

Reliability of tumor-infiltrating lymphocyte and tertiary lymphoid structure assessment in human breast cancer

Laurence Buisseret^{1,2,3,8}, Christine Desmedt^{2,8}, Soizic Garaud^{1,8}, Marco Fornili⁴, Xiaoxiao Wang^{1,2,3}, Gert Van den Eyden⁵, Alexandre de Wind⁶, Sebastien Duquenne^{6,9}, Anais Boisson¹, Céline Naveaux¹, Françoise Rothé², Sandrine Rorive⁷, Christine Decaestecker⁷, Denis Larsimont⁶, Martine Piccart-Gebhart³, Elia Biganzoli⁴, Christos Sotiriou^{2,3} and Karen Willard-Gallo¹

¹Molecular Immunology Unit, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium; ²Breast Cancer Translational Research Laboratory, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium; ³Department of Medicine, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium; ⁴Department of Clinical Sciences and Community Health, Unit of Medical Statistics, Biometry and Bioinformatics 'Giulio A. Maccacaro', University of Milan Campus Cascina Rosa, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy; ⁵Department of Pathology, GZA Ziekenhuizen, Sint-Augustinus Campus, Wilrijk, Belgium; ⁶Department of Pathology, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium and ⁷Center for Microscopy and Molecular Imaging, Gosselies, Belgium

The presence of tumor-infiltrating lymphocytes (TIL), reflecting host immune activity, is frequently correlated with better clinical outcomes, particularly in HER2-positive and triple-negative breast cancer. Recent findings suggest that organization of immune infiltrates in tertiary lymphoid structures also has a beneficial effect on survival. This study investigated inter- and intra-observer variation in TIL assessment using conventional hematoxylin-eosin versus immunohistochemical staining to identify immune cells. Global, intratumoral, and stromal TIL, as well as tertiary lymphoid structures were scored independently by experienced pathologists on full-face tumor sections ($n=124$). The fidelity of scoring infiltrates in core biopsies compared to surgical specimens, and pathological assessment compared to quantitative digital analysis was also evaluated. The inter-observer concordance correlation coefficient was 0.80 for global, 0.72 for intratumoral, and 0.71 for stromal TIL, while the intra-observer concordance correlation coefficient was 0.90 for global, 0.77 for intratumoral, and 0.89 for stromal TIL using immunohistochemical stains. Correlations were lower with hematoxylin-eosin stains, particularly for intratumoral TIL, while global scores had the highest concordance correlation coefficients. Our study concluded that tertiary lymphoid structures are accurately and consistently scored using immunohistochemical but not hematoxylin-eosin stains. A strong association was observed between TIL in core biopsies and surgical samples ($R^2=0.74$) but this did not extend to tertiary lymphoid structures ($R^2=0.26$). TIL scored by pathologists and digital analysis were correlated but our analysis reveals a constant bias between these methods. These data challenge current criteria for TIL and tertiary lymphoid structure assessment in breast cancer and recommend that how pathologists evaluate immune infiltrates be reexamined for future studies.

Modern Pathology (2017) 30, 1204–1212; doi:10.1038/modpathol.2017.43; published online 16 June 2017

Correspondence: Dr K Willard-Gallo, PhD, Molecular Immunology Unit, Institut Jules Bordet, Université Libre de Bruxelles, 121 Blvd. de Waterloo, Brussels B1000, Belgium.

E-mail: kwillard@ulb.ac.be

⁸These authors contributed equally to this work.

⁹Current address: Department of Pathology, Centre Hospitalier Universitaire de la Réunion, Saint-Pierre, La Réunion, France

Received 30 August 2016; revised 30 March 2017; accepted 2 April 2017; published online 16 June 2017

The extent of tumor-infiltrating lymphocytes (TIL) evaluated on hematoxylin-eosin-stained tumor sections is linearly correlated with prognosis in primary breast cancer and particularly robust in the triple-negative and HER2-positive subtypes.¹ TIL scored on diagnostic biopsies have also been correlated with a higher rate of pathological complete response after neoadjuvant chemotherapy in ER-negative as well as in ER-positive breast cancer.^{2,3} These clinical studies

suggest TIL are a new biomarker for breast cancer management decisions and particularly well-suited for treatment decisions associated with new cancer immunotherapies.

Broadening the clinical utility of scoring tumor immune infiltration is currently restricted by the need for standardized evaluation methods. Initial steps to harmonize TIL evaluation for breast cancer were taken by the international working group who published TIL scoring recommendations for hematoxylin-eosin-stained sections in 2014 (ref. 4) and standardized these evaluations in 2016.⁵ Most studies score TIL as ‘stromal,’ which is defined as mononuclear cells located within the border of the invasive tumor but not in direct contact with malignant cells, and as ‘intratumoral,’ which are intraepithelial mononuclear cells in tumor cell nests or in direct contact with tumor cells.² Stromal and intratumoral TIL are correlated but scoring of the latter, generally at lower abundance, is inconsistent on hematoxylin-eosin-stained sections. Current guidelines recommend evaluating immune infiltration on a continuous scale as a percentage of stromal TIL in the invasive tumor area⁴ with this approach effectively applied in the recent ring studies.⁵ The recommendation to score stromal TIL alone is based on the practical experience of many breast pathologists. Comparing inter- and intra-observer variability for global, stromal, and intratumoral TIL on hematoxylin-eosin and immunohistochemically stained sections has not been done. This leaves a number of open questions, including: (1) Are immunohistochemical staining’s more reliable and reproducible than their histochemical counterparts for scoring TIL?; (2) Which of these two stains produces the most consistent results for global, stromal, and intratumoral TIL?; (3) Does tumor

immune infiltration evaluation on core biopsies accurately reflect the surgical specimen?; and (4) Can digital pathology provide reliable help for the quantitative assessment of tumor immune infiltration?

