DOI:
10.6084/m9.figshare.14589063

Document Version
Peer reviewed version

Original article

Title: Systemic immune reaction in axillary lymph nodes adds to tumor infiltrating lymphocytes in triple-negative breast cancer prognostication

Fangfang Liu[1,2,$], Thomas Hardiman[1,3,$], Kailiang Wu[2], Jelmar Quist[1,3,4], Patrycja Gazinska[5], Tony Ng[3], Arnie Purushotham[3], Roberto Salgado[6], Xiaojing Guo[2], Sarah E Pinder[3], Anita Grigoriadis[1,3,4,*]

Affiliations:

[1] Cancer bioinformatics, School of Cancer & Pharmaceutical Sciences, King's College London Faculty of Life Sciences and Medicine, London, UK.

[2] Department of Breast Pathology and Research Laboratory, Key Laboratory of Breast Cancer Prevention and Therapy (Ministry of Education), National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China.

[3] School of Cancer & Pharmaceutical Sciences, King's College London Faculty of Life Sciences and Medicine, London, UK.

[4] Breast Cancer Now Unit, School of Cancer and Pharmaceutical Sciences, King's College London, UK.

[5] Breast Cancer Now Toby Robins Research Center, The Institute of Cancer Research, London, UK.

[6] Division of Research, Peter Mac Callum Cancer Centre, Melbourne, Australia; Department of Pathology, GZA-ZNA Hospitals, Antwerp, Belgium.

29

30 [$] Dr Fangfang Liu and Thomas Hardiman contributed equally to this article.

31

32 [*] Corresponding author:

33 Dr Anita Grigoriadis, Breast Cancer Now Unit, Innovation Hub, Cancer Center at Guy's, Great

34 Maze Pond, London, SE1 9RT, UK

35 Phone: +44 020 7188 2360

36 e-mail: anita.grigoriadis@kcl.ac.uk

37

38 Running title: Systemic immunity in regional lymph nodes and prognosis in breast cancer

39

40

41

Abstract

The level of stromal tumor-infiltrating lymphocytes (sTILs) in triple negative (TNBC) and HER2-positive breast cancers convey prognostic information. The importance of systemic immunity to local immunity is unknown in breast cancer. We previously demonstrated that histological alterations in axillary lymph nodes (LNs) carry clinical relevance. Here, we capture local immune responses by scoring TILs at the primary tumor and systemic immune responses by recording the formation of secondary follicles, also known as germinal centers, in 2,857 cancer-free and involved axillary LNs on haematoxylin and eosin (H&E) stained sections from a retrospective cohort of 161 LN-positive triple-negative and HER2-positive breast cancer patients. Our data demonstrates that the number of germinal center formations across all cancer-free LNs, similar to high levels of TILs, is associated with good prognosis in low TILs TNBC. This highlights the importance of assessing both primary and LN immune responses for prognostication and for future breast cancer research.

Introduction

Triple negative (TNBC) and human epidermal growth factor receptor-2 (HER2)-positive breast cancers display higher prevalence of stromal tumor-infiltrating lymphocytes (sTILs) than estrogen receptor (ER)-positive breast cancers[1-3]. The assessment of sTILs at the primary tumor site via light microscopy of haematoxylin and eosin (H&E) stained sections, has been shown to be superior to classical TNM staging in TNBC and HER2-positive breast cancers in predicting outcome[3], response to chemotherapy[4], anti-HER2 therapy[5] and to immunotherapy[6]. Although sTIL assessment is not, as yet, included in national breast cancer pathological minimum datasets, some clinicians are now requesting this information; the aim being to use the data to advise patients on the appropriateness of systemic therapies for example to de-escalate chemotherapeutic regimens in those patients with very high TILs, who have an excellent prognosis. The St Gallen International Consensus Guidelines 2019 for TNBC recommend evaluation of sTILs in these lesions[7], however, TILs' scoring should currently not be used to take treatment decisions nor to escalate or de-escalate therapy.

The presence and extent of lymph node (LN) metastasis is associated with shorter disease-free and overall survival in breast cancer[8], but LNs, as well as being typically the first site of seeding of many solid tumors, also serve as immunological hubs between the tumor and the patient's systemic immunity. Currently, routine pathological reporting does not extend beyond assessment of the presence and size of metastasis in the LNs and the presence of extra-nodal extension. Recent immunohistochemical and transcriptional studies have examined the immune context of axillary LNs, reporting qualitative changes in certain immune cell populations, such as an increase of CD68+ macrophages in cancer-free LNs associated with disease progression[9,10]. Based on extensive histopathological analyses of immune and stromal features in primary tumors and axillary LNs, we have previously detailed histological changes in cancer-free LNs that are of value in the prediction of risk of developing distant metastasis[11]. In a series of breast cancers, enriched for TNBC, LN-positive patients with

increased germinal center (GC) formation in their cancer-free LNs showed a superior outcome, even compared to LN-negative disease.

In this study, the primary objective was to capture systemic immunity, as identified by histological alterations in cancer-free LNs, and determine whether this carried clinical importance. We conducted an extensive numerical characterisation of GC formation in 2,857 involved and cancer-free axillary LNs from 161 TNBC and HER2-positive patients. sTILs and tertiary lymphoid structures (TLS) were also assessed in the primary tumors on standard diagnostic H&E-stained slides[11]. Our secondary objective was to determine whether systemic immune responses would modify the prognostic effect of local sTILs density, indicating that the assessment of the combination of primary and nodal immune response would aid in prognostication.

Results

Patient characteristics

We selected a cohort of patients with invasive breast carcinoma treated between 2005 and 2010 at Tianjin Medical University Cancer Hospital, China, consisting of 161 grade 3 no special type (IBC-NST) HR-negative carcinomas (HER2-positive or TNBC) with positive LNs (Figure 1). The clinicopathological features of the HER2-positive group were comparable to the TNBC group, with a marginally higher frequency of lympho-vascular invasion (79% versus 62%, Chi-squared test, *P*=.02) and of higher nodal stage (pN3 27% versus 15%, Chi-squared test, *P*=.04) in HER2-positive breast cancer patients (Table 1). For distant disease free survival (dDFS), median follow up was 9.08 years (range, 0.92 to 14.3 years). During follow-up, 34 (21%) patients died of cancer and 47 patients (29%) developed a recurrence, including 17 (11%) local or regional tumor recurrence, and 42 (26%) distant metastasis; of these, 70% developed metastasis within the first 3 years after diagnosis (range, 0.16 to 9.16 years).

sTILs and TLS in the primary tumor

As per the International Immuno-Oncology Biomarker Working Group guidelines[3], sTILs were quantified at the primary tumor site and reported as percentage estimates in increments of 10%. The median sTIL level was 10% (standard deviation 17%, range, 0% to 70%); 47% (75/161) of the carcinomas had ≥20% sTILs. Peritumoral TLS were seen in 24% of cases (38/161) (Table 1, Supplementary Figure 1), with significantly more frequently in those with ≥20% sTILs than <20% sTILs (32% versus 16%, Chi-squared test, *P*=.02, Table 2).

Germinal center formation in cancer-free and involved axillary LNs

A total of 2,212 cancer-free and 645 involved LNs from the 161 breast cancer patients were reviewed; median was 14 cancer-free LNs (range, 2 to 31) and 3 involved LNs (range, 1 to 18) per patient (Table 2). The number of GCs in each LN was assessed and recorded. Cancer-free LNs with more GC numbers showed a weak correlation with larger secondary follicles (Spearman *rho=0.29, P<.001*, Supplementary Figure 2A), and had a predominantly central distribution of the GCs within the LN (peripheral vs predominantly peripheral, Mann-Whitney U test, *P<.001*; peripheral vs predominantly central, Mann-Whitney U test, *P=.001*; Supplementary Figure 2B). No significant correlation with GC size or significant difference in the distribution of GCs was observed in involved LNs (Supplementary Figure 2A and B). Across 2,857 LNs, cancer-free and involved LNs with at least 1 GC were found in 137 (86%) and 122 (76%) patients, respectively. Only 7% (11/161) patients had no GCs in any of their nodes (range of assessed LNs per patient, 10-17).

Patients with tumors with fewer sTILs (<20%) at the primary site had more LNs without any GCs (for all LNs, 12% versus 1%, *P*=.01; for cancer-free LNs 21% versus 7%, *P*=.01; for involved LNs 22% versus 9%, *P*=.04, Chi-squared test, Table 2) and fewer total numbers of GC in their cancer-free LNs (Mann-Whitney U test, *P=.036*, Figure 2A). Considering only patients with any GC formation in their LNs, the median number of cancer-free LNs bearing GCs was statistically higher when sTILs in the primary cancer were ≥20%, compared to those

cases where sTILs were <20% (median 4, range, 1 to 22, versus median 2, range, 1 to 17, Kruskal Wallis test, *P*<.01, Table 2). No difference in the number of cancer-free LNs with GCs, nor between the number of involved LNs with GCs, was observed between the two breast cancer subtypes (Table 2).

Per patient, the total number of GCs in all of the cancer-free LNs was on average 8 (range, 0 to 175) and was 8 (range, 0 to 214) in the total of the involved LNs. In 23/161 (14%) patients ALNC was performed after positive sentinel lymph node biopsies, allowing the comparison of GC formation in sentinel versus other axillary LNs (Supplementary Table 1). In patients with >2 GCs in all assessed cancer-free LNs, the majority of GCs were observed in LNs excised by SLNBs, including involved and cancer-free nodes, in comparisons to nodes obtained by ALNC. In 4/23 patients with SLNB (#20, #21, #22 and #23), neither cancer-free nor involved LNs displayed any GC formation. In patient #19, where a total of 2 GCs were observed amongst all assessed cancer-free LNs, a single GC formation was observed in a node excised by SLNB, whilst the other was in an axillary LN.

When the number of GCs was compared in individual cancer-free and involved LNs, this harboured a median of 3 (range, 0 to 35) and 5 (range, 0 to 43), respectively (Table 2). In the group of carcinomas with ≥20% sTILs: (i) the total GC numbers were higher in both cancer-free and involved LNs compared to those with <20% sTILs; (ii) the maximum GC number in a cancer-free and involved LNs was greater; and (iii) on average any one individual cancer-free or involved LN had more GCs (Table 2). Furthermore, the total number of GCs per patient correlated with the maximum GC number (Spearman *rho=0.95, P<.001,* Figure 2B; Supplementary Figure 2C) and with the number of LNs with GCs in cancer-free LNs (Spearman *rho=0.89, P<.001*, Figure 2B; Supplementary Figure 2C). However, only a moderate correlation was observed between the total number of GCs and the number of assessed LNs, when including both cancer-free and involved LNs (Spearman *rho=0.41, P<.001,* Figure 2C; Supplementary Figure 2D), and when only cancer-free assessed LNs were

tested (Spearman *rho=0.43, P<.001,* Figure 2C, Supplementary Figure 2D). Given, the correlation amongst these different GC assessments, and their independence to the number of assessed LNs, the total number of GCs per patient was used for further analyses.

Association of GC numbers in LNs with clinicopathological features

Patients with TLS adjacent to the primary carcinomas had more GCs in their involved LNs, but not in their cancer-free LNs (Mann-Whitney U test, *P*<.001 and *P*=.21, respectively, Figure 2D). The number of GCs in the total cancer-free LNs per patient decreased slightly with age at diagnosis (Spearman *rho*=-0.32, *P*<.001, Supplementary Figure 2E). The GC number in involved LNs increased with nodal status (Mann-Whitney U test, *P*=.02, Supplementary Figure 2F). No association was observed between GC number (either in involved or cancer-free LNs) with tumor size or the presence of lympho-vascular invasion (Mann-Whitney U test, *P*>.05; Supplementary Figure 2G & H).

Association of GC number in LNs with prognosis

In concordance with recent research[4], an increased sTILs density was associated with improved outcome for all endpoints (invasive Disease Free Survival (iDFS): hazard ratio (HR)=0.96, 95% confidence interval (CI) 0.93-0.98, *P*<.001; dDFS: HR=0.96, 95%CI 0.93-0.98, *P*<.001; overall survival (OS): HR=0.94, 95%CI 0.91-0.98, *P*<.001; Supplementary Table 2). The presence of TLS was also associated with an improved outcome for all endpoints (iDFS: HR=0.25; 95% CI 0.09-0.71, *P* <.001; dDFS: HR=0.21, 95% CI 0.06-0.67, *P*=.001; OS: HR=0.08, 95% CI 0.01-0.59, *P* <.001; Supplementary Table 2). To consolidate whether the number of GCs across all assessed cancer-free LNs is associated with prognosis in this cohort, as we have shown previously[11], we performed an iterative process to determine an optimal cut off point by a minimal P value approach[12] (Supplementary Figure 3), which revealed that patients with ≤2 GCs across all assessed cancer-free LNs had poorer iDFS, dDFS and OS than patients with >2 GCs in all assessed cancer-free LNs (Table 3, Figure 3A-C). In multivariate models, when adjusted for known prognostic factors and TILs, this binary

cut-off for GCs in cancer-free LNs remained statistically associated with dDFS (HR=0.47, 95%CI 0.23-0.94, *P*=.033; Table 3), and increased in significance when only TNBC patients (n = 99) were analysed (iDFS: HR=0.37; 95% CI 0.16-0.84, *P*=.017; dDFS: HR=0.29, 95% CI 0.13-0.67, *P*=.004; Supplementary Table 3, Figure 3D-F). In the subset of HER2-positive patients (n = 62), those patients with >2 GC in all assessed cancer-free LNs had better OS (OS: HR=0.33, 95% CI 0.12-0.92, *P*=.036; Supplementary Table 2), however these significant associations were lost in the multivariate analyses (Supplementary Table 3).

