Predictive and prognostic values of tumor infiltrating lymphocytes in breast cancers treated with neoadjuvant chemotherapy: A meta-analysis

Shiqi Li \textsuperscript{a}, Ying Zhang \textsuperscript{b}, Peigen Zhang \textsuperscript{a}, Shuijing Xue \textsuperscript{a}, Yu Chen \textsuperscript{a}, Lihua Sun \textsuperscript{a,\textsuperscript{*}}, Rui Yang \textsuperscript{c}

\textsuperscript{a} Department of Pharmacy Administration, School of Business Administration, Shenyang Pharmaceutical University, Shenyang, China
\textsuperscript{b} Department of Clinical Pharmacy, School of Life Sciences and Biopharmaceuticals, Shenyang Pharmaceutical University, Shenyang, China
\textsuperscript{c} Clinical Pharmacology Laboratory, The Second Affiliated Hospital, Liaoning University of Traditional Chinese Medicine, Shenyang, 110034, China

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Breast cancer
Tumor infiltrating lymphocytes
Neoadjuvant chemotherapy
Prognosis and predictive
Meta-analysis

\textbf{ABSTRACT}

\textbf{Background:} This meta-analysis assessed the predictive and prognostic value of tumor infiltrating lymphocytes (TILs) in neoadjuvant chemotherapy (NACT) treated breast cancer and an optimal threshold for predicting pathologic complete response (pCR).

\textbf{Methods:} A systematic search of PubMed, EMBASE and Web of Science electronic databases was conducted to identify eligible studies published before April 2022. Either a fixed or random effects model was applied to estimate the pooled hazard ratio (HR) and odds ratio (OR) for prognosis and predictive values of TILs in breast cancer patients treated with NACT. The study is registered with PROSPERO (CRD42020221521).

\textbf{Results:} A total of 29 published studies were eligible. Increased levels of TILs predicted response to NACT in HER2 positive breast cancer (OR = 2.54 95%CI, 1.50–4.29) and triple negative breast cancer (TNBC) (OR = 3.67, 95% CI, 1.93–6.97), but not for hormone receptor (HR) positive breast cancer (OR = 1.68, 95 %CI, 0.67–4.25). A threshold of 20% of H & E-stained TILs was associated with prediction of pCR in both HER2 positive breast cancer (P = 0.035) and TNBC (P = 0.001). Moreover, increased levels of TILs (either iTILs or sTILs) were associated with survival benefit in HER2-positive breast cancer and TNBC. However, an increased level of TILs was not a prognostic factor for survival in HR positive breast cancer (pooled HR = 0.64, 95%CI: 0.03–14.1, P = 0.78).

\textbf{Conclusions:} Increased levels of TILs were associated with increased rates of response to NACT and improved prognosis for the molecular subtypes of TNBC and HER2-positive breast cancer, but not for patients with HR positive breast cancer. A threshold of 20% TILs was the most powerful outcome prognosticor of pCR.

\textbf{1. Introduction}

Worldwide, breast cancer is the most frequently diagnosed malignant tumor, and the leading cause of cancer death in women. Breast cancer is a heterogeneous disease, principally characterized histologically as infiltrating ductal and lobular carcinomas. Based on gene expression profiles, the molecular subtypes include luminal (luminal A and B), the most common subtype constituting the majority of estrogen (ER)-positive breast cancers; HER2-enriched (human epidermal growth factor receptor 2), often negative for ER and progesterone (PR); and basal subtypes, most of which are triple-negative breast cancers (TNBC).

Neoadjuvant chemotherapy (NACT) has been demonstrated to be standard therapeutic strategy in locally advanced, inoperable, and triple negative breast cancers [1,2]. It can convert a locally advanced or inoperable breast cancer into an operable tumor, and for large operable tumors, can downstage tumor size to yield an increase (7%–12%) in breast conservation rates [3–5]. Generally, NACT yields better results in breast cancers that are HER2-negative, with low hormone receptor (HR) expression, and a high proliferative index. NACT of early breast cancer can lead to high clinical response rates ranging from 70% to 90% [6,7], but a complete absence of residual invasive tumor—a pathologic complete response (pCR)—is observed in only 10%–25% of patients.

\textbf{Abbreviations:} TIL, Tumor infiltrating lymphocytes; ITIL, intratumoral; sTIL, stromal; NACT, Neoadjuvant chemotherapy; pCR, Pathologic complete response; TNBC, Triple Negative Breast Cancer; H & E, Hematoxylin and eosin; RCTs, Randomized controlled studies.