Recent studies have shown that TIL in the peritumoral stroma are frequently aggregated in organized tertiary lymphoid structures. These structures may signal local immune responses directed to the tumor, a hypothesis supported by studies showing their prognostic value for several malignancies, including breast cancer (reviewed in Dieu-Nosjean *et al*⁶). Tertiary lymphoid structures were first identified in breast cancer on immunohistochemically stained sections⁷ with recent studies using histochemical staining for scoring.^{8–10} The same questions listed above for TIL can also be asked for tertiary lymphoid structures.

The present study was undertaken to address these issues. We compared scoring of TIL and tertiary lymphoid structures on histologically and immunohistochemically stained full-face tumor sections to determine inter-observer and intra-observer agreement. We also investigated how accurately TIL and tertiary lymphoid structure evaluation on diagnostic core biopsies reflects the corresponding surgical specimen. Finally, we tested digital pathology as a bioimaging processing tool for rigorous and quantitative analysis of the tumor immune infiltrate.

Materials and methods

Tumor Samples

Formalin-fixed and paraffin-embedded breast tumor blocks were selected from the Institut Jules Bordet’s tumor bank, including 124 cases of primary tumors

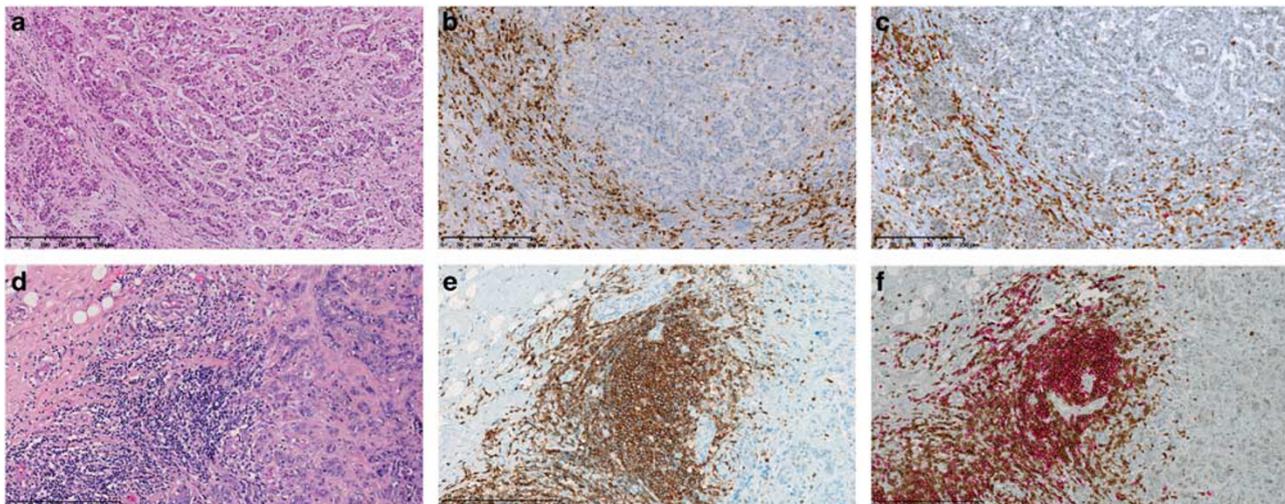


Figure 1 Staining of breast tumor tissues. Representative images of diffuse tumor-infiltrating lymphocytes (TIL) in a breast tumor (a–c) and enlargement of an area where immune cells are organized in tertiary lymphoid structures (d–f) is shown for tissue sections stained with hematoxylin-eosin (a and d) or CD45 immunohistochemistry (b and e; CD45 = total leukocytes) or CD3/CD20 (c and f; CD3 = T cells (brown) and CD20 = B cells (red), respectively). Images (a–c) are at $\times 100$ and (d–f) at $\times 200$ magnification.

(Cohort A) and 65 pairs of diagnostic core biopsies with their corresponding primary tumor surgical specimen (Cohort B). The clinicopathological characteristics for all patients are detailed in Supplementary Tables S1–S3. This study was approved by the Institut Jules Bordet's ethics committee (EC -1943).

Histopathology Staining

Formalin-fixed and paraffin-embedded tissue sections (4 μ m) were histochemically stained with hematoxylin-eosin on a Tissue-Tek Prisma & Coverslipper HQ^{plus} (Sakura Finetek) stainer or immunohistochemically stained to label immune cells on a Ventana Benchmark XT automated staining instrument (Ventana Medical Systems). Antibodies were used to label either all leukocytes (CD45), (Cohort A) or T and B lymphocytes (CD3/CD20), respectively (Cohort B). A detailed protocol for the dual CD3/CD20 immunohistochemical stain is described in Buisseret *et al*.¹¹ Representative stained images are shown in Figure 1. Estrogen receptor, progesterone receptor, Ki67, and HER2 immunohistochemical stains were performed in the routine pathology lab using accredited procedures.

