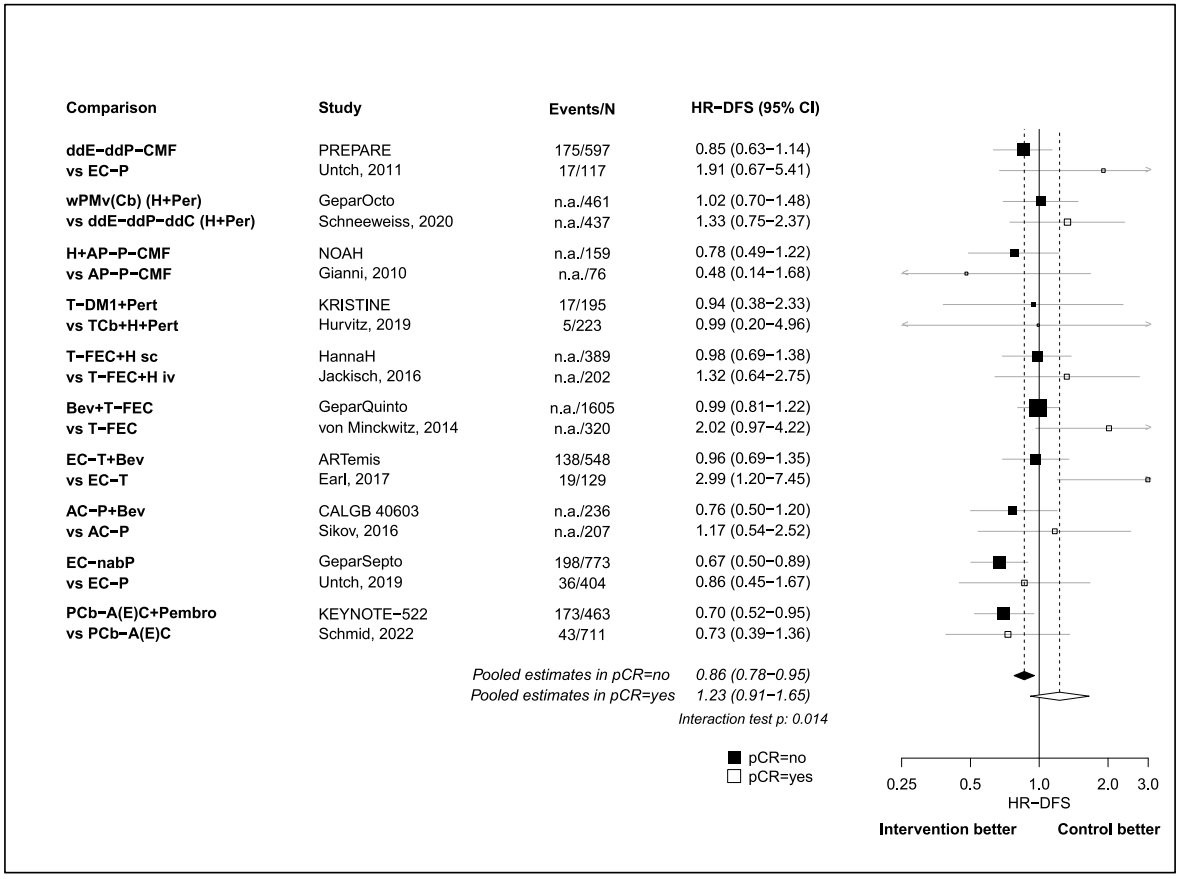

Fig. 1. Effect of treatment (intervention versus control arm) on disease-free survival (DFS) by pathological response type.

Our results highlight another potential mechanism that contributes explaining the poor surrogacy value of pCR at trial-level. Such a mechanism could be a dissociation between the antitumor effect exerted by treatments on primary tumor versus micrometastases [1]. Indeed, it is reasonable to assume that the degree of this divergence increases with higher pCR rates obtained by treatments, which explains why this mechanism affects the intervention arm more than the control arm. The very limited pCR rates reported in older neoadjuvant RCTs could have hidden such mechanism [1,17,18]. In our analysis we included only recent trials, characterized by very high absolute pCR-rates for the intervention arm.