\textsuperscript{*} Corresponding author. Department of pharmacy administration, School of Business Administration, Shenyang Pharmaceutical University, 103 Wen hua Road, Shenyang, 110016, Liaoning Province, PR China.
E-mail address: slb-3632@163.com (L. Sun).

https://doi.org/10.1016/j.breast.2022.10.001
Received 5 April 2022; Received in revised form 11 August 2022; Accepted 3 October 2022
Available online 5 October 2022
0960-9776/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Because NACT constitutes an in vivo chemotherapy-sensitivity test, a pCR strongly indicates benefit from chemotherapy [8]. Patients who achieve pCR have such a clearly improved outcome with a lower risk of recurrence and longer survival [9–11], that the Food and Drug Administration (FDA) has recommended it serve as a surrogate endpoint for accelerated appraisal of new drugs for breast cancer patients treated with NACT. However, there are patients who have achieved a pCR with NACT that still relapse or die, whereas some patients without a pCR can have a good prognosis. Thus, it is essential to find more biomarkers to evaluate the efficacy and prognosis of breast cancer patients treated with NACT.

Tumor infiltrating lymphocytes (TILs) are one of the significant components of the tumor microenvironment, which reflect the intensity of the immune response within the tumor bed [12–14]. According to the site of infiltration, TILs can be divided into intratumoral (iTIL) and stromal (sTIL). Recent evidence supports the correlation between pre-treatment levels of TILs as important predictive (e.g., response to NACT) and prognostic biomarkers (e.g., RFS, OS). According to the results of a meta-analysis in 2014 [15], higher levels of TILs in pre-treatment tumor biopsies indicated higher pCR rates for NACT. However, for breast cancer patients treated with NACT, the value of TIL levels relative to survival outcomes, whether iTILs or sTILs, was not ascertained. Furthermore, to date, no formal recommendation for a clinically relevant TIL threshold(s) level has been given in the International TILs Working Group recommendations. Previous research has defined lymphocyte predominant breast cancer (LPBC) as tumors with >50% TILs [16]. Nevertheless, because of the low frequency of LPBC in clinical practice, lower levels of the TIL biomarker threshold that correlate significantly with prognosis or predictive value of chemotherapy, or eventually immunotherapy, would be valuable.

To address these controversies, we conducted a meta-analysis aimed to evaluate TILs, by site or subtype, as potentially predictive and/or prognostic for patients with breast cancer treated with NACT, and explored the predictive values of different TIL thresholds in terms of pCR.

2. Materials and methods

This meta-analysis was performed using a prespecified protocol. It was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) extension statement. An additional file displays this in more detail (see Supplementary data). The project was registered with the previously defined protocol ROSPERO database of systematic reviews, number CRD42020221521.

2.1. Literature search

Articles were eligible if they reported the predictive and prognostic value of TILs in NACT treated breast cancer. Original studies were identified by a systematic literature search of the PubMed, EMBASE and Web of Science databases, with date restriction up to April 2022. Language was limited to English. We used the following combined text and MeSH terms: "neoadjuvant therapy" and "breast cancer" and "tumor infiltrating lymphocytes". The detailed search strategy used in one database (PubMed) is provided as Supplementary material. The search was independently conducted by two authors, and discrepancies were solved by discussion with a third author. In addition, references identified in other related articles were also scanned to include eligible studies.

2.2. Inclusion and exclusion criteria

Included in the meta-analysis were both randomized controlled studies (RCTs) and case-control studies that evaluated the association between TILs and breast cancer treated with NACT. Specific inclusion criteria were as follows: (a) study population or subgroup consisted of breast cancer patients treated with NACT; (b) studies that investigated the predictive value for pathological complete response and/or long-term survival prognosis of TILs and its subtypes: CD3⁺, CD4⁺, CD8⁺, and Foxp3⁺ lymphocytes; (c) TILs were identified by pathologists using hematoxylin and eosin (H&E) stains. iTILs were defined as intraepithelial mononuclear cells within tumor cell nests or in direct contact with tumor cells and were reported as the percentage of the tumor epithelial nests that contained infiltrating lymphocytes. sTILs were defined as the percentage of the tumor stroma area that contained a lymphocytic infiltrate without direct contact to tumor cells. The cutoff value that defined TILs as high varied among studies; we used the expression values of TILs according to the original articles. And high TILs were defined as the value of total TILs on H&E-stained sections; (d) study statistics included relative risk (RR), odds ratio (OR) or hazard ratio (HR) and 95% confidence intervals (95% CI); and (e) only original research articles published in English with full-text.