Pathologic Assessment

Full-face tumor sections were independently scored for TIL and tertiary lymphoid structures by two (Cohort A) or three (Cohort B) experienced pathologists (GvdE, AdW, and SD). Fifty randomly selected slides from Cohort A were rescored after 3 months for the same immune parameters by the same two pathologists. Global TIL are defined as the percentage of the invasive tumor area (including the tumor bed and peri-tumoral stroma) infiltrated by lymphocytes. Global, intratumoral, and stromal TIL were scored following the 2014 guidelines from the Breast Cancer International Tumor Infiltrating Lymphocyte Working Group.⁴ Tertiary lymphoid structures were identified on hematoxylin-eosin and CD45 immunohistochemically stained tissues as lymphoid aggregates, whereas for CD3/CD20 immunohistochemically stained tissues they were restricted to a B-cell follicle surrounded/adjacent to a T-cell zone. Regions of *in situ* carcinoma, normal glandular epithelium, and necrosis were excluded from evaluation.

Digital Pathology

CD3/CD20 ($n=50$) and the corresponding CD45 ($n=46$) immunohistochemically stained sections from Cohort A and CD45 immunohistochemically stained sections of tumors and biopsies from Cohort B were scanned at a magnification of $\times 20$ on a NanoZoomer slide scanner (Hamamatsu). Images were analyzed using Visiomorph^{DP} software

(Visiopharm) to quantify the CD3⁺CD20⁺ and CD45⁺ areas within the invasive tumor area defined by a pathologist for each digital image. The total positively stained area(s) was scored as a percentage of the defined region.

Statistical Analysis

TIL scores were log-transformed to approximate a normal distribution and stabilize the variance. A value of 0.5 was added to the scores equal to 0. Agreement between staining methods, inter-observer variability, and intra-observer repeatability were assessed using the Bland–Altman method and Passing–Bablok regression analyses. Bland–Altman plots represent the ratio of the TIL scores between the two staining methods or between two assessments (intra- and inter-observer variability) plotted against their geometric mean and report the agreement intervals, within which 95% of the ratios are expected to fall when the ratios are normally distributed.¹² In the Passing–Bablok regression lines, intercepts and slopes are measures, respectively, of constant and proportional bias on the log scale.¹³

The concordance correlation coefficient was used as a summary measure of reproducibility for both the log of continuous TIL scores and binary variable for a tertiary lymphoid structure presence.¹⁴ For binary variables with no repetitions, the concordance correlation coefficient is equivalent to Cohen's kappa. The association between TIL in diagnostic core biopsies and primary tumors was evaluated by a simple linear and Loess regression model of the log geometric means for the two pathologists. A logistic model, adjusting for the pathologist and accounting for intra-patient correlation by the Huber covariance estimator, was fitted to study the relationship between tertiary lymphoid structures in biopsies and tumors. Statistical analyses were performed using R v.3.2.5, exploiting in particular the Agreement, MethComp, and rms packages.¹⁵

Results

Comparative Analysis of TIL Scoring on Histochemically versus Immunohistochemically Stained Tumor Sections

TIL from 124 primary breast tumors (Cohort A) were independently scored by two pathologists on tumor sections stained using hematoxylin-eosin and dual CD3/CD20 immunohistochemistry. Three scores were determined: stromal, intratumoral, and global TIL, which were highly correlated with both staining methods (Supplementary Table S4). The scores from both pathologists were compared by taking the geometric mean and results were reported in Bland–Altman plots depicting the ratio between the two pathologist's scores against their geometric average (Figures 2a–c). The limits of agreement show

