Construction of immune score and its prognostic value in invasive lobular carcinoma of the breast using computational pathology analysis

Liye Wang  
Sun Yat-Sen University Cancer Center

Peng Sun  
Sun Yat-Sen University Cancer Center

Fei Xu  
Sun Yat-Sen University Cancer Center

Qiufan Zheng  
Sun Yat-Sen University Cancer Center

Kuikui Jiang  
Sun Yat-Sen University Cancer Center

Ruoxi Hong  
Sun Yat-Sen University Cancer Center

Shusen Wang ( wangshs@sysucc.org.cn )  
Sun Yat-Sen University Cancer Center

Research Article

Keywords: Invasive lobular breast carcinoma, Tumour-immune microenvironment, Immune score, Prognosis, Computational pathology analysis

Posted Date: March 1st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2616931/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Previous studies have shown that high level of TILs in invasive lobular carcinoma (ILC) is associated with poor prognosis, contrary to that in TNBC and HER2-positive breast cancer. We aimed to systematically explore the expression of various tumor-infiltrating immune cell markers and immune checkpoints in ILC and their relationship with prognosis.

Methods

The expression of immune cell markers CD4, CD8, CD20, CD56, CD68, FOXP3 and immune checkpoints PD-1, PD-L1 and CTLA-4 in the microenvironment were detected by immunohistochemistry and their densities were quantified by computational pathology analysis. Then, the LASSO cox regression model was used to construct an immune score (IS) and further evaluate its prognostic value. We further explored the prognostic value of CD68 and FOXP3 with the transcriptome data of ILC patients in the METABRIC database.

Results

In our ILC cohort, the low density of CD4, CD8, CD20, CD56, CD68, FOXP3, PD-1 and PD-L1 had significantly longer disease-free survival (DFS) and overall survival (OS), however, the low density of CTLA-4 was associated with shorter DFS and OS. Based on this, each patient was given a binary score (0 for low, 1 for high) for each immune cell type (CD68 and FOXP3). Patients with low-IS had significantly prolonged DFS (HR 0.19, 95%CI 0.06-0.64, \( p < 0.0001 \)) and OS (HR 0.14, 95%CI 0.03-0.56, \( p < 0.0001 \)). Multivariate analysis revealed that IS was an independent prognostic indicator for DFS and OS. Further analysis showed that IS may increase the prognostic value of TNM stage. We further explored the prognostic role of CD68 and FOXP3 in the transcriptional level and the corresponding ISm in the METABRIC dataset, and found that low proportion of CD68 and FOXP3 and their ISm were associated with longer OS, and ISm was also an independent prognostic factor for OS. Furthermore, we revealed the possible mechanisms that the level of tumor-infiltrating immune cells influence prognosis.

Conclusions

We systemically evaluated the densities of six immune cells and three immune checkpoints and further constructed an IS model, which is a promising biomarker to distinguish the prognosis in ILC patients.

Background

Invasive lobular carcinoma (ILC) accounting for 5%-15% of all breast cancer cases [1]. ILC is characterized by hormone receptor (HR)-positive with a low proliferation rate [1]. Conventionally, tumor/lymph node/metastasis (TNM) staging has been used to the main tool in determining treatment strategy and assessment of prognosis. However, this anatomy-based system provides the clinical
prognosis may vary significantly among ILC patients with the same TNM stage receive similar treatment, due to its highly heterogeneous nature [2]. Thus, researchers are intensely searching for complementary factors such as tumor-intrinsic genetic features [3] or extrinsic immune factors [4]. Important studies have found the major role played by the immune response in ILC biology [2, 5]. Tumor-infiltrating immune cells and high expression of immune checkpoints by tumor or immune cells are key components of the adaptive immune system with a crucial impact on cancer progression [6]. This indicates that immune markers might play an important role in the development of ILC. However, the role of immune factors as prognostic biomarkers in the tumor microenvironment in patients with ILC remains unclear.

The tumor-immune microenvironment encompasses the surrounding immune cells, lymphocytes, bone marrow-derived inflammatory cells [7]. These cell types include effector CD8+ and CD4+ T cells, B cells, naive and memory lymphocytes, macrophages, natural killer (NK) cells, dendritic cells, mast cells and other immune cell subtypes [8]. In HR-negative invasive breast tumors, higher quantity of tumor-infiltrating lymphocytes (TILs) has strong prognostic associations, particularly in the setting of early-stage triple-negative breast carcinomas (TNBC) [9]and human epidermal growth factor receptor 2 (HER2) - positive breast carcinomas [10] treated with adjuvant chemotherapy, while TILs represent a promising new morphologic biomarker associated with poor outcome of ILC [4]. However, due to the visual assessment of the density of TILs in tumour tissue stained with hematoxylin and eosin is unreliable and has limitations when applied to large-scale populations. Quantification of immune markers using computational pathology and construction of comprehensive immune score (IS) system has been shown to make more accurate prognoses for some cancer types [11-13]. However, an IS system with prognostic significance for ILC patients in clinical settings has not been identified so far.

In recent years, a hallmark of tumor progression is the tumor immune evasion [14], which plays a major role in modulating innate immune and suppressing T-cells [15], leading to tumor growth and progression. Immune checkpoints are the most important signaling pathways mediating tumor immune escape, which are crucial for modulating duration and amplitude of immune response in peripheral tissues and maintaining autoimmune tolerance [16]. The main immune checkpoints for breast cancer include programmed death receptor 1/programmed cell death ligand 1 (PD-1/L1), cytotoxic T-lymphocyte-associated protein-4 (CTLA-4) and other molecules. However, the prognostic value of PD-L1/PD-1 and CTLA-4 in ILC and their interactions with different immune cells remain unclear.

