Original Article

# Use of peripheral lymphocytes and support vector machine for survival prediction in breast cancer patients

**Fang Bai[1#], Chuanchao Wei[2#], Peng Zhang[1], Dexi Bi[3], Meixin Ge[4], Qing Chen[5], Yijun Jia[6], Yunshu Lu[4], Kejin Wu[1]**

[1]Breast Surgery, Obstetrics and Gynaecology Hospital of Fudan University, Shanghai 200011, China; [2]Department of General Surgery, Central Hospital of Minhang District, Shanghai 201100, China; [3]Department of Pathology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai 200072, China; [4]Department of General Surgery, Xinhua Hospital, Affiliated to Shanghai Jiao Tong University, School of Medicine, Shanghai 200092, China; [5]Department of General Surgery, Shanghai Changzheng Hospital, Shanghai 200003, China; [6]Department of Oncology, Fudan University Shanghai Cancer Center, Shanghai 200032, China

*Contributions*: (I) Conception and design: K Wu; (II) Administrative support: Y Lu; (III) Provision of study materials or patients: F Bai, C Wei; (IV) Collection and assembly of data: Q Chen, M Ge; (V) Data analysis and interpretation: Y Jia, P Zhang, D Bi; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

*Correspondence to:* Dr. Kejin Wu. Breast Surgery, Obstetrics and Gynaecology Hospital of Fudan University, No. 419, Fangxie Road, Shanghai 200011, China. Email: kejinwu@163.com; Dr. Yunshu Lu. Department of General Surgery, Xinhua Hospital, Affiliated to Shanghai Jiao Tong University, School of Medicine, No. 1665, Kongjiang Road, Shanghai 200092, China. Email: luyunshu@qq.com.

**Background:** This study aimed to identify the influence of peripheral lymphocytes on prognosis and find prognostic markers for breast cancer patients.

**Methods:** This study enrolled invasive breast cancer patients and they were followed-up for median 4-years over telephone. Distributions of disease-free survival (DFS) and overall survival (OS) between different levels of lymphocytes were estimated with the Kaplan-Meier (K-M) method. Support vector machine (SVM) methods were used to develop a prognostic classifier for breast cancer.

**Results:** A total of 190 patients were enrolled. Patients with low level of cluster of differentiation (CD)3+ lymphocytes had worse DFS and OS (P<0.05). Strong association was reported between SVM-DFS model and DFS (sensitivity, 97%; specificity, 75%); whereas the SVM-OS model was strongly associated with OS (sensitivity, 67%; specificity, 100%).

**Conclusions:** Patients with low level of CD3+ lymphocytes could have a poorer survival and the SVM method could predict prognosis in breast cancer patients.

**Keywords:** Peripheral lymphocytes; breast cancer; cluster of differentiation (CD)3+; support vector machine (SVM); prognosis

Submitted Dec 23, 2017. Accepted for publication Apr 27, 2018.
doi: 10.21037/tcr.2018.07.08
**View this article at:** http://dx.doi.org/10.21037/tcr.2018.07.08

## Introduction

Burden of breast cancer is high among women worldwide (1), manifested with a great heterogeneity across the range of clinical patterns, biologic behaviors and prognostic characteristics (2). Immune system plays a crucial role in the pathogenesis, development and progression of cancer (3). T-lymphocytes, B-lymphocytes and natural killer (NK) cells comprise the peripheral lymphocyte system (4). Furthermore, T-cell-related immune responses are significant to immunotherapy as they protect the host from tumorigenesis and tumor progression (5). An active cellular immunity may induce apoptosis of tumor cells and thereby maintain homeostasis. Imbalance of cellular or humoral immunity may trigger tumor progression and subsequently

lead to treatment failure (6-10).

Previous research reported that breast cancer patients have impaired immune system with decreased T cell proliferation, reduced cluster of differentiation (CD)3+ and CD4+ count but increased CD8+ count, and low CD4+/CD8+ ratio, compared with normal individuals (11,12). However, Caras *et al.* showed decreased CD8+ cells compared to healthy controls (13). Different tumor staging and varied immune status may attribute to these differences. However, whether the compromised immune function influences the prognosis of breast cancer patients remains elusive. Hence, we retrospectively investigated the patient medical records to evaluate the levels of the lymphocyte subsets in peripheral blood lymphocyte samples of breast cancer patients who visited the Xinhua Hospital affiliated to Shanghai Jiao Tong University, School of Medicine and Central Hospital of Minhang District, Shanghai, in order to analyze the relation between the immune status and prognosis of breast cancer in these patients. Furthermore, support vector machine (SVM)-based prognostic classifiers were established for clinical diagnosis of breast cancer.

## Methods

### *Patients*

Patients who had been diagnosed with invasive breast cancer (without metastatic diseases) at Xinhua Hospital affiliated to Shanghai Jiao Tong University, School of Medicine, between January 2010 and December 2012 and Central Hospital of Minhang District, Shanghai between June 2015 and December 2016 were enrolled. Patients who had confirmation of primary breast cancer by clinical manifestation and pathological examination of breast; no previous chemotherapy or radiotherapy; who were informed of the study procedures and signed informed consent forms; no older than 70 years of age were included in the study. Patients with abnormal liver or kidney function; metastatic disease or other tumors; immune system-related disease or recent history of infection/inflammation were excluded. The study was approved by the Research Ethics Committee of Xinhua Hospital affiliated to Shanghai Jiao Tong University, School of Medicine and Central Hospital of Minhang District, Shanghai.

Venous blood (10 mL) was obtained and enriched from the patients before surgery. CD3+ (clone: UCHT1), CD4+ (clone: SK3), CD8+ (clone: RPA-T8), and NK cells (CD56+, clone: NCAM16.2), CD19+ (clone: HIB19), CD20+ (clone: 2H7) were analyzed and counted using the flow cytometer (BD Bioscience, USA), as described previously (14). The patients were followed up by telephone till April 2017. The information of patients' survival and recurrence was recorded in case report forms.

### *Study objectives*

The primary objective of our study was to identify if peripheral lymphocytes could indicate prognosis of breast cancer patients. Secondary objective was to establish SVM classifiers for prognosis of breast cancer.