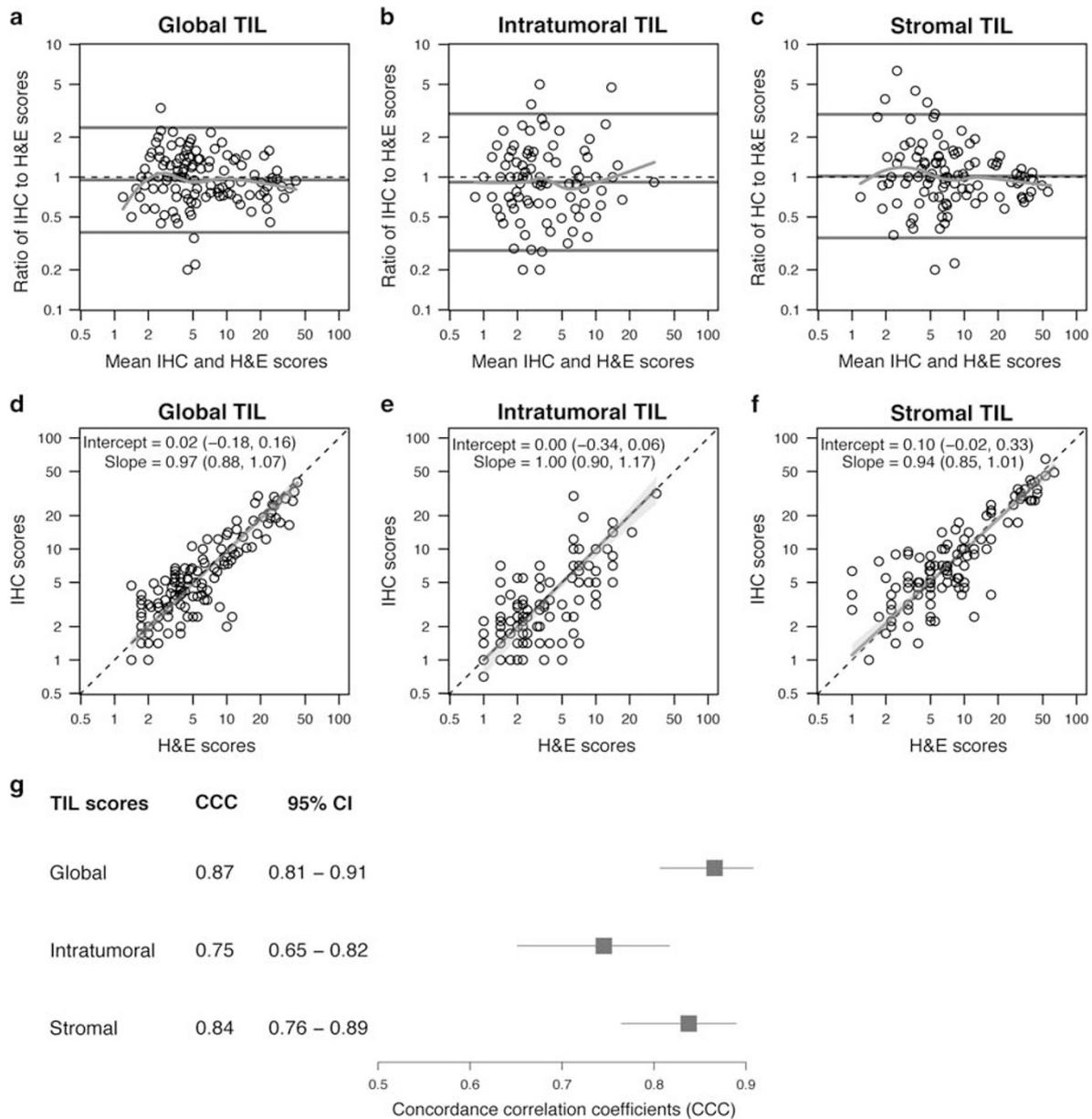


Figure 2 Hematoxylin and eosin (H&E) versus immunohistochemistry. Bland–Altman analyses of tumor-infiltrating lymphocytes (TIL) score agreement using H&E versus immunohistochemical (IHC) stains for global (a), intratumoral (b), and stromal (c) TIL. For each comparison, the geometric mean value for the two scores (histochemical H&E and IHC) from the same tumor sample (x axis) is plotted against the ratio between the same two scores (y axis). The horizontal lines represent the overall geometric mean of the ratios (center line) and the approximate 95% limits of agreement. The Loess fitted curve is also shown. Passing–Bablok regression of H&E versus IHC log-transformed scores for global (d), intratumoral (e), and stromal TIL (f) are shown. Each comparison is represented by a scatter diagram where the regression line and 95% pointwise confidence bands are superimposed with the identity line (dashed line). The intercept and slope are reported with their 95% confidence intervals (CI). Forest plots showing estimated concordance correlation coefficients with 95% CI between H&E and IHC stains are depicted for global, intratumoral, and stromal TIL (g).

a relatively high level of imprecision between measurements; however, there is no major constant (intercept) or proportional (slope) drift between the two methods shown in the Passing–Bablok regression analysis (Figures 2d–f) and the estimated concordance correlation coefficient values confirmed there is a good correlation between the two methods, which is better for global and stromal compared to intratumoral TIL scores (Figure 2g).

The data from Cohort A were used to examine inter-observer agreement. Bland–Altman and Passing–Bablok analyses demonstrate that inter-observer bias is stronger when scoring global and stromal TIL, while this trend is not apparent for intratumoral TIL (Figures 3a–f; Supplementary Figure S1). In terms of precision, immunohistochemistry produces more consistent results between pathologists, reflected by the higher estimated