In the current study, we evaluated the density of immune cells (CD4, CD8, CD20, CD56, CD68 and FOXP3) and immune checkpoints (PD-1, PD-L1 and CTLA-4) for ILC patients using computational pathology. We then explored the prognostic values of these immune biomarkers. In addition, we developed an IS based on multivariate analysis of immune cells to predict clinical outcomes of ILC patients. This approach can divide ILC patients into different risk subgroups and might add prognostic value to the TNM staging system. We further explored the prognostic value of CD68 and FOXP3 and their ISm with the transcriptome data of ILC patients in the METABRIC database. Based on gene expression and genomic alterations data, we further analyzed the molecular characteristics of CD68 and FOXP3, aiming to reveal the potential molecular mechanism of different subgroups of CD68 and FOXP3 prognosis difference.
Methods

Patients and database

We retrospectively collected formalin-fixed paraffin-embedded (FFPE) tumor specimens and clinical data from 172 non-metastatic ILC patients between May 2003 and December 2017 in our hospital. No patients had received any antitumor therapy before surgery, and all of the patients were pathologically diagnosed with ILC. All patients were restaged according to the 8th AJCC TNM staging system [17]. The METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) dataset [18] containing 1904 tumor cases was downloaded from the cBioPortal database (http://www.cbioportal.org/) (access date: November 30, 2021). A total of 139 ILC samples with full clinical characteristics and transcriptome data and performed the following data exploration. The lists of immune-related genes were downloaded from the ImmPort (HTTPS://www.immport.org/shared/home) [19]. The study has been approved by the ethical committee of Sun Yat-Sen University Cancer Center (Ethics approval number of clinical study project: B2021-061). All the patients provided written informed consent prior to inclusion into the study.

Immunohistochemistry (IHC)

Sequential histological tumor sections of 4µm thick were obtained from a representative FFPE tumor block and used for IHC analysis. The IHC assay was performed as previously described [20]. IHC staining was performed for: 6 immune cells (CD4⁺ helper T cells, CD8⁺ cytotoxic T cells, CD20⁺ B cells, CD56⁺ NK cells, CD68⁺ total macrophages and FOXP3⁺ regulatory T cells) and 3 immune checkpoints (PD-1, PD-L1 and CTLA-4). The following primary antibodies were used: anti-CD4 (ZA-0519, 1:200; ZS), anti-CD8 (ZA-0508, 1:200; ZS), anti-CD20 (kit-0001; MXB), anti-CD56 (ZM-0057, 1:200; ZS), anti-CD68 (ZM-0060, 1:800; ZS), anti-FOXP3 (UM800140, 1:100; ZS, Ultra-MAB), anti-PD-1 (ZM-0381, 1:200; ZS), anti-PD-L1 (clone E1L3N, 1:100 dilution; Cell Signaling Technology, USA), anti-CTLA-4 (SP355, 1:100; Abcam).

Computational pathology analysis

All IHC slides were examined by two independent pathologists, only slides with good staining quality would be included in image acquisition. In addition, all normal and necrosis tissues were excluded from the assessment through manual annotation by pathologists. Then, a full view of each IHC slide was digitally scanned at 20x magnification, with an image resolution of 0.25 µm/pixel (Aperio, ScanScope AT2, Leica). All images were auto-examined using computational pathology analysis, and the expression was quantified on tumour cells (TCs) or tumour-associated immune cells (TAICs) expressing the immune checkpoints. Representative IHC staining images with high and low positive numbers for these immune markers are shown in Supplemnetary Figure S1.

The computational pathology analysis was performed as previously described and showed good performance [21, 22], and the methods are briefly described here. First, the individual cell nuclei were manual annotated as a TC or TAIC by two independent pathologists. Then, each nucleus in the hematoxylin channel was segmented using stain deconvolution and achieved good performance [11, 23].
Computational pathology analysis can automatically segment of the nuclei in the hematoxylin channel, and automatically classify of the cells into TCs or TAICs based on Xception deep learning model [24]. The computational pathology analysis showed a high reproducibility and consistency with pathological classification. Finally, the density of TCs or TAICs could be quantified as the total number of TCs or TAICs divided by the entire nonnecrotic area.

**Construction of the Immune Score (IS)**

We adopted a least absolute shrinkage and selection operator (LASSO) Cox regression model [25] to select the most useful prognostic features out of all 6 immune cell markers and then constructed an IS for predicting survival. The analysis was performed by using the “glmnet” package in R software [26]. Ten-time cross validations with the Lambda.min criteria were used to determine the optimal values of $\lambda$, and a value of $\lambda = 0.047$ with $\log(\lambda) = -3.121$ was chosen. Based on this value, CD4, CD20, CD68 and FOXP3 were selected. Among 4 immune cell markers, the CD68 and FOXP3 were identified significantly affecting OS and DFS and used to construct an IS by multivariate Cox regression analysis.

The optimal cutoff threshold immune cell density was used to stratify patients into groups based on the degree of tumor infiltration. This method has been independently validated in evaluating IS in multiple solid tumors [11, 13, 27], and was thus adopted for the current study. We then used X-tile software (version 3.6.1; Yale University, New Haven, CT, USA) [28] to determine the optimal cut-off values for high and low density regarding CD68 and FOXP3 based on the associations with patient disease-free survival (DFS). Based on the threshold, each patient was given a binary score (0 for low, 1 for high) for each immune cell type (CD68 and FOXP3). IS for each patient was obtained by adding the two binary score values, the scale being from 0 to 2. Three patient groups were defined: patients with low densities of CD68 and FOXP3 were classified as IS-0; patients with one high density for one marker were classified as IS-1; and patients with two high densities of these two markers were stratified as IS-2. Patients with a high degree of immune cell infiltration (IS: 2) were assigned as the high-IS group, and patients with a low degree of immune cell infiltration (IS: 0–1) were assigned as the low-IS group.