All this means that in some patients the new experimental

neoadjuvant treatments, characterized by very high antitumor activity, could convert partial responses obtainable with standard treatments into pCR, but ultimately fail to eradicate micrometastases. The biological mechanism underpinning such observation can be easily hypothesized for antivascular agents, which might induce deep shrinkage of large primary tumors to an undetectable level through angiogenesis inhibition, without eradicating distant micrometastases, that are less dependent on neoangiogenesis for survival as compared with macroscopic primary tumors [1].

It is therefore urgently necessary to identify who are these patients with substantial residual risk of relapse despite pCR after neoadjuvant treatments. Recent evidence from the CTNeoBC project suggested that they could be those patients with more advanced tumor stage and nodal involvement before treatment start [19].

Stated in other words, our results and those of the CTNeoBC project, taken together, show that the baseline tumors' features and type of neoadjuvant treatments received, meaningfully affect the long-term prognosis of patients achieving pCR.

We acknowledge several limitations of our work to be considered. Our analysis is based on aggregate data (AD) from trials, and not on individual patient data (IPD).

IPD analyses allow checking the plausibility of randomization sequences, verifying data integrity and consistency and adjusting the analyses for baseline prognostic covariates. Nevertheless, the specific aim of our analysis was to assess pCR-surrogacy at trial-level and we used only data from RCTs of high quality, making it unlikely that an IPD analysis would substantially change our conclusions [20].

Furthermore, no significant heterogeneity of results according to type of treatment administered in the intervention arm was found in our subgroup analysis, but this might be due to lack of power.

In conclusion, we reported new evidence shedding light on potential mechanisms underpinning the poor surrogacy value of pCR for long-term patients' outcome at trial level. A better understanding of such mechanisms is useful to identify new surrogate endpoints that overcome the limits of pCR.

Ethical approval

Not applicable.

Financial support

This research was financially supported by the Italian Ministry of University and Research (PRIN 2022Y7HHNW).

Funding

There was no funding source for this study.

Data availability

Detailed extracted data on all included studies are available upon reasonable request to the corresponding author.

Authors' disclosures of potential conflicts of interest

All the authors declared: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work.

CRediT authorship contribution statement

Laura Pala: Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. **Isabella Sala:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Writing – review & editing. **Eleonora Pagan:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – review & editing. **Tommaso De Pas:** Validation, Visualization, Writing – review & editing. **Emma Zattarin:** Validation, Visualization, Writing – original draft, Writing – review & editing. **Chiara Catania:** Validation, Visualization, Writing – review & editing. **Emilia Cocorocchio:** Validation, Visualization, Writing – review & editing. **Giovanna Rossi:** Validation, Visualization, Writing – review & editing. **Daniele Laszlo:** Validation, Visualization, Writing – review & editing. **Giovanni Ceresoli:** Validation, Visualization, Writing – review & editing. **Jacopo Canzian:** Validation, Visualization, Writing – review & editing. **Elena Valenzi:** Validation, Visualization, Writing – review & editing. **Vincenzo Bagnardi:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Fabio Conforti:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.breast.2024.103672.

