

Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer

Sidra Nawaz^{1,2,3}, Andreas Heindl^{1,2,3}, Konrad Koelble^{4,5} and Yinyin Yuan^{1,2,3}

¹Division of Molecular Pathology, The Institute of Cancer Research, London, UK; ²Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK; ³Centre for Molecular Pathology, Royal Marsden Hospital, London, UK; ⁴Functional Genomics Laboratory, Institute of Cancer Research, London, UK and ⁵Department of Histopathology, The Royal Marsden Hospital, London, UK

The abundance of tumor-infiltrating lymphocytes has been associated with a favorable prognosis in estrogen receptor-negative breast cancer. However, a high degree of spatial heterogeneity in lymphocytic infiltration is often observed and its clinical implication remains unclear. Here we combine automated histological image processing with methods of spatial statistics used in ecological data analysis to quantify spatial heterogeneity in the distribution patterns of tumor-infiltrating lymphocytes. Hematoxylin and eosin-stained sections from two cohorts of estrogen receptor-negative breast cancer patients (discovery: $n=120$; validation: $n=125$) were processed with our automated cell classification algorithm to identify the location of lymphocytes and cancer cells. Subsequently, hotspot analysis (Getis–Ord G_i^*) was applied to identify statistically significant hotspots of cancer and immune cells, defined as tumor regions with a significantly high number of cancer cells or immune cells, respectively. We found that the amount of co-localized cancer and immune hotspots weighted by tumor area, rather than number of cancer or immune hotspots, correlates with a better prognosis in estrogen receptor-negative breast cancer in univariate and multivariate analysis. Moreover, co-localization of cancer and immune hotspots further stratified patients with immune cell-rich tumors. Our study demonstrates the importance of quantifying not only the abundance of lymphocytes but also their spatial variation in the tumor specimen for which methods from other disciplines such as spatial statistics can be successfully applied.

Modern Pathology (2015) 28, 766–777; doi:10.1038/modpathol.2015.37; published online 27 February 2015

Estrogen receptor (ER) status is one of the most important clinical parameters in breast cancer, with ER-negative patients generally having a poorer 10-year prognosis than ER-positive patients.^{1–4} However, the considerable variation in clinical outcome of ER-negative patients has fuelled the quest for additional markers to better stratify this disease subset and eventually provide more effective personalized treatment.⁵ Tumor infiltration by various cells of the immune system has been associated with a favorable clinical outcome in ER-negative breast cancer.⁶ Methods to assess tumoral immune infiltration include traditional pathological examination and gene expression profiling.^{7,8} Although the latter can indicate the immune cell type and their relative abundance, a more comprehensive analysis of their

number, lineage and spatial distribution requires pathological assessment.^{9–12}

Spatial information on the tumor immune micro-environment is of clinical relevance, as not only the abundance and type but also the spatial locations of immune cells have been shown to be associated with clinical outcome in colorectal¹³ and breast cancer.^{10,11,14} In the former, a high density of $CD3^+$ T-lymphocytes at the invasive margin was found to be significantly associated with disease-free survival in three independent cohorts of patients, whereas in breast tumors the total and distant stromal counts of $CD8^+$ cytotoxic T-cells were independently capable of predicting breast cancer-specific survival.¹⁴ More recently, in ER-negative/HER2-negative¹⁰ and HER2-negative¹¹ breast cancer patients, a high degree of stromal immune infiltration was found to be associated with increased disease-specific survival and pathological complete response rates, respectively.

Although the intra-specimen location of lymphocytes is often accounted for on an individual basis during evaluation of immune infiltration,^{9,10,15} there has been no systematic investigation in larger

Correspondence: Dr Y Yuan, BSc, MSc, PhD, Division of Molecular Pathology, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, London SM2 5NG, UK.

E-mail: yinyin.yuan@icr.ac.uk

Received 2 September 2014; revised 17 December 2014; accepted 18 December 2014; published online 27 February 2015

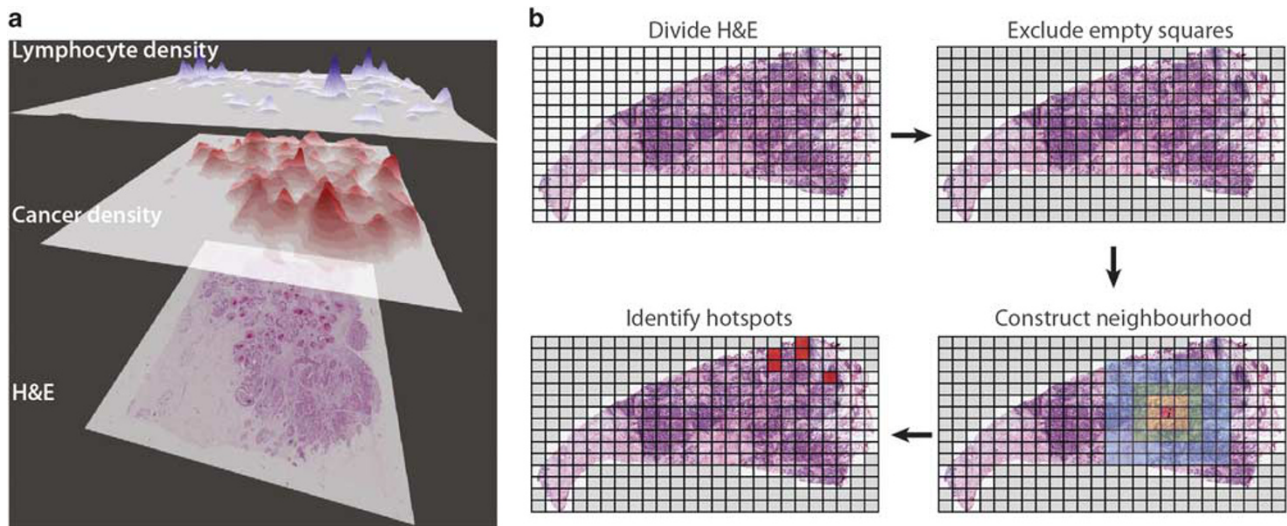


Figure 1 Spatial heterogeneity in lymphocytic infiltration and the use of Getis–Ord hotspot analysis. (a) 3D landscape of lymphocyte and cancer cell density within an H&E-stained tumor section showing spatial heterogeneity in the pattern of lymphocytic infiltration. (b) Illustration of the concept of Getis–Ord hotspot analysis, which detects statistically significant cancer and immune hotspots tested for every square tumor region. An example tumor section with a grid overlay to divide the image into squares (top-left, not drawn to scale), exclusion of empty (grey) squares (top-right), construction of neighborhood for one grid square (bottom-right, first order consists of squares in yellow, second order includes also the green squares and fourth order includes yellow, green and blue squares) and detection of hotspots (bottom-left) are displayed as an example.

cohorts of their spatial relation to one another and the degree of uniformity or clustering in the pattern of their distribution (Figure 1a). Given the occasionally marked spatial heterogeneity of immune cell distribution in breast cancer, we aimed to: (i) develop a robust methodology to elucidate the spatial heterogeneity of cancer and immune cells and their spatial relationship, and (ii) investigate whether this spatial heterogeneity is of clinical significance in ER-negative breast cancer patients.

