

<https://doi.org/10.1038/s41523-025-00730-1>

Machine learning-based spatial characterization of tumor-immune microenvironment in the EORTC 10994/BIG 1-00 early breast cancer trial



Ioannis Zerdas^{1,2,15}✉, Alexios Matikas^{1,3,15}, Artur Mezheyeuski^{4,5}, Georgios Manikis^{1,6}, Balazs Acs^{1,7}, Hemming Johansson¹, Ceren Boyaci^{1,7}, Caroline Boman^{1,3}, Coralie Poncet⁸, Michail Ignatiadis⁹, Yalai Bai^{10,11}, David L. Rimm^{10,11}, David Cameron¹², Hervé Bonnefoi¹³, Jonas Bergh^{1,3}, Gaetan MacGrogan¹⁴ & Theodoros Foukakis^{1,3}

Breast cancer (BC) represents a heterogeneous ecosystem and elucidation of tumor microenvironment components remains essential. Our study aimed to depict the composition and prognostic correlates of immune infiltrate in early BC, at a multiplex and spatial resolution. Pretreatment tumor biopsies from patients enrolled in the EORTC 10994/BIG 1-00 randomized phase III neoadjuvant trial (NCT00017095) were used; the CNN11 classifier for H&E-based digital TILs (dTILs) quantification and multiplex immunofluorescence were applied, coupled with machine learning (ML)-based spatial features. dTILs were higher in the triple-negative (TN) subtype, and associated with pathological complete response (pCR) in the whole cohort. Total CD4⁺ and intra-tumoral CD8⁺ T-cells expression was associated with pCR. Higher immune-tumor cell colocalization was observed in TN tumors of patients achieving pCR. Immune cell subsets were enriched in *TP53*-mutated tumors. Our results indicate the feasibility of ML-based algorithms for immune infiltrate characterization and the prognostic implications of its abundance and tumor-host interactions.

Breast cancer (BC) represents a clinically and biologically heterogeneous disease ecosystem. Further characterization of the tumor microenvironment (TME), its components and molecular drivers could elucidate the complexity of tumor-host interactions and provide rationale for biomarker development. One of the principal TME components, tumor-infiltrating lymphocytes (TILs), are both prognostic for outcomes and predictive for response to chemotherapy^{1–5}.

Our understanding of the immune cell composition and spatial interactions within the TME is evolving. Towards this end, artificial-intelligence (AI) and machine-learning (ML) methods—which have revolutionized digital pathology and diagnostics—could represent valuable tools for studying the TME^{6,7}. Indeed, recent studies have reported on the performance, prognostic implications, advantages, and challenges of digital-assisted scoring of TILs and multiplex immunofluorescence methods for

¹Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden. ²Theme Cancer, Karolinska Comprehensive Cancer Center and University Hospital, Stockholm, Sweden. ³Breast Center, Theme Cancer, Karolinska Comprehensive Cancer Center and University Hospital, Stockholm, Sweden. ⁴Department of Immunology, Genetics, and Pathology, Uppsala University, Uppsala, Sweden. ⁵Molecular Oncology Group, Vall d'Hebron Institute of Oncology, Barcelona, Spain. ⁶Computational BioMedicine Laboratory (CBML), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Greece. ⁷Department of Clinical Pathology and Cancer Diagnostics, Karolinska University Hospital, Stockholm, Sweden. ⁸European Organisation for Research and Treatment of Cancer Headquarters, Brussels, Belgium. ⁹Department of Medical Oncology, Institut Jules Bordet and L'Université Libre de Bruxelles (U.L.B), Brussels; Academic Trials Promoting Team (ATPT), Institut Jules Bordet, Brussels, Belgium. ¹⁰Department of Pathology, Yale School of Medicine, New Haven, CT, USA. ¹¹Yale Cancer Center, Yale School of Medicine, New Haven, CT, USA. ¹²Edinburgh University Cancer Centre, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, UK. ¹³Department of Medical Oncology, Institut Bergonié Unicancer, INSERM U1312, Université de Bordeaux, Bordeaux, France. ¹⁴Department of Biopathology, Institut Bergonié Unicancer, INSERM U1312, Bordeaux, France. ¹⁵These authors contributed equally: Ioannis Zerdas, Alexios Matikas. ✉e-mail: ioannis.zerdas@ki.se

TME characterization in BC^{8,9}. However, the spatial relationships between the various components of immune infiltrate and tumor cells still warrant further investigation.