Next, we asked whether the positive prognostic effect of the systemic immune response in cancer-free LNs differs in patients with different sTILs at the primary lesion. In patients with high sTILs tumors, the frequency of GCs in cancer-free LNs had no influence on disease trajectories. However, in univariate and multivariate models, patients with low sTIL tumors and >2 GCs in cancer-free LNs in comparison to those with ≤2 GC frequency had superior dDFS (HR=0.34, 95% CI 0.17-0.77, *P*=.009), and iDFS (HR=0.41, 95% CI 0.19-0.89, *P*=.023), and a tendency in OS (HR=0.48, 95% CI 0.2-1.17, *P*=.106; Figure 4A-C and Table 3A). Subgroup analyses demonstrated that this association was driven by the subset of TNBC (n = 99), in which patients >2 GC in all assessed cancer-free LNs had better dDFS (HR=0.21, 95% CI 0.08-0.55, *P*=.001), iDFS (HR=0.26, 95% CI 0.1-0.64, *P*=.003), and OS (HR=0.32, 95% CI 0.11-0.93, *P*=.036; Figure 4D-F and Table 3B).

The five-year iDFS, dDFS and OS in patients with <20% sTILs was 39%, 39% and 52% respectively for those with ≤2 GCs whilst those with >2 GCs had five-year iDFS, dDFS and OS of 73%, 76% and 85%, respectively. As 66/75 (88%) patients with high sTILs tumors have >2 GC in cancer-free LNs, the five-year iDFS, dDFS and OS could only be estimated in this subgroup and was 89%, 89% and 94%, respectively (Table 4A). In the subset of TNBC with <20% sTILs, patients with ≤2 GCs in their cancer-free LNs had five-year iDFS, dDFS and OS of 25%, 25% and 52% respectively, in comparison to patients with >2 GCs in their cancer-free

LNs who had five-year iDFS, dDFS and OS of 75%, 77% and 82%, respectively (Table 4B), illustrating a prognostic value for the number of GC formation in low TILs TNBCs.

Discussion

We describe here, in TNBC and HER2-positive cancer patients, the largest set to date of cancer-free and involved axillary LNs with matched primary tumors and show that humoral, systemic immune responses at the time of primary surgery have prognostic value. Thus, this study supports and extends our previous findings[11], since particularly in TNBC patients with low sTIL tumors, time to progression of disease was prolonged when their LNs displayed some indications of immune response. The better outcome in patients with GC formation in their cancer-free LNs, even when stromal TILs are low in the primary lesion, alludes to a systemic anticancer immune response. This data indicates that pathological assessment of GCs in cancer-free LNs, in conjunction with TILs, is of value for prognostication in high-risk patients.

All patients in this series had primary therapeutic breast surgery and axillary LN clearance, so that any anti-tumor immune response beyond that at the primary tumor site could be examined. Other models have already highlighted the importance of this systemic response; for example, successful tumor eradication after immunotherapy in genetically engineered cancer models required immune activation in the periphery[13]; and recently, Hollern and colleagues have elegantly illustrated how T follicular helper (Tfh) cell activation of B cells can facilitate anti-tumor responses to immune checkpoint inhibitors[14]. A productive GC response requires the collaboration of multiple cell types. Although the underlying stimuli that results in GC formation in breast cancer are incompletely understood, after infection or vaccination, GCs are transiently formed as B cell follicles of secondary lymphoid tissues[15] with clonal expansion of B cells, ensuring the development of long-lived pathogen-specific humoral immunity.

We observed an inverse relationship between the number of GCs in LNs and the age of the patient at diagnosis, which is in alignment with a decreased GC prevalence and volume in LNs in elderly patients, potentially resulting in a decrease in LN's reactivity[16]. While B cells still retain the ability to migrate in aging LNs and produce immunoglobulin, the number of follicular dendritic cells in LNs and the ability to hold on to immune complexes is significantly impaired, potentially as a result of poor humoral immunity in the older patients[17]. In alignment with previous reports, patients with high sTILs in the primary tumor had not only more TLS but also more GCs[18,19]. Both of these lymphoid structures may potentially indicate an effective humoral immune response in these patients, who, in general, have a better prognosis. Deciphering the fundamental drivers of GC formation in LNs in breast cancer patients may reveal mechanisms underpinning the generation of robust humoral immunity and thus identify strategies to potentially target the modulation of GCs in cancer.

Increased pathological complete response is reported in clinical trials of TNBC patients when immune checkpoint blockade immunotherapies (*e.g.* anti-PD1/PDL1) are combined with chemotherapy[20,21], and in patients with high sTILs[6]. In particular, LN-positive patients showed a greater benefit to immune checkpoint inhibitors with neoadjuvant chemotherapy in the randomised Phase III KEYNOTE-522 trial, than patients with lower risks (Δ21% for node-positive and Δ25% for stage IIIA/B disease breast cancer patients)[22]. We postulate that the systemic immune responses in node-positive breast cancer patients may be advantageous for immune checkpoint inhibitors therapy response. By further exploring these systemic immune responses (*i.e*. in LNs), we will expand on our understanding of why some patients are more likely respond to these immunotherapies.

In the present study, a significant survival improvement for LN-positive patients with low TILs was observed when cancer-free LNs harboured >2 GCs for all patient outcomes examined. In particular, the presence of numerous GCs may indicate immune responses in a patient that are not captured by their sTILs levels at the primary tumor site at the time when the tumor is

histopathologically assessed. We cannot comment on whether immune responses were previously present, however the reactivity of these secondary follicles indicates the patient's ability to mount an immune response, and potentially represents a component contributing to the better disease trajectory for these patients compared to patients without any local and systemic immune responses (*i.e.* with both low sTILs & low GC numbers). A functional influence on lymphocytes at the primary cancer by immune checkpoints in LNs has already been proposed[19], also corroborating a close connection between the primary tumor and adjacent LNs.

Of note 38% patients in the present study had HER2-positive tumors, and it is possible that an assessment of systemic immune response by examination of GCs in addition to TILs may be of predictive importance for these patients; in the A TRYPHAENA substudy those with low TILs had an inferior response to trastuzumab/pertuzumab-based chemotherapy[5]. However, our study was not intended to analyse interactions with chemotherapy or targeted agents and further research is needed to determine whether the assessment of GCs in cancer-free LNs provides additive value for prediction of immunotherapy or anti-HER2 treatment response. Recent studies have brought attention to the role of B cells, especially within TLS, which act akin to LNs within a tumor, and have noted that B cell presence is critical for response to checkpoint blockade, thereby pointing to a dynamic interaction between several components of the immune system[23]. Thus, understanding the bipartite nature of the immune system may then help to identify patient subgroups for whom targeting both T cells and B cells could improve treatment response.

Given the retrospective nature of this study, further analytical and clinical validation, as well as evaluation of reproducibility of assessment of GCs, is required. Ideally this would be undertaken on samples from patients in clinical trials, with uniform management and follow-up, but the LNs (involved or cancer-free) from such women are not typically curated in clinical trials tissue banks; this should be considered in future. Assessment of the LNs from patients

within neoadjuvant chemotherapy trials for GC numbers would provide evidence of value in this setting. Indeed, TILs have been examined in this setting and residual cancer burden (RCB) used as an endpoint[24] and this approach would similarly provide an excellent opportunity to consolidate our results.

In 14% of our study cohort, SLNB was performed, suggesting that capturing data on GC formation in SLN can reflect on the frequency of GC formation overall in axillary LNs in these patients. However, further studies are warranted to evaluate the minimum number of nodes required and whether the cut-point for GC numbers are the same. The proposed cut-offs for GC numbers in cancer-free LNs may also then need revision. Conversely, the examination and counting of GCs in all LNs in an axillary clearance requires additional pathology time and resources. Convolutional neural networks applied to digitised whole slide images can detect LN metastasis with high accuracy in some studies[25] and digital pathological approaches to the quantification of TILs have also been described[26]. The histology of GCs is suited to be captured by machine learning methods[27] and will potentially facilitate assessment in large cohorts and additional numbers of cases of all breast cancer sub-types.

In conclusion, we show that systemic immune response at the time of primary surgery, by the recording of GC formation in the cancer-free LNs, has prognostic value. This highlights that axillary LN assessment, above and beyond the presence and size of cancer cell deposits, in conjunction with sTILs, carries prognostic value in high-risk patients.

Methods

Patients

Patient selection and data analyses are reported according to Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) criteria[28]. Ethical clearance was obtained from the local research ethics committee (Medical Ethics Committee of Tianjin Medical University Cancer Institute and Hospital, Ek2020021). This is a retrospective study of 161 patients with invasive breast carcinoma of no special type (IBC-NST) treated between 2005 and 2010 at Tianjin Medical University Cancer Hospital, China, consisting of HR-negative patients (HER2-positive or TNBC) with positive LNs and of histological grade 3 (Figure 1). The median age at diagnosis was 52 years (range, 23 to 75). All patients underwent modified radical mastectomy or breast-conserving surgery and had axillary LN dissection. None of the patients had prior history of breast or axillary surgery, or suffered from Small Lymphocytic Lymphoma, Chronic Lymphocytic Leukemia, dermatopathic lymphadenopathy, benign inflammatory disease of the breast or upper limb. None had neoadjuvant systemic therapy. Postoperatively, all patients received adjuvant chemotherapy; 85% anthracycline plus taxane, 12% anthracycline-based (and another 3% taxane only-based chemotherapy (Table 1). In this period HER2-positive patients in China did not receive any anti-HER2 therapy.
Clinicopathological data are recorded in Table 1.

Histopathological assessment of primary tumor and LNs

Routine H&E-stained sections of formalin-fixed paraffin embedded tissue from the primary invasive breast carcinoma and involved and cancer-free LNs were scanned at 40x magnification using a NanoZoomer HT Digital Pathology Scanning System (Hamamatsu, Japan). All sections were reviewed by two breast pathologists (FL and XG) who assessed the presence and number of GCs, TILs and TLSs. A total of 2,857 axillary LNs from 161 patients were obtained, with an average of 5 sections per primary tumor and 10 to 37 (median, 17) LNs per patient.

As per the International Immuno-Oncology Biomarker Working Group guidelines[3], sTIL density was quantitatively assessed and reported as a percentage estimate, in increments of 10%. Patient groups were dichotomised into those with <20% or ≥20% sTIL, in keeping with recent literature[24,29]. TLS were defined as a follicular structure in the peritumoral stroma on H&E stains[30], and were reported as present or absent (Supplementary Figure 1). No immunohistochemical stains for immune cells were used, so this may represent an underestimation of TLS numbers, but represents day-to-day pathology practice. Under conditions of antigenic stimulation, LNs develop secondary follicles composed of a peripheral area of closely packed, small lymphocytes and a centrally located GC. We defined GCs in H&E-stained sections as lighter areas within the small mature lymphoid population composed of both larger lymphoid cells and cells of a non-lymphoid nature. The pathologist chose one of the LN slices with the most GCs and recorded the number of GCs in one LN. Using the NDP.view software of the NanoZoomer Scanning System, the size of each GC, defined as the maximum dimension, was recorded as a continuous variable. The localisation of GCs within LNs was classified as peripheral, predominantly peripheral (more GCs close to the capsule), central and predominantly central (more GCs in the centre of the LN), as previously described[11].

Statistical analysis

Standard summary statistics were performed, to establish if there were associations between GC number, sTILs, TLS and clinicopathological characteristics and with patient outcome. The primary endpoint was distant Disease Free Survival, defined as the date of first distant recurrence or death from any cause. Invasive Disease Free Survival was defined as the date of first invasive recurrence, or second primary, or death from any cause[31]. Overall Survival was defined as the date of death from any cause. For all these analyses patients still alive were censored at the date of the last visit.

A Kaplan-Meier method was used to visualise survival curves and the log-likelihood test to compare survival curves across groups. Follow-up was curtailed at 10 years because of the declining numbers of patients after this time point. Cox regression proportional hazards models were performed to estimate the hazard ratios according to clinicopathological and histological-assessed features across all endpoints in univariate and multivariate analyses. Statistical significance of features was assessed using the log-likelihood test whereby a two-sided P <0.05 was considered significant. Statistical analyses were performed in the statistical environment R 3.5.1.

Code availability

Available upon request.

Data availability

De-identified data, including clinical covariates are available upon request.

References

1 Denkert, C. *et al.* Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol* **19**, 40-50, doi:10.1016/S1470-2045(17)30904-X (2018).

2 Loi, S. *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* **31**, 860-867, doi:10.1200/JCO.2011.41.0902 (2013).

3 Salgado, R. *et al.* The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol* **26**, 259-271, doi:10.1093/annonc/mdu450 (2015).

4 Loi, S. *et al.* Tumor-Infiltrating Lymphocytes and Prognosis: A Pooled Individual Patient Analysis of Early-Stage Triple-Negative Breast Cancers. *J Clin Oncol* **37**, 559-569, doi:10.1200/JCO.18.01010 (2019).

5 Ignatiadis, M. *et al.* Tumor-Infiltrating Lymphocytes in Patients Receiving Trastuzumab/Pertuzumab-Based Chemotherapy: A TRYPHAENA Substudy. *J Natl Cancer Inst* **111**, 69-77, doi:10.1093/jnci/djy076 (2019).