To achieve these aims, we developed a computational approach combining high-throughput pathological image analysis with spatial statistics. Objective assessment of a large number of pathological samples requires the use of automated methods to identify cells in their thousands. The emerging field of computational pathology utilizes the advanced processing capabilities of modern-day computers to aid in the pathological assessment of tissue samples by reducing subjectivity and improving reproducibility. Fully automated cell, vessel and tumor region classification algorithms have been published that use features such as object size and morphology to classify them with high accuracy.^{15–23} We have previously carried out a fully automated, quantitative assessment of immune infiltration in ER-negative breast cancer to report that a greater abundance of lymphocytes ascertained by image analysis is associated with a better prognosis.¹²

To understand the complexity in the resulting large amount of spatial data following image analysis, we can employ methods from other disciplines such as ecology, where the spatial interactions among species are routinely studied. Geographically mapped regions populated by different species

correspond to the two-dimensional representation of tumor architecture in histological sections, infiltrated by different cell types. One method widely used on ecological and demographical data to analyze spatial interactions is the Getis–Ord hotspot analysis,²⁴ which pinpoints the location of regions where there is a statistically significant spatial clustering of high magnitudes of a variable, independent of the number of observations. For example, de Souza *et al*²⁵ used Getis–Ord hotspot analysis to identify zones of significantly high prevalence of malaria-carrying mosquitoes in Ghana. Setiadi *et al*²⁶ carried out spatial analysis using the L-function statistic to study clustering differences in T- and B-lymphocytes in healthy and tumor draining lymph nodes, but we are thus far unaware of the use of such methods in a large-scale investigation that relates to clinical outcome.

Therefore, this paper is based on two key concepts: (i) the use of high-throughput image analysis to generate spatial data denoting locations of cancer and immune cells within tumors, and (ii) the introduction of spatial statistics to detect and quantify cancer/immune spatial clusters or hotspots, ie, spatial regions where the population of cancer cells or lymphocytes is significantly high.

Materials and methods

Sample Set

To investigate the clinical implications of immune hotspots, we chose a sample of 245 ER-negative breast cancer patients from the METABRIC consortium.²⁷ METABRIC is a large-scale investigation of breast

cancer heterogeneity, for which 1026 pretreatment primary breast tumors with hematoxylin and eosin (H&E)-stained frozen tumor sections, molecular profiles and 10-year disease-specific survival data are available for further analysis. The tumor sections were scanned using ScanScope TX Scanner from Aperio Technologies with $\times 20$ magnification and digitized for image analysis as described in Yuan *et al.*¹² We have previously reported the association between lymphocyte abundance (*lym*, calculated as the percentage of immune cells in the total sample cell count) and a good prognosis in 115 ER-negative breast cancers in METABRIC.¹² For this study, we expanded our sample size to 245 ER-negative breast cancer patients for a larger investigation. The data were collected independently at three hospitals: A (120 patients), B (62 patients) and C (63 patients). We chose patients from A as our discovery cohort and patients from B and C collectively as our validation cohort, in order to have a similar population size in each set.

Pathological Image Analysis for Cell Classification

H&E-stained frozen tumor section images for 245 ER-negative breast cancer patients were analyzed using our automated cell classification pipeline CRImage.¹² This tool first segmented cell nuclei and then analyzed morphological features, such as size and circularity, of each cell nucleus detected in the image. These quantitative features were input into a support vector machine for supervised classification into cancer, lymphocyte and stromal cell nuclei. Lymphocytes have a typical morphology of small, round and homogeneously basophilic nuclei, thus can be reliably differentiated from other cell types in breast cancer. The accuracy of classification was evaluated by cross-validation in the initial training set of 871 cells (90.1% accuracy), in comparison to pathological tumor scores, including tumor cellularity and lymphocytic infiltration, and gene expression data.¹² In addition, visual scoring of 10 000 cells by a pathologist was compared with automated classification, and they were found to be highly correlated ($R^2=0.98$). As a result of image analysis, we identified, on average, 77 200 cancer cells (s.d. 75 400) and 14 000 lymphocytes (s.d. 22 300) as well as their spatial locations in each tumor in our sample set.

Getis–Ord Spatial Statistics for Hotspot Identification

The cell classification and location data were used as input for Getis–Ord hotspot analysis to enable the automated detection of statistically significant spatial clusters. Hotspot mapping is carried out on spatial data to identify locations where a variable of interest is found to be clustered. This means that, in relation to the entire area of study, the frequency or magnitude of this variable at these locations is greater than expected and, importantly, that the difference between the actual and expected value (determined from the area

mean) is statistically significant. Tumor section images contain areas devoid of tissue, which were excluded from our analysis using a binary tissue mask constructed by CRImage that classifies each pixel as tissue or non-tissue. Pixels corresponding to each region in the tumor were then summed up to determine the extent of tissue cover. Regions with $< 50\%$ tissue cover were excluded from hotspot analysis. From the remaining regions, we also excluded those that contained no cancer cells, as our primary interest was to study tumor regions in close proximity to cancer.

To perform Getis–Ord hotspot analysis on pathological samples, a grid of square size s and neighborhood size N_R needed to be determined, as input data are required in the form of spatial points with associated values and neighbors. In our exploratory analysis, three region and three neighborhood sizes were chosen for comparison. The region size s was defined as the side length of grid squares, with $s=50, 100$ or $250 \mu\text{m}$. The neighborhood sizes N_R were first, second and fourth order, where first order refers to the region being analyzed and the regions that share an edge or vertex with it, second order refers to first-order neighbors and regions that share an edge or vertex with first-order neighbors and so on. Figure 1b illustrates application of the grid and the different neighborhood arrangements. The combination of three region sizes and three neighborhood sizes were tested on the discovery cohort samples, with the exception that for $s=250 \mu\text{m}$ only N_R =first order was considered, as we wanted to keep the neighborhood radius in all cases to be no more than twice the cell–cell communication distance ($250 \mu\text{m}$) proposed by Francis *et al.*²⁸

These parameter combinations allowed us to investigate clustering of a variable across small, medium and large spatial configurations in order to determine the extent to which they affect any potential correlation that hotspots may have with clinical outcome of the patient. The z score for every region i is computed as,

$$z_i = \sum_{j=1}^n w_{i,j} c_j - \frac{\bar{c} \sum_{j=1}^n w_{i,j}}{SU},$$

where S and U are two normalizing factors, given by:

$$S = \sqrt{\frac{\sum_{j=1}^n c_j^2}{n} - (\bar{c})^2},$$

$$U = \sqrt{\frac{\left[n \sum_{j=1}^n w_{i,j}^2 - \left(\sum_{j=1}^n w_{i,j} \right)^2 \right]}{n-1}},$$

where n is the total number of grid regions (excluding those with $< 50\%$ tissue cover or no

cancer cells); c_j is the cancer or lymphocyte count for region j ; \bar{c} is the mean value of c for all regions in the image and w_{ij} is element (i,j) of the matrix of weights, w , which indicates the influence of two regions on each other and is used to check whether or not the two regions are neighbors:

$$w_{ij} = \begin{cases} 1 & \text{if } j \text{ is a neighbour of } i, \\ 0 & \text{if } j \text{ is not a neighbour of } i. \end{cases}$$

The P -value corresponding to each evaluated z score is determined using the thresholds specified in Getis and Ord.²⁴ For each region i , we evaluated its potential as a hotspot of cancer or immune cells, ie, a region with statistically significant clustering of these cells, as determined by its associated $p_i < 0.05$. Distributions of cancer, immune and co-localized cancer-immune hotspots for an example tumor specimen are given in Figure 2.