The aims of this study were i) to comprehensively and objectively characterize the composition of tumor-immune microenvironment landscape at a spatial level, ii) explore its distribution according to *TP53* mutational status and iii) evaluate relevant prognostic implications by using automated, ML-based and multispectral spatial imaging approaches in the context of a large randomized phase III trial at the neoadjuvant setting¹⁰.

Results

Patient characteristics

Out of the 1856 patients initially enrolled in the trial, 697 had available tissue and were eligible for this study. Upon staining and tissue quality assessment, 587 and 478 patients had available digital TIL (dTIL) and multiplex immunofluorescence data, respectively and included in the final analysis. The patient disposition presenting the eligibility criteria and data availability is shown in Fig. 1. The demographic/clinicopathologic characteristics of the different patient populations according to the methods used are presented in Supplementary Table 1. At the time of the latest data cutoff (9th October 2018), median follow-up was 11.4 years (interquartile range = 10.16–12.41 years).

Digital image analysis-based enumeration of TILs and correlation with outcomes

Using the previously validated CNN11 image-analysis algorithm on hematoxylin & eosin (H&E)-stained tissue microarrays (TMA), we calculated different digital TILs (dTILs) metrics based on the various cell annotations (Fig. 2A). Among the dTILs variables, the median

expression of easTILs was 11% (range: 0.01–76.7) (Fig. 2B, Supplementary Table 2). The different dTILs metrics were overall strongly and statistically significantly correlated with each other (Fig. 2C). DTILs abundance was generally higher in the triple-negative subtype compared to the HER2+ and HR+/HER2- (Fig. 2D, Supplementary Table 2).

Regarding associations with patient outcomes, easTILs were associated with pCR rates both in univariate ($OR_{unadjusted} = 1.73$, 95% CI 1.10–2.69, $p = 0.016$) and multivariable ($OR_{adjusted} = 1.59$, 95% CI 1.00–2.54, $p = 0.05$) analysis (Fig. 2E, F). There was a significant interaction between easTIL and IHC-based subtypes for pCR ($p = 0.045$), with the association mainly observed in the HR+/HER2- subtype. However, no significant association with PFS was noted (Supplementary Fig. 1).

Immune cell landscape composition at a multiplex resolution and correlation with outcomes

To further characterize the specific composition of the immune infiltrate, we performed multiplex immunofluorescence in the available TMA using an immune-related antibody panel. We evaluated the abundance (i.e., defined as normalized cell densities = number of marker positive cells/mm²) and localization (total, tumor and stroma area) of the different immune cell subpopulations based on i) the expression of single markers (CD4+ for T-helpers, CD8+ for cytotoxic T-cells, FoxP3+ for T-regulatory cells, CD68+ for macrophages) and ii) the respective co-expression of PD-L1 and PD-1 checkpoint markers on these cells (Fig. 3A, Supplementary Table 3). Upon the application of an established workflow and tissue/staining quality control, immune cell quantification could be performed in a total of 478 patients ($n = 468$ for tumor area analysis; 10 patients had only stroma area for analysis).