Exclusion criteria were the following: (a) TILs were not clearly defined with using the median, quartiles or various scores and related statistics; (b) studies lacking key information such as odds ratio (OR), hazard ratio (HR), 95 %CI and P value; (c) non-English language; (d) in vitro and animal studies; and (e) reviews, commentaries, editorials, protocols, case reports, qualitative research, or letters.

2.3. Selection of studies and data extraction

Two of the authors independently assessed the eligibility of studies identified in the literature search, and resolved any disagreements by discussion with the third reviewer. All data and information were recorded in pre-designed tables. Information extracted included: name of first author, publication date and country, study design, number of participants, NACT program, median follow-up time, breast cancer type, subtypes of TILs, threshold values, clinicopathological characteristics, and outcome indicators.

2.4. Assessment of the risk of bias and quality of evidence

Two authors independently evaluated the risk of bias in RCTs using the Cochrane Collaboration’s risk of bias assessment tool [17], including the following seven modules: random sequence generation; allocation concealment; blinding (participants, personnel, and outcome assessment); incomplete outcome data; and selective reporting and other bias. These domains were categorized as having a high, low, or unclear risk of bias. Qualified case-control studies were assessed according to the Newcastle-Ottawa scale (NOS) [18]. The NOS contains eight items, which were categorized into the three dimensions of selection, comparability, and exposure (case control studies). The quality scores in NOS ranged from 0 to 9 and studies with scores of 6 or more were rated as high quality. Any discrepancies were resolved by consensus and arbitration by a panel of investigators within the review team.

2.5. Outcomes

TILs were analyzed as a continuous parameter or in two predefined groups of low and high TILs. Thresholds to discriminate high/low level groups of TILs varied among studies. We defined high expression values of TILs and type of lymphocytes (CD3⁺, CD4⁺, Foxp3⁺) according to the original articles. Our outcomes were the predictive significance of TILs and/or subsets of TILs for pCR, the optimal threshold of TILs to predict pCR, and the prognostic value of TILs for disease-free survival (DFS) and overall survival (OS). Pathological complete response was defined as the absence of all invasive disease in the breast and lymph node metastasis [19]. Overall survival was defined as the time from randomisation to death irrespective of cause, and disease-free survival was defined as the interval from the start of treatment to the first recurrence, or to death without any reason [20]. To avoid bias from studies contributing very long-term follow-up data compared with other studies, both OS and DFS
rates were standardized. Biomarkers can be prognostic or predictive. Predictive biomarkers predict a response to a given treatment [21]. Prognostic biomarkers correlate with the natural progression or aggressiveness of a cancer and can be quite useful for informing about the risk of recurrence or survival for particular tumor types [22]. This meta-analysis explored the predictive value of TILs for pCR and the prognostic value of TILs for DFS and OS for breast cancer treated with NACT.

2.6. Data synthesis and statistical analysis

This meta-analysis calculated the pooled OR or HR with its corresponding 95% CI to assess the association of TILs in terms of pCR or DFS, OS. Study heterogeneity was measured using the Q test and $I^2$ test. Fixed-effects models (Mantel-Haenszel, $P > 0.05$ and $I^2 < 50\%$) assume that the differences between the results of various studies are due to chance. Random-effects models (M-H heterogeneity, $P < 0.05$ or $I^2 > 50\%$) assume that the results can genuinely differ between studies. When heterogeneity is present, the random-effects model is considered to be more appropriate than a fixed-effects model, resulting in wider intervals and a more conservative estimate of effect. The likelihood of publication bias was assessed by visual inspection of funnel plot for study size against treatment effect. Begg’s and Egger’s tests were used to detect publication bias. To further investigate heterogeneity, sensitivity analyses, meta-regressions and subgroup analyses were performed to assess the outcomes data from standardized meta-analyses and associations. The STATA software version 15.0 or Revman5.3 were used for all statistical analyses.

3. Results

3.1. Search results and characteristics of eligible studies

The systematic literature search returned 2366 records. After all exclusions, 12 eligible RCTs [23–34] and 17 case-control studies [35–51] were included, consisting of approximately 9145 participants (Fig. 1). The population of patients in each study varied from 35 to 1060 cases, and the follow-up time ranged from 3.4 to 120 months. Eight and 14 publications had available data for OS and DFS analyses, respectively. Twenty-six studies provided evidence of the predictive value of TILs for pCR. The threshold values of TILs chosen were 10% (n = 3), 20% (n = 2), 30% (n = 3), 40% (n = 2), 50% (n = 4), 60% (n = 7) or 70% (n = 1). The continuous variable most used per 10% increment. The majority of NACT regimens contained anthracycline and taxane. Trastuzumab or lapatinib were typically used in HER2 positive patients. The basic characteristics are listed in Table 1.