### *Survival analyses*

Based on the lymphocyte count (percentage), patients were divided into three groups, low, median and high, according to clinical practice and guidelines (14). Kaplan-Meier (K-M) curves were plotted to estimate the disease-free survival (DFS) and overall survival (OS) among different groups which were compared using the log-rank test. In addition, univariate and multivariate analysis of prognostic factors were performed by Cox's regression model. The levels of CD3+, CD4+, CD8+, CD4+/CD8+ ratio and NK cells in peripheral blood between preoperative (n=73) and those of 7 days after surgery (n=73) are analyzed by independent *t*-test (P value, two-tailed). The levels of CD3+, CD4+, CD8+, CD4/CD8+ ratio and NK cells in peripheral blood among preoperative (n=24), 7 days after surgery (n=24) and 1-month follow-up (n=24) are analyzed by analysis of variance (ANOVA). A two-sided P value less than 0.05 was considered statistically significant. All statistical analyses were performed on the SPSS 18.0 (SPSS Inc., Chicago, IL, USA) software.

### *Selection of cutoff scores*

For peripheral lymphocytes, the sensitivity and specificity of each outcome (DFS or OS) of breast cancer within 5 years were plotted to generate a receiver operating characteristic (ROC) curve. These curves were used to select cutoff scores for dichotomizing each predictor based on maximum cases under the ROC curve (i.e., score nearest to point on curve (0.0, 1.0) with maximum sensitivity and specificity).

### *Prognosis prediction using SVM-based methods*

The SVM was introduced by Vapnik for data classification

and function approximation in 1999 (15). Apart from linear classification, SVM can efficiently perform non-linear classification with a class of algorithms for pattern analysis that is known as kernel trick, to implicitly map inputs into high-dimensional feature spaces, which could achieve robust classification when traditional statistical methods could not. It has recently been used in medical researches to deal with big and complex data when traditional statistical methods are not suitable. The machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Two SVM models, SVM-OS and SVM-DFS were developed to predict OS and DFS, respectively. We addressed the prognostic prediction of breast cancer at two-class classification levels (i.e., whether a patient can have a DFS or OS for more than 5 years). Patients' basic information, including age, family history, menopausal status, complications, tumor grade, TNM stage, axillary nodes and cancer subtype, peripheral blood lymphocytes level and follow-up results was used as inputs. Patients were randomly divided into training cohort (n=95) and validation cohort (n=95). Same method was used to scale both training and testing data. Radial basis function (RBF) kernel, a popular kernel function used in various kernelized learning algorithms, was employed here. The RBF kernel was used: $k(x) = \exp[-x^2/(2\times \text{sigma})]$. The best values of C and γ, two key RBF kernal-associated parameters that govern SVM performance were found using five-fold cross-validation according to standard protocol. Then, the best parameter C and γ (C= and γ= for SVM-OS; C= and γ= for SVM-DFS) to were used to train SVMs with the training dataset. In SVM-DFS and SVM-OS, the values of sigma were 0.038 and 0.040, respectively, which were automatic optimized. The sensitivity and specificity of each SVM were calculated by ROC curve. The programs were coded using Matlab software (MathWorks, Natick, MA, USA) and Matlab scripts are available on request (16,17).

## Results

### *Patient characteristics*

A total of 190 invasive breast cancer patients were enrolled in this study. The characteristics of the patients divided into training cohort (n=95) and validation cohort (n=95) are listed in *Table 1*. The characteristics included: sex, age, menopausal status, TNM stage, tumor grade, axillary nodes, cancer subtype, follow-up time, leukocyte cell count and CD3+ range. Generally, the distributions of the above

**Table 1** Characteristics of patients

| Characteristic | Training cohort (n=95) | | Validation cohort (n=95) | |
|---|---|---|---|---|
| | No. | % | No. | % |
| Sex | 95 | | 95 | |
| Female | 94 | 98.95 | 93 | 97.89 |
| Male | 1 | 1.05 | 2 | 2.11 |
| Age | | | | |
| Median | 54 | | 53 | |
| Range | 32–70 | | 26–70 | |
| Menopausal status* | | | | |
| Premenopausal | 35 | 37.23 | 35 | 37.63 |
| Postmenopausal | 59 | 62.77 | 58 | 62.37 |
| TNM stage | | | | |
| I | 31 | 32.63 | 28 | 29.47 |
| II | 45 | 47.36 | 49 | 51.58 |
| III | 14 | 14.74 | 13 | 13.68 |
| IV | 4 | 4.21 | 5 | 5.27 |
| Unknown | 1 | 1.06 | 0 | 0 |
| Tumor grade | | | | |
| I | 2 | 2.11 | 1 | 1.05 |
| II | 35 | 36.84 | 39 | 41.05 |
| III | 36 | 37.90 | 37 | 38.95 |
| Unknown | 22 | 23.15 | 18 | 18.95 |
| Axillary nodes | | | | |
| 0 | 46 | 48.42 | 49 | 51.58 |
| 1–3 | 17 | 17.90 | 16 | 16.84 |
| >3 | 20 | 21.05 | 18 | 18.95 |
| Unknown | 12 | 12.63 | 12 | 12.63 |
| Subtype | | | | |
| Luminal A | 28 | 29.47 | 28 | 29.47 |
| Luminal B | 42 | 44.21 | 42 | 44.21 |
| Her2 like | 14 | 14.74 | 15 | 15.79 |
| TNBC | 9 | 9.48 | 8 | 8.43 |
| Unknown | 2 | 2.10 | 2 | 2.10 |
| Follow-up time (months) | | | | |
| Median | 49 | | 47 | |
| Range | 5–91 | | 5–91 | |

**Table 1** (*continued*)

**Table 1** (*continued*)

| Characteristic | Training cohort (n=95) | | Validation cohort (n=95) | |
|---|---|---|---|---|
| | No. | % | No. | % |
| Leukocyte cell count ($1\times10^9$) | | | | |
| Median | 6.0 | | 6.1 | |
| Range | 1.8–13.9 | | 2.2–16.6 | |
| CD3+ range (%) | | | | |
| Median | 71.0 | | 72.7 | |
| Range | 47.7-88.2 | | 44.9-85.7 | |

*, male patients are excluded in this analyze. TNBC, triple negative breast cancer.

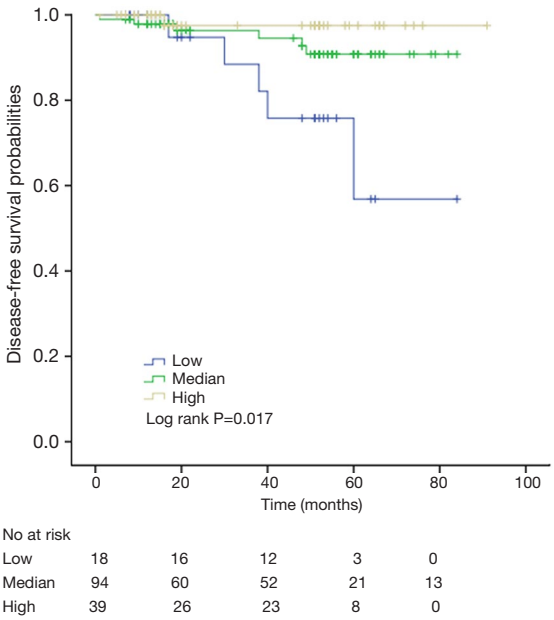



**Figure 1** The K-M curve of CD3+ lymphocytes in peripheral blood of breast cancer patients' DFS. Data was analyzed using K-M plotters (n=190). K-M, Kaplan-Meier; DFS, disease-free survive.

aspects of two groups are similar.