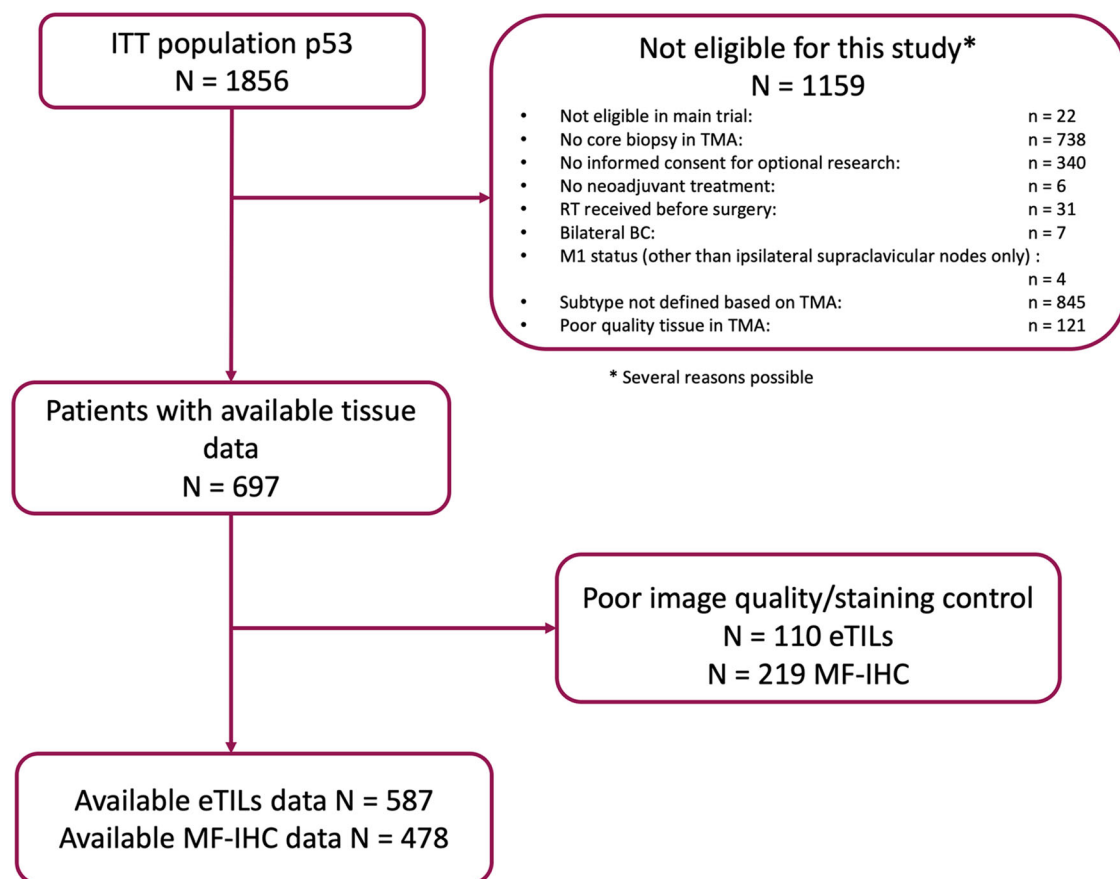