Parameter Selection and Validation

After applying hotspot analysis to our discovery cohort, we processed the results to select optimum values of the two spatial parameters, s and N_R , and a threshold from one of our three proposed hotspot measures, f_C , f_I and f_{C-I} , that best differentiated between good and poor clinical outcome (see Results section). This was achieved by recursively selecting every ten-thousandth value of each of our hotspot measures in the 15–85% quantiles to dichotomize these parameters in univariate survival analysis using the available 10-year disease-specific survival data. Following this, we constructed a multivariate Cox proportional hazards model for each of the cohorts, using the known clinical predictors node metastasis, tumor size and *lym*. Tumor grade was not found to be significant for survival in univariate analysis in the discovery ($P=0.196$) or validation ($P=0.852$) cohort and was therefore excluded in multivariate analysis.

Other Statistical Methods

Statistical tests were carried out in R. Survival curves were fitted using the Kaplan–Meier method, and univariate and multivariate analyses were performed using Cox proportional hazards models,²⁹ with P -values obtained from the Wald test. Associations between continuous variables were tested by evaluation of Pearson's correlation coefficient and between discrete and continuous variables using the Jonckheere–Terpstra (JT) trend test.³⁰ Student's t -test was used to determine if two group means are significantly different. In all cases, a P -value of < 0.05 was taken to be a significant result.

Results

Statistical Assessment of Cancer and Immune Hotspots in Primary ER-Negative Breast Tumors

To identify cancer cells and lymphocytes in breast whole-tumor sections, the H&E-stained images were first analyzed with our previously validated image analysis pipeline¹² for automated classification and detection of the spatial location of each cell nucleus (see Materials and methods section, Figures 2a and b). Subsequently, Getis–Ord spatial analysis was carried out to identify the precise location of the statistically significant clusters of cancer cells, ie, cancer hotspots, given a set of predefined parameters for spatial configuration (see Materials and methods section, Figure 2c). Specifically, Getis–Ord analysis computes a z -score for each region, which is a measure of the difference between observations in that region and the overall area being studied. Each z -score has an associated P -value, which can be used to determine statistical significance. Similarly, clusters or hotspots of lymphocytes in the same samples were identified (Figure 2d). To generate the summary statistics of spatial heterogeneity of immune infiltration for each tumor, we proposed three scores:

1. Cancer hotspot fractional: the fraction of cancer hotspots that are also immune hotspots (f_C);
2. Immune hotspot fractional: the fraction of immune hotspots that are also cancer hotspots (f_I); and
3. Cancer-immune hotspot co-localization: the fractional area of tumor where cancer and immune hotspots were co-localized (f_{C-I}).

The rationale for choosing these measurements is to investigate how cancer and immune hotspots are spatially related to each other, as it is only meaningful to study immune cells within the spatial context of cancer cells, and how the scores should be normalized, ie, by the number of cancer cell clusters in the sample (f_C), the number of immune cell clusters in the sample (f_I) or the total area of tumor analyzed for hotspot detection (f_{C-I}).

Histopathological review of representative scans yielding the highest co-localized hotspot scores confirmed the presence of a mixed population of mononuclear cells in close proximity to invasive tumor cells within the areas outlined by the program, as shown in an example in Figure 3.

Co-localization of Cancer and Immune Hotspots Is an Independent Prognostic Marker

To test the association between hotspot analysis results and prognosis in ER-negative breast cancer, we carried out univariate survival analysis in our discovery cohort for each of the seven spatial configuration results sets (see Materials and methods section), using Cox regression with a single predictor: f_C , f_I or f_{C-I} (Table 1). Three of the seven spatial configurations tested gave a significant result