### *Survival analyses*

The K-M curves are depicted in *Figures 1*,*2*. Compared to the median and high CD3+ lymphocytes count groups, patients of lower lymphocyte count group had lower DFS (mean 66.8 months *vs.* medium and high group 79.1 and 89.1 months) and OS (mean 66.5 months *vs.* medium and high 78.0 and 82.5 months) (log rank P<0.05). However, of the DFS and OS did not vary significantly between patients with low, median and high CD4+, CD8+, CD4/CD8+ ratio, and NK count (*Figures S1-S8*).

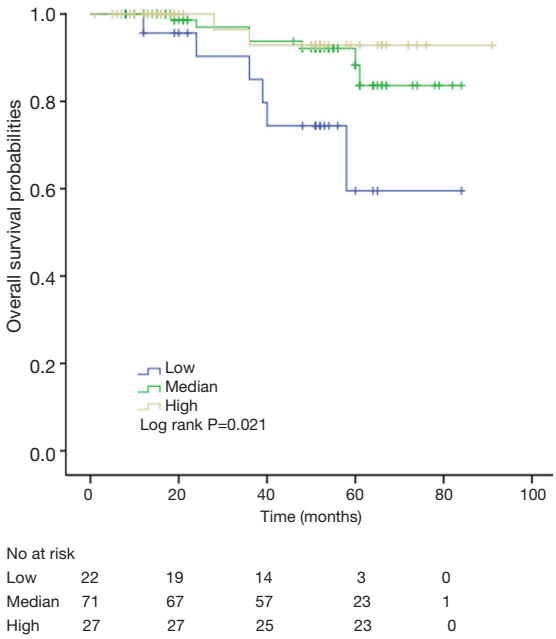



**Figure 2** The K-M curve of CD3+ lymphocytes in peripheral blood of breast cancer patients' OS. Data was analyzed using K-M plotters (n=190). K-M, Kaplan-Meier; OS, overall survival.

In a series of univariate analyses that included covariates such as CD3+ count, CD4+ count, CD8+ count, CD4+/CD8+ radio, NK count, tumor grade, TNM stage, axillary nodes, history and menopausal status, only CD3+ count and stage showed with a significant higher risk of DFS and OS prognosis, as shown in *Table 2*. The RR of poor DFS prognosis was nearly 10-fold higher in the low CD3+ count group compared to the group with high CD3+ count (P<0.045). These results were reproducible in the multivariate analysis as shown in *Table 3*.

Among patients with low level of CD3+, K-M survival analyses by log-rank test were performed between patients with different tumor grade, TNM stage, axillary nodes or cancer subtype. Only axillary nodes and TNM stage showed a significant association with DFS. In addition, only TNM stage showed a significant association with OS, as is shown in *Figures 3-5* (log rank, P<0.05).

The levels of CD3+, CD4+, CD8+, CD4+/CD8+ ratio and NK cells in peripheral blood between preoperative and those of 7 days after surgery are shown in *Table 4*. The levels

**Table 2** Prognostic value of clinic-pathologic factors on DFS and OS of 190 patients with invasive breast cancer

| Characteristics | DFS | | OS | |
|---|---|---|---|---|
| | RR (95% CI) | P value | RR (95% CI) | P value |
| CD3+ | | 0.045 | | 0.012 |
| Low | 9.993 (1.166–85.638) | 0.036 | 10.922 (1.312–90.092) | 0.027 |
| Median | 3.137 (0.377–26.101) | 0.290 | 2.419 (0.282–20.756) | 0.420 |
| High | 1 (Ref.) | | 1 (Ref.) | |
| CD4+ | | 0.331 | | 0.810 |
| Low | 6.633 (0.414–106.175) | 0.181 | 2.056 (0.214–19.799) | 0.533 |
| Median | 4.268 (0.546–33.373) | 0.167 | 1.073 (0.284–4.048) | 0.917 |
| High | 1 (Ref.) | | 1 (Ref.) | |
| CD8+ | | 0.934 | | 0.965 |
| Low | 1,0250.929 (0–9.792e$^{99}$) | 0.935 | 8,050.456 (0–3.73e$^{105}$) | 0.940 |
| Median | 8,301.450 (0–7.926e$^{99}$) | 0.936 | 9,409.537 (0–4.36e$^{105}$) | 0.939 |
| High | 1 (Ref.) | | 1 (Ref.) | |
| CD4+/CD8+ radio | | 0.753 | | 0.984 |
| Low | 0 | 0.983 | 0 | 0.983 |
| Median | 1.555 (0.492–4.909) | 0.452 | 1.108 (0.357–3.436) | 0.860 |
| High | 1 (Ref.) | | 1 (Ref.) | |
| NK | | 0.288 | | 0.291 |
| Low | 0.232 (0.27–2.030) | 0.187 | 0.257 (0.30–2.221) | 0.217 |
| Median | 0.477 (0.145–1.565) | 0.222 | 0.455 (0.139–1.495) | 0.195 |
| High | 1 (Ref.) | | 1 (Ref.) | |
| Tumor grade | | 0.952 | | 0.397 |
| I | 0 | 0.986 | 0 | 0.989 |
| II | 0.832 (0.262–2.642) | 0.755 | 0.414 (0.116–1.475) | 0.174 |
| III | 1 (Ref.) | | 1 (Ref.) | |
| TNM stage | | 0.035 | | 0.049 |
| I | 0.065 (0.011–0.754) | 0.016 | 0.178 (0.015–0.978) | 0.017 |
| II | 0.231 (0.011–1.646) | 0.123 | 0.457 (0.089–1.427) | 0.146 |
| III | 1 (Ref.) | | 1 (Ref.) | |
| Axillary nodes | | 0.163 | | 0.127 |
| 0 | 0.358 (0.109–1.177) | 0.091 | 0.233 (0.052–1.045) | 0.057 |
| 1–3 | 0.229 (0.027–1.962) | 0.179 | 0.879 (0.197–3.932) | 0.866 |
| >3 | 1 (Ref.) | | 1 (Ref.) | |
| History | | | | |
| No | 0.604 (0.788–4.702) | 0.630 | 21.825 (0–1967845.104) | 0.596 |
| Yes | 1 (Ref.) | | 1 (Ref.) | |
| Menopausal status | | | | |
| Premenopausal | 0.471 (0.102–2.169) | 0.334 | 0.029 (0–6.436) | 0.200 |
| Postmenopausal | 1 (Ref.) | | 1 (Ref.) | |

DFS, disease-free survival; OS, overall survival; RR, risk ratio; CI, confidence interval; NK, natural killer.