6 Loi, S. *et al.* n *Cancer Research.* Relationship between tumor-infiltrating lymphocytes (TILs) and outcomes in the KEYNOTE-119 study of pembrolizumab vs chemotherapy for previously treated metastatic triple-negative breast cancer (mTNBC). In: *Cancer Research*. *San Antonio*, *2020*: Abstract 80, pp. PD5-03.

7 Burstein, H. J. *et al.* Estimating the benefits of therapy for early-stage breast cancer: the St. Gallen International Consensus Guidelines for the primary therapy of early breast cancer 2019. *Ann Oncol* **30**, 1541-1557, doi:10.1093/annonc/mdz235 (2019).

8 Dent, R. *et al.* Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res* **13**, 4429-4434, doi:10.1158/1078-0432.CCR-06-3045 (2007).

9 Chatterjee, G. *et al.* Molecular patterns of cancer colonisation in lymph nodes of breast cancer patients. *Breast Cancer Res* **20**, 143, doi:10.1186/s13058-018-1070-3 (2018).

10 Lopez, C. *et al.* The Immune Response in Nonmetastatic Axillary Lymph Nodes Is Associated with the Presence of Axillary Metastasis and Breast Cancer Patient Outcome. *Am J Pathol* **190**, 660-673, doi:10.1016/j.ajpath.2019.11.002 (2020).

11 Grigoriadis, A. *et al.* Histological scoring of immune and stromal features in breast and axillary lymph nodes is prognostic for distant metastasis in lymph node-positive breast cancers. *J Pathol Clin Res* **4**, 39-54, doi:10.1002/cjp2.87 (2018).

12 Camp, R. L., Dolled-Filhart, M. & Rimm, D. L. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* **10**, 7252-7259, doi:10.1158/1078-0432.CCR-04-0713 (2004).

13 Spitzer, M. H. *et al.* Systemic Immunity Is Required for Effective Cancer Immunotherapy. *Cell* **168**, 487-502 e415, doi:10.1016/j.cell.2016.12.022 (2017).

14 Hollern, D. P. *et al.* B Cells and T Follicular Helper Cells Mediate Response to Checkpoint Inhibitors in High Mutation Burden Mouse Models of Breast Cancer. *Cell* **179**, 1191-1206 e1121, doi:10.1016/j.cell.2019.10.028 (2019).

15 Victora, G. D. & Nussenzweig, M. C. Germinal centers. *Annu Rev Immunol* **30**, 429-457, doi:10.1146/annurev-immunol-020711-075032 (2012).

16 Lazuardi, L. *et al.* Age-related loss of naive T cells and dysregulation of T-cell/B-cell interactions in human lymph nodes. *Immunology* **114**, 37-43, doi:10.1111/j.1365-2567.2004.02006.x (2005).

17 Turner, V. M. & Mabbott, N. A. Structural and functional changes to lymph nodes in ageing mice. *Immunology* **151**, 239-247, doi:10.1111/imm.12727 (2017).

18 Buisseret, L. *et al.* Tumor-infiltrating lymphocyte composition, organization and PD-1/ PD-L1 expression are linked in breast cancer. *Oncoimmunology* **6**, e1257452, doi:10.1080/2162402X.2016.1257452 (2017).

19 Quintana, A. *et al.* Immune Analysis of Lymph Nodes in Relation to the Presence or Absence of TILs in TNBC. In: *ESMO*, Abstract 31(suppl_2) (2020).

20 Nanda, R. *et al.* Pembrolizumab plus standard neoadjuvant therapy for high-risk breast cancer (BC): Results from I-SPY 2., Abstract 55, p. 506-506. Journal of Clinical Oncology (2017).

21 Schmid, P. *et al.* Pembrolizumab plus chemotherapy as neoadjuvant treatment of high-risk, early-stage triple-negative breast cancer: results from the phase 1b open-label, multicohort KEYNOTE-173 study. *Ann Oncol* **31**, 569-581, doi:10.1016/j.annonc.2020.01.072 (2020).

22 Schmid, P. *et al.* Pembrolizumab for Early Triple-Negative Breast Cancer. *N Engl J Med* **382**, 810-821, doi:10.1056/NEJMoa1910549 (2020).

23 Bruno, T. C. New predictors for immunotherapy responses sharpen our view of the tumour microenvironment. *Nature* **577**, 474-476, doi:10.1038/d41586-019-03943-0 (2020).

24 Luen, S. J. *et al.* Prognostic implications of residual disease tumor-infiltrating lymphocytes and residual cancer burden in triple-negative breast cancer patients after neoadjuvant chemotherapy. *Ann Oncol* **30**, 236-242, doi:10.1093/annonc/mdy547 (2019).

25 Litjens, G. *et al.* 1399 H&E-stained sentinel lymph node sections of breast cancer patients: the CAMELYON dataset. *Gigascience* **7**, doi:10.1093/gigascience/giy065 (2018).

26 Amgad, M. *et al.* Report on computational assessment of Tumor Infiltrating Lymphocytes from the International Immuno-Oncology Biomarker Working Group. *NPJ Breast Cancer* **6**, 16, doi:10.1038/s41523-020-0154-2 (2020).

27 Grigoriadis, A. *et al.* Assessments of cancer-free lymph nodes for the prediction of disease progression, in *Molecular Analysis for Personalised Therapy (MAP).* vii28 Elsevier, (2019).

28 McShane, L. M. *et al.* Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst* **97**, 1180-1184, doi:10.1093/jnci/dji237 (2005).

29 Sharma, P. *et al.* Validation of the DNA Damage Immune Response Signature in Patients With Triple-Negative Breast Cancer From the SWOG 9313c Trial. *J Clin Oncol* **37**, 3484-3492, doi:10.1200/JCO.19.00693 (2019).

30 Hiraoka, N., Ino, Y. & Yamazaki-Itoh, R. Tertiary Lymphoid Organs in Cancer Tissues. *Front Immunol* **7**, 244, doi:10.3389/fimmu.2016.00244 (2016).

31 Gourgou-Bourgade, S. *et al.* Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials)dagger. *Ann Oncol* **26**, 873-879, doi:10.1093/annonc/mdv106 (2015).

Acknowledgements

The authors would like to thank all members of the Cancer bioinformatics team at King's College London for their helpful suggestions. We would like to thank Professors Elinor Sawyer and Agamemnon Grigoriadis for critical reading and their suggestions.

Author contributions

Study concept and design (F.L, S.E.P, A.G.); acquisition, analysis, or interpretation of data (all authors); drafting of the paper (F.L, R.S., S.E.P, A.G.); critical revision of the manuscript for important intellectual content (all authors); administrative, technical, or material support.

Competing interests

None

Materials and Correspondence

Dr Anita Grigoriadis, Breast Cancer Now Unit, Innovation Hub, Cancer Center at Guy's, Great Maze Pond, London, SE1 9RT, UK

Phone: +44 020 7188 2360

e-mail: anita.grigoriadis@kcl.ac.uk

Funding

F. Liu was awarded a K.C. Wong Postdoctoral Fellowship to study at King's College London. This work was supported by Breast Cancer Research Trust (now part of Breast Cancer Now), Cancer Research UK and Breast Cancer Now (and their legacy charity Breakthrough Breast Cancer) [CRUK/07/012, KCL-BCN-Q3]. R.S. is supported by a grant from the Breast Cancer Research Foundation [BCRF, grant No. 17-194]. X. Guo received a grant from the National Natural Science Foundation of China [grant No. 81772840]. The authors acknowledge financial support from Mrs Alison Millford.

Tables

Table 1. Clinicopathological features and immune features of the primary tumor

	All cases	HER2	TNBC		sTILs<20%	sTILs≥20%	
	n=161(%)	n=62(%)	n=99(%)	P_value	n=86(%)	n=75(%)	P_value
sTILs							
< 20%	86 (53.4)	34 (54.8)	52 (52.5)				
≥20%	75 (46.6)	28 (45.2)	47 (47.5)	0.775*			
Tertiary lymphoid structures							
Absent	123 (76.4)	46 (74.2)	77 (77.8)		72 (83.7)	51 (68.0)	
Present	38 (23.6)	16 (25.8)	22 (22.2)	0.602*	14 (16.3)	24 (32.0)	0.019*
Age at diagnosis							
<50	66 (41)	20 (32.3)	46 (46.5)		35 (40.7)	31 (41.3)	
≥50	95 (59)	42 (67.7)	53 (53.5)	0.074*	51 (59.3)	44 (58.7)	0.935*
Tumor size							
pT1	33 (20.5)	11 (17.7)	22 (22.2)		12 (14.0)	21 (28.0)	
pT2	116 (72)	46 (74.2)	70 (70.7)		65 (75.6)	51 (68.0)	
pT3	10 (6.2)	4 (6.5)	6 (6.1)		8 (9.3)	2 (2.7)	
pT4	2 (1.2)	1(1.6)	1 (1)	0.906*	1 (1.2)	1 (1.3)	0.071*
Histological grade							
III	161 (100)	62 (100)	99 (100)		86 (100)	75 (100)	
Lymphovascular invasion							
Absent	51 (31.7)	13 (21)	38 (38.4)		24 (27.9)	27 (36.0)	
Present	110 (68.3)	49 (79)	61 (61.6)	0.021*	62 (72.1)	48 (64.0)	0.271*
Lymph node status							
pN1 (1-3)	90 (55.9)	27 (43.6)	63 (63.6)		47 (54.7)	43 (57.3)	
pN2 (4-9)	39 (24.2)	18 (29)	21 (21.2)		20 (23.3)	19 (25.3)	
pN3 (>=10)	32 (19.9)	17 (27.4)	15 (15.2)	0.037*	19 (22.1)	13 (17.3)	0.748*
Chemotherapy							
Anthracycline + taxane	137 (85.1)	55 (88.7)	82 (82.8)		77 (89.5)	60 (80)	
Anthracycline	19 (11.8)	5 (8.1)	14 (14.1)		7 (8.1)	12 (16)	
Taxane	5 (3.1)	2 (3.2)	3 (3)	0.508*	2 (2.3)	3 (4)	0.236*
Local or regional tumor recurrence							
Absent	144 (89.4)	55 (88.7)	89 (89.9)		73 (84.9)	71 (94.7)	
Present	17 (10.6)	7 (11.3)	10 (10.1)	0.811*	13 (15.1)	4 (5.3)	0.044*
Distant metastasis							
Absent	119 (73.9)	46 (74.2)	73 (73.7)		52 (60.5)	67 (89.3)	
Present	42 (26.1)	16 (25.8)	26 (26.3)	0.949*	34 (39.5)	8 (10.7)	<0.001*
Breast Cancer-Specific Death							
Absent	127 (78.9)	47 (75.8)	80 (80.8)		57 (66.3)	70 (93.3)	
Present	34 (21.1)	15 (24.2)	19 (19.2)	0.449*	29 (33.7)	5 (6.7)	<0.001*

* chi-squared test

Table 2. Germinal centers in involved and cancer-free lymph nodes

	All cases	HER2	TNBC		sTILs<20%	sTILs≥20%	
	n=161	n=62	n=99	P value	n=86	n=75	P value
LN assessment							
All LNs, median (range)	17 (10-37)	17 (10-29)	17 (10-37)		17 (10-31)	17 (10-37)	
Cancer-free LNs, median (range)	14 (2-31)	13 (2-24)	16 (3-31)		14 (2-26)	16 (3-31)	
Involved LNs, median (range)	3 (1-18)	4 (1-18)	2 (1-18)		3 (1-17)	3 (1-18)	
GC assessment in LNs per patient basis							
All LNs, n (%)							
GC absent	11 (6.8)	5 (8.1)	6 (6.1)		10 (11.6)	1 (1.3)	
GC present	150 (93.2)	57 (91.9)	93 (93.9)	0.624*	76 (88.4)	74 (98.7)	0.010*
Cancer-free LNs							
GC NA[&]	1		1			1	
GC absent	23 (14.4)	10 (16.1)	13 (13.3)		18 (20.9)	5 (6.8)	
GC present	137 (85.6)	52 (83.9)	85 (86.7)	0.615*	68 (79.1)	69 (93.2)	0.011*
Involved LNs							
GC NA[&&]	13 (8.1)	4 (6.5)	9 (9.1)		5 (5.8)	8 (10.7)	
GC absent	26 (16.1)	8 (12.9)	18 (18.2)		19 (22.1)	7 (9.3)	
GC present	122 (75.8)	50 (80.6)	72 (72.7)	0.333*	62 (72.1)	60 (80)	0.038*
LN number GC present							
Cancer-free LN, median (range)	3 (1-22)	3 (1-13)	3 (1-22)	0.552[$]	2 (1-17)	4 (1-22)	0.002[$]
Involved LN, median (range)	1 (1-16)	2 (1-16)	1 (1-12)	0.294[$]	1 (1-7)	1 (1-16)	0.598[$]
Total number of GCs across all assessed LNs per patient							
Cancer-free LN, median (range)	8 (0-175)	6 (0-142)	9 (0-175)	0.139[$]	6 (0-145)	12 (0-175)	0.002[$]
Involved LN, median (range)	8 (0-214)	9 (0-198)	7 (0-214)	0.508 [$]	5 (0-198)	14 (0-214)	0.002[$]
Max GC number in a LN across all assessed LNs per patient							
Cancer-free LN, median (range)	5 (0-63)	4 (0-59)	5 (0-63)	0.076[$]	4 (0-59)	6 (0-63)	0.002[$]
Involved LN, median (range)	7 (0-76)	7 (0-76)	6 (0-54)	0.611[$]	3 (0-76)	10 (0-54)	0.003[$]
Average GC number [^]							
Cancer-free LN, median (range)	3 (0-35)	3 (0-19)	3 (0-35)	0.091[$]	3 (0-17)	4 (0-35)	0.001[$]
Involved LN, median (range)	5 (0-43)	5 (0-40)	5 (0-43)	0.942[$]	3 (0-40)	8 (0-43)	0.001[$]

* chi-squared test, [$]Mann-Whitney U test, [&] uninterpretable LN slide, [&&] whole LN involved, [^] mean GC number on a lymph node level

Table 3. Univariate and multivariate Cox regression analyses of germinal center numbers in cancer-free LNs for iDFS, dDFS, and OS of HR-negative, their TILs sub-groups, all TNBC and low TILs TNBC.