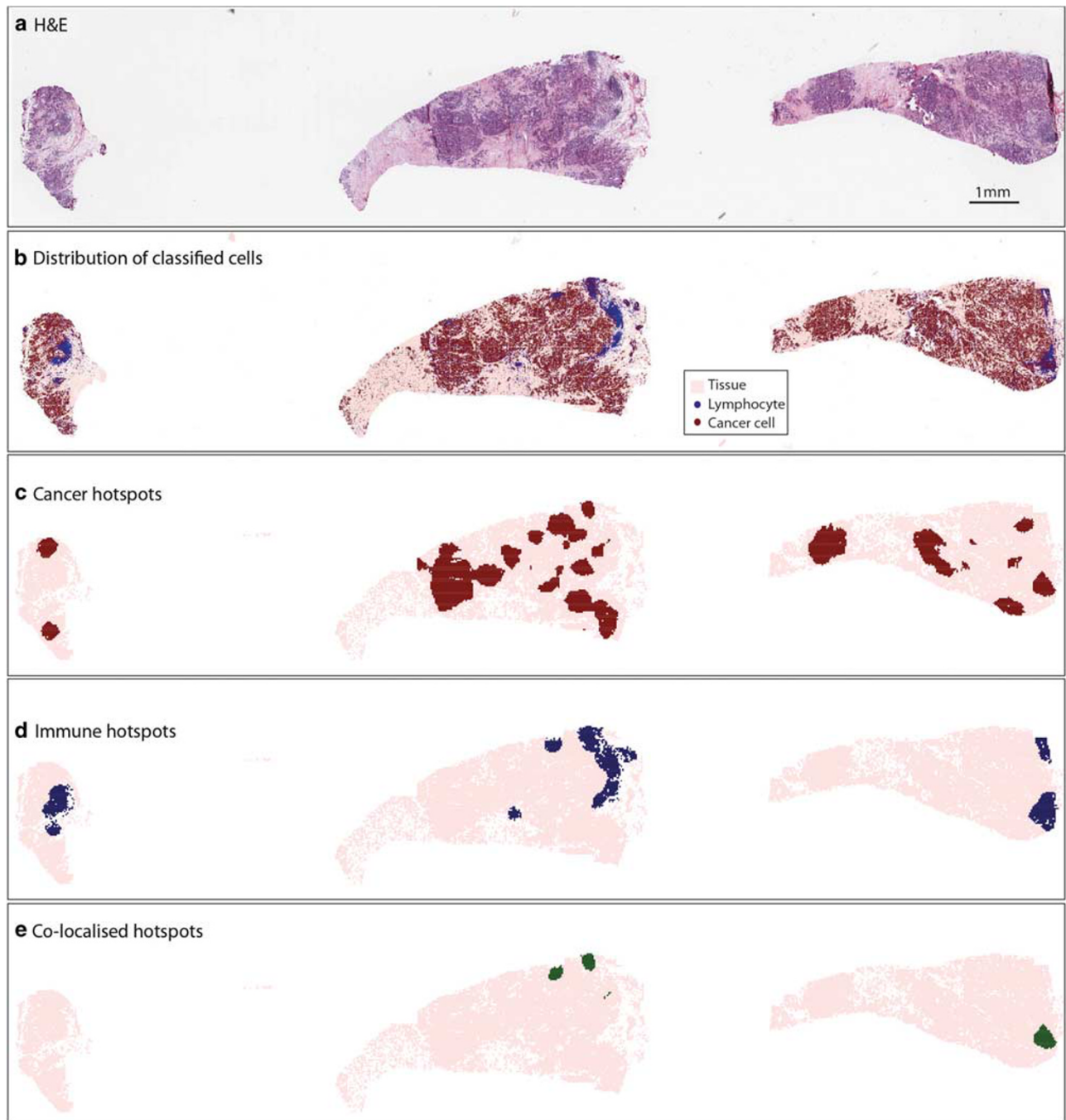


Figure 2 Mapping cancer and lymphocyte hotspots. (a) A H&E-stained tumor sample consisting of three sections. (b) Image analysis result, classifying nuclei in the image into cancer and lymphocyte and displaying their distribution in the sample. (c) Cancer and (d) immune (lymphocyte) hotspot detection using Getis–Ord hotspot analysis. (e) Co-localized cancer and immune hotspots, ie, regions classified as both cancer and immune hotspots.

in differentiating between good and poor disease-specific survival rate, for one or more of the three predictors tested. The most significant result was obtained for a spatial configuration of $s = 50 \mu\text{m}$ and $N_R = \text{fourth order}$, which, using a threshold in f_{C-I} of 1.91% for dichotomization, was able to differentiate between good and poor clinical outcome ($P = 0.009$, hazard ratio (HR) = 3.204, 95% confidence interval =

1.343–7.646, Figure 4a). This was further confirmed in the validation cohort ($P = 0.017$, HR = 2.207, HR 95% confidence interval = 1.138–3.747, Table 2, Figure 4a). Thus, best results can be obtained using a spatial configuration of small regions ($s = 50 \mu\text{m}$) and a large neighborhood ($N_R = \text{fourth order}$, see Materials and methods section), suggesting that the analysis benefits from abundant data points with

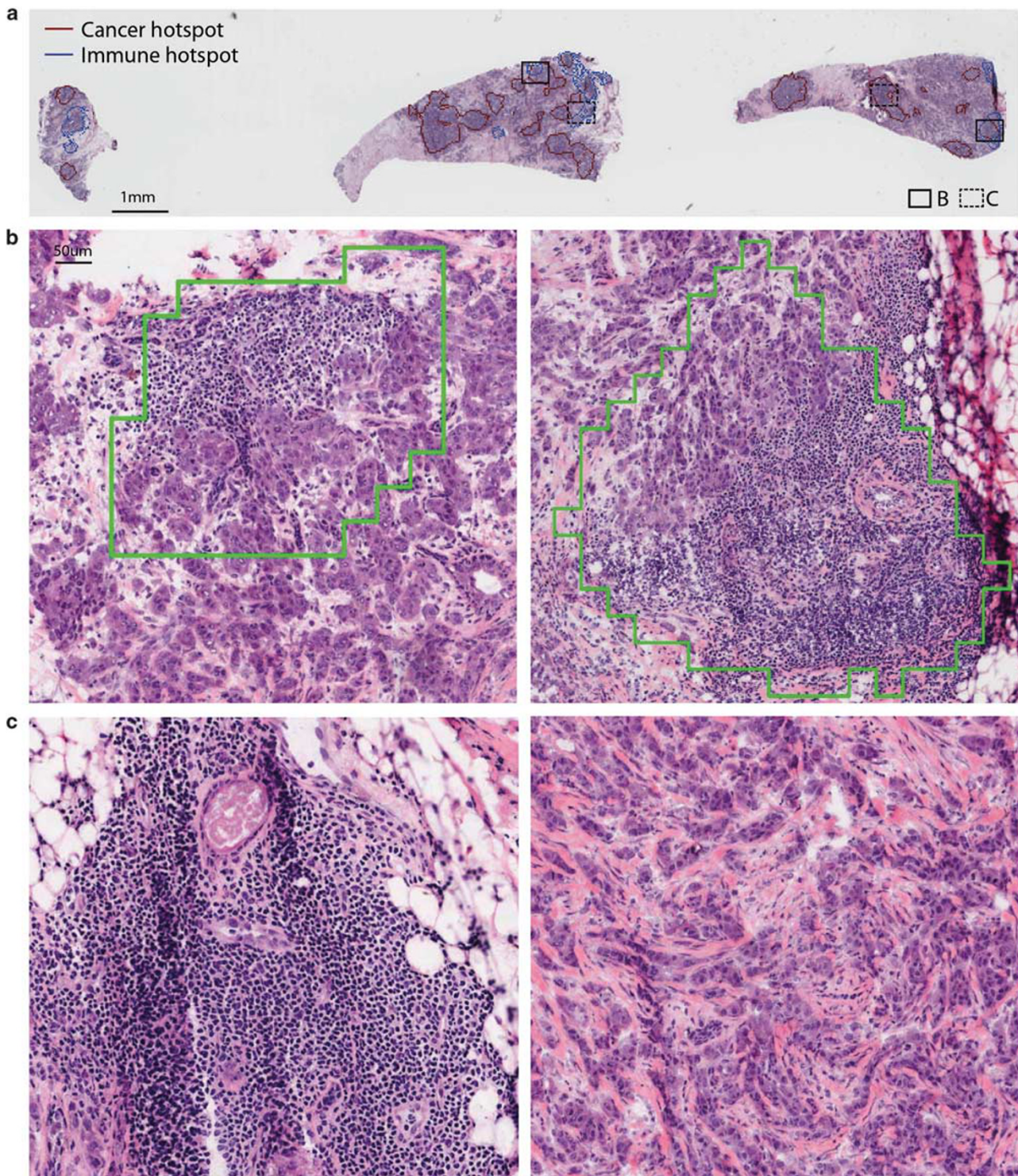


Figure 3 Representative examples of co-localized hotspots. (a) A low-resolution image to show locations of cancer (red) and immune (blue) hotspots. (b) High-resolution images to show two examples of co-localized hotspots (outlined in green) in the corresponding regions of panel (a). (c) Example of an immune hotspot (left) and a cancer hotspot (right).

large neighborhood coverage. Taken together, we found that f_{C-I} , the fractional tumor area where cancer and immune hotspots co-localized (co-localized hotspot score), is significantly associated with

good disease-specific survival in ER-negative breast cancer.

We proceeded to test our proposed co-localized hotspot measure in multivariate analysis with the

Table 1 The minimum P -values found in univariate survival analysis in the discovery cohort using our three measures for overlapping hotspots, f_C , f_I and f_{C-I} , given for the different spatial configurations tested to detect hotspots

Grid cell size, s (μm)	f_C			f_I			f_{C-I}		
				Neighborhood size, N_R					
	First	Second	Fourth	First	Second	Fourth	First	Second	Fourth
50	×	×	×	×	0.031	×	×	0.034	0.009
100	×	×	×	×	×	×	×	×	0.030
250	×	—	—	×	—	—	×	—	—

Where the minimum P -value found was >0.05 , it is indicated by an '×'. Only first order neighbors were considered when a grid cell size s of 250 μm was used.

clinical variables node metastasis, tumor size and *lym* to assess its potential as an independent prognostic indicator for ER-negative breast cancer. Tumor grade was found to be statistically insignificant for both cohorts in univariate analysis, hence it was excluded from the multivariate model. We constructed a Cox proportional hazards model and found that all three of the clinical variables as well as co-localized hotspots ($P=0.019$, HR = 2.910, HR 95% confidence interval = 1.189–7.118) could predict survival independently in our discovery cohort (Table 2). In our validation cohort, node metastasis and *lym* as a continuous variable failed to predict survival independently; however, our proposed hotspot index ($P=0.036$, HR = 1.907, HR 95% confidence interval = 1.042–3.489) remained independent of tumor size (Table 2).