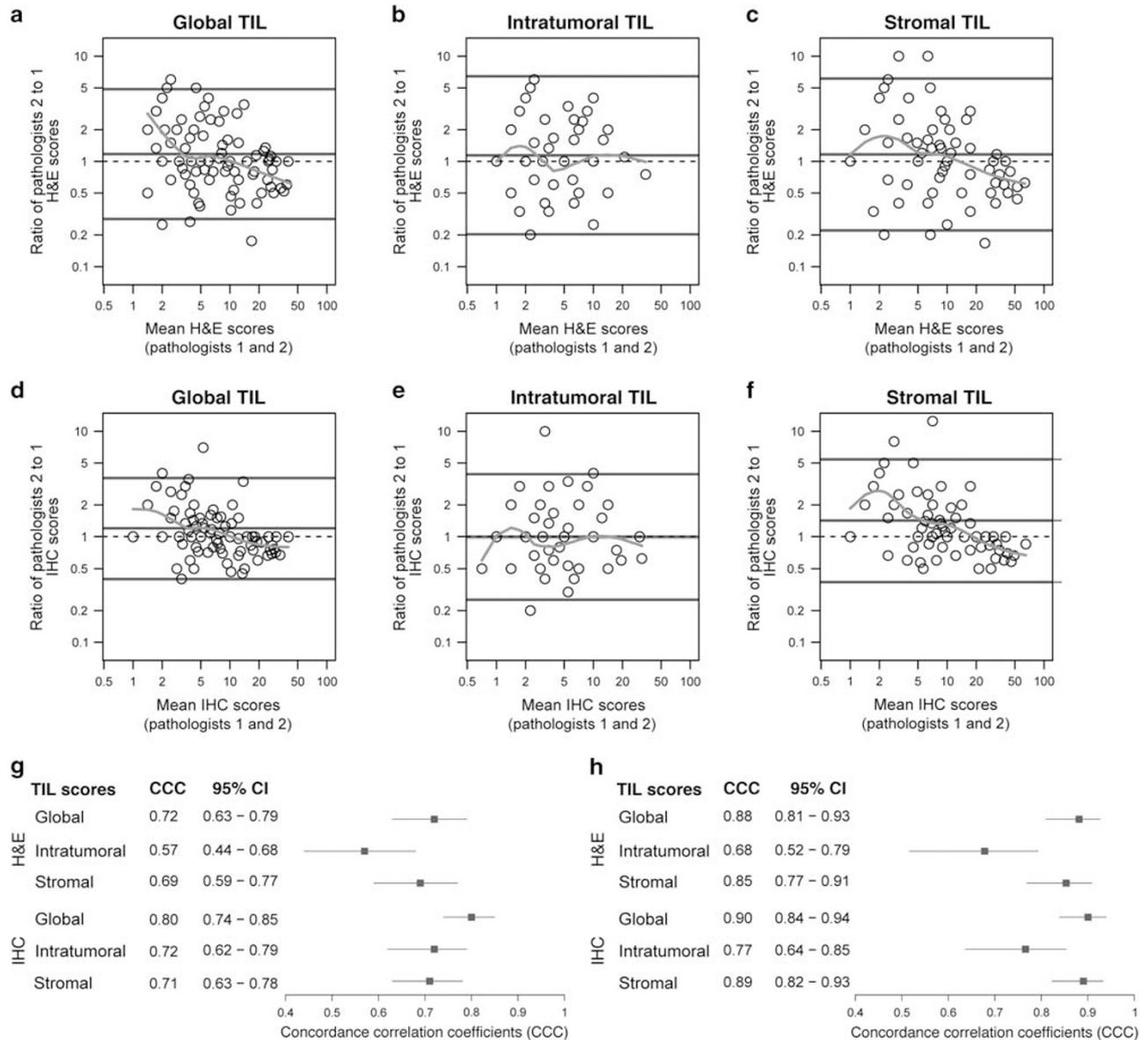


Figure 3 Inter-observer and Intra-observer agreement. Bland–Altman plots showing inter-observer agreement between two pathologists using tumor-infiltrating lymphocytes (TIL) scores from hematoxylin-eosin (H&E) (a–c) or immunohistochemical (IHC) (d–f) stains for global (a and d), intratumoral (b and e), and stromal TIL (c and f). For each comparison, the geometric mean of the two scores (pathologist 1 and 2) from the same histochemically or immunohistochemically stained tumor section (x axis) is plotted against the ratio between pathologist 2 to pathologist 1 scores (y axis). The horizontal lines represent the overall geometric mean of the ratios (center line) and the approximate 95% limits of agreement. The Loess fitted curve is also shown. Forest plots representing estimated concordance correlation coefficients with 95% confidence interval (CI) for each pairing between pathologists (inter-observer agreement in g) and between the first and second reading (intra-observer agreement in h).

concordance correlation coefficient values for all immunohistochemical measurements compared to their histochemical counterparts (Figure 3g). Scoring intratumoral TIL on hematoxylin-eosin-stained tissues is relatively inaccurate with a concordance correlation coefficient of only 0.57 (95% confidence interval (CI): 0.44–0.68).

This data were also used to explore intra-observer agreement by having the same pathologists re-score TIL on 50 randomly selected histochemically and immunohistochemically stained sections from

Cohort A after a 3-month interval. Regression analyses, together with the intra-observer concordance correlation coefficient values show that TIL evaluation using immunohistochemical stains has a slightly higher degree of reliability than their histochemical counterparts (Figure 3h; Supplementary Figures S2–S5). Importantly, these data show that scoring global TIL is more accurate than intratumoral or stromal TIL both for inter-observer and intra-observer variability with both staining methods.