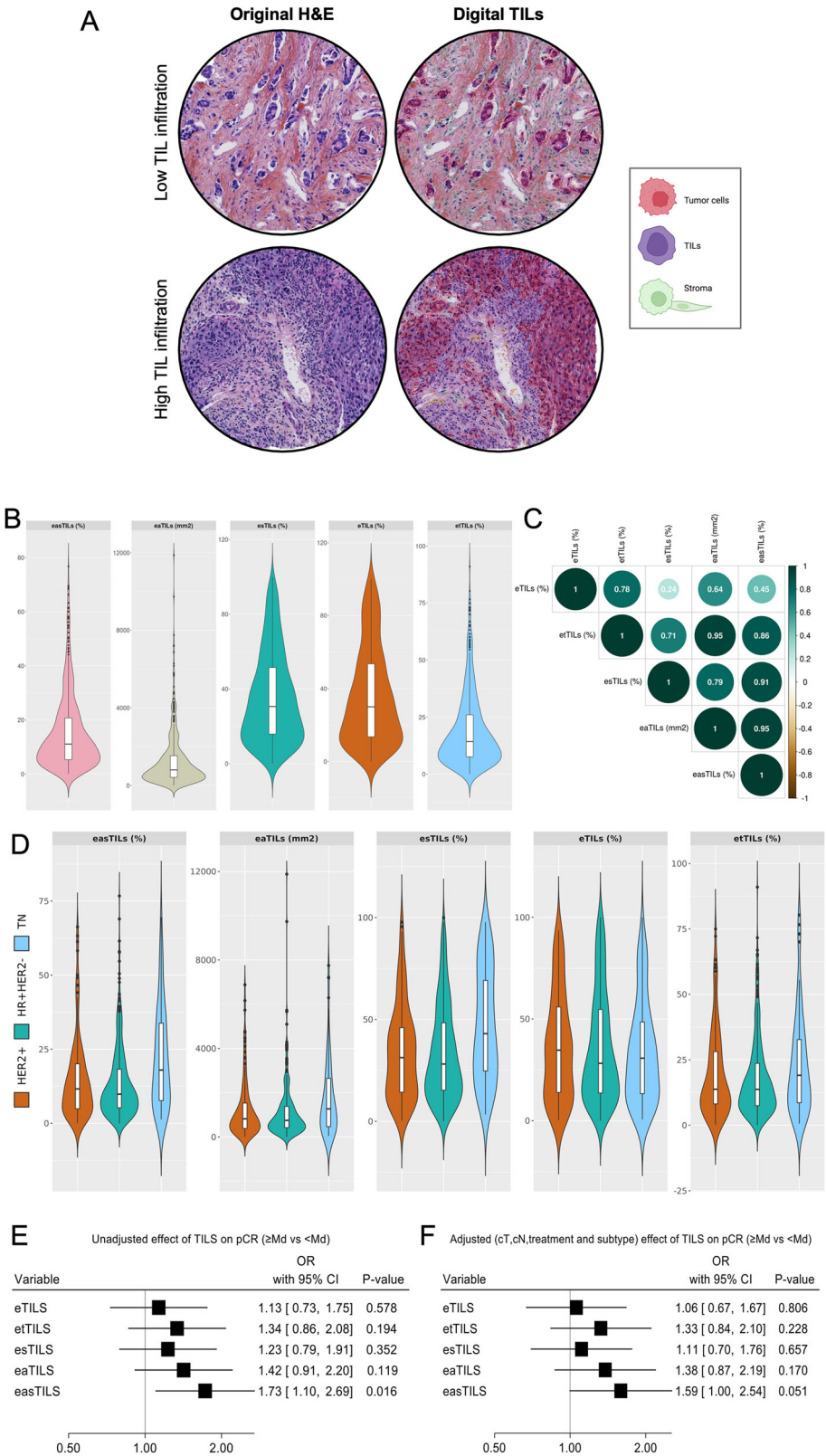