We compared our proposed prognostic indicator with six clinical factors known to be associated with survival in ER-negative breast cancer: node metastasis, tumor size, HER2 status, tumor area, *lym*, and patient age (Figure 4b), using the JT trend test for the categorical variables and Pearson's correlation test for the continuous parameters. The strongest correlation with the co-localized hotspot index found was *lym*, where a Pearson's correlation coefficient r of 0.212 ($P=0.017$) and 0.237 ($P=0.007$) was obtained for the discovery and validation cohorts, respectively. The association between our measure and the remaining factors was found to be weak ($-0.07 < r < 0.17$) or insignificant (JT > 0.05).

Cancer–Immune Hotspots are Associated with Organ-Specific Metastasis

We next investigated whether any significant differences were present between patients with low and high co-localized hotspot scores and time to local and distant metastasis. Excluding patients whose tumors did not metastasize in our records, we found for the remaining patients that the mean time in months to local metastasis was significantly lower in patients with a low co-localized hotspot score ($< 1.91\%$) (63.8 vs 91.0, $P=0.021$). Similarly, the mean time in months to distant metastasis was found

to be significantly lower in patients with a low co-localized hotspot score (26.9 vs 38.9, $P=0.030$). Analysis of metastasis location showed a pronounced difference between patients with a low and high co-localized hotspot score for metastasis in the bone (16.4 and 8.5%, respectively), liver (15.1 and 8.5%, respectively) and, to a lesser extent, lung (8.2 and 11.3%, respectively) and brain (8.2 and 10.2%, respectively) (Figure 5).

Immune Spatial Heterogeneity Further Stratifies Patients with High Lymphocyte Abundance

As there is a significant correlation between *lym* and our proposed co-localized hotspot index, we further compared their performances as prognostic factors in ER-negative breast cancer. Low levels of co-localized hotspots and *lym* both identified a group of patients with a more aggressive form of ER-negative breast cancer (Figures 6a and b). Here we identified 8.52% as the optimal cutoff for *lym* in all 245 samples and used this cutoff for the purpose of comparison. This cutoff is slightly different to the 8% we previously reported,¹² which also has borderline significance in this sample set ($P=0.078$, Supplementary Materials). We discovered, first, that patients with a high abundance of immune cells do not necessarily have a high number of cancer–immune hotspots, while patients with a high number of hotspots can have a low level of immune cell abundance in the tumor (Figure 6c). This is expected as our hotspot score measures spatial heterogeneity of cancer and immune cell distribution; thus, despite high abundance, more uniformly distributed immune cells will receive a lower immune hotspot score. Second, our proposed score can further stratify patients with high *lym* into good and poor prognosis groups ($P=0.007$, Figure 6d). Finally, the prognosis of patients with low co-localized hotspot scores does not differ according to their *lym* ($P=0.555$).

Discussion

This study has brought together spatial statistics, image processing and pathological and clinical data

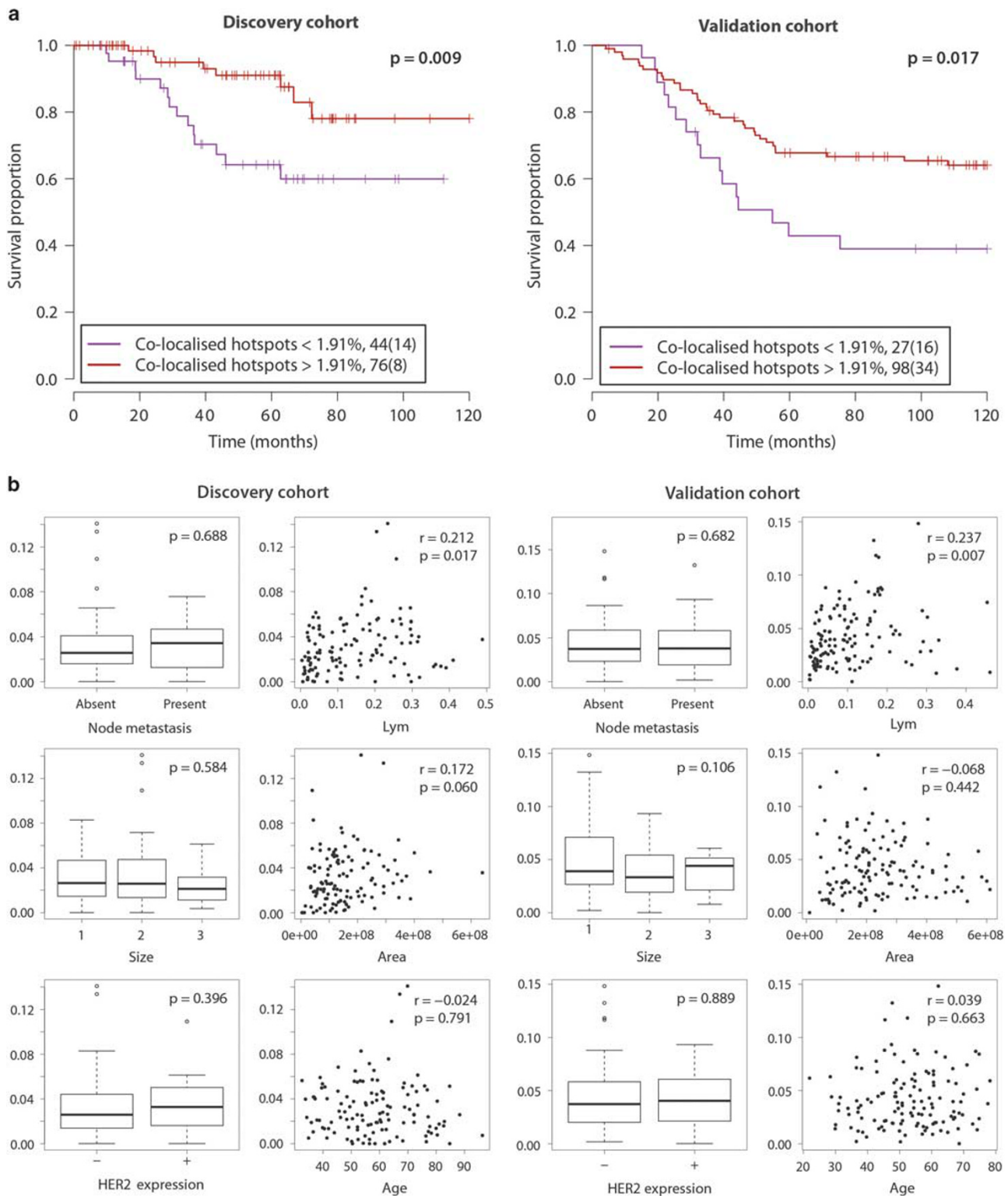


Figure 4 Prognostic value of the co-localized hotspot score. (a) Kaplan-Meier curves displaying disease-specific survival rates over 10 years in the discovery and validation cohorts. Patients in both cohorts are stratified by the co-localized hotspot score using a selected threshold of 1.91%. The numbers and the numbers in brackets (bottom-right) indicate the numbers of patients and the numbers of breast cancer-specific deaths in each group, respectively. (b) Plots illustrating how the cancer-immune hotspot measure (y axis) relates to six factors known to be associated with clinical outcome in breast cancer, in both the discovery (left) and validation (right) cohorts. Jonckheere-Terpstra trend test P -value is given for lymph node metastasis, tumor size and HER2 expression and Pearson's correlation coefficient r and its P -value are given for lymphocytic abundance (*lym*), tumor section area and patient age, at the top-right of each plot.