References

[1] Conforti F, Pala L, Bagnardi V, et al. Urrogacy of pathologic complete response in trials of neoadjuvant therapy for early breast cancer: critical analysis of strengths, weaknesses, and misinterpretations. JAMA Oncol 2022 Nov 1;8(11):1668–75.
[2] Conforti F., Pala L., Sala I., et al. Evaluation of pathological complete response as surrogate endpoint in neoadjuvant randomized clinical trials for early breast cancer. A systematic review and meta-analysis of RCTs. BMJ: 375:e066381.
[3] Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 2014;384:164–72.
[4] BMJ (OPEN ACCESS), Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. https://doi.org/10.1136/bmj.n71.
[5] Rohatgi A. WebPlotDigitizer. https://automeris.io/WebPlotDigitizer/2011.
[6] Guyot P, Ades AE, Ouwens MJ, et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol 2012;12:9.
[7] Untch M, Fasching PA, Konecny GE, et al. PREPARE trial: a randomized phase III trial com- paring preoperative, dose-dense, dose-intensified chemotherapy with epirubicin, paclitaxel and CMF versus a standard-dosed epirubicin/cyclophosphamide followed by paclitaxel +/- darbepoetin alfa in primary breast cancer—results at the time of surgery. Ann Oncol 2011;22:1988–98.
[8] Schneeweiss A, Mobus V, Tesch T, et al. Survival analysis of the randomized phase III GeparOcto trial comparing neoadjuvant chemotherapy (NACT) of iddEPC versus weekly paclitaxel, liposomal doxorubicin (plus carboplatin in triple- negative breast cancer, TNBC) (PM(Cb)) for patients (pts) with high- risk early breast cancer (BC). Ann Oncol supplement 2020;4:S303–4.
[9] Gianni L, Eiermann W, Semiglazov V, et al. Neoadjuvant chemotherapy with trastuzumab followed by adjuvant trastuzumab versus neoadjuvant chemotherapy alone, in patients with HER2-positive locally advanced breast cancer (the NOAH trial): a randomised controlled superiority trial with a parallel HER2-negative cohort. Lancet 2010;375:377–84.
[10] Hurvitz SA, Martin M, Hae Jun K, et al. Neoadjuvant trastuzumab emtansine and pertuzumab in human epidermal growth factor receptor 2–positive breast cancer: three-year outcomes from the phase III KRISTINE study. J Clin Oncol 2019;37:2206–16.
[11] Jackisch C, Hegg R, Stroyakovskiy D, et al. HannaH phase III randomised study: association of total pathological complete response with event-free survival in HER2-positive early breast cancer treated with neoadjuvanteadjuvant trastuzumab after 2 years of treatment-free follow-up. EJC 2016;62:62–75.
[12] Earl HM, Hiller L, Dunn JA, et al. ARTemis Investigators Group. Disease-free and overall survival at 3.5 years for neoadjuvant bevacizumab added to docetaxel followed by fluorouracil, epirubicin and cyclophosphamide, for women with HER2 negative early breast cancer: ARTemis Trial. Ann Oncol 2017;28:1817–24.
[13] von Minckwitz G, Loibl S, Untch M, et al. Survival after neoadjuvant chemotherapy with or without bevacizumab or everolimus for HER2-negative primary breast cancer (GBG 44-GeparQuinto). Ann Oncol 2014;25:2363–72.

[14] Sikov WM, Berry DA, Perou CM, et al. Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (alliance). J Clin Oncol 2014; 33:13–21.
[15] Untch M, Jackisch C, Schneeweiss A, et al. NAB-paclitaxel improves disease-free survival in early breast cancer: GBG 69–GeparSepto. J clin oncol 37. J Clin Oncol 2019;37:2226–34. 2226-2234, 2019 Untch M, Jackisch C, Schneeweiss A, et al : NAB-Paclitaxel Improves Disease-Free Survival in Early Breast Cancer: GBG 69–GeparSepto.
[16] Schmid P, Cortes J, Dent R, Pusztai L, McArthur H, Kümmel S, O'shaughnessy J. Event-free survival with pembrolizumab in early triple-negative breast cancer. N Engl J Med 2022;386(6):556–67.
[17] Hayes FD, Schott FA. Neoadjuvant chemotherapy: what are the benefits for the patient and for the investigator? J Natl Cancer Inst Monogr 2015;51:36–9.
[18] Rose B, Winer E, Mamon H. Perils of the pathologic complete response. J Clin Oncol 2016;34:3959–62.
[19] Van Mackelenbergh M, Loibl S, Untch M, et al. Pathologic complete response and individual patient prognosis after neoadjuvant chemotherapy plus anti–human epidermal growth factor receptor 2 therapy of human epidermal growth factor receptor 2–positive early breast cancer. J Clin Oncol 2023;41(16):2998–3008.
[20] Korn EL, Albert PS, McShane LM. Assessing surrogates as trial endpoints using mixed models. Stat Med 2005;24(2):163–82.