**Comprehensive analysis of immune and molecular characteristics in different CD68 and FOXP3 subgroups in the METABRIC dataset**

To identify immune characteristics of 139 ILC samples in the METABRIC dataset, their expression data were imported into CIBERSORT (HTTPS://cibersort.stanford.edu/) and iterated 1000 times to estimate the relative proportion of 22 types of immune cells. The 22 types of infiltrating immune cells inferred by CIBERSORT include B cells, T cells, natural killer cells, macrophages, dendritic cells, eosinophils, and neutrophils. The relative proportions of the same immune markers were added together to obtain the total ratio. The CD68$^+$ total macrophages include M0, M1 and M2 macrophages. We used X-tile software to determine optimal cut-off values for high and low proportion regarding CD68 and FOXP3 cells for OS.

We further explored the prognostic value of immune cell markers CD68 and FOXP3 and their ISm constructed from this dataset. According to the IS grouping method outlined above (0 for low proportion,
1 for high proportion), ILC patients were also divided into three groups: ISm-0, ISm-1, and ISm-2. Patients with a high degree of immune cell infiltration (ISm: 2) were assigned as the high-ISm group, and patients with a low degree of immune cell infiltration (ISm: 0–1) were assigned as the low-ISm group. Then, we compared the relative proportions of 22 types of immune cells in patients with different clinicopathological factors and ISm subgroups.

Spearman correlation method was used to estimate the correlations between continuous variables. In biological function and signaling pathway analysis, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of the genes that correlated with high and low proportion CD68 and FOXP3 groups were performed using clusterProfiler package of R. In the gene mutation analysis, the quantity of gene mutations was analyzed in high and low proportion CD68 and FOXP3 subgroups by using the ComplexHeatmap package of R.

**Statistical analyses**

Data are presented as whole numbers and proportions for categorical variables, and medians or means (interquartile ranges (IQRs)) for continuous variables. Clinicopathological variables associated with IS were analyzed using the χ² test or Fisher's exact test. Differences between groups were assessed using the Student t-test or Mann–Whitney U test for continuous variables. DFS was calculated from the date of surgery to the date of disease relapse (local or distant relapse or death from any cause). Overall survival (OS) was calculated from the date of surgery to the date of death or the latest follow-up. The Kaplan-Meier method was used to estimate OS and DFS and differences were compared using the log-rank test. Calculating the exponential of the regression coefficients from the Cox model provided an estimate of the hazard ratios (HRs) and the 95% confidence interval (CI). ROC curves were used to compare IS prognostic validity with immune checkpoints, and to compare the prognostic validity of IS in different molecular subtypes. Multivariate Cox regression analysis with backward selection was performed to test the independent significance of different factors. Multivariate analysis was performed using variables with \( p < 0.1 \) in the univariate analysis, and only independent prognostic factors were retained in the multivariate model. In addition, we established a prognostic score model combining the IS, ER and TNM stage. Moreover, nomograms predicting 3 years or 5 years OS were established. The model performance was evaluated by the accuracy of point estimates of the survival function (calibration). The performance of the nomograms was evaluated using the concordance index (C-index) [29]. In addition, bootstraps with 1000 resamples were applied to internal validation to provide an unbiased estimate of model performance.

Statistical analyses were performed with software programs (SPSS version 26.0 (IBM); R version 4.1.2; GraphPad Prism 8). All statistical tests were two-sided and considered significant when the \( p \) value was less than 0.05.

**Results**

**Patient characteristics and immune markers**
The clinicopathological characteristics of ILC patients are summarized in Table 1. The majority of patients were HR-positive: Luminal-A (40.1%, n = 69) and Luminal-B (54.7%, n = 94), only 2.3% were HER2-positive and 2.9% were TNBC. Together, 48.8% patients had lymphatic invasion, and 29.1% patients belonged to stage III according to the criteria released by the 8th AJCC. The median follow-up time was 84.6 months (range: 13.8 – 224.4 months). At the end of follow-up, 34 (19.8%) patients had disease progression and 25 (14.5%) patients had died from ILC.

Based on computational pathology analysis, we investigated the numeration of 6 immune cell markers and 3 immune checkpoints per mm$^2$. Next, we analyzed the distribution of immune markers and their correlations (Figure 1). We observed that the CTLA-4 expression (median density: 175.07 cells/mm$^2$) was the highest among immune markers (Figure 1A). We found that CD$^+$ T cells was strongly correlated with CD$^+$ T cells and CD20$^+$ B cells, while CTLA-4 expression was negatively correlated with immune cells and other immune checkpoints except for CD56$^+$ NK cells (Figure 1B). We also observed that CD$^+$ T cells and CD68$^+$ macrophages were upregulated in the stage III group, while CTLA-4 and PD-1 expression were upregulated in the stage II group (Figure 1C). Meanwhile, we found that CD$^+$ T cells and CD68$^+$ macrophages were upregulated in the TNBC subtype, while CD8$^+$ T cells was upregulated in the HER2-enriched subtype (Figure 1D).