Table 2 Univariate and multivariate survival analysis results using known clinical parameters and our proposed prognostic indicator, co-localized hotspot score, as predictors, and 10-year disease-specific survival data as the response variable

	Univariate analysis			Multivariate analysis		
	P-value	HR	HR 95% CI	P-value	HR	HR 95% CI
<i>Discovery cohort</i>						
Node metastasis	0.002	4.504	1.74–11.66	0.008	3.898	1.421–10.693
Tumour size	0.003	2.674	1.387–5.156	0.116	1.837	0.860–3.921
Lym	0.031	164.6	1.591–17020	0.029	231	1.733–30800
Co-localized hotspot score	0.009	3.204	1.343–7.646	0.019	2.910	1.189–7.118
<i>Validation cohort</i>						
Node metastasis	0.076	1.801	0.940–3.451	0.294	1.431	0.733–2.796
Tumour size	0.007	2.027	1.218–3.372	0.017	1.916	1.122–3.272
Lym	0.271	0.168	0.248–143.5	0.574	2.465	0.106–57.2
Co-localized hotspot score	0.017	2.207	1.138–3.747	0.036	1.907	1.042–3.489

Abbreviations: CI: confidence interval; HR: hazard ratio; Lym: lymphocytic abundance.

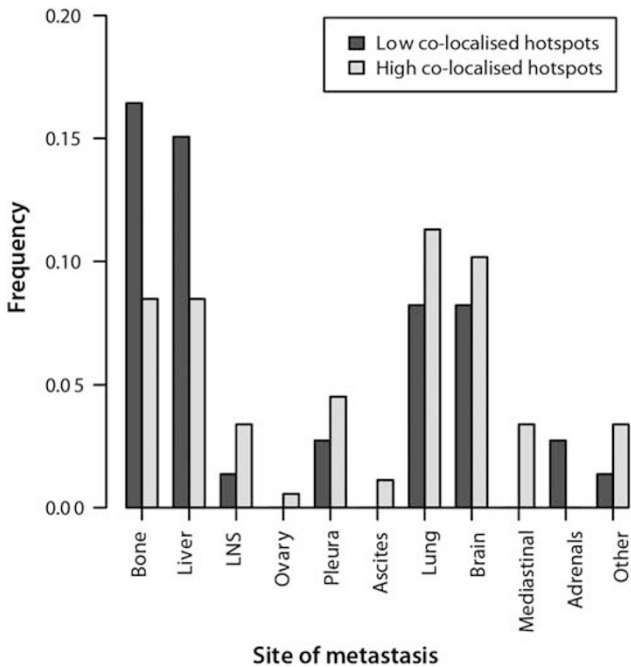


Figure 5 The location of distant metastasis in patients who relapsed, grouped by the co-localized hotspot score. The y axis displays the fractional occurrence of relapse for a given site.

to analyze 245 H&E-stained tumor sections from ER-negative breast cancer patients in the search for new understanding of the spatial heterogeneity of lymphocytic infiltration. The use of Getis–Ord hotspot analysis has been reported in a wide range of ecological studies, but we are thus far unaware of any previous application of this powerful statistical tool to analyze pathological images of breast tumors. Here it was used to describe spatial clusters or hotspots of cancer and immune cells, defined as regions in the tumor with a significantly high number of these cells given the global cell

distribution. The most important findings of this study are that: (i) there is a significant correlation between a high fractional tumor area where cancer and immune hotspots are co-localized (high f_{C-I} or co-localized hotspot score) and a favorable clinical outcome, (ii) this score, as our proposed prognostic indicator, can predict patient survival independently from known clinical parameters, and (iii) there is no significant correlation between our prognostic indicator and the routinely used clinical parameters. Furthermore, among patients whose tumors metastasized, those with low co-localized hotspot scores had significantly less time to local and distant metastasis and presented a twofold increased occurrence of metastasis to the bone and liver compared with those with high co-localized hotspot scores. This observation raises the question of whether cancer–immune hotspot co-localization is a feature of the microenvironment that provides organ-specific niche for successful metastasis.^{31–33}

Notably, a weak correlation was found between the degree of cancer and immune hotspot co-localization and immune abundance, a parameter of clinical significance reported by us and other groups^{7,10,12} who showed that a high immune abundance is correlated with better survival outcome in ER-negative breast cancer. We were able to demonstrate that our hotspot score can further stratify patients with high immune abundance. Remarkably, patients with low hotspot scores, regardless of their immune abundance, had a similar, poor prognosis. This highlights the need to study heterogeneity of spatial distribution in addition to measuring abundance of microenvironmental components.

Our proposed approach scores a type of immune feature in pathological samples that is different from published assessment of lymphocytic infiltration,^{10,11,14} where tumor nests and stromal areas were scored based on whether infiltration by mononuclear inflammatory cells is present, thus evaluating the amount of intraepithelial or stromal immune