**Table 3** Independent predictors of DFS and OS in multivariate analysis of 190 patients with breast cancer

| Characteristics | DFS | | OS | |
|---|---|---|---|---|
| | RR (95% CI) | P value | RR (95% CI) | P value |
| CD3 | | 0.053 | | 0.025 |
| Low | 8.632 (0.972–76.629) | 0.053 | 9.501 (1.110–81.356) | 0.040 |
| Median | 2.574 (0.302–21.492) | 0.390 | 2.314 (0.267–20.034) | 0.446 |
| High | 1 (Ref.) | | 1 (Ref.) | |
| TNM stage | | 0.042 | | 0.037 |
| I | 0.799 (0.009–0.826) | 0.005 | 0.316 (0.068–1.235) | 0.051 |
| II | 0.427 (0.089–1.211) | 0.062 | 0.422 (0.134–1.366) | 0.113 |
| III | 1 (Ref.) | | 1 (Ref.) | |

DFS, disease-free survival; OS, overall survival; RR, risk ratio; CI, confidence interval.

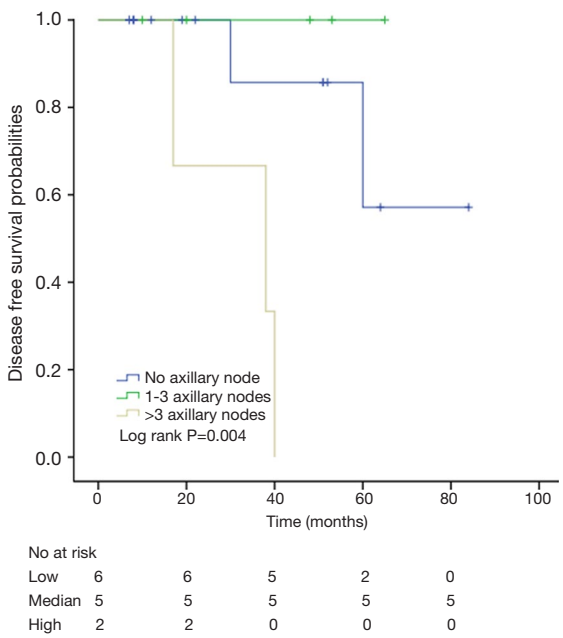



**Figure 3** The patients whose CD3+ level in peripheral blood is low were selected, and the DFS of patients in different groups were compared (n=23). Data was analyzed using K-M plotters. K-M, Kaplan-Meier; DFS, disease-free survive.

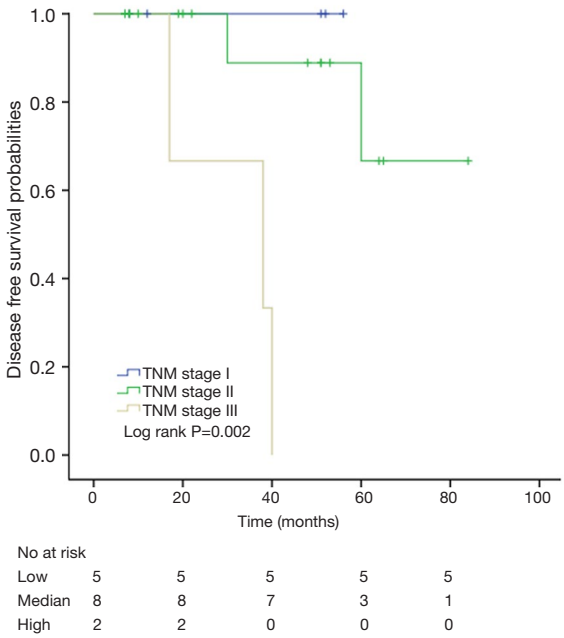



**Figure 4** The patients whose CD3+ level in peripheral blood is low were selected, and the DFS of patients in different groups were compared (n=25). Data was analyzed using K-M plotters. The K-M curve of TNM stage of breast cancer patients' DFS is shown. K-M, Kaplan-Meier; DFS, disease-free survive.

of CD3+, CD4+, CD8+, CD4/CD8+ ratio and NK cells in peripheral blood among preoperative, 7 days after surgery and 1-month follow-up are shown in *Table 5*. These results demonstrated that there is no significant difference between different time points.

### *SVM-based prognostic classifiers for prediction of breast cancer survival*

The SVM-DFS model had an area under curve (AUC) in ROC of 0.8578 which could efficiently predict a 5-year DFS (sensitivity, 97%; specificity, 75%). In addition, SVM-OS

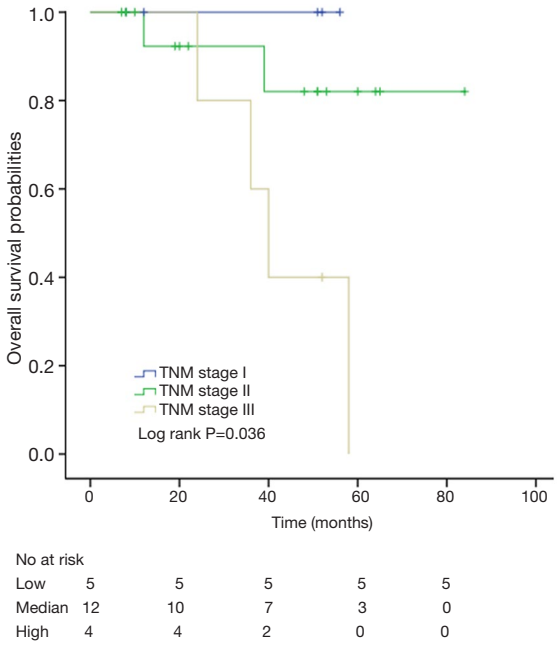



**Figure 5** The patients whose CD3+ level in peripheral blood is low were selected, and the OS of patients in different groups were compared (n=28). Data was analyzed using K-M plotters. K-M, Kaplan-Meier; OS, overall survival.