**Prognostic value of each biomarker**

We then explored the prognostic value of each marker. X-tile was used to determine optimal cut-off values for high and low density regarding 6 immune cell markers and 3 immune checkpoints for DFS (Supplementary Table S1). As shown in Figure 2, patients with low-infiltrating CD$^+$ T cells ($p = 0.0001$), CD$^+$ T cells ($p = 0.005$), CD20$^+$ B cells ($p = 0.024$), CD68$^+$ macrophages ($p = 0.0001$), or FOXP3$^+$ T cells ($p = 0.043$) had better DFS than patients with high-infiltrating. The patients with low expression of PD-L1 ($p = 0.033$) was associated with better DFS than high expression, while low expression of CTLA-4 ($p = 0.005$) was associated with poorer DFS than high expression (Figure 2). The associations between the 9 immune markers and OS are shown in Supplementary Figure S2.

**IS and association with prognosis**

To construct an IS, we identified 2 immune cell markers CD68 and FOXP3 that were significantly associated with DFS using penalized LASSO Cox regression models (Supplementary Figure S3). According to the definition of IS (see method), 49.4%, 37.2% and 13.4% of the patients were classified as IS-0, IS-1, IS-2, respectively. We aimed to explore the prognostic value of IS. The Kaplan-Meier curves revealed three distinct patient groups with statistically significant differences in DFS ($p = 0.0001$, Figure 3A) and OS ($p = 0.0001$, Figure 3B). Patients with an IS-0 experienced the best post-operative outcome, while patients with an IS-2 had the worst outcome.

Patients with a high degree of immune cell infiltration (IS: 2) were assigned as the high-IS group, and patients with a low degree of immune cell infiltration (IS: 0–1) were assigned as the low-IS group. Based
on this, we assigned 149 (86.6%) patients into the low-IS group and 23 (13.4%) patients into the high-IS group. Patients with low-IS had longer DFS ($p = 0.0001$) and OS ($p = 0.0001$) compared with patients with high-IS (Figure 3C, D). IS could be as a prognostic factor for DFS (AUC = 0.618) and OS (AUC = 0.656) in the ILC cohort (Figure 3E, F).

We performed univariate analysis showed that IS was significantly associated with DFS (Figure 4A) and OS (Figure 4C). We then performed multivariate Cox regression analysis, which showed that the IS remained significant for DFS ($p = 0.003$ Figure 4B) and OS ($p = 0.0001$ Figure 4D). In addition, the TNM stage and ER levels were also significantly associated with DFS and OS in multivariate analysis (Figure 4B, D).

**Relationship of the IS and immune checkpoints**

We further analyzed the composition of immune markers in different IS subgroups (Supplementary Figure S4). We then used ROC analysis to compare the sensitivity and specificity of the prognostic value of IS with immune checkpoints. IS showed better prognostic value than immune checkpoints for overall DFS (AUC: 0.618 vs PD-L1 0.588, PD-1 0.508, CTLA-4 0.454), 3-year DFS (0.644 vs 0.588, 0.593, 0.437) and 5-year DFS (0.646 vs 0.581, 0.489, 0.423), which similar results were observed for overall OS and 5-year OS (Supplementary Figure S5). With the same method, we compared the prognostic value of IS in HR positive subtypes. IS showed better prognostic value in the luminal A subtype than in the luminal B subtype (Supplementary Figure S6).

**Development of nomogram with IS**

We constructed nomogram A combining IS, TNM stage and ER to predict the 3 years and 5 years OS of ILC patients based on the multivariate Cox regression analysis (Figure 5A). In addition, we also constructed nomogram B using TNM stage and ER. The C-index for the nomogram A to predict OS was 0.827 (95%CI 0.794-0.859), while the C-index for the nomogram B was 0.786 (95% CI 0.75 to 0.822). The addition of IS to nomogram A slightly enhanced the accuracy compared with nomogram B. Subsequently, the calibration plot of nomogram A for the probability of 3 years or 5 years OS showed good agreement between the prediction by nomogram and actual observation for nomogram (Figure 5B, C). We used ROC analysis to compare the sensitivity and specificity of the prognostic score model with the TNM stage combined with ER, TNM stage or IS alone model. Combination of the IS, TNM stage and ER showed better prognostic value than the TNM stage and ER (AUC: 0.889 vs 0.819), TNM stage alone (AUC: 0.889 vs 0.767) for 5 years OS and overall OS (AUC: 0.824 vs 0.788 vs 0.753) (Figure 5D, E).

**ISm and association with prognosis in the METABRIC dataset**

We further explored the prognostic value of CD68 and FOXP3 and their ISm with the transcriptome data of ILC patients (n=139) in the METABRIC database. The clinical characteristics of ILC patients in the METABRIC database are summarized in Supplementary Table S2. To analyze the composition of immune cells in the METABRIC database, we used the Wilcoxon test to compare the relative proportions of 22
types of immune cells among different clinicopathological factors and ISm subgroups (Supplementary Figure S7). We then explored the prognostic value of CD68$^+$ macrophages and FOXP3$^+$ T cells. Patients with low-infiltrating CD68$^+$ macrophages ($p = 0.022$) and FOXP3$^+$ T cells ($p = 0.012$) had better OS than patients with high-infiltrating (Figure 6A, B).