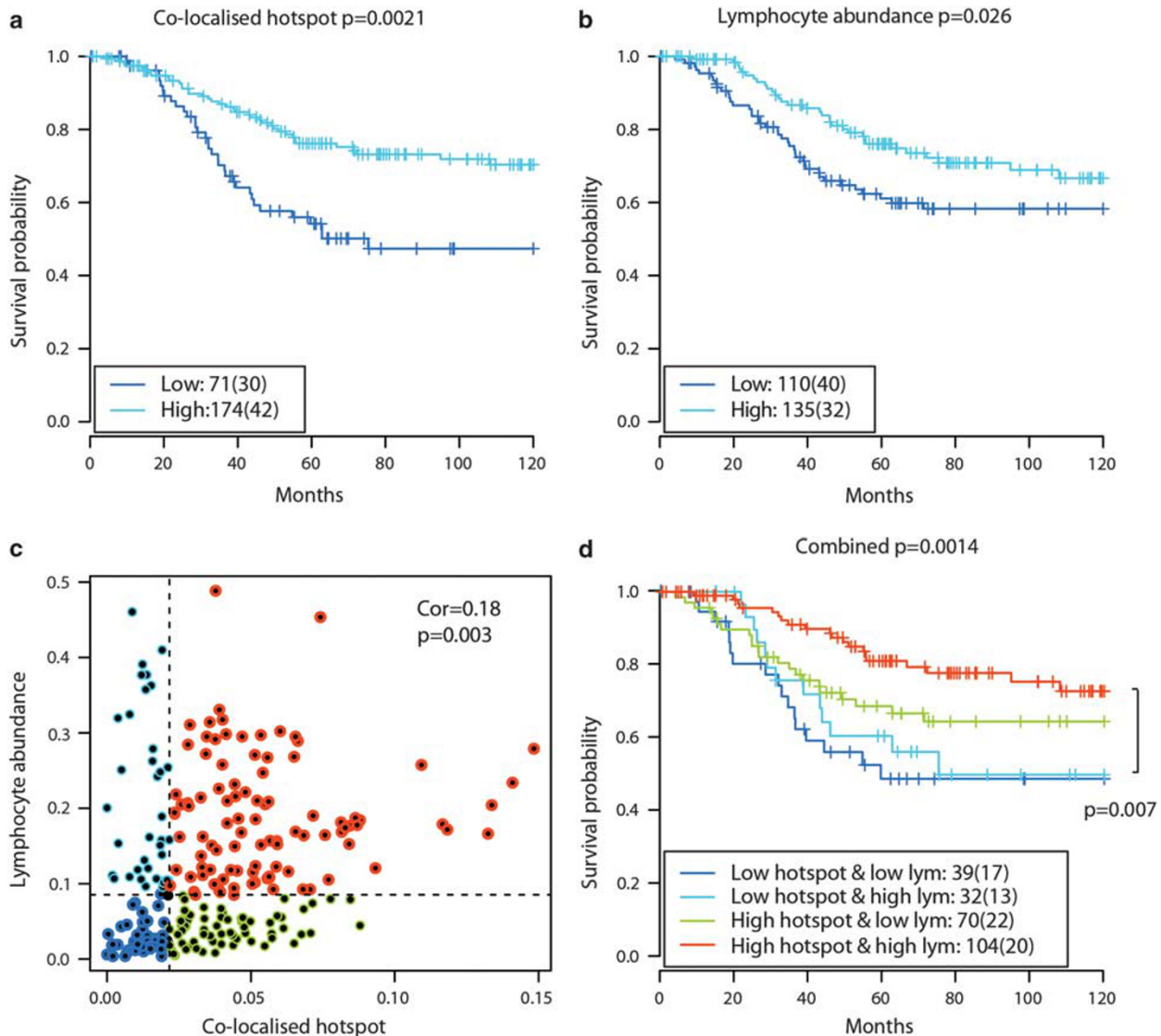


Figure 6 Comparison of cancer-immune hotspots with lymphocyte abundance (*lym*). (a, b) Kaplan-Meier curves illustrating disease-specific survival differences between patient groups stratified by co-localized hotspots and *lym*, respectively, using all 245 ER-negative breast tumor samples. (c) Scatter plot to show correlation (quantified using Pearson's method, top-right) between these two parameters. Dashed lines marked the critical thresholds for the parameters. Colors of the points correspond to groupings in panel (d). (d) Kaplan-Meier curves showing patient stratification using both indices. Cancer-immune hotspot score was able to further stratify patients with high *lym* (cyan and orange groups, $P = 0.007$) but not those with low *lym* (dark blue and green groups, $P = 0.155$). *Lym* could not stratify patients with a low cancer-immune hotspot score (cyan and dark blue groups, $P = 0.555$).

infiltrates. In these studies, the authors did not find evidence of association between the degree of intra-tumor lymphocytic infiltration and clinical outcome. In contrast, our focus is on spatial aggregation of cancer cells and lymphocytes, thereby capturing spatial heterogeneity instead of abundance and frequency. The amount of heterogeneity was summarized as a ratio of spatial aggregation to tissue area, the quantification of which required the use of automated whole-section analysis, as it is impractical to visually account for every lymphocyte present. In samples we inspected visually, the majority of cancer-immune hotspots were located

adjacent to tumor nests. ER-negative breast tumors are known to have a pushing growth pattern,³⁴ and a future systematic investigation on distribution of hotspots according to tumor growth patterns could further our understanding of these hotspots.

The main strength of Getis-Ord hotspot analysis is that the global distribution of cells in the whole tumor is accounted for when assigning a local P -value indicative of cell clustering. Thus the identified clusters are not merely regions of high cell density but regions of significantly high cell density, which indicates spatial heterogeneity. Frequency mapping would not suffice for studying this

heterogeneity, as frequency measures would not indicate how different regions compare with each other in terms of statistical significance. Compared with other methods in spatial statistics such as Ripley's *K*, which gives a global *P*-value to indicate the trend of clustering,³⁵ the distinct advantage of Getis–Ord hotspot analysis lies in its ability to pinpoint local regions of cell clustering. These regions will be important targets for follow-up experimental investigation to understand the molecular aberrations underlying cancer–immune co-localization—a clinically important phenomenon reported here. In our subsequent studies, immunohistochemical staining can be used to determine whether particular types of immune cells dominate these hotspots and whether cancer cells in these areas have a common marker that can explain the greater immune infiltration in their surrounding tissue. DNA sequencing and gene expression analysis of these select areas can further shed light on the biochemical framework underpinning the cancer–immune relationship. In addition, whether these hotspots form a particular biological structure is an interesting question, albeit beyond the scope of our present study.

The limitations of analyzing a three-dimensional structure on a two-dimensional scale inherent in all pathological assessment cannot be ignored. Any current sampling procedure can only approximate the state of the entire tumor. We evaluated cancer and immune hotspots based on an average of three sections from different locations within a tumor, thereby effectively approximating three-dimensional and inter-section heterogeneity. Furthermore, the association of our cancer–immune hotspot measure with patient survival is a significant one, having been discovered and validated across two independent patient cohorts, and thus there is undoubtedly valuable information to be found in two-dimensional tumor sections despite the approximation. In addition, H&E-stained tumor samples with $\times 20$ magnification that were used in our study do not allow further sub-classification of lymphocytes; we identified lymphocytes in our image processing algorithm as nuclei with a small, round and dark morphology. However, determining lymphocytic infiltration in breast cancer based on H&E-stained samples has produced significant results in previous studies,^{9–12} suggesting that, for the purpose of developing a prognostic biomarker, H&E samples may be sufficient in specific breast cancer subtypes. To enhance the potential clinical utility of our approach, we believe that it should be further tested on much larger sample sets, as well as extended to other cancer types where immune infiltration is often observed. In particular, following guidelines presented by Salgado *et al*³⁶ (unpublished results), recommending exclusion of DCIS foci in visual assessment of tumor-infiltrating lymphocytes, we are investigating suitable approaches for implementing this in our computational image analysis pipeline.

In conclusion, the application of multidisciplinary techniques to understand the complex interplay between different cell types in a tumor is an emerging and promising area. With many image processing algorithms developed for automated detection of different tissue types and constituents, including cancerous and stromal tissue, automation allows for high-throughput processing while avoiding observer bias; for example, our hotspot score for a slide remains unchanged upon re-evaluation. Our study presents an alternative yet complementary method for studying tumor immune infiltration to the one presented by Salgado *et al*³⁶ (unpublished), which is a powerful collaborative effort to standardize tumor immune infiltration scoring in routine diagnosis via visual inspection. Fully automated cell classification algorithms are not in mainstream clinical use at present, but we believe that our study demonstrates clear advantages of implementing computational methods to complement visual assessment by pathologists. In particular, we have shown that detecting cell locations using image analysis can allow the application of spatial statistical tools from other disciplines to effectively account for spatial heterogeneity in the tumor and its microenvironment, which can add valuable prognostic information.