Fig. 1 | Patient disposition and data availability of the study. Flowchart depicting the data availability for the translational analyses in the present study; ITT intention-to-treat, eTILs digitally-assessed tumor-infiltrated lymphocytes, MF-IHC multiplex fluorescent immunohistochemistry, TMA tissue microarrays.

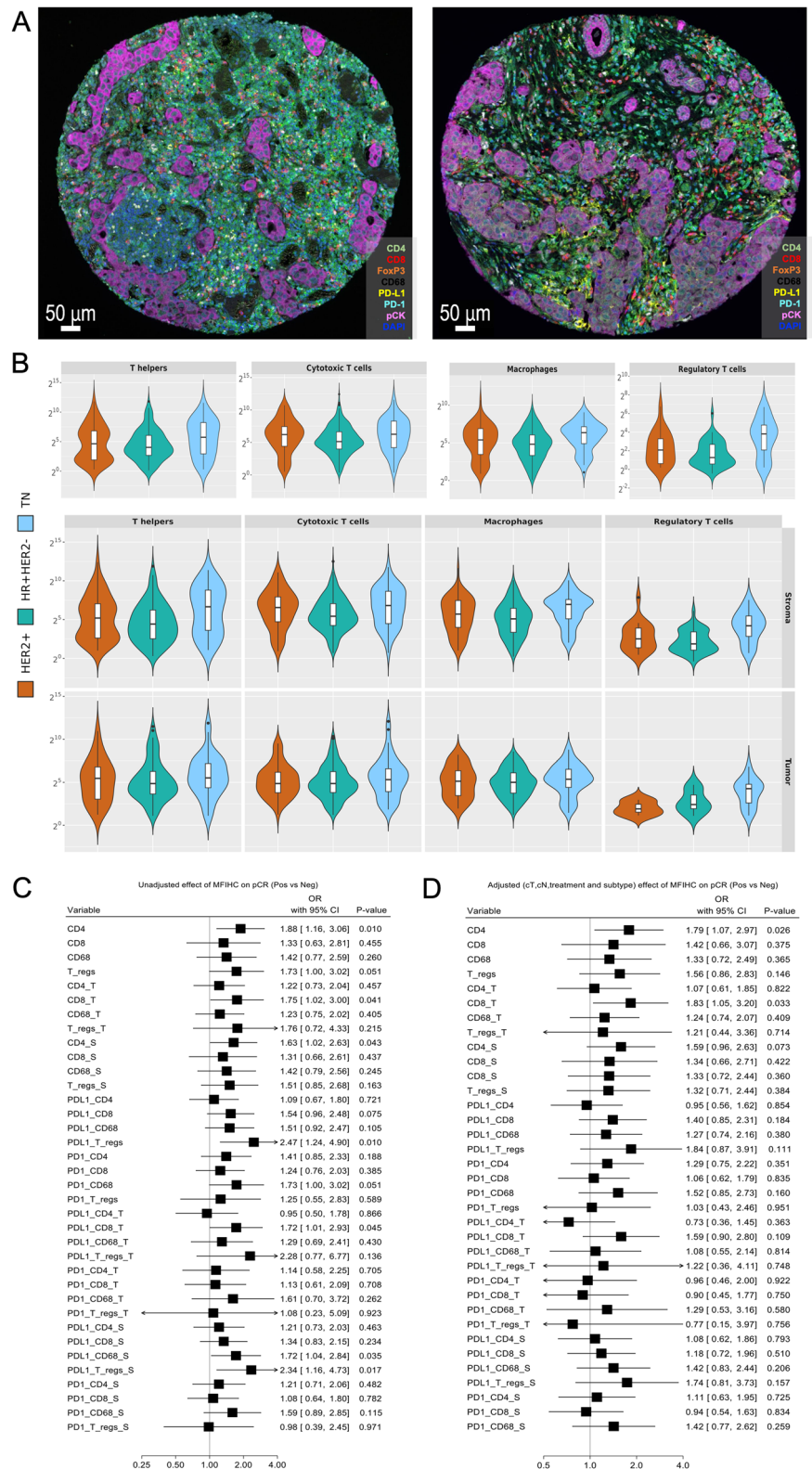
Fig. 2 | Digital TILs evaluation using a machine-learning algorithm in archival H&E-stained FFPE breast cancer tissue microarrays. **A** Representative TMA images of digital TILs enumeration in patients with low and high immune infiltration, created with BioRender.com; **B** Distribution of the digital TILs variables in the whole population (*n* = 587); **C** Correlation matrix for the different dTILs variables, Spearman's rank correlation coefficient; **D** Distribution of dTILs variables within IHC-based subtypes; Forest plots on prognostic effect of dTILs on pCR in the univariate (**E**) and multivariable (**F**) logistic regression model (adjusted for tumor size, nodal status, treatment and stratified by subtype) in the whole cohort.



In the whole patient population, CD8+ cytotoxic T-cells were the most abundant cell subset in the total, intra-tumoral and stromal areas. Both PD-1 and PD-L1 checkpoint expression was highest in CD4+ T-cells (T-helpers), followed by cytotoxic T-cells (CD8+) in all tissue compartments (Supplementary Table 3). Within IHC-based subtypes, the mean cell densities of all immune cell subpopulations (i.e., T-cell and macrophages) and

also the respective co-expression of the PD-1/PD-L1 markers were significantly enriched in the triple-negative compared to the other subtypes, regardless of tissue localization (Supplementary Table 4, Fig. 3B). Furthermore, the different immune cell subsets derived from the multiplex immunofluorescence analysis were correlated both with each other and with digital TILs (Supplementary Fig. 2).

Fig. 3 | Characterization of the immune infiltrate at a multiplex resolution. **A** Representative stained images with the multiplex immunofluorescence antibody panel; **B** Distribution of cell densities/immune cell subsets for total area (upper panel), stroma (middle panel) and tumor (lower panel) compartments within IHC-based subtypes. All values are log₂-transformed; Forest plots on prognostic effect of multiplex immunofluorescence immune cell subpopulations on pCR in the univariate (C) and multivariable (D) logistic regression model (adjusted for tumor size, nodal status, treatment and stratified by subtype) in the whole cohort.



Among the different cell subpopulations, CD4 + T-helpers in the total area (OR_{adj} = 1.79, 95% CI 1.07–2.97, p = 0.026) and intra-tumoral CD8+ cytotoxic T-cells (OR_{adj} = 1.83, 95% CI 1.05–3.20, p = 0.033) were associated with improved pCR (Fig. 3C, D). No significant association with PFS was observed for any marker (Supplementary Fig. 3).