HER2: human epidermal growth factor receptor-2; TNBC: triple negative breast cancer; HR: hormone receptor, including progesterone and estrogen receptor; NA: types of breast cancer not clearly classified in the literature; pCR: pathologic complete response; DFS: disease-free
| First Author publication year [reference] | Study Type | Type of lymphocytes | Participants Number | Country of origin | Duration of follow-up (months) | Breast subtypes | TIL evaluation method | Threshold value | Evaluation indicator | Quality score |
|------------------------------------------|------------|---------------------|---------------------|-------------------|-------------------------------|----------------|---------------------|----------------|---------------------|-------------|
| Denkert 2010 [23]                        | Prospective | All                 | 1058                | Germany           | –                             | HER2 positive, HR positive | H&E               | 60%              | pCR             | –           |
| Loil 2014 [24]                           | Prospective | All                 | 1010                | Australia         | 62                            | HER2 positive, HR positive | H&E               | 50%              | OS, DFS         | –           |
| Salgado 2015 [25]                        | Prospective | All                 | 455                 | –                 | 50.64                         | HER2 positive, TNBC         | H&E               | 10% INC          | DFS, pCR        | –           |
| Ingold 2016 [26]                         | Prospective | All                 | 1060                | Germany           | 60.39                         | HER2 positive              | H&E               | 60%              | DFS, pCR        | –           |
| Ingold Heppnerl 2016 [26]                | Prospective | All                 | 1060                | Germany           | 60.39                         | HER2 positive              | H&E               | 60%              | DFS, pCR        | –           |
| Dieci 2016 [27]                          | Prospective | All                 | 121                 | Italy             | –                             | HER2 positive, TNBC         | H&E               | 60%              | DFS, pCR        | –           |
| Ignatiadis 2018 [28]                     | Prospective | CD4⁺, Foxp3⁺       | 1010                | Australia         | 62                            | HER2 positive, TNBC         | H&E               | 50%              | DFS, pCR        | –           |
| Schmidt 2018 [29]                        | Retrospective | All                 | 50                  | Multicentre       | –                             | HER2 positive, TNBC         | H&E               | 10% INC          | DFS, pCR        | 9           |
| Loibl 2017 [30]                          | Retrospective | All                 | 146                 | Germany           | –                             | NA                        | H&E               | 50%              | pCR             | –           |
| Würtfel 2018 [31]                        | Retrospective | All                 | 580                 | Germany           | –                             | NA                        | H&E               | 60%              | pCR             | –           |
| Nummer 2014 [32]                         | Retrospective | All                 | 313                 | Germany           | –                             | HR positive                | H&E               | 60%              | pCR             | –           |
| Guarnieri 2021 [34]                      | Retrospective | All                 | 121                 | Italy             | 108                           | HER2 positive, TNBC         | H&E               | –                | OS, pCR, DFS    | 8           |
| Ochi 2019 [35]                           | Retrospective | All                 | 209                 | Japan             | 120                           | HER2 positive, TNBC         | H&E               | 10%              | DFS, pCR        | 9           |
| Lee 2013 [36]                            | Retrospective | CD8⁺, Foxp3⁺       | 175                 | Korea             | –                             | NA                        | H&E               | 40%              | pCR             | 6           |
| Dieci 2014 [37]                          | Retrospective | All                 | 278                 | France            | 76                            | TNBC                      | H&E               | 60%              | OS, DFS         | 8           |
| Russo 2019 [38]                          | Retrospective | All                 | 187                 | Venezuela         | 62.5                          | HER2 positive, TNBC         | H&E               | 30%              | OS, pCR         | 8           |
| Asano 2018 [39]                          | Retrospective | All                 | 177                 | Japan             | 3.4                           | HER2 positive, TNBC, HR positive | H&E               | 10%              | DFS, pCR        | 8           |
| Song 2017 [40]                           | Retrospective | CD8⁺                | 108                 | Korea             | 34.9                          | TNBC                      | H&E               | 10% INC          | DFS, pCR        | 6           |
| Cerbelli 2017 [41]                       | Retrospective | All                 | 54                  | Italy             | –                             | TNBC                      | H&E               | 50%              | pCR             | 8           |
| Khoury 2017 [42]                         | Retrospective | All                 | 331                 | Canadian          | –                             | HER2 positive, TNBC         | H&E               | 10% INC          | pCR             | 6           |
| Ruan 2018 [43]                           | Retrospective | All                 | 166                 | China             | –                             | HER2 positive, TNBC         | H&E               | 10% INC          | pCR             | 8           |
| Yang 2018 [44]                           | Retrospective | All                 | 143                 | China             | 53                            | HER2 positive, TNBC         | H&E               | 10% INC          | OS, DFS, pCR    | 8           |
| de Grootl 2019 [45]                      | Retrospective | CD8⁺, Foxp3⁺       | 196                 | Netherlands       | 55.2                          | HER2 positive, TNBC         | H&E               | –                | DFS, pCR        | 6           |
| Giuseppe 2021 [46]                       | Retrospective | All                 | 445                 | Belgium           | 91.56                         | HER2 positive, TNBC         | H&E               | 30%              | pCR             | 7           |
| Dieci 2020 [47]                          | Retrospective | All                 | 224                 | Italy             | 81.6                          | TNBC                      | H&E               | 30%              | pCR             | 7           |
| Yuan 2021 [48]                           | Retrospective | All                 | 67                  | USA               | 43.7                          | TNBC                      | H&E               | 10%              | pCR, DFS, OS    | 6           |
| Van Bockstal 2020 [49]               | Retrospective | All                 | 35                  | Belgium           | 8                             | HER2 positive, TNBC         | H&E               | 40%              | pCR             | 6           |
| Ha 2021 [50]                             | Retrospective | All                 | 121                 | Italy             | 108                           | HER2 positive, TNBC         | H&E               | 20%              | pCR             | 8           |
| Jimenez 2022 [51]                        | Retrospective | All                 | 80                  | USA               | –                             | HER2 positive, TNBC         | H&E               | 20%              | pCR             | 6           |
survival; OS: overall survival; INC: increment. H&E: hematoxylin and eosin-stained (H&E) sections.