Acknowledgments

SN is a recipient of an EPSRC doctoral scholarship. The authors acknowledge funding from the Institute of Cancer Research. This work was supported by the Wellcome Trust [105104/Z/14/Z].

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

- 1 Knight WA, Livingston RB, Gregory EJ, *et al*. Estrogen receptor as an independent prognostic factor for early recurrence in breast cancer. *Cancer Res* 1977;37: 4669–4671.
- 2 Miglietta L, Morabito F, Provinciali N, *et al*. A prognostic model based on combining estrogen receptor expression and Ki-67 value after neoadjuvant chemotherapy predicts clinical outcome in locally advanced breast cancer: extension and analysis of a previously reported cohort of patients. *Eur J Surg Oncol* 2013;39:1046–1052.
- 3 Hoefnagel LD, Moelans CB, Meijer SL, *et al*. Prognostic value of estrogen receptor alpha and progesterone receptor conversion in distant breast cancer metastases. *Cancer* 2012;118:4929–4935.
- 4 Weigelt B, Reis-Filho JS. Back to the basis: breast cancer heterogeneity from an etiological perspective. *J Natl Cancer Inst* 2014;106:1.
- 5 Brenton JD, Carey LA, Ahmed AA, *et al*. Molecular classification and molecular forecasting of breast

- cancer: ready for clinical application? *J Clin Oncol* 2005;23:7350–7360.
- 6 Teschendorff AE, Miremadi A, Pinder SE, *et al*. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. *Genome Biol* 2007;8:R157.
 - 7 Calabro A, Beissbarth T, Kuner R, *et al*. Effects of infiltrating lymphocytes and estrogen receptor on gene expression and prognosis in breast cancer. *Breast Cancer Res Treat* 2009;116:69–77.
 - 8 Ascierto ML, Kmiecik M, Idowu MO, *et al*. A signature of immune function genes associated with recurrence-free survival in breast cancer patients. *Breast Cancer Res Treat* 2012;131:871–880.
 - 9 Denkert C, Loibl S, Noske A, *et al*. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 2010;28:105–113.
 - 10 Loi S, Sirtaine N, Piette F, *et al*. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. *J Clin Oncol* 2013;31:860–867.
 - 11 Issa-Nummer Y, Darb-Esfahani S, Loibl S, *et al*. Prospective validation of immunological infiltrate for prediction of response to neoadjuvant chemotherapy in HER2-negative breast cancer—a substudy of the neoadjuvant GeparQuinto trial. *PLoS One* 2013;8:e79775.
 - 12 Yuan Y, Failmezger H, Rueda OM, *et al*. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med* 2012;4:157ra43.
 - 13 Galon J, Costes A, Sanchez-Cabo F, *et al*. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006;313:1960–1964.
 - 14 Mahmoud SMA, Paish EC, Powe DG, *et al*. Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. *J Clin Oncol* 2011;29:1949–1955.
 - 15 Kruger JM, Wemmert C, Sternberger L, *et al*. Combat or surveillance? Evaluation of the heterogeneous inflammatory breast cancer microenvironment. *J Pathol* 2013;229:569–578.
 - 16 Balsat C, Blacher S, Signolle N, *et al*. Whole slide quantification of stromal lymphatic vessel distribution and peritumoral lymphatic vessel density in early invasive cervical cancer: a method description. *ISRN Obstet Gynecol* 2011;2011:354861.
 - 17 Balsat C, Signolle N, Goffin F, *et al*. Improved computer-assisted analysis of the global lymphatic network in human cervical tissues. *Mod Pathol* 2014;27:887–898.
 - 18 Singanamalli A, Sparks R, Rusu M, *et al*. Identifying *in vivo* DCE MRI parameters correlated with *ex vivo* quantitative microvessel architecture: a radiohistomorphometric approach. *Proc SPIE* 8676, Medical Imaging 2013: Digital Pathology, 867604 (March 29, 2013); doi:10.1117/12.2008136.
 - 19 Doyle S, Feldman M, Tomaszewski J, *et al*. A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Trans Biomed Eng* 2012;59:1205–1218.
 - 20 Janowczyk A, Chandran S, Madabhushi A. Quantifying local heterogeneity via morphologic scale: distinguishing tumoral from stromal regions. *J Pathol Inform* 2013;4:8.
 - 21 Basavanthally AN, Ganesan S, Agner S, *et al*. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. *IEEE Trans Biomed Eng* 2010;57:642–653.
 - 22 Beck AH, Sangoi AR, Leung S, *et al*. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci Transl Med* 2011;3:108ra13.
 - 23 Belien JA, Somi S, de Jong JS, *et al*. Fully automated microvessel counting and hot spot selection by image processing of whole tumor sections in invasive breast cancer. *J Clin Pathol* 1999;52:184–192.
 - 24 Getis A, Ord JK. The analysis of spatial association by use of distance statistics. *Geogr Anal* 1992;24:189–206.
 - 25 de Souza D, Kelly-Hope L, Lawson B, *et al*. Environmental factors associated with the distribution of *Anopheles gambiae* s.s in Ghana; an important vector of lymphatic filariasis and malaria. *PLoS One* 2010;5:e9927.
 - 26 Setiadi AF, Ray NC, Kohrt HE, *et al*. Quantitative, architectural analysis of immune cell subsets in tumor-draining lymph nodes from breast cancer patients and healthy lymph nodes. *PLoS One* 2010;5:e12420.
 - 27 Curtis C, Shah SP, Chin SF, *et al*. The genomic and transcriptomic architecture of 2,000 breast tumors reveals novel subgroups. *Nature* 2012;486:346–352.
 - 28 Francis K, Palsson BO. Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion. *Proc Natl Acad Sci USA* 1997;94:12258–12262.
 - 29 Cox DR. Regression models and life-tables. *J R Stat Soc Series B Stat Methodol* 1972;34:187–220.
 - 30 Jonckheere AR. A distribution-free k-sample test against ordered alternatives. *Biometrika* 1954;41:133–145.
 - 31 Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. *Cell* 2011;147:275–292.
 - 32 Wels J, Kaplan RN, Rafii S, *et al*. Migratory neighbors and distant invaders: tumor-associated niche cells. *Genes Dev* 2008;22:559–574.
 - 33 Gupta GP, Massague J. Cancer metastasis: building a framework. *Cell* 2006;127:679–695.
 - 34 Putti TC, El-Rehim DMA, Rakha EA, *et al*. Estrogen receptor-negative breast carcinomas: a review of morphology and immunophenotypical analysis. *Mod Pathol* 2004;18:26–35.
 - 35 Ripley BD. The second-order analysis of stationary point processes. *J Appl Probab* 1976;13:255–266.
 - 36 Salgado R, Denkert C, Demaria S, *et al*. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol* 2014;26:259–271.

Supplementary Information accompanies the paper on Modern Pathology website (<http://www.nature.com/modpathol>)