**Table 4** Independent t test of peripheral lymphocyte levels between before surgery and that of 7 days after surgery (n=73)

| Lymphocytes | t value | P value (two-tailed) |
|---|---|---|
| CD3+ | 0.232 | 0.817 |
| CD4+ | –0.157 | 0.875 |
| CD8+ | 0.501 | 0.617 |
| CD4+/CD8+ | –0.379 | 0.705 |
| NK | 0.104 | 0.918 |

NK, natural killer.

**Table 5** One-way ANOVA of peripheral lymphocyte levels between before surgery, 7 days after surgery and 1 month follow up (n=24)

| Lymphocytes | F | P value | SNK test P value |
|---|---|---|---|
| CD3+ | 0.269 | 0.765 | 0.752 |
| CD4+ | 1.647 | 0.200 | 0.237 |
| CD8+ | 1.740 | 0.183 | 0.175 |
| CD4+/CD8+ ratio | 0.718 | 0.492 | 0.510 |
| NK | 0.824 | 0.443 | 0.417 |

ANOVA, analysis of variance; SNK test, Student-Newman-Keuls test; NK, natural killer.

had an AUC of 0.8333 which was also strongly associated with 5-year prognosis of OS (sensitivity, 67%; specificity, 100%), as is shown in *Figures 6*,7.

## Discussion

The findings of this study reveal that breast cancer patients with low level of CD3+ lymphocytes in peripheral blood have poorer DFS and OS, compared to median and high level of CD3+ groups. While the levels of CD4+, CD8+, CD4+/CD8+ ratio and NK are of no statistical significance. In addition, the SVM-based prognostic classifiers established in this study could predict 5-year-DFS or 5-year-OS for breast cancer patients when medical information of age, menopausal status, family history, complications, TNM stage, tumor grade, axillary nodes, cancer subtype and peripheral lymphocytes are input. According to the ROC curve, the sensitivity and specificity of the SVM-DFS and SVM-OS are relatively high.

Lymphocytopenia is an unfavorable prognostic factor for survival of breast cancer patients. The attenuation of immune-surveillance, mediated by NK cells and/or cells with lymphokine-activated killer activity, was believed to play an important role in tumor dissemination mechanisms and could exacerbate the ability of circulating tumor cells (CTCs) to promote metastasis (18,19).

Furthermore, studies have shown that high neutrophil-to-lymphocyte ratio (NLR) may be indicative of inflammation and it may associate with poor prognosis in breast cancer patients (20,21). Neutrophils have been shown to inhibit the immune system and promote tumor growth by suppressing the activity of lymphocytes and T-cell response (22,23). By contrast, increased lymphocytic tumor infiltration could improve survival, especially in ER-negative/HER2-negative breast cancer (24). As NLR is an inexpensive and readily available prognostic marker, it may bring refinement to risk estimates.

Previous studies showed that compared to normal people, cancer patients had decreased CD3+, CD4+ count, but elevated CD8+ count (11,12,25-27). In addition, stage IV breast cancer and the Her-2/VEGF-positive breast cancer patients have worsened immune function (12) and poor survival prognosis. Furthermore, level of CD4+ CD25+ Foxp3 and CD8+ CD28-in peripheral blood of

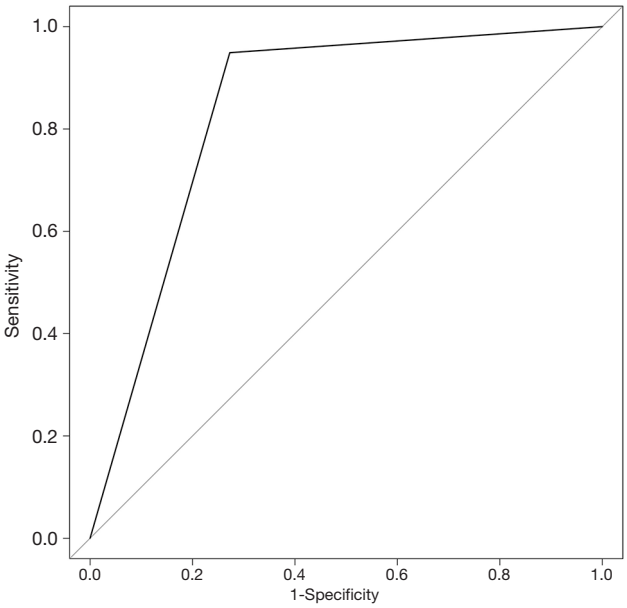



**Figure 6** The ROC curve of SVM-DFS. SVM-DFS indicates the prognostic classifiers for prediction of DFS in breast cancer patients. ROC, receiver operating characteristic; SVM-DFS, support vector machine-disease free survival; DFS, disease-free survive.

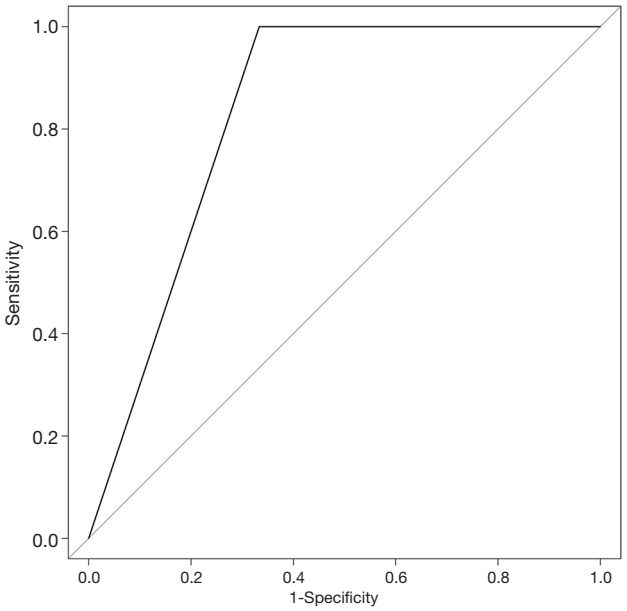



**Figure 7** The ROC curve of SVM-OS. SVM-OS indicates the prognostic classifiers for prediction of OS in breast cancer patients. ROC, receiver operating characteristic; SVM-OS, support vector machine-overall survival; OS, overall survival.

patients could be a prognostic indicator of OS (11,27).