A) All HR-negative cases

	iDFS				dDFS				OS			
All cases (n=161)												
Univariate	Model P		HR	CI	Model P		HR	CI	Model P		HR	CI
Total GCs number (≤2 / >2)	<0.001		0.33	0.19 - 0.59	<0.001		0.26	0.14 - 0.48	<0.001		0.28	0.14 - 0.55
Adjusted for	Age, pTstage, pNstage, LVI, sTILs & TLS											
Multivariate	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI
Total GCs number (≤2 / >2)	0.110	<0.001	0.58	0.30 - 1.12	0.033	<0.001	0.47	0.23 - 0.94	0.351	<0.001	0.69	0.32 - 1.50
<20% sTILs (n=86)												
Univariate	Model P		HR	CI	Model P		HR	CI	Model P		HR	CI
Total GCs number (≤2 / >2)	0.002		0.36	0.19 - 0.69	<0.001		0.28	0.14 - 0.56	0.006		0.36	0.17 - 0.75
Adjusted for	Age, pTstage, pNstage, LVI & TLS											
Multivariate	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI
Total GCs number (≤2 / >2)	0.023	0.004	0.41	0.19 - 0.89	0.009	<0.001	0.34	0.17 - 0.77	0.106	0.001	0.48	0.20 - 1.17
≥20% sTILs (n=75)												
Univariate	Model P		HR	CI	Model P		HR	CI	Model P		HR	CI
Total GCs number (≤2 / >2)	0.804		1.29	0.16 - 10.16	0.986		1.02	0.13 - 8.29	0.652		0.59	0.07- 5.25
Adjusted for	Age, pTstage, pNstage, LVI & TLS											
Multivariate	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI
Total GCs number (≤2 / >2)	0.949	0.034	0.93	0.11 - 7.93	0.665	0.031	0.61	0.07 - 5.64	Group size too small / too few events			

492

493

B) Triple-negative breast cancers

	iDFS				dDFS				OS			
All cases (n=99)												
Univariate	Model P		HR	CI	Model P		HR	CI	Model P		HR	CI
Total GCs number (≤2 / >2)	<0.001		0.25	0.12 - 0.52	<0.001		0.20	0.09 - 0.44	0.004		0.24	0.10 - 0.60
Adjusted for	Age, pTstage, pNstage, LVI, sTILs & TLS											
Multivariate	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI
Total GCs number (≤2 / >2)	0.017	<0.001	0.37	0.16 – 0.84	0.004	<0.001	0.29	0.13 – 0.67	0.119	<0.001	0.46	0.17 - 1.22
<20% sTILs (n=52)												
Univariate	Model P		HR	CI	Model P		HR	CI	Model P		HR	CI
Total GCs number (≤2 / >2)	<0.001		0.25	0.11 - 0.57	<0.001		0.21	0.09- 0.49	0.013		0.29	0.11 - 0.76
Adjusted for	Age, pTstage, pNstage, LVI & TLS											
Multivariate	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI	Covariate P	Model P	HR	CI
Total GCs number (≤2 / >2)	0.003	0.016	0.26	0.1- 0.64	0.001	0.004	0.21	0.08- 0.55	0.036	0.005	0.32	0.11 - 0.93

Table 4. 5-year outcome for patients by TILs in primary cancers & germinal center subgroups

A) All HR-negative cases

	Number (%)	5-Year iDFS, (95%CI)	5-Year dDFS, (95%CI)	5-Year OS, (95%CI)
Low sTILs (n=86)				
≤2 GCs	31 (36)	39 (22-55)	39 (22-55)	52 (33-68)
>2 GCs	55 (64)	73 (59-83)	76 (62-85)	85 (72-92)
High sTILs (n=75)				
≤2 GCs	9 (12)	100 (-)	100 (-)	100 (-)
>2 GCs	66 (88)	89 (79-95)	89 (79-95)	94 (84-98)

B) Triple-negative breast cancers

	Number (%)	5-Year iDFS, (95%CI)	5-Year dDFS, (95%CI)	5-Year OS, (95%CI)
Low sTILs (n=52)				
≤2 GCs	16 (31)	25 (8-47)	25 (8-47)	52 (25-74)
>2 GCs	36 (69)	75 (58-86)	77 (60-88)	82 (66-92)
High sTILs (n=47)				
≤2 GCs	5 (11)	100 (-)	100 (-)	100 (-)
>2 GCs	42 (89)	90 (76-96)	90 (76-96)	95 (82-99)

Figure legends

Figure 1. CONSORT diagram. IBC, Invasive breast cancer; NAT, neoadjuvant therapy; IBC-NST, invasive breast cancer of no special type; ER, estrogen receptor; PR, progesterone receptor; sTILs, stromal tumor-infiltrating lymphocytes.

Figure 2. Association between germinal centre formation in lymph nodes, stromal TILs and tertiary lymphoid structures. (A) Violin plots, showing the distribution of germinal centers (GCs) compared to sTILs with 20% cut off (X axis), in cancer-free LNs (left), and involved LNs (right); **(B)** Scatter plot of the number of GCs compared to the maximum number of GCs in a single LN (left side); and compared to the number of LNs which contain GCs (right side); **(C)** Scatter plots of the number of GCs in all assessed LNs (left) and all cancer-free assessed LNs (right) compared to the number of LNs; **(D)** Violin plots, showing the distribution of GCs compared to peritumoral TLS, in cancer-free LNs (left), and involved LNs (right). Mann Whitney U tests were used to calculate P values.

Figure 3. Kaplan-Meier survival analyses predicting: (A) invasive Disease-Free Survival (iDFS), **(B)** distant Disease-Free Survival (dDFS), **(C)** Overall Survival (OS), **(D)** invasive Disease-Free Survival (iDFS) in TNBC, **(E)** distant Disease-Free Survival (dDFS) in TNBC and **(F)** Overall Survival (OS) in TNBC. Patients were dichotomized into those with ≤ 2GCs versus >2 GCs in all assessed cancer-free LNs. P values correspond to likelihood ratio tests.

Figure 4. Association between germinal center formation in lymph nodes and prognosis in HR-negative breast cancers. Kaplan-Meier curves: **(A)** invasive Disease-Free Survival (iDFS), **(B)** distant Disease-Free Survival (dDFS), **(C)** Overall Survival (OS), **(D)** invasive Disease-Free Survival (iDFS) in TNBC, **(E)** distant Disease-Free Survival (dDFS) in TNBC, and **(F)** Overall Survival (OS) in TNBC, according to stromal tumor-infiltrating lymphocytes (TILs) and germinal center (GC) number. Patient groups were stratified by TILs

(≥20%, <20%) and the number of GCs (≤ 2GCs, > 2GC) in all assessed cancer-free LNs, as categorical variables. P values correspond to likelihood ratio tests.

Data Supplement

Supplementary Figure 1. Examples of tertiary lymphoid structures (TLS) in primary tumor and germinal centers (GCs) in cancer-free lymph nodes. (A, B) presence of TLS (with lymphoid GC formation, arrows) in the peritumoral stroma (inset showing a higher power view of TLS); **(C)** few GCs (arrow) located close to the LN capsule (predominantly in periphery) (inset showing a higher power view of GC); **(D)** numerous GC formation (arrow); larger, rounder and fuller GCs throughout the whole LN (predominantly in the central part of the LN) (inset showing a higher power view of GC).

Supplementary Figure 2. Association of the number of germinal centers in cancer-free and involved lymph nodes with clinicopathological features. (**A**) Scatter plots of the number of GCs compared to the maximum GC size. Cancer-free and involved LNs analysis is shown on the left and right, respectively. (**B**) Violin plot of maximum GC size in GC with regards to their predominant location in the LN. P values displayed calculated via Mann-Whitney U tests. Cancer-free and involved LNs analysis is shown on the left and right, respectively. (**C**) Scatter plot of the log10 number of GCs compared to the maximum number of GCs in a single LN (left side); and compared to the number of LNs which contain GCs (right side). (**D**) Scatter plots of the log10 number of GCs in all assessed LNs (left) or cancer-free LNs (right) compared to the number of LNs (including those with zero GCs). (**E**) Scatter plots of the number of GCs in cancer-free LNs (left) or involved LNs (right) compared to age at diagnosis; (**F**) Violin plot of the number of GCs in cancer-free LNs (left) or involved LNs (right) involved LNs (Y axis) compared to LN stage; (**G**) Violin plot, showing the distribution of GC formations in cancer-free LNs (left) or involved LNs (right) compared to the size of the primary tumor. P values displayed calculated via Mann-Whitney U test; (**H**) Violin plots, displaying the

number of GCs in cancer-free LNs (left) or involved LNs (right) compared to patients with and without lympho-vascular invasion.

Supplementary Figure 3. Cut off selection for defining the number of germinal centers in cancer-free lymph nodes. The optimal germinal center number cut off was determined by taking all patients (including those with zero GCs) and applying an iterative process using a minimal p value approach. For all three endpoints a single cut-off was identified (dashed black lines) at 2 germinal centers.

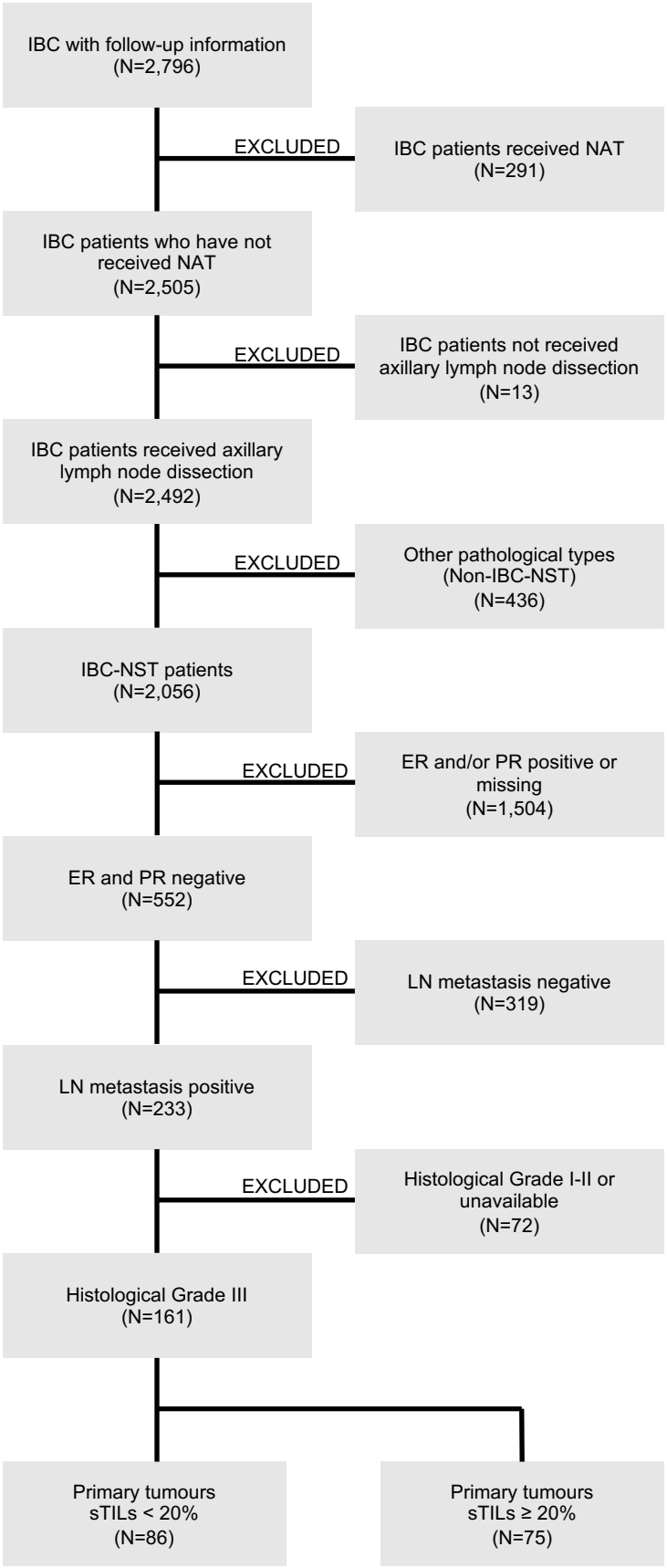
Supplementary Table 1. The distribution of total GCs in LNs excised by SLNB and ALNC. Patients are listed who had a SLNB followed by an ALNC. For each patient, the numbers of LNs and numbers of GCs in cancer-free and involved LNs are listed separately. Within cancer-free and involved LNs, the number of LNs excised by SLNB and by ALNC (called non-SLN) are showed. The number of GCs in SLN and non-SLN, as well as the percentage of GCs formation given the total number of GCs observed in all cancer-free or involved LNs (referred to as % of total). The median number of LNs, GC and percentage of GC formation is shown.