According to the ISm grouping method outlined above, 39.6%, 32.4% and 28.1% of the patients were classified as ISm-0, ISm-1 and ISm-2, respectively. We observed the same prognostic effect of ISm in METABRIC ILC cohort as IS in our own cohort. The Kaplan-Meier curves revealed three distinct patient groups with statistically significant differences in OS times ($p = 0.011$, Figure 6C). Patients with an ISm-0 experienced the best outcome, while patients with an ISm-2 had the worst outcome. Patients with a high degree of immune cell infiltration (ISm: 2) were assigned as the high-ISm group, and patients with a low degree of immune cell infiltration (ISm: 0–1) were assigned as the low-ISm group. Based on this, 71.9% (n=100) patients were assigned to the low-ISm group and 28.1% (n=39) patients to the high-ISm group. The patients with low-ISm also had longer OS (HR 0.44, 95%CI 0.22 to 0.87, $p = 0.005$) than those with high-ISm (Figure 6D). We also performed univariate and multivariate Cox regression analysis, which showed that the ISm remained significant for OS (Supplementary Figure S8A, B). ISm also could be as a prognostic factor for 10-year OS (AUC = 0.684) and overall OS (AUC = 0.605) in the METABRIC cohort (Supplementary Figure S8C, D).

**Molecular characteristics of different CD68 and FOXP3 subgroups in the METABRIC cohort**

TNBC and HER2-overexpressing patients with higher levels of total TILs tend to exhibit better treatment outcomes and prognosis, however, ILC patients with high-infiltrating immune cells were associated with poorer DFS and OS. Therefore, we characterized the molecular and immune profile of CD68$^+$ macrophages and FOXP3$^+$ T cells subgroups in the METABRIC dataset, and examined its prognostic ability.

To further explore CD68 and FOXP3 subgroups related biological processes in ILC, a total of 308 genes and 238 genes have strongly correlation with low-CD68 and low-FOXP3 (Supplementary Figure S9A, C), respectively, according to Spearman correlation analysis ($|R| > 0.3$ and $p < 0.05$). Next, GO enrichment analyses were performed to investigate CD68 and FOXP3 subgroups potential biological functions. We found that low-CD68 subgroup related genes were mainly involved in immune-related pathways and inflammatory pathways (Supplementary Figure S9B), including T cells activation, positive regulation of lymphocyte differentiation and leukocyte proliferation, antigen processing and presentation. Meanwhile, low-FOXP3 related genes were mainly involved in negative regulation of epithelial cell differentiation, thereby inhibiting cell proliferation (Supplementary Figure S9D). In the contrary, using the same method, 504 genes and 404 genes have strongly correlation with high-CD68 and high-FOXP3, respectively, and were not enriched for relevant biological processes by GO enrichment analysis.

To further investigate different CD68 and FOXP3 subgroups related immune functions in ILC, 2483 immunologically related genes were retrieved from the ImmPort. By intersecting these genes with the
expressed genes obtained from METABRIC dataset, 1197 immune-related genes were obtained, of which 17 genes and 36 genes have strongly correlation with low-CD68 and high-CD68 (Supplementary Figure S10A, D), respectively, according to Spearman correlation analysis ($|R| > 0.2$ and $p < 0.05$). The functional enrichment analysis was performed, and the top ten significantly enriched GO terms and KEGG pathways were shown in Supplementary Figure S10. The gene sets of the low-CD68 samples were enriched in inhibiting cell growth and angiogenesis-related pathways (Supplementary Figure S10B, C), while the gene sets of the high-CD68 samples were enriched in cell proliferation and metastasis related pathways (Supplementary Figure S10E, F). With the same method, the FOXP3 subgroups related immune functions are shown in Supplementary Figure S11. Next, we analyzed gene mutations to gain further biological insight into the molecular characteristics of the CD68 and FOXP3 subgroups (Supplementary Figure S12, S13).

**Discussion**

In the present study, we determined the density of 9 immunological variables derived from 6 immune cells and 3 immune checkpoints using computational pathology analysis, and evaluated their prognostic value in ILC patients. Next, we developed an IS based on the density of CD68$^+$ macrophages and FOXP3$^+$ T cells, which was an important factor in predicting prognosis in ILC patients. Furthermore, we found that IS was a powerfully independent predictor of DFS and OS. Then, we developed a nomogram to predicted the 3 years and 5 years OS of ILC patients combining all independent variables IS, TNM stage and ER. In addition, we further explored the prognostic role of CD68 and FOXP3 and their ISm in the METABRIC dataset, and our results demonstrated that the ISm was a valid prognostic immune-related biomarker for ILC. Furthermore, we characterized the molecular and immune profile of CD68 and FOXP3 subgroups in METABRIC dataset, and further revealed the possible mechanisms that the level of tumor-infiltrating immune cells in ILC influence prognosis. To the best of our knowledge, this is the first study to simultaneously measure 9 different immunologic variables derived from 6 immune cells and 3 immune checkpoints in the TME using computational pathology analysis, as well as to construct an IS for ILC, and to evaluate its prognostic significance.

Computational pathology analysis has recently produced encouraging results due to its quantitative, automated, and reproducible evaluation of whole-slide sets [30]. This permits automatic and accurate large-scale analysis without subjective bias. In addition, algorithms based on “deep learning” neural networks have translated well to digital pathology, where they have demonstrated outstanding performance in tasks like tissue segmentation, prognostication, and computational TILs assessment [11, 21, 31]. Particularly for IS, it demands accurate and quantitative evaluation of immune cells in the nonnecrotic invasive carcinoma area. However, due to factors of breast cancer, such as the spatial distribution of TILs, the tumor-stroma ratio, histologic subtypes, TILs in ductal carcinomas in situ, and intra-tumoral heterogeneity, which may increase the interobserver and intra-observer variability of visual TIL assessments. Thus, it is necessary to deeply explore the immune markers properties of ILC using computational pathology analysis. In this study, we evaluated the density of 6 immune cells and 3
immune checkpoints through computational pathology analysis, which was developed based on the Xception model, achieved good performance in identifying TC and TAIC nuclei and accurate classification of cell types. Moreover, our digital pathology can obtain a large amount of quantitative information with high speed, which may quickly determine the IS of each patient.