3.2. Literature quality evaluation results

We evaluated the risk of bias for all included studies. For RCTs, bias is principally derived from the method of blinding (participants and researchers). Six studies adopted open labels with a high risk of implementing bias and detecting bias, while five studies did not describe in detail the blind method adopted. The quality evaluation results of retrospective studies were all greater than 6 points. The scoring results are listed in Table 1. The risk of bias assessments for each cohort and evaluations for each domain across full reported studies are shown in Fig. 2A and B.

3.3. The value of TILs for predicting response to NACT in breast cancer

An increased proportion of TILs predicted a higher pCR rate for NACT in total breast cancer, pooled OR = 3.18, 95% CI, 2.55–3.97, P = 0.000, with medium heterogeneity (I² = 49.1%), Fig. 3. According to subgroup analysis of NACT regimens (combined therapy with anthracyclines and paclitaxel or anthracycline-based treatment) and study type (retrospective and prospective studies), results showed that the source of heterogeneity derived from retrospective study (I² = 42.9%), Table 2.

We analyzed the predictive value of TILs in the different molecular subtypes of breast cancer, Fig. 4. The analysis indicated that prior to NACT treatment, TILs had predictive values in HER2-enriched breast cancer and TNBC, pooled OR = 2.54, 95% CI, 1.50–4.29, P = 0.000 and pooled OR = 3.67, 95% CI, 1.93–6.97, P = 0.000. However, when stratified by the hormone receptor (HR) positive breast cancer, no statistical differences in pCR were found in the subgroup analysis, pooled OR = 1.68, 95% CI, 0.67–4.25, P = 0.271. The results of subgroup analysis showed that there was statistically significant heterogeneity among the studies, so meta-regression analysis was performed, and this showed that the source of heterogeneity arose from the different thresholds of TILs across the included studies (P = 0.007).

According to subgroup analysis of TIL subsets, high levels of TIL subtypes (CD8+ and Foxp3+) also predicted better pathological response to NACT, Table 2. However, limited studies analyzed the relationship between CD3+, CD4+ and the pCR rate, and more prospective studies are needed in the future.

3.4. The optimal threshold of TILs to predict pCR

Because the included studies had a wide range of TIL thresholds, we were able to analyze the predictive value of TILs at different thresholds. For patients with HER2 positive breast cancer treated with NACT, a TIL threshold of 10% of the cells was not associated with the pCR outcome prediction, P = 0.813. However, if the studies were separated according to a threshold of ≥20%, and compared with those that used a lower threshold, the high-level group of TILs was correlated with a statistically improved pCR rate, pooled OR = 2.19, 95% CI: 1.06–4.52 (P = 0.035,
Moreover, considering the linear relationship between TIL and pCR (reported in the original literature), in patients with HER2-positive breast cancer treated with NACT, a TIL threshold of $\geq 20\%$ was associated with the most powerful outcome prediction. For patients with TNBC on NACT, TIL thresholds of 10$\%$ and 30$\%$ were not predictive of pCR, $P = 0.342$ and $P = 0.066$, respectively (Fig. 5B). However, the overall compilation of data statistically supported TIL thresholds of 20$\%$, 40$\%$ and 50$\%$ as predictive of pCR in TNBC. Nevertheless, only a single study was included in the 20$\%$ threshold, which indicated that the evidence was not convincing to predict pCR of TNBC. Therefore, the 10$\%$, 20$\%$, and 30$\%$ thresholds were not sufficiently predictive of pCR for TNBC, and the 40$\%$ and 50$\%$ thresholds need to be further verified. Considering the relatively low proportion of breast cancer patients with higher TIL levels in clinical practice, we suggest choosing a TIL threshold of at least 20$\%$ (higher for TNBC) to discriminate pCR from non-pCR subgroups in future clinical trials.