Supplementary Table 2. Univariate & Multivariate Cox regression for iDFS, dDFS and OS of all cases or sTILs subgroups (< 20%, ≥ 20%).

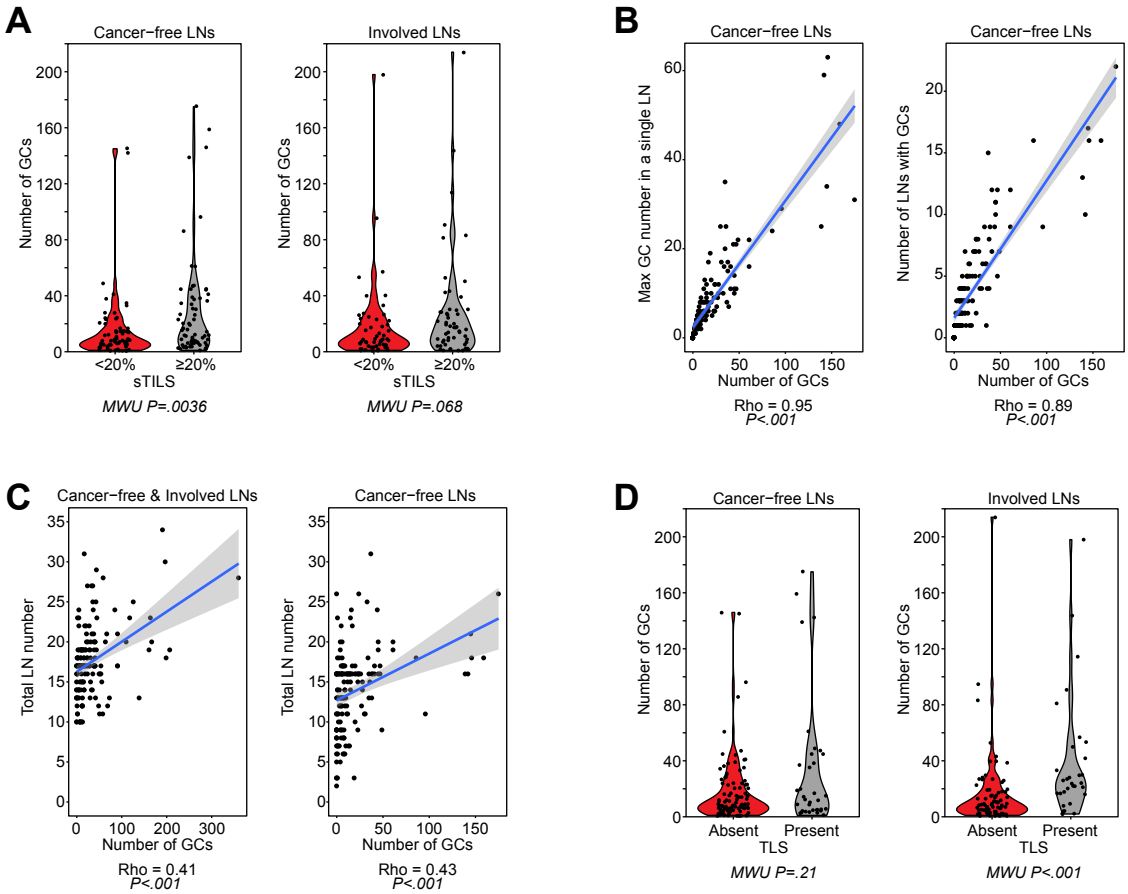
Supplementary Table 3. Univariate and Multivariate Cox regression analyses for iDFS, dDFS and OS of TNBC and HER2-positive subgroups.

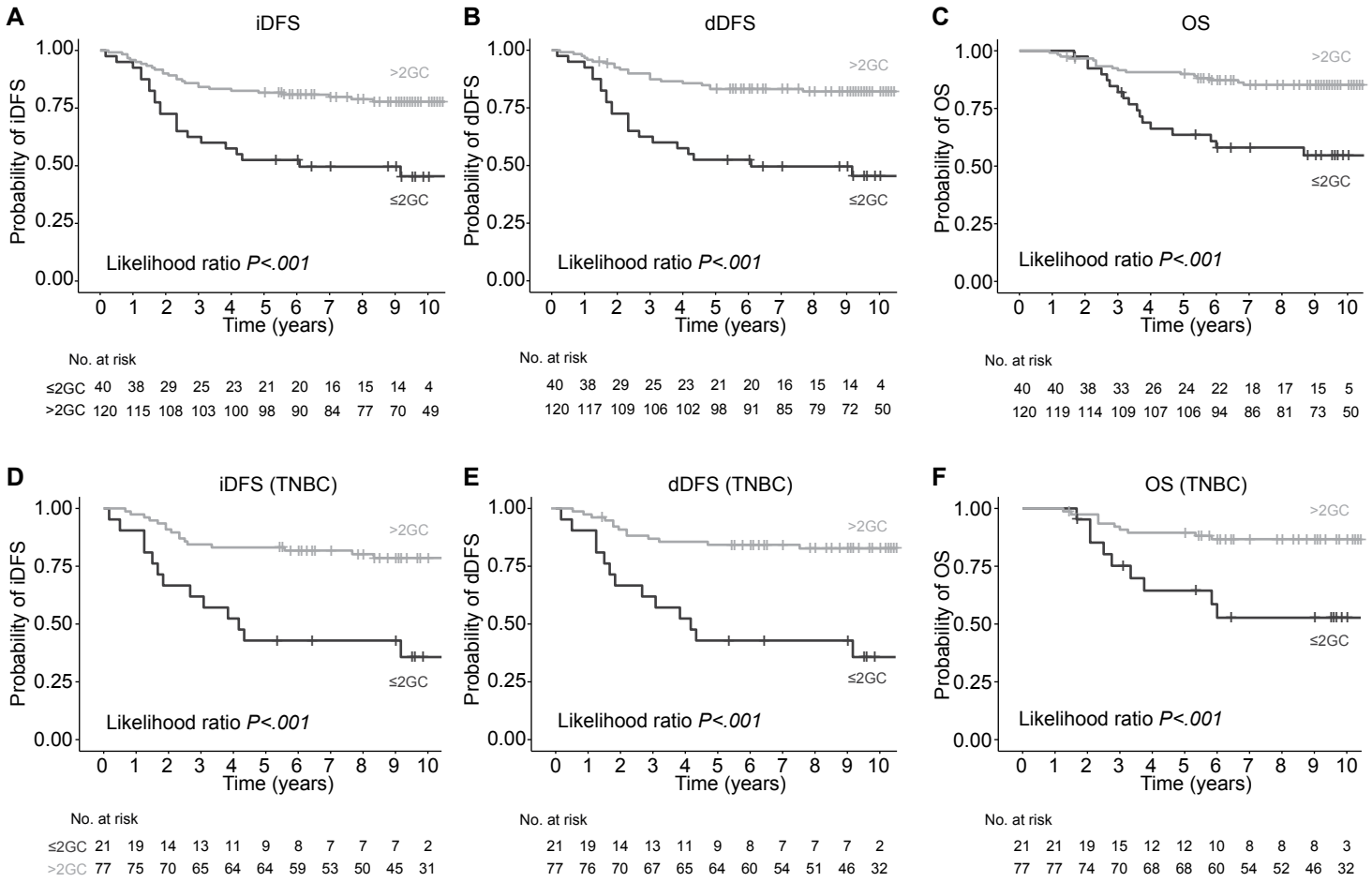
Supplementary Table 4: Multivariate Cox regression analyses for iDFS, dDFS and OS comparing the prognostic value of the proportion of LNs with GCs and the ratio of GCs in LNs.