But few researchers have focused on the comparisons between the prognoses of patients with different peripheral lymphocyte levels. According to our results, we find that patients with low level of CD3+ lymphocytes in peripheral blood tend to have poor DFS and OS. However, the results of CD4+, CD8+, CD4+/CD8+ ratio and NK fail to show statistical significance. The reason could be that, firstly, these indexes may not indicate prognosis indeed. Secondly, the study sample may not big enough to support this idea. Thirdly, the follow-up time is not long enough. We hope that more conclusions could be reached during our future further research by a 10- or 15-year follow-up and more patients included in. In addition, the lymphocyte level may vary during time. However, to minimize the interference of this on the findings of study we excluded patients with inflammation which could influence the level of lymphocytes and lead to selection bias. According to the related papers, the venous blood was drawn once before surgery from each patient to test lymphocytes (11,12,26,27). The lymphocytes levels before surgery are the best to clarify in the study. The levels of lymphocytes may vary and be influenced by inflammation after surgery. Moreover, the comparisons between the peripheral lymphocytes of patients before surgery, 7 days after surgery, and 1-month follow-up were made by One-way ANOVA. The results indicate that there is no statistical significance ($P>0.05$). Furthermore, the SVM models we built could be beneficial to predict breast cancer patients' prognosis.

The readers should infer the conclusions in light of some limitations of the study. First, the lymphocyte subsets were measured in the peripheral blood of patients and lymphocytes in blood account only a part of that in the total lymphatic circulation (28) and hence, there may be a bias in the cell count which could have influenced the results of the study. Second, the small sample size may not be adequately powered to generalize the findings to a wider population. Third, the data pertaining to the training and test groups are limited and only we report the establishment of SVM classifiers. Fourth, telephonic follow up was chosen for this study, which would have hindered accurate collection of patient wellness, had the patient visited the clinic for follow up would have given one more reliable information. Fifth, we studied a relatively shorter median follow-up time of 4 years. Finally, the SVM-based prognostic classifiers could predict survival of cancer patients, however,

with limited sensitivities and specificities. However, we need to note that no SVM model is developed so far to achieve 100% precision.

To the best of our knowledge, it is the first article to report association between different levels of peripheral blood lymphocytes and prognosis of cancer patients and it is the first SVM-based prognostic classifiers established to predict survival of breast cancer patients. We should know that, the lymphocytes level could not be the only ingredient that has an effect on survival, so we included all the medical information available in our SVM-based model. Given a shorter follow up time in this study, we warrant future long-term studies to further validate the findings.

It is demonstrated that reduced immune function may lead to carcinoma. So effective measures should be taken to control immune responses against cancer in order to achieve durable responses and if possible, complete eradication of cancer in patients safely. However, the incompletely understood human immunology imposes a major challenge to researches. By understanding the biologic behaviors of individual patients, immune-specific anti-cancer treatment can be strategically tailored. During the decade, antitumor cytotoxic T cell responses are activated by blocking cytotoxic T lymphocyte antigen-4 and programed cell death protein-1 (6). Therefore, targeting the immune system function may be a best approach in treating breast cancer patients (29,30).

## Acknowledgments

*Funding:* The research was funded by Science and Technology Commission of Shanghai Municipality (No. 14411950206 and No. 17411961100).

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr.2018.07.08). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Research Ethics Committee of Xinhua Hospital affiliated to Shanghai Jiao Tong University: XHEC-D-2018-026, School of Medicine and Central Hospital of Minhang District, Shanghai: MCHEC-2015-040. Patients signed informed consent forms.



## References

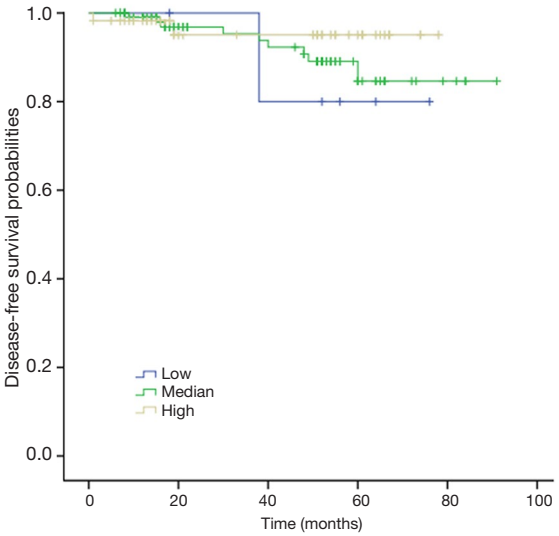
1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359-86.
2. Voduc KD, Cheang MC, Tyldesley S, et al. Breast cancer subtypes and the risk of local and regional relapse. J Clin Oncol 2010;28:1684-91.
3. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. Cell 2010;140:883-99.
4. Rankin L, Groom J, Mielke LA, et al. Diversity, function, and transcriptional regulation of gut innate lymphocytes. Front Immunol 2013;4:22.
5. Mockler MB, Conroy MJ, Lysaght J. Targeting T cell immunometabolism for cancer immunotherapy; understanding the impact of the tumor microenvironment. Front Oncol 2014;4:107.
6. Swart M, Verbrugge I, Beltman JB. Combination Approaches with Immune-Checkpoint Blockade in Cancer Therapy. Front Oncol 2016;6:233.
7. Pearce EL, Poffenberger MC, Chang CH, et al. Fueling immunity: insights into metabolism and lymphocyte function. Science 2013;342:1242454.
8. Voutsadakis IA. Immune Blockade Inhibition in Breast Cancer. Anticancer Res 2016;36:5607-22.
9. Molon B, Cali B, Viola A. T Cells and Cancer: How Metabolism Shapes Immunity. Front Immunol 2016;7:20.
10. Falkenberg KJ, Johnstone RW. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. Nat Rev Drug Discov 2014;13:673-91.
11. Song G, Wang X, Jia J, et al. Elevated level of peripheral CD8(+)CD28(-) T lymphocytes are an independent predictor of progression-free survival