3.5. The prognostic value of TILs for DFS in NACT-treated breast cancer

When TILs were analyzed as a continuous parameter (with incremental increases of 10$\%$), patients with increased TILs concentrations had significantly longer DFS than did patients with lower TIL concentrations, whether they were detected as intra-tumoral (pooled HR $= 0.91$, 95 $\%$CI, 0.84–0.98, $P = 0.020$), stromal (pooled HR $= 0.96$, 95 $\%$CI, 0.93–0.98, $P = 0.003$), or both sites (pooled HR $= 0.95$, 95 $\%$CI, 0.92–0.98, $P = 0.0003$, Fig. 6 A). We also analyzed the prognostic value of TILs in different molecular subtypes of breast cancer. Meta-analysis showed that in HER2 positive breast cancer (pooled HR $= 0.96$, 95 $\%$CI, 0.94–0.98, $P = 0.008$) and TNBC (pooled HR $= 0.89$, 95 $\%$CI: 0.82–0.96, $P = 0.002$), each 10$\%$ increase in TILs was significantly associated with improved DFS. Both iTILs and sTILs were associated with better DFS for HER2 positive breast cancer patients and TNBC patients, Fig. 6 B and C. By contrast, in HR positive breast cancer patients, TILs were not significantly associated with DFS, pooled HR $= 0.64$, 95 $\%$CI: 0.03–14.1, $P = 0.78$. However, this result was limited to only two studies that reported the association between TILs and survival in HR-positive breast cancer. Using subgroup analyses according to TIL subtypes, there was a notably poor DFS for any 10$\%$ increase of

![Fig. 3. Forest plots of the meta-analysis for the efficacy of tumor-infiltrating lymphocytes for predicting response to neoadjuvant chemotherapy in breast cancer.](image)

**Abbreviations:** OR, odds ratio; CI, confidence intervals.

### Table 2

| Subgroup | Sample size | OR (95% CI) | $P$-value | $I^2$ (%) |
|----------|-------------|-------------|-----------|-----------|
| **Subtypes of TILs** | | | | |
| CD8$^+$ | 164 367 | 3.722 (2.038–6.796) | 0.000$^a$ | 0.0 |
| Foxp3$^+$ | 136 362 | 2.35 (1.273–4.347) | 0.006$^a$ | 0.0 |
| **Study type** | | | | |
| Retrospective | 436 621 | 4.011 (2.861–5.623) | 0.000$^a$ | 42.9 |
| Prospective | 337 2291 | 2.821 (2.230–3.568) | 0.000$^a$ | 0.0 |
| **Neoadjuvant chemotherapy regimens** | | | | |
| Mainly | 68 784 | 8.072 (4.225–15.423) | 0.000$^a$ | 0.0 |
| Anthracycline | 469 2006 | 2.818 (2.288–3.470) | 0.000$^a$ | 0.0 |

$^a$ Statistical results are significantly different. High invasion, or high level, was defined as the percentage of Foxp3$^+$ and CD8$^+$ cells infiltrating the peritumoral area that exceeded a predefined threshold. We assessed CD8$^+$ and Foxp3$^+$ as predictive markers for pCR from 3 studies.
Foxp3+ (pooled HR = 1.11, 95% CI: 0.76–1.62). Limited studies analyzed the relationship between CD3+, CD4+, CD8+ and DFS. To assess the impact of each included study, a sensitivity analysis was performed. After excluding each study, the results did not change (P value).

3.6. The prognostic value of TILs for OS in NACT-treated breast cancer

We assessed TILs as a prognostic marker for OS from 6 studies, and the results showed that high levels of TILs showed a favorable OS following treatment with NACT (pooled HR = 0.90, 95% CI: 0.85–0.95, P < 0.0001); both iTILs and sTILs achieved similar results (Fig. 7A). In a subgroup analysis of types of breast cancer, with each 10% increase in TIL level, patients with TNBC (pooled HR = 0.86, 95% CI: 0.81–0.91, P < 0.0001) and HER2 positive breast cancer (pooled HR = 0.93, 95% CI: 0.87–0.99, P = 0.010) had an improved OS (Fig. 7B and C).