Liu et al., Figure 1

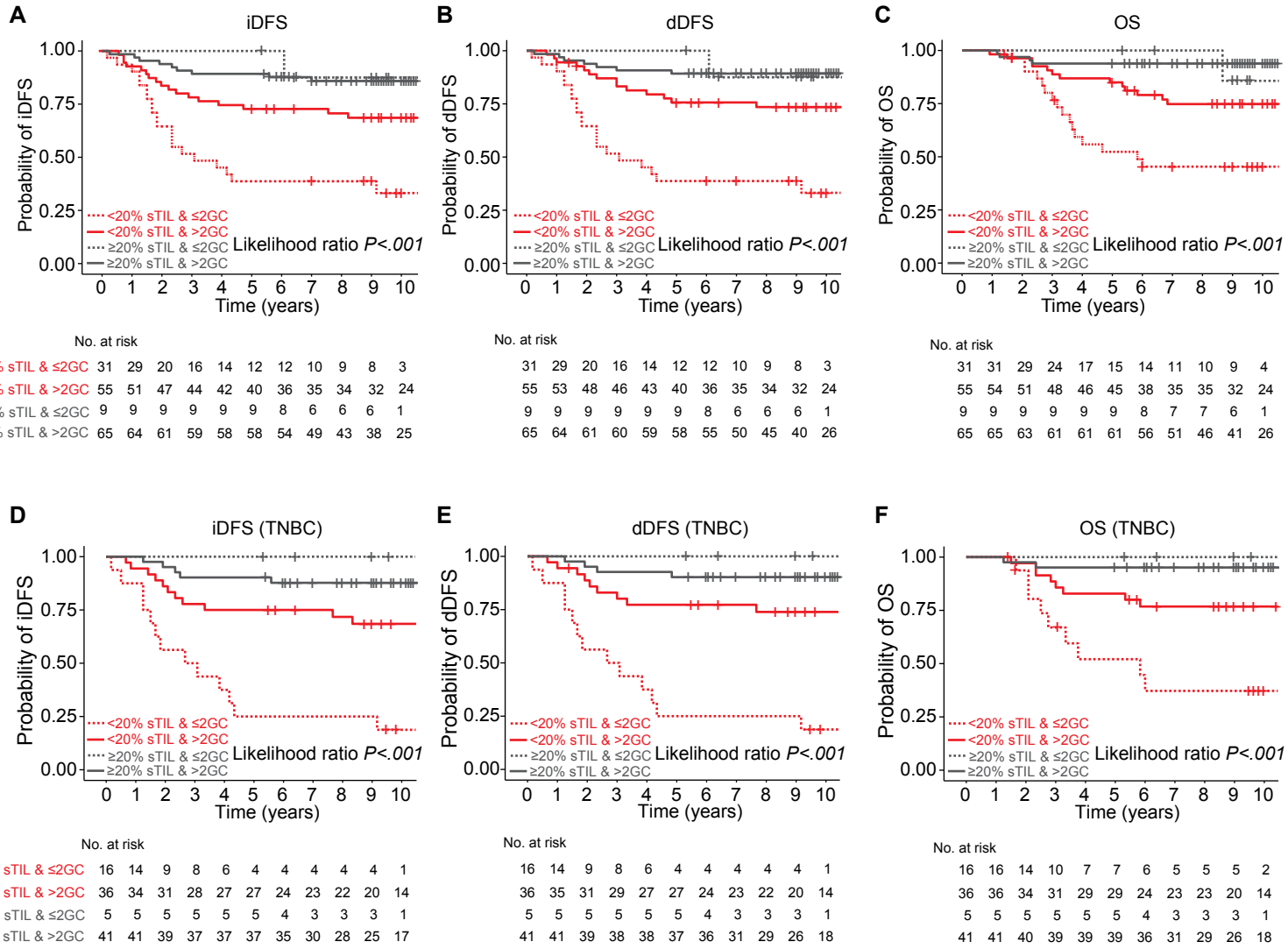


Liu et al., Figure 2





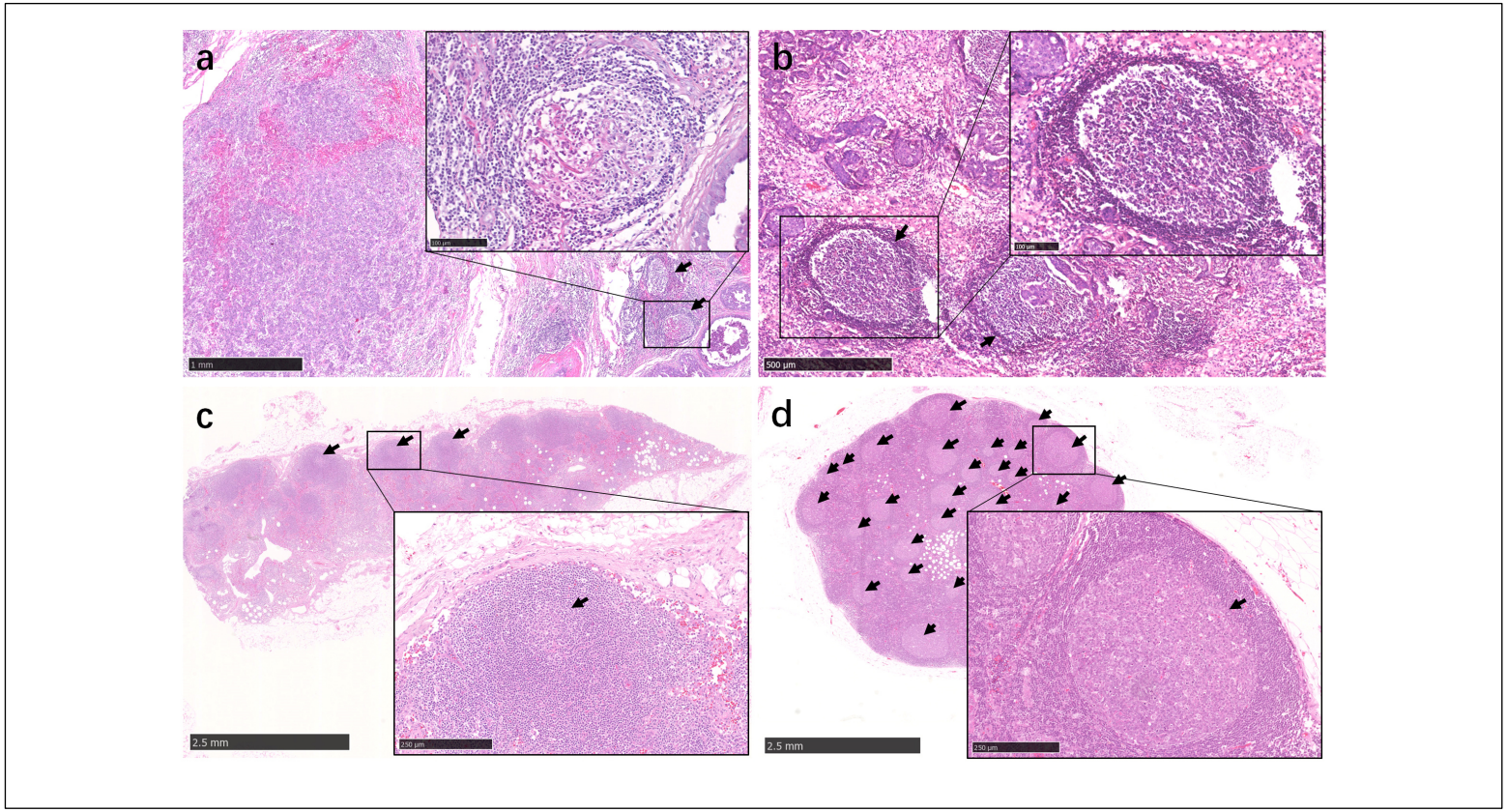
Liu et al., Figure 3



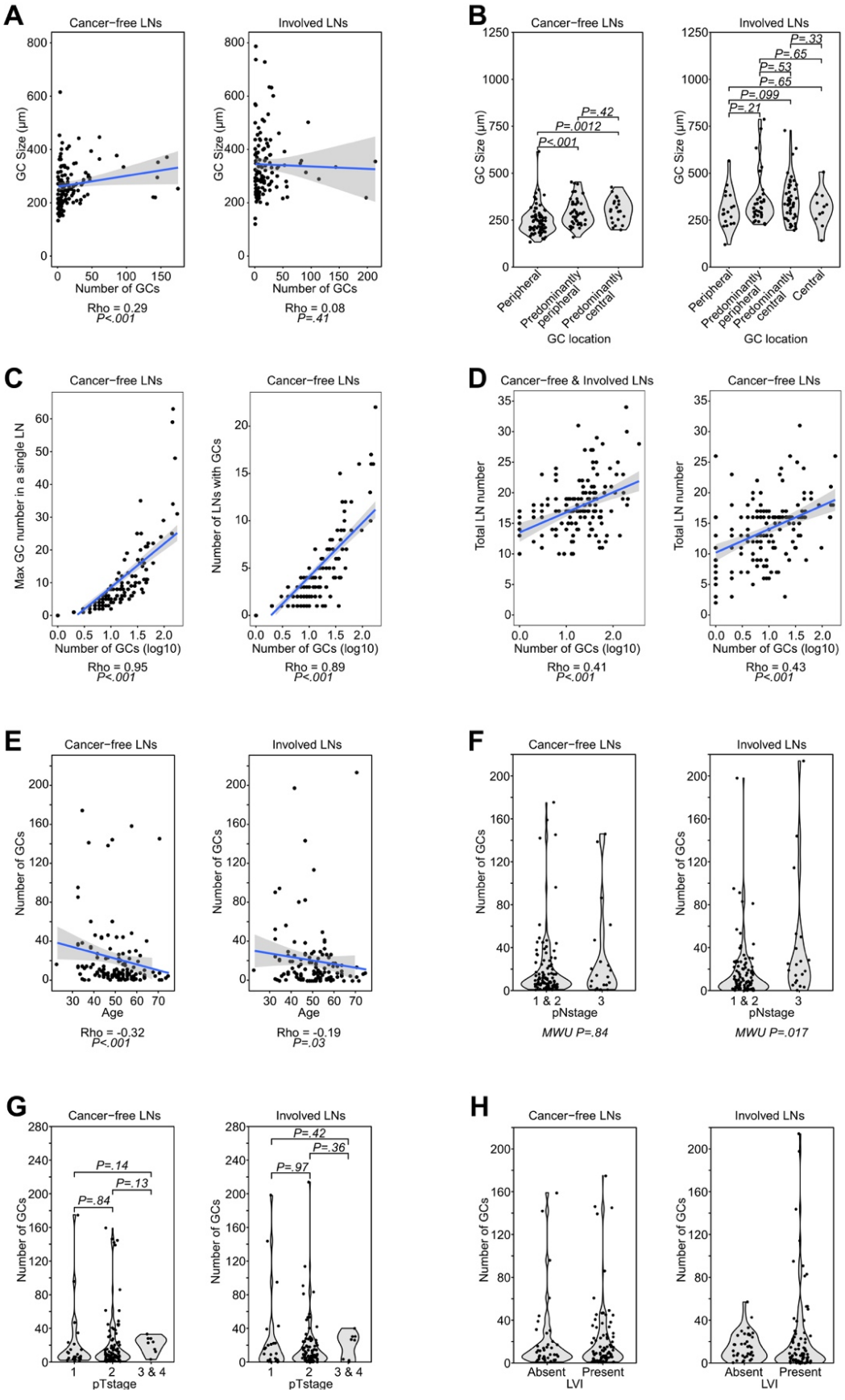
Liu et al., Figure 4

Data Supplement

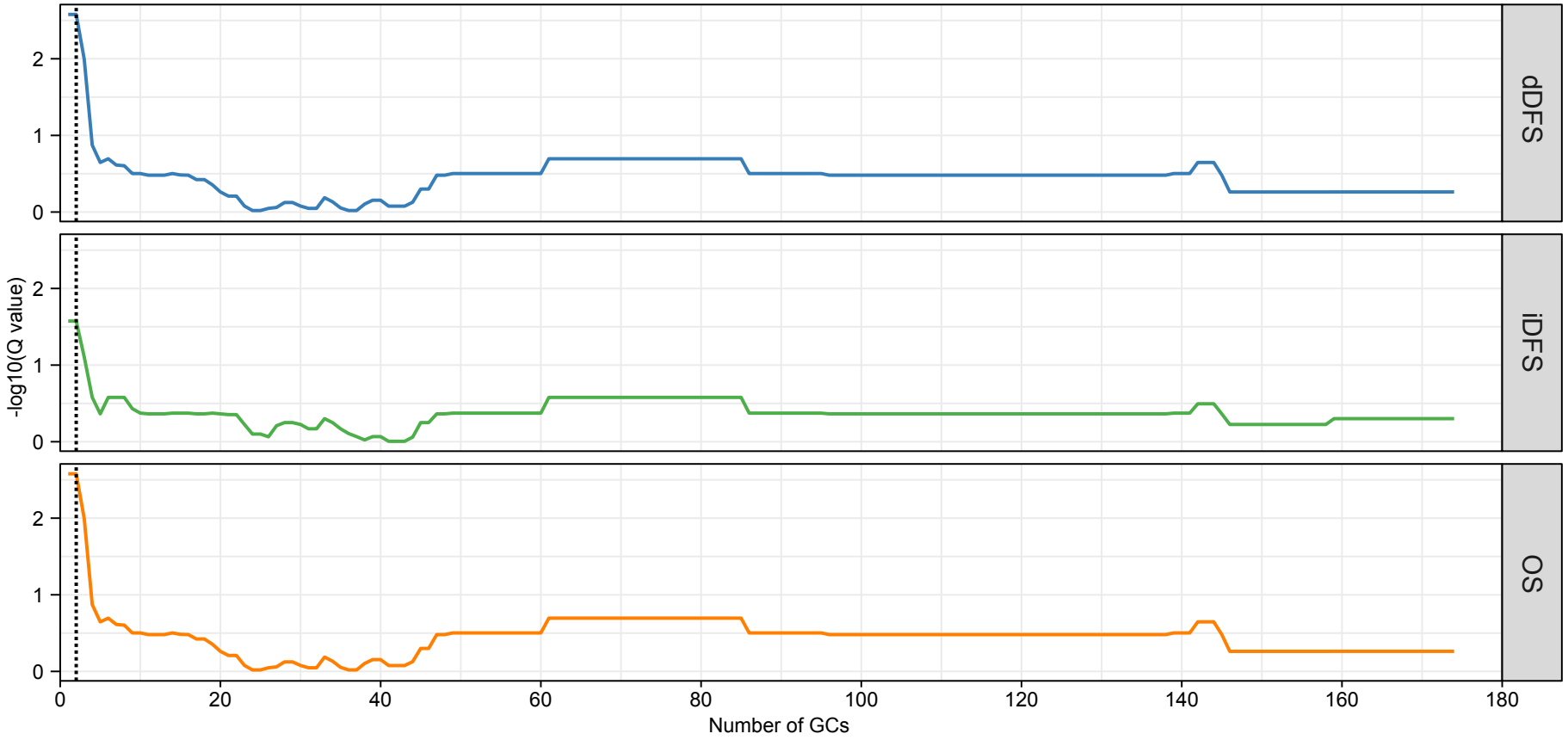
Supplementary Figure 1. Examples of tertiary lymphoid structures (TLS) in primary tumor and germinal centers (GCs) in cancer-free lymph nodes. (A, B) presence of TLS (with lymphoid GC formation, arrows) in the peritumoral stroma (inset showing a higher power view of TLS); **(C)** few GCs (arrow) located close to the LN capsule (predominantly in periphery) (inset showing a higher power view of GC); **(D)** numerous GC formation (arrow); larger, rounder and fuller GCs throughout the whole LN (predominantly in the central part of the LN) (inset showing a higher power view of GC).



Supplementary Figure 2. Association of the number of germinal centers in cancer-free and involved lymph nodes with clinicopathological features. (**A**) Scatter plots of the number of GCs compared to the maximum GC size. Cancer-free and involved LNs analysis is shown on the left and right, respectively. (**B**) Violin plot of maximum GC size in GC with regards to their predominant location in the LN. P values displayed calculated via Mann-Whitney U tests. Cancer-free and involved LNs analysis is shown on the left and right, respectively. (**C**) Scatter plot of the log10 number of GCs compared to the maximum number of GCs in a single LN (left side); and compared to the number of LNs which contain GCs (right side). (**D**) Scatter plots of the log10 number of GCs in all assessed LNs (left) or cancer-free LNs (right) compared to the number of LNs (including those with zero GCs). (**E**) Scatter plots of the number of GCs in cancer-free LNs (left) or involved LNs (right) compared to age at diagnosis; (**F**) Violin plot of the number of GCs in cancer-free LNs (left) or involved LNs (right) involved LNs (Y axis) compared to LN stage; (**G**) Violin plot, showing the distribution of GC formations in cancer-free LNs (left) or involved LNs (right) compared to the size of the primary tumor. P values displayed calculated via Mann-Whitney U test; (**H**) Violin plots, displaying the number of GCs in cancer-free LNs (left) or involved LNs (right) compared to patients with and without lympho-vascular invasion.


A
Cancer−free LNs
Involved LNs
GC Size (μm)
Number of GCs
Rho = 0.29
P<.001
Rho = 0.08
P=.41
B
Cancer−free LNs
Involved LNs
P=.42
P=.0012
P<.001
P=.33
P=.65
P=.53
P=.65
P=.099
P=.21
Peripheral
Predominantly peripheral
Predominantly central
Central
GC location
C
Cancer−free LNs
Max GC number in a single LN
Number of LNs with GCs
Number of GCs (log10)
Rho = 0.95
P<.001
Rho = 0.89
P<.001
D
Cancer−free & Involved LNs
Cancer−free LNs
Total LN number
Rho = 0.41
P<.001
Rho = 0.43
P<.001
E
Cancer−free LNs
Involved LNs
Age
Rho = -0.32
P<.001
Rho = -0.19
P=.03
F
Cancer−free LNs
Involved LNs
pNstage
1 & 2
3
MWU P=.84
MWU P=.017
G
Cancer−free LNs
Involved LNs
P=.14
P=.13
P=.84
P=.42
P=.36
P=.97
pTstage
1
2
3 & 4
H
Cancer−free LNs
Involved LNs
Absent
Present
LVI

Supplementary Figure 3. Cut off selection for defining the number of germinal centers in cancer-free lymph nodes. The optimal germinal center number cut off was determined by taking all patients (including those with zero GCs) and applying an iterative process using a minimal p value approach. For all three endpoints a single cut-off was identified (dashed black lines) at 2 germinal centers.