in patients with metastatic breast cancer during the course of chemotherapy. Cancer Immunol Immunother 2013;62:1123-30.
12. Wang ZK, Yang B, Liu H, et al. Regulatory T cells increase in breast cancer and in stage IV breast cancer. Cancer Immunol Immunother 2012;61:911-6.
13. Caras I, Grigorescu A, Stavaru C, et al. Evidence for immune defects in breast and lung cancer patients. Cancer Immunol Immunother 2004;53:1146-52.
14. Jia Y, Xu L, Lin Q, et al. Levels of lymphocyte subsets in peripheral blood prior treatment are associated with aggressive breast cancer phenotypes or subtypes. Med Oncol 2014;31:981.
15. Vapnik VN. An overview of statistical learning theory. IEEE Trans Neural Netw 1999;10:988-99.
16. Zhu ZH, Sun BY, Ma Y, et al. Three immunomarker support vector machines-based prognostic classifiers for stage IB non-small-cell lung cancer. J Clin Oncol 2009;27:1091-9.
17. Wang HY, Sun BY, Zhu ZH, et al. Eight-signature classifier for prediction of nasopharyngeal carcinoma survival. J Clin Oncol 2011;29:4516-25.
18. De Giorgi U, Mego M, Scarpi E, et al. Relationship between lymphocytopenia and circulating tumor cells as prognostic factors for overall survival in metastatic breast cancer. Clin Breast Cancer 2012;12:264-9.
19. Tredan O, Manuel M, Clapisson G, et al. Patients with metastatic breast cancer leading to CD4+ T cell lymphopaenia have poor outcome. Eur J Cancer 2013;49:1673-82.
20. Ethier JL, Desautels D, Templeton A, et al. Prognostic role of neutrophil-to-lymphocyte ratio in breast cancer: a systematic review and meta-analysis. Breast Cancer Res 2017;19:2.
21. Asano Y, Kashiwagi S, Onoda N, et al. Predictive Value of Neutrophil/Lymphocyte Ratio for Efficacy of Preoperative Chemotherapy in Triple-Negative Breast Cancer. Ann Surg Oncol 2016;23:1104-10.
22. De Larco JE, Wuertz BR, Furcht LT. The potential role of neutrophils in promoting the metastatic phenotype of tumors releasing interleukin-8. Clin Cancer Res 2004;10:4895-900.
23. el-Hag A, Clark RA. Immunosuppression by activated human neutrophils. Dependence on the myeloperoxidase system. J Immunol 1987;139:2406-13.
24. Loi S, Sirtaine N, Piette F, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. J Clin Oncol 2013;31:860-7.
25. He L, Liang JH, Wu JZ, et al. Low absolute CD4+ T cell counts in peripheral blood are associated with inferior survival in follicular lymphoma. Tumour Biol 2016;37:12589-95.
26. Liu HZ, Deng W, Li JL, et al. Peripheral blood lymphocyte subset levels differ in patients with hepatocellular carcinoma. Oncotarget 2016;7:77558-64.
27. Ke X, Zhang S, Xu J, et al. Non-small-cell lung cancer-induced immunosuppression by increased human regulatory T cells via Foxp3 promoter demethylation. Cancer Immunol Immunother 2016;65:587-99.
28. Ilyas S, Yang JC. Landscape of Tumor Antigens in T Cell Immunotherapy. J Immunol 2015;195:5117-22.
29. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity 2013;39:1-10.
30. Mahmoud SM, Paish EC, Powe DG, et al. Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. J Clin Oncol 2011;29:1949-55.

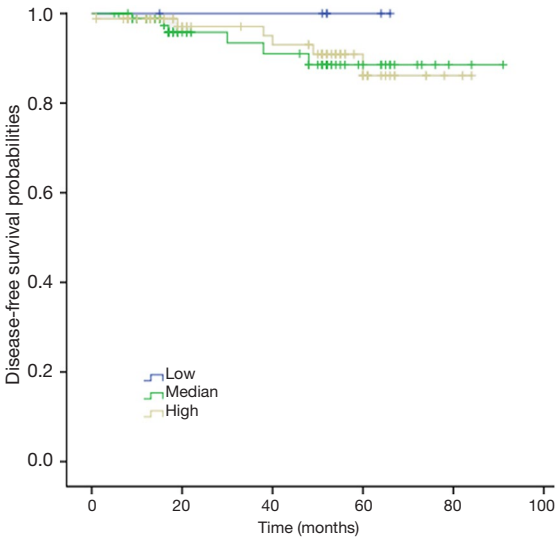
**Cite this article as:** Bai F, Wei C, Zhang P, Bi D, Ge M, Chen Q, Jia Y, Lu Y, Wu K. Use of peripheral lymphocytes and support vector machine for survival prediction in breast cancer patients. Transl Cancer Res 2018;7(4):978-987. doi: 10.21037/tcr.2018.07.08



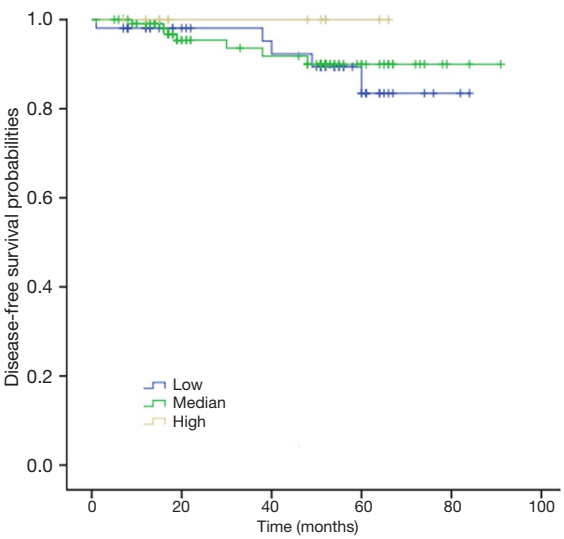

| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 1 | 4 | 4 | 3 | 1 |
| Median | 107 | 72 | 61 | 19 | 3 |
| High | 56 | 26 | 23 | 12 | 0 |

**Figure S1** The K-M curve of CD4+ lymphocytes in peripheral blood of breast cancer patients' DFS. Data was analyzed using Kaplan-Meier Plotters (n=190, P>0.05). K-M, Kaplan-Meier; DFS, disease-free survival.



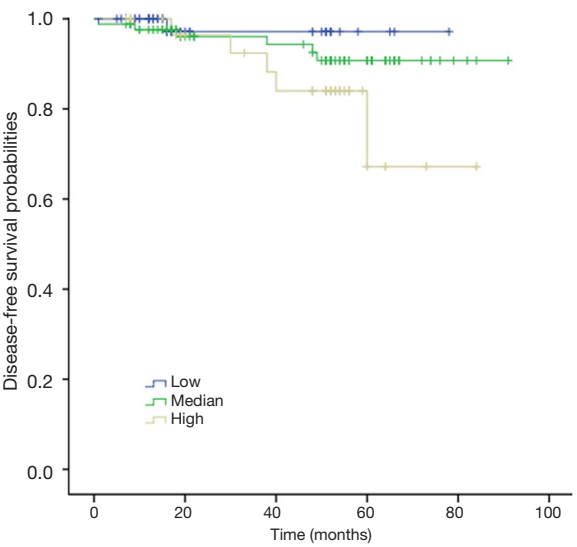

| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 6 | 6 | 6 | 6 | 6 |
| Median | 83 | 45 | 37 | 13 | 1 |
| High | 83 | 53 | 46 | 18 | 1 |

**Figure S3** The K-M curve of CD4+/CD8+ ratio lymphocytes in peripheral blood of breast cancer patients' DFS. Data was analyzed using K-M plotters (n=190, P>0.05). K-M, Kaplan-Meier; DFS, disease-free survival.