3.7. Publication bias

Funnel plot analysis was performed to assess publication bias of the selected studies for the pooled pCR rate. Funnel plot analysis of potential publication bias was presented in Supplementary material. Visual inspection of analysis indicated some evidence of asymmetry, but Egger’s tests indicated that there was no significant publication bias, P > 0.05. Only limited data were available for OS outcome indicators, so the funnel plot was not performed.

4. Discussion

There is strong evidence that high TIL expression of the tumor microenvironment modulates the cancer cell killing effect of NACT. TILs can not only effectively reflect the interaction between the immune microenvironment of the body and tumor cells, but also predict outcome and treatment effect, to provide rational guidance for the formulation and adjustment of clinical treatment plans for breast cancer patients. Elevated levels of TILs are also known as having powerful predictive value in breast cancer. In breast cancers treated with NACT, its predictive implication, especially prognostic effects, have not been fully verified. In an earlier review [15], higher levels of TILs in pre-treatment tumor biopsies demonstrated higher pCR rates in patients treated with NACT, but an analysis of the prognostic role of TILs in breast cancer patients on NACT was not reported. Subsequently, Gao Z-h et al. [52] pointed out that high TILs were predictive of pCR after NAC in only HER2pos and TNBC and were also linked to improved survival. However, the methodological design of this meta-analysis did not specifically limit the treatment option for breast cancer to neoadjuvant chemotherapy. This led to the inclusion of both neoadjuvant and adjuvant
Fig. 5. Forest plots from the meta-analysis for the efficacy of tumor-infiltrating lymphocytes on the neoadjuvant chemotherapy response stratified by different cut-off values. A: HER2 positive breast cancer; B: Triple negative breast cancer. Abbreviations: OR, odds ratio; CI, confidence intervals.
Fig. 6. The forest plot of HRs was assessed for association between TILs and its subtypes and breast cancer on neoadjuvant chemotherapy for disease-free survival. A: total TILs and all breast cancer; B: TILs and HER2 positive breast cancer; C: TILs and triple negative breast cancer. Abbreviations: HR, hazard ratio; CI, confidence intervals; TIL, tumor infiltrating lymphocytes; sTIL, stromal tumor infiltrating lymphocytes; iTIL, intratumoral tumor infiltrating lymphocytes.
The results of our meta-analysis complement, update, and strengthen the findings of the previous reviews. The present meta-analysis reinforced that higher levels of TILs predicted a better response to NACT in HER2 positive breast cancers and TNBC, whereas this was not seen in HR positive breast cancers. The differences in the predictive values of TILs comparing HR positive breast cancer with TNBC may be explained by different immune cell compositions and the predictive effect of each type of immune cell in each molecular subtype of breast cancer. Higher infiltration levels of most immune cell types were observed in TNBC than in HR positive breast cancer, and most immune cells including T cells, B cells, and macrophages were associated with favorable predictive outcomes in TNBC. In contrast, the only cell types associated with improved outcome in HR positive breast cancer were B cells and myeloid dendritic cells. This supports the hypothesis that the cellular composition of immune infiltration in tumors is different among each breast cancer type, which influences clinical outcomes and NACT response [53].

TILs have previously been investigated as predictive factors in breast cancers treated with NACT, but the optimal threshold of TILs to predict pCR in breast cancers treated with NACT has yet to be defined. No formal recommendation for a clinically relevant TIL threshold(s) has been provided by the TILs Working Group [54], nor did the prior...
meta-analyses [15] derive optimal thresholds of TILs to predict pCR or survival in breast cancers treated with NACT. Our study shows that in patients with HER2-positive breast cancer treated with NACT, a TIL threshold of $\geq 20\%$ is associated with the most powerful outcome prediction. But considering the higher level of TIL infiltration in TNBC, we recommend setting a higher threshold to distinguish pCR. Because only a relatively low proportion of breast cancer patients have higher levels of TILs in clinical practice, and the preponderance of evidence suggests that a TIL threshold of $20\%$ can usually be relied upon to discriminate pCR from non-pCR subgroups, this level seems reasonable as a minimum threshold, noting that the predictive value increases linearly with each $10\%$ increment.