Supplementary Table 1. The distribution of total GCs in LNs excised by SLNB and ALNC. Patients are listed who had a SLNB followed by an ALNC. For each patient, the numbers of LNs and numbers of GCs in cancer-free and involved LNs are listed separately. Within cancer-free and involved LNs, the number of LNs excised by SLNB and by ALNC (called non-SLN) are shown. The number of GCs in SLN and non-SLN, as well as the percentage of GC formation given the total number of GCs observed in all cancer-free or involved LNs (referred to as % of total) are reported. The median number of LNs, GCs and percentage of GC formation is shown.

		Cancer-free LNs								Involved LNs							
		Total LNs		SLN			non-SLN			Total LNs		SLN			non-SLN		
					GCs			GCs					GCs			GCs	
Patient ID		Number of LNs	Number of GCs	Number of LNs	Number	% of total	Number of LNs	Number	% of total	Number of LNs	Number of GCs	Number of LNs	Number	% of total	Number of LNs	Number	% of total
case 1	83787	16	142	2	101	71%	14	41	29%	2	57	2	57	N/A	0	N/A	
case 2	84658	26	24	2	12	50%	24	12	50%	1	12	1	12	N/A	0	N/A	
case 3	93074	16	7	2	6	86%	14	1	14%	1	2	1	2	N/A	0	N/A	
case 4	90419	18	159	5	124	78%	13	35	22%	1	2	1	2	N/A	0	N/A	
case 5	90507	16	6	1	6	100%	15	0	0%	1	0	1	0	N/A	0	N/A	
case 6	91212	18	6	2	5	83%	16	1	17%	1	7	1	7	N/A	0	N/A	
case 7	92028	11	12	2	8	67%	9	4	33%	2	30	2	30	N/A	0	N/A	
case 8	94096	13	7	3	4	57%	10	3	43%	2	8	2	8	N/A	0	N/A	
case 9	96285	9	5	2	4	80%	7	1	20%	3	7	3	7	N/A	0	N/A	
case 10	100631	9	6	1	5	83%	8	1	17%	1	0	1	0	N/A	0	N/A	
case 11	101812	20	61	5	47	77%	15	14	23%	1	30	1	30	N/A	0	N/A	
case 12	103272	20	45	3	29	64%	17	16	36%	6	81	3	61	75%	3	20	25%
case 13	105096	18	34	3	24	71%	15	10	29%	2	7	2	7	N/A	0	N/A	
case 14	108069	13	38	3	23	61%	10	15	39%	3	53	1	42	79%	2	11	21%
case 15	108630	20	4	2	2	50%	18	2	50%	2	17	1	14	82%	1	3	18%
case 16	108631	17	7	1	6	86%	16	1	14%	1	9	1	9	N/A	0	N/A	
case 17	103813	24	16	3	11	69%	21	5	31%	1	27	1	27	N/A	0	N/A	
case 18	107221	11	6	3	4	67%	8	2	33%	1	6	1	6	N/A	0	N/A	
case 19	81543	16	2	1	1	50%	15	1	50%	2	0	1	0	N/A	1	0	N/A
case 20	93805	11	0	1	0	N/A	10	0	N/A	2	0	1	0	N/A	1	0	N/A
case 21	108731	16	0	2	0	N/A	14	0	N/A	1	0	1	0	N/A	0	N/A	
case 22	101463	9	0	1	0	N/A	8	0	N/A	1	0	1	0	N/A	0	N/A	
case 23	106265	9	0	1	0	N/A	8	0	N/A	1	0	1	0	N/A	0	N/A	
Median		16	7	2	6	71%	14	2	29%	1	7	1	7	79%	0	3	21%

Supplementary Table 2. Univariate & Multivariate Cox regression for iDFS, dDFS and OS of all cases or sTILs subgroups (< 20%, ≥ 20%)

Supplementary Table 2A. Univariate and Multivariate Cox regression analyses for iDFS

iDFS												
Univariate	**All cases**						**<20% sTILs**			**≥20% sTILs**		
Clinicopathological features	**Model P**		**HR**	**CI**			**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**
Age at diagnosis	0.123		1.02	0.99 - 1.05			0.110	1.03	0.99-1.06	0.983	1.00	0.94 - 1.07
Tumor size												
pTstage 1	0.056		Reference				0.078	Reference		Group size too small / too few events		
pTstage 2			1.53	0.68 - 3.45				1.44	1.48 - 5.57			
pTstage 3 & 4			3.76	1.32 - 10.74				3.71	0.26-14.56			
Lymph node stage												
pNstage 1 & 2	0.005		Reference				0.154	Reference		0.008	Reference	
pNstage 3			2.52	1.38 - 4.60				1.71	0.84 - 3.46		5.84	1.69-20.21
Lymphovascular invasion	0.194		1.54	0.78 - 3.03			0.459	1.33	0.61 - 2.92	0.407	1.33	0.34 - 5.15
HER2 status	1.000		1.00	0.56 - 1.80			0.538	0.81	0.41 - 1.59	0.407	1.69	0.49 - 5.85
sTILs												
20% cut off	<0.001		0.26	0.13 - 0.51			\			\		
Increments	<0.001		0.96	0.93 - 0.98								
TLS	0.002		0.25	0.09 - 0.71			0.010	0.23	0.05 - 0.94	0.355	0.50	0.11-2.37
LN characteristics	**Cancer-free LN**				**Involved LN**			**Cancer-free LN**			**Cancer-free LN**	
Total number of GCs across all assessed LNs per patient	**Model P**	**HR**	**CI**		**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**
Total GCs number	0.769	1.00	0.99-1.01		0.459	1.00	0.98-1.01	0.427	0.99	0.97 - 1.01	0.019	1.01	1.00 - 1.02
Total GCs number (≤2 / >2)	<0.001	0.33	0.19-0.59		0.023	0.48	0.26-0.88	0.002	0.36	0.19 - 0.69	0.804	1.29	0.16-10.16
Multivariate	**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**							**Corrected for: Age, pTstage, pNstage, LVI & TLS**					
LN characteristics	**Cancer-free LN**				**Involved LN**			**Cancer-free LN**			**Cancer-free LN**		
Across all assessed LNs per patient	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**
Total GCs number	0.010	<0.001	1.01	1.00 - 1.02	0.362	<0.001	1.00	0.99 - 1.01	0.832	0.025	1.00	0.98-1.02	0.086	0.013	1.01	1.00-1.03
Total GCs number (≤2 / >2)	0.110	<0.001	0.58	0.30 - 1.12	0.161	<0.001	0.63	0.33 - 1.20	0.023	0.004	0.41	0.19 -0.89	0.949	0.034	0.93	0.11-7.93

Supplementary Table 2B. Univariate and Multivariate Cox regression analyses for dDFS

dDFS														
Univariate	**All cases**						**<20% sTILs**			**≥20% sTILs**				
Clinicopathological features	**Model P**		**HR**	**CI**			**Model P**	**HR**	**CI**	**Model P**		**HR**	**CI**	
Age at diagnosis	0.055		1.03	1.00 - 1.06			0.041	0.103	1.00-	0.949		1.00	0.93 - 1.07	
Tumor size														
pTstage 1	0.017		Reference				0.034	Reference		Group size too small / too few events				
pTstage 2			1.95	0.76 - 5.03				1.76	0.53 - 5.84					
pTstage 3 & 4			5.43	1.72 - 17.13				5.12	1.32 -19.88					
Lymph node stage														
pNstage 1 & 2	0.003		Reference				0.074	Reference		0.018		Reference		
pNstage 3			2.75	1.46 - 5.18				1.99	0.97 - 4.08			5.88	1.47-23.55	
Lymphovascular invasion	0.219		1.54	0.76 - 3.13			0.324	1.50	0.65 - 3.44	0.953		0.95	0.23 - 4.01	
HER2 status	0.953		0.98	0.53 - 1.83			0.512	0.79	0.39 - 1.60	0.437		0.44	0.43 - 6.96	
sTILs							\			\				
20% cut off	<0.001		0.22	0.10 - 0.48										
Increments	<0.001		0.96	0.93 - 0.98										
TLS	0.001		0.21	0.06 - 0.67			0.013	0.23	0.06 - 0.98	0.183		0.18	0.04 - 2.38	
LN characteristic	**Cancer-free LN**				**Involved LN**			**Cancer-free LN**			**Cancer-free LN**			
Across all assessed LNs per patient	**Model P**		**HR**	**CI**	**Model P**		**HR**	**CI**	**Model P**		**HR**	**CI**	**Model P**		**HR**	**CI**
Total GCs number	0.618		1.00	0.99 - 1.01	0.253		0.98	0.98 - 1.01	0.483		0.99	0.97 - 1.01	0.009		1.02	1.01 - 1.03
Total GCs number (≤2 / >2)	<0.001		0.26	0.14 - 0.48	0.038		0.25	0.25 - 0.93	<0.001		0.28	0.14 - 0.56	0.986		1.02	0.13 - 8.29
Multivariate	**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**								**Corrected for: Age, pTstage, pNstage, LVI,& TLS**							
LN characteristics	**Cancer-free LN**				**Involved LN**				**Cancer-free LN**				**Cancer-free LN**			
Across all assessed LNs per patient	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**
Total GCs number	0.002	<0.001	1.01	1.00 - 1.02	0.565	<0.001	1.00	0.99 - 1.01	0.665	0.005	1.00	0.98 - 1.03	0.044	0.007	1.02	1.00 - 1.03
Total GCs number (≤2 / >2)	0.033	<0.001	0.47	0.23 - 0.94	0.21	<0.001	0.64	0.32 - 1.28	0.009	<0.001	0.34	0.17 - 0.77	0.665	0.031	0.61	0.07 – 5.64

Supplementary Table 2C. Univariate and Multivariate Cox regression analyses for OS

OS									
Univariate	**All cases**			**<20% sTILs**			**≥20% sTILs**		
Clinicopathological features	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**
Age at diagnosis	0.019	1.04	1.01 - 1.08	0.018	1.04	1.01-1.08	0.711	1.02	0.93 - 1.12
Tumor size									
pTstage 1	0.004	Reference		0.053	Reference		Group size too small / too few events		
pTstage 2		4.23	1.00 - 17.84		2.30	0.54 - 9.82			
pTstage 3 & 4		10.80	2.18 - 53.54		5.96	1.20 -29.69			
Lymph node stage									
pNstage 1 & 2	0.001	Reference		0.087	Reference		0.001	Reference	
pNstage 3		3.34	1.69 - 6.63		2.01	0.94 - 4.34		22.18	2.47-198.7
Lymphovascular invasion	0.054	2.22	0.92 - 5.37	0.429	1.42	0.58 - 3.49	Group size too small / too few events		
HER2 status	0.564	1.22	0.62 - 2.40	0.980	0.99	0.47 - 2.08	0.310	2.50	0.42 -14.96
sTILs				\			\		
20% cut off	<0.001	0.17	0.06 - 0.43						
Increments	<0.001	0.94	0.91 - 0.98						
TLS	<0.001	0.08	0.01 - 0.59	0.005	0.13	0.02 - 0.97	Group size too small / too few events		

LN characteristics	**Cancer-free LN**			**Involved LN**			**Cancer-free LN**			**Cancer-free LN**		
Total number of GCs across	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**
Total GCs number	0.951	1.00	0.99-1.01	0.488	1.00	0.98 - 1.01	0.869	1.00	0.98 - 1.02	0.152	1.01	1.00 - 1.03
Total GCs number (≤2 / >2)	<0.001	0.28	0.14 - 0.55	0.033	0.44	0.22 - 0.91	0.006	0.36	0.17 - 0.75	0.652	0.59	0.07 - 5.25

Multivariate	**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**								**Corrected for: Age, pTstage, pNstage, LVI & TLS**							
LN characteristics	**Cancer-free LN**				**Involved LN**				**Cancer-free LN**				**Cancer-free LN**			
Across all assessed LNs per patient	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**
Total GCs number	0.005	<0.001	1.02	1.00 - 1.03	0.092	<0.001	1.01	1.00 - 1.02	0.163	0.002	1.01	1.00 - 1.03	0.372	<0.001	0.96	0.96 - 1.02
Total GCs number (≤2 / >2)	0.351	<0.001	0.69	0.32 - 1.50	0.193	<0.001	0.60	0.28 - 1.29	0.106	0.001	0.48	0.20 – 1.17	Group size too small / too few events			

Supplementary Table 3. Univariate and Multivariate Cox regression analyses for iDFS, dDFS and OS of TNBC and HER2-positive subgroups

Supplementary Table 3A. Univariate and Multivariate Cox regression analyses for iDFS of TNBC and HER2-positive subgroups

	iDFS											
Univariate	**TNBC**						**HER2**					
Clinicopathological features	**Model P**		**HR**		**CI**		**Model P**		**HR**		**CI**	
Age at diagnosis	0.076		1.03		1 - 1.07		0.921		1.00		0.95 - 1.06	
Tumor size												
pTstage 1	0.260		Reference				0.206		Reference			
pTstage 2			1.47		0.55 - 3.92				1.66		0.38 - 7.38	
pTstage 3 & 4			3.16		0.85 - 11.8				4.97		0.83 - 29.83	
Lymph node stage												
pNstage 1 & 2	0.043		Reference				0.049		Reference			
pNstage 3			2.48		1.1 - 5.6				2.64		1.04 - 6.7	
Lymphovascular invasion	0.218		1.64		0.73 - 3.71		0.616		1.36		0.39 - 4.7	
sTILs												
20% cut off	<0.001		0.19		0.07 - 0.49		0.064		0.40		0.14 - 1.11	
Increments	0.002		0.96		0.93 - 0.99		0.010		0.96		0.92 - 1	
TLS	0.010		0.22		0.05 - 0.92		0.065		0.30		0.07 - 1.32	
LN characteristics	**Cancer-free LN**			**Involved LN**			**Cancer-free LN**			**Involved LN**		
	Model P	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**
Total GCs number	0.161	1.01	1 - 1.02	0.740	1.00	0.99 - 1.01	0.023	0.95	0.9 - 1.01	0.079	0.98	0.94 - 1.01
Total GCs number (≤2 / >2)	<0.001	0.25	0.12 - 0.52	0.004	0.32	0.15 - 0.67	0.139	0.49	0.19 - 1.24	0.849	1.13	0.32 - 3.96