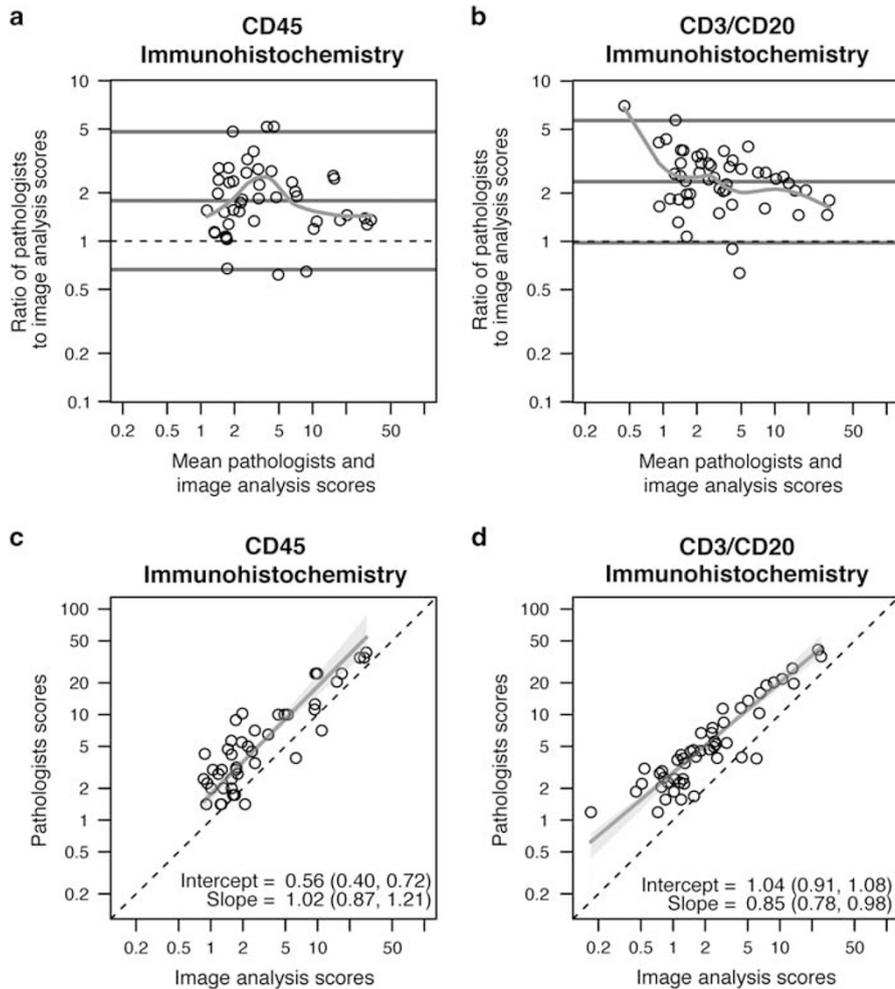


Figure 4 Pathological assessment *versus* digital image analysis. Bland–Altman plots showing the agreement between pathologist’s scores and data from digital image analysis for CD45 (a) and CD3/CD20 (b) immunohistochemical stains. For each comparison, the geometric mean value of the two scores (CD45⁺ or CD3⁺CD20⁺) (x axis) is plotted against the ratio between the same two scores (y axis). The horizontal lines represent the overall geometric mean of the ratios (center line) and the approximate 95% limits of agreement. The Loess fitted curve is also shown. Passing–Bablok regression of log-transformed tumor-infiltrating lymphocytes (TIL) scores assessed by digital image analysis (x axis) or by two pathologists (y axis) using CD45 (c) or CD3/CD20 immunohistochemistry (d). Each comparison is represented by a scatter diagram where the regression line and 95% pointwise confidence bands are superimposed with the identity line (dashed line). The intercept and slope are reported with their 95% confidence intervals (CI).

Accuracy of Identifying Tertiary Lymphoid Structures on Hematoxylin-Eosin *versus* Immunohistochemically Stained Tumor Tissues

We used tumors from Cohort A to compare the accuracy of tertiary lymphoid structure identification on sections stained with hematoxylin-eosin or immunohistochemistry. Our results show that tertiary lymphoid structure identification is underestimated on hematoxylin-eosin-stained tissues as the number of tertiary lymphoid structure-positive tumors and the global tertiary lymphoid structure count increases when scored using their immunohistochemically stained counterparts (Supplementary Table S5). These data show relatively weak agreement between the two staining methods for scoring tertiary lymphoid structures (inter-method concordance correlation coefficient: 0.36, 95% CI: 0.24–0.47).

An analysis of inter-observer and intra-observer agreement for tertiary lymphoid structures scores clearly shows that sections stained with hematoxylin-eosin are inaccurate (Supplementary Table S6). Alternatively, scoring immunohistochemically stained tumor sections produces consistent results between pathologists and duplicate readings from individual pathologists (Supplementary Table S6).

Pathological Assessment *versus* Digital Pathology

Digital pathology was used as a tool to quantify global TIL on scanned CD3/CD20 and CD45 immunohistochemically stained tissues (Cohort A). The resulting data were then compared with the mean pathologist’s scores. Regression analysis clearly shows a constant bias between the two approaches with digital image

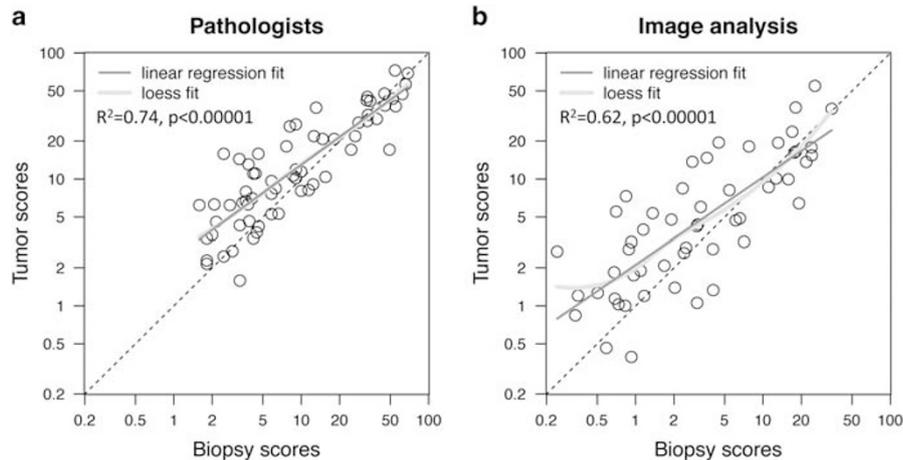


Figure 5 Diagnostic core biopsies and tumors. (a and b) show the linear regression line (dark gray) and loess fitted curve (light gray) of log-transformed tumor-infiltrating lymphocytes (TIL) scores between diagnostic biopsies (x axis) and their matched full tumor sections from surgical specimens (y axis), for pathological assessment (a) or digital image analysis (b). The identity line (dashed line) is also shown.

analysis reporting lower percentages than pathologists, both for CD45 and CD3/CD20 immunohistochemical staining (Figure 4). We further compared both stains and detected a good correlation when comparing pathologist's scores with those obtained from image analysis (Supplementary Figure S6).