As addressed previously, some studies have reported the relationship between TILs and prognosis following treatment with NACT in breast cancer. An earlier meta-analysis [15] only explored the relationship between levels of TILs and the pCR rate of patients with breast cancer treated with NACT; the prognostic value of TILs for this group of patients remained unclear. Our study indicated that per $10\%$ incremental increase in TILs translated to further improvement in DFS and OS, specifically for patients with either TNBC or HER2 positive breast cancer. By contrast, in HR positive breast cancer patients, TILs were not significantly associated with DFS. Regarding the specific site of TILs, the number of iTILs were correlated with the number of sTILs, but typically had a much lower density and therefore were less suitable as a biomarker. Our meta-analysis showed that the value of TILs for prognostic implication was not affected by the specific site of TILs, as both iTILs and sTILs contributed to better prognosis.

Although the usual overall population of TILs is dominated by CD8$^+$ cytotoxic T cells, a variety of subgroups are also included. Differences in the proportions of various lymphocytes in the tumor microenvironment can affect the balance of immune response, leading to different outcomes. Some studies have reported that Foxp3$^+$ TILs can suppress antitumor immune response and lead to escape immune clearance [55]. Therefore, patients with high expression levels of Foxp3$^+$ TILs have a reduced likelihood of obtaining survival benefit from NACT. On the contrary, recruitment of activated CD8$^+$ cytotoxic T cell of NACT is associated with an improved outcome [56–59]. However, our meta-analysis results showed the higher pCR rate of breast cancer patients with high expression levels of Foxp3$^+$ TILs or CD8$^+$ TILs, but a significant association between Foxp3$^+$ TILs and DFS was not observed.

As reported in the results of our meta-analysis, lymphocytes that infiltrate tumors modulate the cancer cell killing effect of chemotherapy, which provides a strong correlation between pre-treatment TIL level and pathological response to NACT. High T cell levels have been demonstrated to be predictive of higher pCR rates and longer survival. However, neoadjuvant chemotherapeutic strategies that have reportedly led to significantly increased pCR rates have not uniformly led to significantly increased survival [60]. Therefore, a composite neoadjuvant clinical trial endpoint that encompasses both residual disease (Residual Cancer Burden; RCB) and TILs, as well as immune response to therapy has been suggested as a better predictor of survival. To develop such a surrogate marker would necessitate large RCTs to collect pCR status, TIL evaluation, event-free and OS data, to understand how such changes would correlate with survival. Unfortunately, to our knowledge, there are no current RCTs to investigate the correlation of TIL levels with outcome, irrespective of pCR status.

5. Limitations

Despite careful attention to the design and performance of the meta-analysis, there still exist some limitations. Some individual study results had considerable heterogeneity, which may be related to differences in the age of patients, as well as the specific clinical research methodology. The present study included only literature published in English which might have led to language bias and study heterogeneity. Additionally, some of the included studies were retrospective, which would inevitably allow bias inherent in that study design. As with all studies, the accuracy of the meta-analysis results was dependent on the accuracy of the original literature research. The TIL thresholds used in the series were those reported in the original research publications; there is no definitive consensus among pathologists to date on the optimal TIL threshold. Due to the limited amount of literature and the wide variation of thresholds reported, an optimal threshold will require validation in subsequent larger study cohorts.

Our data demonstrated a strong correlation between pCR and a TIL threshold of $>20\%$ for HER2$^+$ patients that received NACT, but a higher threshold would be recommended for pCR in TNBC patients. Unfortunately, the original studies included in this meta-analysis did not refine their TIL percentages beyond $10\%$ increments. However, with increasing $10\%$ increments, the data showed increasing correlation of TILs with pCR and correspondingly, with survival outcomes.

6. Conclusions

In summary, increased levels of TILs were associated with an increased frequency of response to NACT and longer survival for patients with TNBC and HER2-positive breast cancer. In contrast, this was not true for patients with HR positive breast cancer. TIL subtypes played different roles in predicting response to NACT. TILs represent a reliable biomarker for both disease prognosis and predictive tumor response in the TNBC and HER2 molecular subgroups of breast cancer. Therefore, in the management of such breast cancer patients, specifically regarding treatment with NACT, TILs should be monitored and reasonably stratified, to adjust treatment strategies.

Until definitive evidence for a TIL threshold has been established, our data would support a $20\%$ threshold to discriminate pCR from non-pCR subgroups in future clinical trials. But for TNBC, we recommend a higher threshold.

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors’ contributions

SQL and YZ conceived of the study and participated in its design. SQL, YZ, PGZ, SJX and YC performed the search and extracted the data. SQL performed the meta-analysis and drafted the manuscript. YC, YZ and PGZ contributed to the final manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Informed consent

Informed consent was obtained from all individual participants included in the study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the other members of the research group for providing language assistance. The authors would like to express their gratitude to EditSprings (https://www.editsprings.cn/) for the expert linguistic services provided.