Breast cancer was previously considered a relatively weakly immunogenic tumor compared to other tumor types. Recent evidence has suggested that TILs have prognostic and predictive capabilities for TNBC and HER2-positive breast cancers [32, 33]. However, the density of TILs and immune checkpoints in the tumour-immune microenvironment of HR-positive ILC is still unclear. Desmedt et al [34] showed that TILs levels were statistically significantly lower in ILC compared with invasive ductal cancer, and high TILs levels were associated with worse prognosis in ILC. However, the ability to visual TILs assessment on the basis of H&E slides is highly subjective, less reproducible, and interobserver and intra-observer variability. Furthermore, Pagès et al [12] conducted a study comparing the differences between TILs and Immunoscore, the results showed that Immunoscore was highly objective, reproducible and had stronger ability of prognostic. Wang et al [35] applied computational pathology to quantify the densities of CD3, CD8, and CD45RO in nasopharyngeal carcinoma, and then developed the IS based on the density of these three markers, and found that IS had an independent prognostic effect. In this study, we determined the density of 6 immune cells and 3 immune checkpoints using computational pathology analysis. In addition, immune cell markers CD68 and FOXP3 density was converted to a binary score to construct the IS, and found that IS has significant prognostic and predictive value in ILC. The results also showed that patients with high-IS were associated with larger tumor size, lymph node metastasis, later stage and faster proliferation, and were significantly associated with poorer DFS and OS.

Our study has shown that ILC patients with low density of CD4, CD8, CD20, CD56, CD68 and FOXP3 had significantly longer DFS and OS, in contrary to TILs in TNBC and HER2-positive breast cancer [32, 33]. Previous studies have shown that the level of TILs in ILC is significantly lower than that in invasive ductal carcinoma, and high level of TILs in ILC is associated with poor prognosis [4, 34]. Studies have shown that the prognostic role of TIL subsets in breast cancer depends upon HR status and immune cells distribution [35, 36]. Liu et al [37] have reported that FOXP3+ regulatory TILs are a poor prognostic indicator in ER+ breast cancer, but a favorable prognostic factor in the HER2+/ER- subtype. Mahmoud et al [38] found that higher numbers of CD68 macrophages were significantly associated with worse breast cancer-specific survival and shorter disease-free interval. Therefore, ILC patients with low density of immune cells have a worse outcome, which may be related to the fact that ILC is mainly HR-positive, and it may also be related to the underlying molecular features and biological mechanisms that are different between ILC and IDC.

Our study showed that CTLA-4 was the highest expression among the densities of immune markers, and negatively correlated with other immune checkpoints and infiltrating immune cells. This suggested that ILC may be an immunosuppressive-dominant tumor, which was consistent with a previous study [2]. Moreover, our results demonstrated that the density of different immune checkpoints (PD-1, PD-L1) could predict prognosis of ILC patients, and patients with high density had shorter DFS and OS.
Conversely, low expression of CTLA-4 was associated with poorer DFS and OS than high expression, which implied that anti-CTLA-4 therapy may be negatively correlated with response and survival in ILC patients. Recently, Santa-Maria et al [39] designed a single-arm pilot research to determine the overall response-rate (ORR) of durvalumab plus tremelimumab in metastatic ER-positive or TNBC, and found that only three TNBC patients had a response (ORR = 17%). Tille et al [4] reported that TILs were associated with larger tumors, lymph node involvement, HER2 amplification, and poor OS and iDFS was significantly associated with increasing TILs in ILC patients. The results of the two research were consistent with ours, that ILC was characterized by an immune-suppressive phenotype with elevated expression of CTLA4, and may benefit less from anti-CTLA4 therapy.

Currently, immunoscore has become a clinically useful prognostic marker in a variety of cancers, such as colorectal cancer, non-small-cell lung cancer and nasopharyngeal carcinoma [11, 12, 40]. Until now, the prognostic significance of IS was unknown in ILC patients. In this study, ILC patients with low-IS had significantly longer DFS and OS than those with high-IS. Importantly, our results demonstrated that IS was an independent prognostic factor for DFS and OS. Furthermore, we further used ROC analysis to compare the prognostic value of IS with immune checkpoints. The results showed that IS had better prognostic value than immune checkpoints. Meanwhile, we also found IS had better prognostic value in Luminal A subtype than in Luminal B subtype. These results showed that IS was a promising prognostic classifier, which could be widely used to predict the prognosis of ILC patients. In addition, a prognostic score model combined the IS, TNM stage and ER was constructed and had a better prognostic value than the TNM stage and ER, which could be an attractive tool to help in guiding treatment selection. The comprehensive immune score system can help to understand the immune state of tumours in individuals and improve the accuracy of TNM staging for predicting survival. Therefore, we believe it’s more significant to identify IS for ILC patients based on the density of CD68 and FOXP3.