Accuracy of Core Biopsies as a Reflection of the Primary Tumor

We compared mean global TIL scores (three pathologists) from 65 immunohistochemically stained (CD45), matched pairs of diagnostic core biopsies, and primary tumor sections (surgical specimens from untreated patients). This analysis investigated the accuracy of TIL and tertiary lymphoid structure scores (Cohort B), and showed a marked linear dependence between the two measurements (Figure 5). These data reveal that TIL scores from immunohistochemically stained biopsies predict immune infiltration in the tumor at surgery with a slight proportional underestimation effect that disappears as the TIL scores increase. Tertiary lymphoid structure scores from core biopsies were next investigated for their ability to predict scores from surgical specimens. The odds ratio estimate from the logistic model was 10.0 (95% CI: 2.9–35.2, $P < 0.00001$) with a Nagelkerke coefficient of determination R^2 of 0.26, indicating that tertiary lymphoid structure scoring of biopsies does not accurately reflect the tumor at surgery. TIL scoring for matched biopsy tumor pairs was also measured using image analysis. These analyses reiterate the pathologist's scores showing that TIL in core biopsies predict tumor immune infiltration at surgery with the same effect (Figure 5b).

Discussion

This study draws several conclusions based on our comparative analysis of conventional hematoxylin-eosin

versus immunohistochemical stains for quantifying global, stromal, and intratumoral TIL in primary breast cancer. TIL assessment on sections stained with hematoxylin-eosin is relatively well-correlated with immunohistochemistry and without systematic bias. This correlation is better for global and stromal compared with intratumoral TIL. A comparison of inter- and intra-observer agreement between pathologists using the two different stains reveals higher concordance for immunohistochemistry. However, hematoxylin-eosin staining is suitably reproducible and accurate except for scoring intratumoral TIL. Most of the current studies investigating TIL as a breast cancer prognostic biomarker scored stromal TIL as a continuous variable on sections stained with hematoxylin-eosin.^{4,16–18} A big advantage of this stain is that the sections prepared for routine clinical analysis are readily available; however, distinguishing TIL and apoptotic tumor cells on these stained tissues can be ambiguous. A recent study of inter-observer variability for assessment of breast cancer TIL using hematoxylin-eosin-stained sections found that tumor cell necrosis and apoptosis were specific factors causing discrepancies between pathologists.¹⁹ In our study, the weakest inter-observer agreement was detected for intratumoral TIL scores from hematoxylin-eosin-stained tissues, which supports recommendations from the International Working Group to avoid scoring intratumoral TIL (ie, global TIL or stromal TIL) on these slides.⁴ A recently published ring study organized by this group show that using a software-guided image evaluation approach for scoring hematoxylin-eosin-stained tissues improves reproducibility between pathologists.⁵ Using immunohistochemically stained sections for specific immune markers has the advantage of counteracting these limitations and making speeding up the evaluation by pathologists but require additional tissue sections and staining. The dual CD3/CD20 immunohistochemical stain used here is highly

advantageous because TIL and tertiary lymphoid structures are easily scored on the same section. The highest concordance values for inter- and intra-observer comparisons were obtained for global TIL scores, suggesting that this measurement could be used instead of stromal TIL, the current recommendation.⁴ All three TIL measures were well-correlated both in our study and previous publications from others;^{2,16} however, biological and clinical evidence that identifies the most meaningful parameter for breast cancer is still lacking. A recent study of triple-negative breast cancer identified three subgroups based on the immune infiltrate's location, finding that intratumoral infiltration was associated with a better prognosis.²⁰

Tertiary lymphoid structure detection was also examined in this study because these mini lymph nodes forming at the tumor site have been described in several types of solid tumors (reviewed in Dieu-Nosjean *et al*⁶). Our previous work demonstrated that tertiary lymphoid structures are specifically associated with CD4⁺ Tfh cells and positively correlated with breast cancer survival.⁷ The data presented here show that tertiary lymphoid structure detection is underestimated on hematoxylin-eosin compared to immunohistochemically stained tissues and that inter- and intra-observer agreement is also clearly superior using immunohistochemistry. The potential clinical relevance of tertiary lymphoid structures at the tumor site requires further investigation into their functional contributions to anti-tumor immunity; however, our data clearly show that tertiary lymphoid structures should only be evaluated using immunohistochemical stains specific for immune cells.

Examination of whether immune infiltrate quantification in core biopsies accurately reflects surgical specimens is critically relevant for studies that evaluate TIL, particularly for situations where tissues are limited, such as pre-therapeutic biopsies or tissue microarrays. Our data show a good correlation between lymphocyte infiltration scores from core biopsies and tumors at surgery. This suggests that in the neoadjuvant setting, TIL assessed on core biopsies can be used as a surrogate for the primary tumor. Unfortunately, tertiary lymphoid structure evaluation on core biopsies does not accurately reflect all tumors, with 50% of those found to contain these structures in the surgical specimen negative on the core biopsy. Tertiary lymphoid structures are principally resident in the stroma, with their position often non-concentric and thus biopsy sampling would only randomly bisect these tumor regions.