Multivariate	**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**								**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**							
LN characteristics	**Cancer-free LN**				**Involved LN**				**Cancer-free LN**				**Involved LN**			
Across all assessed LNs per patient	**Covariate P**	**Model**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**
Total GCs number	<0.001	<0.001	1.01	1.01 - 1.03	0.110	<0.001	1.00	1.00 - 1.02	0.127	0.046	0.94	0.87 – 1.02	0.260	0,169	0.93	0.93 – 1.02
Total GCs number (≤2 / >2)	0.017	<0.001	0.37	0.16 – 0.84	0.092	<0.001	0.50	0.22 - 1.12	0.959	0.144	0.97	0.26 – 3.59	0.713	0.261	1.27	0.35 – 4.63

Supplementary Table 3B. Univariate and Multivariate Cox regression analyses for dDFS of TNBC and HER2-positive subgroups

	dDFS											
Univariate	**TNBC**						**HER2**					
Clinicopathological features	**Model P**		**HR**		**CI**		**Model P**		**HR**		**CI**	
Age at diagnosis	0.035		1.04		1 - 1.08		0.780		1.01		0.95 - 1.07	
Tumor size												
pTstage 1	0.082		Reference				0.168		Reference			
pTstage 2			2.37		0.7 - 8.02				1.37		0.30 - 6.20	
pTstage 3 & 4			5.41		1.21 - 24.21				5.11		0.85 - 30.71	
Lymph node stage												
pNstage 1 & 2	0.021		Reference				0.060		Reference			
pNstage 3			2.90		1.26 - 6.69				2.68		0.99 - 7.21	
Lymphovascular invasion	0.404		1.42		0.62 - 3.26		0.310		2.02		0.46 - 8.90	
sTILs												
20% cut off	<0.001		0.16		0.06 - 0.47		0.051		0.35		0.11 - 1.09	
increments	0.002		0.96		0.93 - 0.99		0.012		0.95		0.91 - 1.00	
TLS	0.022		0.25		0.06 - 1.05		0.018		0.16		0.02 - 1.19	
LN characteristics	**Cancer-free LN**				**Involved LN**			**Cancer-free LN**			**Involved LN**	
	Model P	**HR**	**CI**		**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**
Total GCs number	0.120	1.01	1.00 - 1.02		0.962	1.00	0.99 - 1.01	0.015	0.94	0.88 - 1.01	0.024	0.96	0.91 - 1.01
Total GCs number (≤2 / >2)	<0.001	0.20	0.09 - 0.44		0.015	0.36	0.17 - 0.79	0.057	0.38	0.14 - 1.01	0.910	0.93	0.26 - 3.33
Multivariate	**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**							**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**					
LN characteristics	**Cancer-free LN**				**Involved LN**			**Cancer-free LN**			**Involved LN**		
Across all assessed LNs per patient	**Covariate P**	**Model**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**
Total GCs number	<0.001	<0.001	1.02	1.01 - 1.03	0.429	<0.001	1.01	0.99 - 1.02	0.200	0.011	0.96	0.90 - 1.02	0.089	0.007	0.95	0.90 – 1.01
Total GCs number (≤2 / >2)	0.004	<0.001	0.29	0.13 – 0.67	0.034	<0.001	0.43	0.20 – 0.94	0.375	0.037	0.60	0.2 - 1.84	0.970	0.069	0.98	0.27 – 3.57

Supplementary Table 3C. Univariate and Multivariate Cox regression analyses for OS of TNBC and HER2-positive subgroups

	OS											
Univariate	**TNBC**						**HER2**					
Clinicopathological features	**Model P**		**HR**		**CI**		**Model P**		**HR**		**CI**	
Age at diagnosis	0.012		1.06		1.01- 1.11		0.714		1.01		0.95 - 1.07	
Tumor size												
pTstage 1	0.040		Reference				0.095		Reference			
pTstage 2			5.50		0.73 - 41.63				2.91		1.01 - 22.56	
pTstage 3 & 4			11.39		1.18 - 109.58				9.74		0.38 - 93.94	
Lymph node stage												
pNstage 1 & 2	0.009		Reference				0.052		Reference			
pNstage 3			3.89		1.53 - 9.89				2.84		1.03 - 7.85	
Lymphovascular invasion	0.267		1.74		0.63 - 4.84		0.086		4.22		0.55 - 32.1	
sTILs												
20% cut off	<0.001		0.11		0.03 - 0.47		0.019		0.26		0.07 - 0.91	
increments	0.002		0.95		0.9 - 0.99		0.004		0.94		0.89 - 0.99	
TLS	Group size too small / too few events						0.026		0.17		0.02 - 1.28	
LN characteristics	**Cancer-free LN**				**Involved LN**			**Cancer-free LN**			**Involved LN**	
Total number of GCs across all assessed LNs per patient	**Model P**	**HR**	**CI**		**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**	**Model P**	**HR**	**CI**
Total GCs number	0.301	1.01	1.00 - 1.02		0.565	1.00	0.99 - 1.02	0.019	0.94	0.88 - 1.01	0.027	0.96	0.91 - 1.01
Total GCs number (≤2 / >2)	0.004	0.24	0.10 - 0.60		0.013	0.31	0.12 - 0.75	0.036	0.33	0.12 - 0.92	0.848	0.88	0.24 - 3.20
Multivariate	**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**							**Corrected for: Age, pTstage, pNstage, LVI, sTILs & TLS**					
LN characteristics	**Cancer-free LN**				**Involved LN**			**Cancer-free LN**			**Involved LN**		
Across all assessed LNs per patient	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**	**Covariate P**	**Model P**	**HR**	**CI**
Total GCs number	0.002	<0.001	1.02	1.01 - 1.03	0.080	<0.001	1.01	1.00 - 1.03	0.401	0.015	0.96	0.88 - 1.05	0.236	0.058	0.96	0.91 - 1.02
Total GCs number (≤2 / >2)	0.119	<0.001	0.46	0.17 - 1.22	0.077	<0.001	0.40	0.15 – 1.10	0.851	0.021	1.14	0.28 - 4.72	0.751	0.102	1.24	0.33 - 4.71

Supplementary Table 4. Multivariate Cox regression analyses for iDFS, dDFS and OS comparing the prognostic value of the proportion of LNs with GCs and the ratio of GCs in LNs

Multivariate analysis									
corrected for Age, pTstage, pNstage, LVI & TLS									
	<20% sTILs					≥20% sTILs			
Covariate	Covariate P value	Model P values	HR	CI		Covariate P value	Model P values	HR	CI
iDFS									
Proportion of LNs with GC (Involved and cancer-free LNs)	0.236	0.015	0.20	0.01 - 2.88		0.058	0.010	33.34	0.89 - 1255.32
Proportion of LNs with GC (cancer-free LNs)	0.607	0.023	0.54	0.05 - 5.58		0.037	0.007	14.92	1.18 - 187.97
Proportion of LNs with GC (Involved LNs)	0.136	0.021	0.47	0.17 - 1.27		0.490	0.039	0.44	0.04 - 4.55
Total GCs / Total LNs (Involved and cancer-free)	0.938	0.026	0.99	0.73 - 1.33		0.069	0.012	1.20	0.99 - 1.46
Total GCs / Total LNs (cancer-free)	0.823	0.025	0.96	0.64 - 1.42		0.009	0.004	1.37	1.08 - 1.73
Total GCs / Total LNs (Involved)	0.250	0.026	0.95	0.87 - 1.04		0.603	0.041	0.98	0.93 - 1.04
dDFS									
Proportion of LNs with GC (Involved and cancer-free LNs)	0.129	0.002	0.10	0.01 - 1.94		0.040	0.007	66.23	1.22 - 3602.85
Proportion of LNs with GC (cancer-free LNs)	0.542	0.004	0.46	0.04 - 5.49		0.024	0.004	32.03	1.58 - 650.92
Proportion of LNs with GC (Involved LNs)	0.095	0.004	0.40	0.13 - 1.17		0.844	0.040	0.77	0.05 - 10.95
Total GCs / Total LNs (Involved and cancer-free)	0.585	0.004	0.90	0.62 - 1.31		0.068	0.012	1.20	0.99 - 1.46
Total GCs / Total LNs (cancer-free)	0.890	0.005	0.97	0.64 - 1.47		0.025	0.006	1.33	1.04 - 1.70
Total GCs / Total LNs (Involved)	0.034	0.001	0.87	0.76 - 0.99		0.143	0.011	0.85	0.69 - 1.06
OS									
Proportion of LNs with GC (Involved and cancer-free LNs)	0.765	0.003	0.63	0.03 - 13.42		0.443	<0.001	0.05	0 - 116.27
Proportion of LNs with GC (cancer-free LNs)	0.470	0.003	2.51	0.21 - 30.25		0.521	<0.001	0.19	0 - 30.14
Proportion of LNs with GC (Involved LNs)	0.154	0.005	0.42	0.13 - 1.39		0.541	<0.001	0.25	0 - 20.97
Total GCs / Total LNs (Involved and cancer-free)	0.602	0.003	1.10	0.78 - 1.55		0.431	<0.001	0.80	0.46 - 1.39
Total GCs / Total LNs (cancer-free)	0.305	0.002	1.20	0.85 - 1.70		0.424	<0.001	0.80	0.46 - 1.39
Total GCs / Total LNs (Involved)	0.066	0.002	0.88	0.76 - 1.01		0.677	<0.001	0.93	0.65 - 1.32

Multivariate analysis									
corrected for Age, pTstage, pNstage, LVI, TLS & sTILs									
	TNBC					HER2			
Covariate	Covariate P value	Model P values	HR	CI		Covariate P value	Model P values	HR	CI
iDFS									
Proportion of LNs with GC (Involved and cancer-free LNs)	0.178	0.001	4.41	0.51 - 38.29		0.547	0.129	0.37	0.01 - 9.5
Proportion of LNs with GC (cancer-free LNs)	0.055	<0.001	5.77	0.96 - 34.56		0.734	0.139	0.57	0.02 - 14.08
Proportion of LNs with GC (Involved LNs)	0.212	0.001	0.50	0.17 - 1.48		0.826	0.268	1.23	0.2 - 7.55
Total GCs / Total LNs (Involved and cancer-free)	0.002	<0.001	1.25	1.08 - 1.45		0.277	0.092	0.76	0.46 - 1.25
Total GCs / Total LNs (cancer-free)	<0.001	<0.001	1.41	1.17 - 1.7		0.699	0.137	0.91	0.57 - 1.46
Total GCs / Total LNs (Involved)	0.692	0.001	0.98	0.91 - 1.07		0.719	0.261	0.99	0.94 - 1.04
dDFS									
Proportion of LNs with GC (Involved and cancer-free LNs)	0.124	<0.001	5.99	0.61 - 58.74		0.154	0.029	0.07	0 - 2.74
Proportion of LNs with GC (cancer-free LNs)	0.030	<0.001	8.08	1.23 - 53.1		0.588	0.054	0.38	0.01 - 12.86
Proportion of LNs with GC (Involved LNs)	0.356	<0.001	0.58	0.19 - 1.83		0.549	0.156	0.55	0.08 - 3.88
Total GCs / Total LNs (Involved and cancer-free)	0.003	<0.001	1.27	1.08 - 1.49		0.073	0.015	0.52	0.25 - 1.06
Total GCs / Total LNs (cancer-free)	<0.001	<0.001	1.46	1.19 - 1.78		0.124	0.021	0.48	0.19 - 1.23
Total GCs / Total LNs (Involved)	0.186	<0.001	0.91	0.8 - 1.04		0.228	0.078	0.92	0.8 - 1.05
OS									
Proportion of LNs with GC (Involved and cancer-free LNs)	0.109	6.80E-06	9.83	0.6 - 161.44		0.530	0.02	0.30	0.01 - 13.32
Proportion of LNs with GC (cancer-free LNs)	0.023	2.51E-06	14.40	1.45 - 142.92		0.747	0.02	1.87	0.04 - 82.66
Proportion of LNs with GC (Involved LNs)	0.286	1.35E-05	0.49	0.13 - 1.82		0.604	0.10	0.58	0.08 - 4.47
Total GCs / Total LNs (Involved and cancer-free)	0.020	2.29E-06	1.29	1.04 - 1.59		0.298	0.01	0.71	0.37 - 1.36
Total GCs / Total LNs (cancer-free)	0.004	1.18E-06	1.44	1.12 - 1.85		0.454	0.02	0.73	0.33 - 1.65
Total GCs / Total LNs (Involved)	0.490	1.81E-05	0.95	0.81 - 1.1		0.302	0.06	0.93	0.82 - 1.06