We further explored the prognostic role of CD68 and FOXP3 and their ISm in the METABRIC dataset, and found that low proportion of CD68 and FOXP3 and their ISm were associated with longer OS, and ISm was also an independent prognostic factor for OS. In the highly proliferative subtypes, usually the triple-negative and HER2-overexpressing breast cancers, immune infiltrates are most frequently seen and has prognostic and predictive capabilities [32, 33]. However, in this study, low proportion of CD68 and FOXP3 patients had longer OS than those with high proportion. Thus, we further explore the molecular characteristics of different CD68 and FOXP3 subgroups. The results suggested that the low proportion of CD68 and FOXP3 subgroups were characterized by immune-active pathways and gene sets, such as expression of HLA-B is associated with T cell activation and identifies immune activated [41], TMSB4Y plays a negative regulatory role in the initiation and progression of human breast cancer cells [42], high expression of NR3C2 was statistically associated with prolonged OS and disease-specific survival for breast cancer patients [43]. By contrast, the high proportion of CD68 and FOXP3 subgroups were characterized by tumor-promoting gene sets and was associated with poor outcome, such as IL33 in facilitating breast cancer lung metastasis by modifying the immune microenvironment at the metastatic niche toward type 2 inflammation [44], Rac3-induced Rac3/ERK-2/NF-κB signaling pathway triggers
breast cancer cell aggressiveness [45]. The CD68+ macrophages and FOXP3+ T cells proved to be valid prognostic immune-related biomarkers for ILC, with better survival in low-proportion patients and worse survival in high-proportion patients.

There were several limitations which should be considered while interpreting the study findings. Firstly, the density of immune cells detected by IHC, while the relative proportions of immune cells measured by transcriptome data, this difference may introduce bias in the analysis procedure. Secondly, immune markers were not assessed separately from the tumor and stroma, as the distribution of different regions may affect the prognosis and treatment of ILC. Thirdly, this was a single-center retrospective small-sample study, should be further verified in a larger multicenter cohort.

Conclusions

Our study demonstrated that IS, which was based on the density of CD68 and FOXP3, was a promising immune-related prognostic biomarker of ILC patients and might be an attractive option to help guide treatment selection. Moreover, we revealed the possible mechanisms that the level of tumor-infiltrating immune cells in ILC influence prognosis, but further studies are needed to clarify this point.

Abbreviations

DFS: disease-free survival; HER2: human epidermal growth factor receptor 2; ILC: invasive lobular carcinoma; IS: immune score; OS: overall survival; TNBC: triple-negative breast carcinomas; TCs: tumour cells; TAICs: tumour-associated immune cells.

Declarations

Ethics approval and consent to participate

The study has been approved by the ethical committee of Sun Yat-Sen University Cancer Center (B2021-061). All procedures in this study were conducted in accordance with ethical principles.

Patient consent for publication Not required.

Competing interests None declared.

Availability of data and material

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Funding

The work was funded by the Sun Yat-sen University Clinical Research 5010 Program (2017011).
Authors’ contributions

Conception and design of the study: L-YW, R-XH and S-SW. Development of methodology: L-YW, P-S, R-XH and S-SW. Acquisition of data: L-YW and S-P. Analysis and interpretation of data: L-YW, P-S, F-X, Q-FZ, K-KJ, R-XH and S-SW. Writing, review and/or revision of the manuscript: L-YW, R-XH and S-SW. Administrative, technical or material support: L-YW, P-S, R-XH and S-SW. All authors read and approved the final manuscript.

Acknowledgements

We thank Weimin Zhang for assistance in digital scanning of IHC slides, who is employed by Peking University Cancer Hospital and Institute. We would like to thank the staff members of the Department of Pathology, Sun Yat-sen University Cancer Center for their assistance with the preparation and preservation of paraffin-embedded breast cancer samples.

Provenance and peer review Not commissioned; externally peer reviewed.