We also investigated the potential for using image analysis and digital pathology to score TIL. These technologies are rapidly growing with high-throughput, high-resolution microscopic scanners and powerful image analysis software currently under development. Digital pathology has recently shown real promise in translational trials.^{21,22} Our approach coupled CD45 immunohistochemically stained

leukocytes with Visiomorph software to measure positivity in defined areas of invasive carcinoma. Our analyses show a constant bias between automated digital and pathological assessment, indicating that pathologists may overestimate TIL as a percentage of the tissue area because they are influenced by the density of TIL. Image analysis was limited to global lymphocytic infiltration within the 'invasive carcinoma' region manually defined by a pathologist. Other studies have shown that tumor and stromal areas can be discriminated digitally even on sections stained with hematoxylin-eosin.^{23,24}

This study is, to the best of our knowledge, the first to formally compare hematoxylin-eosin and immunohistochemical staining, differential scoring (global, stromal, and intratumoral TIL) and different types of tumor specimens (core biopsies *versus* surgical specimens) for quantification of breast cancer TIL. We show that scoring on immunohistochemically stained sections delivers higher inter- and intra-observer agreement compared to hematoxylin-eosin stains. These data further demonstrate that quantification of global TIL on histochemically or immunohistochemically stained sections is best, but the former should be avoided for quantification of intratumoral TIL. The added value of immunohistochemistry for the predictive and prognostic significance of TIL remains to be demonstrated. Our results also reveal that tertiary lymphoid structures should only be scored using full-face immunohistochemically stained sections surgical specimens to ensure accuracy, reproducibility, and distinguish them from lymphoid aggregates. Finally, image analysis with digital pathology is potentially interesting for efficient and reproducible TIL quantification; however, further work is needed before it can be used in routine pathology.

Acknowledgments

This work was supported by grants from the Belgian Fund for Scientific Research (FRS-FNRS 3.4633.10F), Les Amis de l'Institut Bordet, Operation Télévie (7.4647.11F), Cancer Plan of Belgium (KPC_29_048), MEDIC Foundation, the Italian Association for Leukemia and Lymphoma (AIL), and the Italian Association for Cancer Research (AIRC). The CMMI is supported by the European Regional Development Fund and the Walloon Region. LB is a fellow and CS is a research director of the FRS-FNRS.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

- 1 Dushyanthen S, Beavis PA, Savas P, *et al*. Relevance of tumor-infiltrating lymphocytes in breast cancer. *BMC Med* 2015;13:202.

- 2 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.
- 3 Denkert C, von Minckwitz G, Darb-Esfahani S *et al*. Evaluation of Tumor-Infiltrating Lymphocytes (TILs) as predictive and prognostic biomarker in different subtypes of breast cancer treated with neoadjuvant therapy—a metaanalysis of 3771 patients. *Cancer Res* 2017;77:S1–09.
- 4 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* 2015;26:259–271.
- 5 Denkert C, Wienert S, Poterie A, *et al*. Standardized evaluation of tumor-infiltrating lymphocytes in breast cancer: results of the ring studies of the international immuno-oncology biomarker working group. *Mod Pathol* 2016;29:1155–1164.
- 6 Dieu-Nosjean M-C, Goc J, Giraldo NA, *et al*. Tertiary lymphoid structures in cancer and beyond. *Trends Immunol* 2014;35:571–580.
- 7 Gu-Trantien C, Loi S, Garaud S, *et al*. CD4⁺ follicular helper T cell infiltration predicts breast cancer survival. *J Clin Invest* 2013;123:2873–2892.
- 8 Lee HJ, Kim JY, Park IA, *et al*. Prognostic significance of tumor-infiltrating lymphocytes and the tertiary lymphoid structures in HER2-positive breast cancer treated with adjuvant trastuzumab. *Am J Clin Pathol* 2015;144:278–288.
- 9 Figenschau SL, Fismen S, Fenton KA, *et al*. Tertiary lymphoid structures are associated with higher tumor grade in primary operable breast cancer patients. *BMC Cancer* 2015;15:101.
- 10 Lee HJ, Park IA, Song IH, *et al*. Tertiary lymphoid structures: prognostic significance and relationship with tumour-infiltrating lymphocytes in triple-negative breast cancer. *J Clin Pathol* 2016;69:422–430.
- 11 Buisseret L, Garaud S, de Wind A, *et al*. Tumor-infiltrating lymphocyte composition, organization and PD-1/ PD-L1 expression are linked in breast cancer. *Oncoimmunology* 2017;6:e1257452.
- 12 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310.
- 13 Bablok PH. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem* 1983;21:709–720.
- 14 Lin LH, Hedayat A, Wu W. *Statistical Tools for Measuring Agreement*. Springer: New-York, 2012.
- 15 R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0, Available at: <http://www.R-project.org>.
- 16 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.
- 17 Loi S, Michiels S, Salgado R, *et al*. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. *Ann Oncol* 2014;25:1544–1550.
- 18 Adams S, Gray RJ, Demaria S, *et al*. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. *J Clin Oncol* 2014;32:2959–2966.
- 19 Swisher SK, Wu Y, Castaneda CA, *et al*. Interobserver agreement between pathologists assessing tumor-infiltrating lymphocytes (TILs) in breast cancer using methodology proposed by the International TILs Working Group. *Ann Surg Oncol* 2016;23:2242–2248.
- 20 Gruosso T, Gigoux M, Bertos N, *et al*. Abstract P4-03-08: mechanisms of CD8+ T cell immunosuppression in triple negative breast cancer. *Cancer Res* 2017;77:P4-03-08; Abstract.
- 21 Ali HR, Dariush A, Provenzano E, *et al*. Computational pathology of pre-treatment biopsies identifies lymphocyte density as a predictor of response to neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res* 2016;18:21.
- 22 Laurinavicius A, Plancoulaine B, Laurinaviciene A, *et al*. A methodology to ensure and improve accuracy of Ki67 labelling index estimation by automated digital image analysis in breast cancer tissue. *Breast Cancer Res* 2014;16:R35.
- 23 Beck AH, Sangoi AR, Leung S, *et al*. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci Transl Med* 2011;3:108ra113.
- 24 Yuan Y, Failmezger H, Rueda OM, *et al*. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med* 2012;4:157ra143.

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