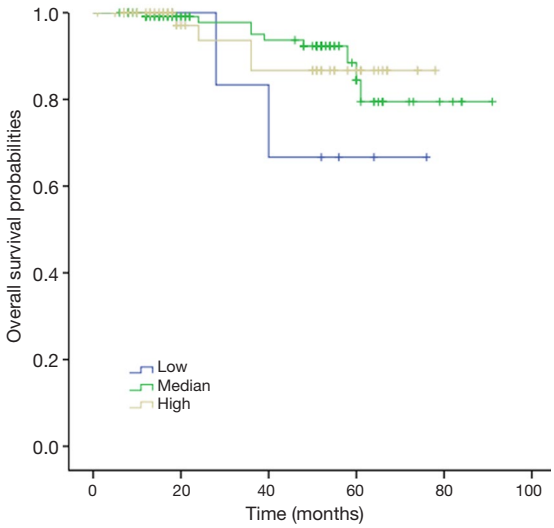

| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 51 | 36 | 32 | 14 | 1 |
| Median | 106 | 62 | 50 | 17 | 1 |
| High | 12 | 12 | 12 | 12 | 12 |

**Figure S2** The K-M curve of CD8+ lymphocytes in peripheral blood of breast cancer patients' DFS. Data was analyzed using K-M plotters (n=190, P>0.05). K-M, Kaplan-Meier; DFS, disease-free survival.



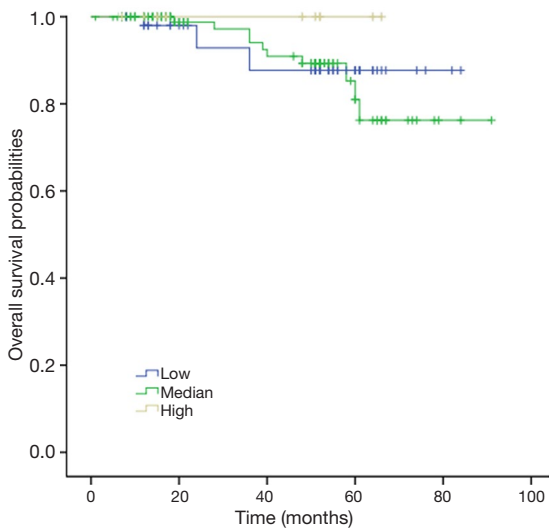

| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 34 | 19 | 14 | 2 | 0 |
| Median | 84 | 59 | 53 | 26 | 1 |
| High | 27 | 24 | 20 | 4 | 0 |

**Figure S4** The K-M curve of NK lymphocytes in peripheral blood of breast cancer patients' DFS. Data was analyzed using K-M plotters (n=190, P>0.05). NK, natural killer; K-M, Kaplan-Meier; DFS, disease-free survival.



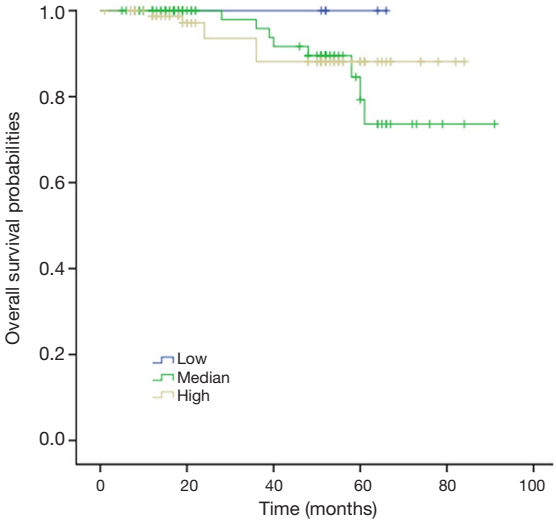

| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 5 | 5 | 4 | 1 | 0 |
| Median | 111 | 81 | 68 | 21 | 3 |
| High | 33 | 29 | 24 | 12 | 0 |

**Figure S5** The K-M curve of CD4+ lymphocytes in peripheral blood of breast cancer patients' OS. Data was analyzed using K-M plotters (n=190, P>0.05). K-M, Kaplan-Meier; OS, overall survival.



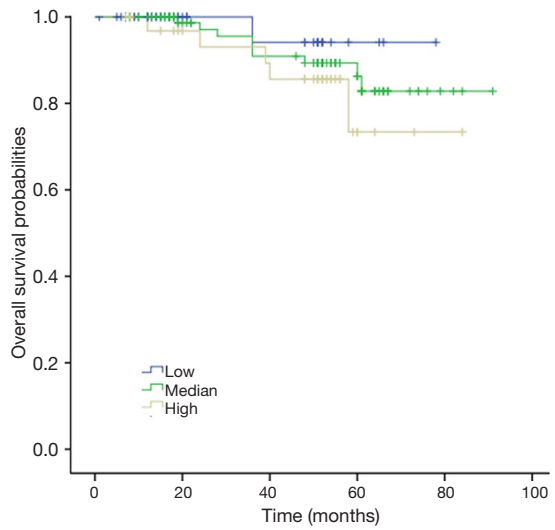

| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 49 | 40 | 33 | 14 | 1 |
| Median | 79 | 70 | 58 | 19 | 1 |
| High | 12 | 12 | 12 | 12 | 12 |

**Figure S6** The K-M curve of CD8+ lymphocytes in peripheral blood of breast cancer patients' OS. Data was analyzed using K-M plotters (n=190, P>0.05). K-M, Kaplan-Meier; OS, overall survival.




| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 6 | 6 | 6 | 6 | 6 |
| Median | 47 | 47 | 44 | 15 | 1 |
| High | 77 | 58 | 48 | 26 | 26 |

**Figure S7** The K-M curve of CD4+/CD8+ ratio lymphocytes in peripheral blood of breast cancer patients' OS. Data was analyzed using K-M plotters (n=190, P>0.05). K-M, Kaplan-Meier; OS, overall survival.




| No at risk | | | | | |
|---|---|---|---|---|---|
| Low | 16 | 16 | 15 | 2 | 0 |
| Median | 72 | 68 | 58 | 28 | 2 |
| High | 30 | 26 | 23 | 4 | 0 |

**Figure S8** The K-M curve of NK lymphocytes in peripheral blood of breast cancer patients' OS. Data was analyzed using K-M plotters (n=190, P>0.05). K-M, Kaplan-Meier; NK, natural killer; OS, overall survival.