References

7. Arneth B: Tumor Microenvironment. Medicina (Kaunas) 2019, 56(1).


Tables

Table 1 Clinicopathological characteristics of the patients stratified by immune score
<table>
<thead>
<tr>
<th></th>
<th>All (n%)</th>
<th>Low (n%)</th>
<th>High (n%)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total population</strong></td>
<td>172 (100)</td>
<td>149 (86.6)</td>
<td>23 (13.4)</td>
<td>0.653</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50 years</td>
<td>99 (57.6)</td>
<td>87 (58.4)</td>
<td>12 (52.2)</td>
<td></td>
</tr>
<tr>
<td>&gt;50 years</td>
<td>73 (42.4)</td>
<td>62 (41.6)</td>
<td>11 (47.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Menopausal status</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.819</td>
</tr>
<tr>
<td>Pre-menopausal</td>
<td>109 (63.4)</td>
<td>95 (63.8)</td>
<td>14 (60.9)</td>
<td></td>
</tr>
<tr>
<td>Post-menopausal</td>
<td>63 (36.6)</td>
<td>54 (36.2)</td>
<td>9 (39.1)</td>
<td></td>
</tr>
<tr>
<td><strong>T stage</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>T1</td>
<td>83 (48.3)</td>
<td>79 (53.0)</td>
<td>4 (17.4)</td>
<td></td>
</tr>
<tr>
<td>T2-T3</td>
<td>89 (51.7)</td>
<td>70 (47.0)</td>
<td>19 (82.6)</td>
<td></td>
</tr>
<tr>
<td><strong>N stage</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>N0</td>
<td>88 (51.2)</td>
<td>83 (55.7)</td>
<td>5 (21.7)</td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td>36 (20.9)</td>
<td>32 (21.5)</td>
<td>4 (17.4)</td>
<td></td>
</tr>
<tr>
<td>N2-N3</td>
<td>48 (27.9)</td>
<td>34 (22.8)</td>
<td>14 (60.9)</td>
<td></td>
</tr>
<tr>
<td><strong>TNM stage</strong></td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>I</td>
<td>54 (31.4)</td>
<td>52 (34.9)</td>
<td>2 (8.7)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>68 (39.5)</td>
<td>62 (41.6)</td>
<td>6 (26.1)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>50 (29.1)</td>
<td>35 (23.5)</td>
<td>15 (65.2)</td>
<td></td>
</tr>
<tr>
<td><strong>ER expression</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.625</td>
</tr>
<tr>
<td>Yes</td>
<td>162 (94.2)</td>
<td>141 (94.6)</td>
<td>21 (91.3)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>10 (5.8)</td>
<td>8 (5.4)</td>
<td>2 (8.7)</td>
<td></td>
</tr>
<tr>
<td><strong>PR expression</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.517</td>
</tr>
<tr>
<td>Yes</td>
<td>149 (86.6)</td>
<td>130 (87.2)</td>
<td>19 (82.6)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23 (13.4)</td>
<td>19 (12.8)</td>
<td>4 (17.4)</td>
<td></td>
</tr>
<tr>
<td><strong>HER2 expression</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td>Yes</td>
<td>14 (8.1)</td>
<td>12 (8.1)</td>
<td>2 (8.7)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>158 (91.9)</td>
<td>137 (91.9)</td>
<td>21 (91.3)</td>
<td></td>
</tr>
<tr>
<td><strong>KI67</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>≤14</td>
<td>&gt;14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>92</td>
<td>84</td>
<td>8 (34.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>65</td>
<td>15 (65.2)</td>
<td></td>
</tr>
<tr>
<td>Subtype (based on receptor status)</td>
<td>0.069</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luminal A</td>
<td>69 (40.1)</td>
<td>64 (43.0)</td>
<td>5 (21.7)</td>
<td></td>
</tr>
<tr>
<td>Luminal B</td>
<td>94 (54.7)</td>
<td>78 (52.3)</td>
<td>16 (69.6)</td>
<td></td>
</tr>
<tr>
<td>HER2-positive</td>
<td>4 (2.3)</td>
<td>4 (2.7)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TNBC</td>
<td>5 (2.9)</td>
<td>3 (2.0)</td>
<td>2 (8.7)</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>0.150</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>116 (67.4)</td>
<td>97 (65.1)</td>
<td>19 (82.6)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>56 (32.6)</td>
<td>52 (34.9)</td>
<td>4 (17.4)</td>
<td></td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>0.336</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>52 (30.2)</td>
<td>43 (28.9)</td>
<td>9 (39.1)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>120 (69.8)</td>
<td>106 (71.1)</td>
<td>14 (60.9)</td>
<td></td>
</tr>
<tr>
<td>Endocrine therapy</td>
<td>0.453</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>156 (90.7)</td>
<td>136 (91.3)</td>
<td>20 (87.0)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>16 (9.3)</td>
<td>13 (8.7)</td>
<td>3 (13.0)</td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td>&lt; 0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25 (14.5)</td>
<td>15 (10.1)</td>
<td>10 (43.5)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>147 (85.5)</td>
<td>134 (89.9)</td>
<td>13 (56.5)</td>
<td></td>
</tr>
<tr>
<td>Disease progression</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>34 (19.8)</td>
<td>23 (15.4)</td>
<td>11 (47.8)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>138 (80.2)</td>
<td>126 (84.6)</td>
<td>12 (52.2)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: TNM, tumor-node-metastasis; ER, Estrogen receptor; PR, Progesterone receptor; HER2, human epidermal growth factor receptor 2.

**Figures**
Figure 1

Composition and distribution of tumor-associated immune markers in ILC.

Composition of tumor-associated immune markers (A) and correlation between immune markers (B) in ILC. Distribution of infiltrating immune cell markers and immune checkpoints density according to tumor stage (C) and different molecular subtypes (D). The thick lines represent the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range), respectively. The scattered dots represent the corresponding subgroups in the graph. Significant statistical differences between the two subgroups were assessed using the Mann–Whitney test (ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001).
Figure 2

Kaplan-Meier curves for disease-free survival according to different immune markers groups. Plots show the Kaplan-Meier curves of CD4 (A), CD8 (B), CD20 (C), CD56 (D), CD68 (E), FOXP3 (F), CTLA-4 (G), PD-1 (H) and PD-L1 (I).
Figure 3

Kaplan-Meier curves for DFS and OS according to the immune score.

Kaplan-Meier curves comparing DFS (A) and OS (B) in patients with different IS subgroups. Patients were stratified in high and low IS groups using the optimal cut-off value. Kaplan-Meier curves comparing DFS
(C) and OS (D) in patients with high and low IS groups. Receiver operating characteristics (ROC) curves for the prediction of DFS (E) and OS (F).

Figure 4

Univariate and multivariate analysis of factors associated with DFS and OS.

Plots show univariate (A) and multivariate (B) analysis of DFS; univariate (C) and multivariate (D) analysis of OS.
Figure 5

Nomogram and calibration plots for predicting 3-year and 5-year OS.

(A) Nomogram A including IS, TNM stage and ER; (B, C) showed the calibration plots for predicting 3-year and 5-year OS in ILC patients; comparisons of the sensitivity and specificity for the prediction of 5-year
OS (D) and overall OS (E) by the combined IS, TNM stage and ER model, the TNM stage and ER model, the TNM stage alone model, and the IS alone model.

**Figure 6**

Kaplan-Meier curves for OS according to the different CD68$^+$ macrophages, FOXP3$^+$ T cells and immune score groups. Plots show the Kaplan-Meier curves of CD68 (A), FOXP3 (B) and different ISm subgroups (C). Patients were stratified in high and low ISm groups using the optimal cut-off value. Kaplan-Meier curves comparing 10-year OS (C) and overall OS (D) in patients with high and low ISm groups.

**Supplementary Files**

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

- [SupplementaryFigures.docx](#)
- [SupplementaryTables.docx](#)