Spatial composition of immune infiltrate, cell-to-cell interactions, and correlation with pathologic complete response in the triple-negative subtype

To further explore the spatial distribution and complexity of immune-tumor host interactions, we calculated two different ML-based metrics using the multiplex immunofluorescence cell data, namely i) the

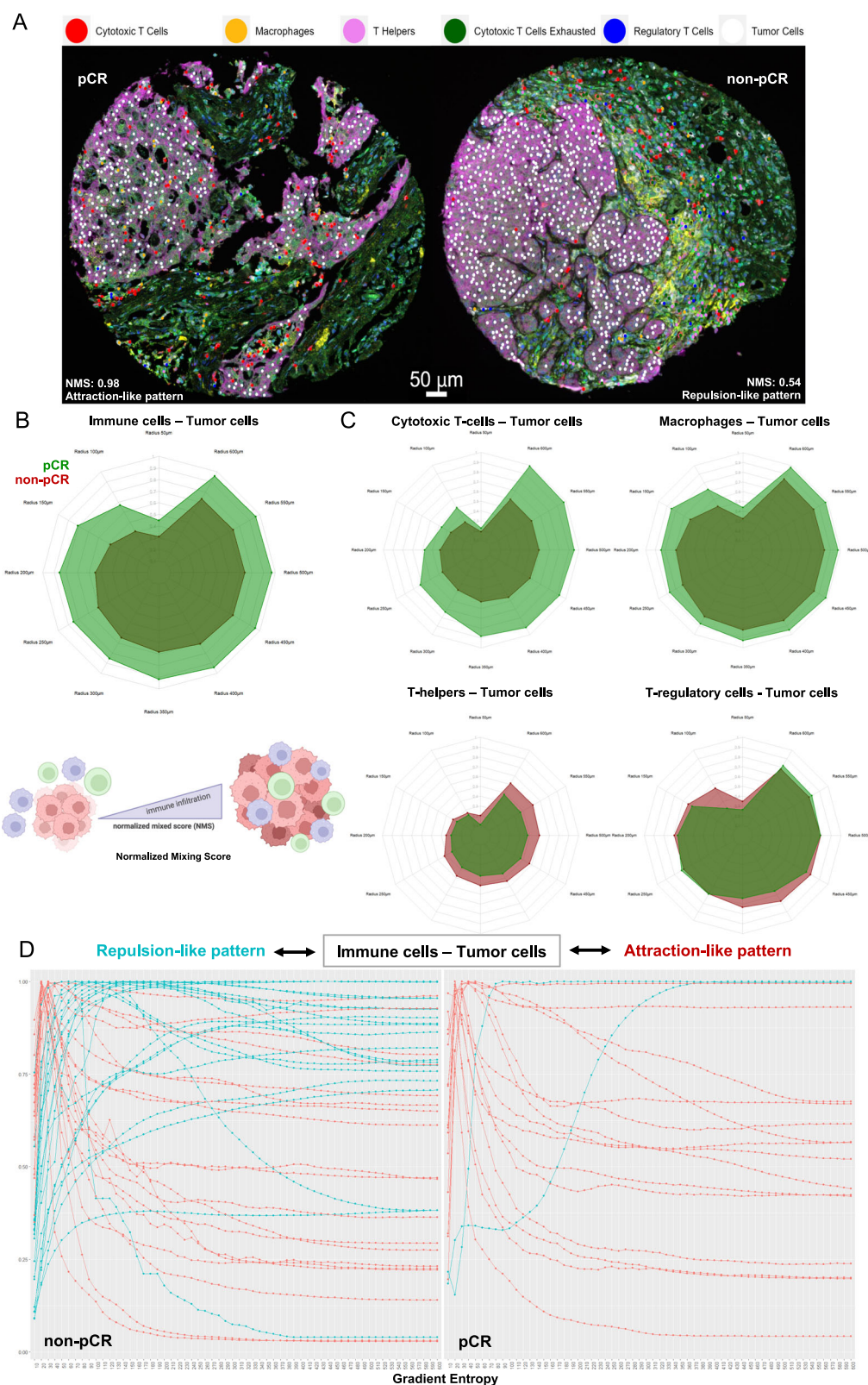


Fig. 4 | Spatial characterization of the tumor-immune cell interactions (using ML-based metrics) in the triple-negative subtype in patients that achieved pCR versus non-pCR. A Representative images depicting the spatial phenotyping and interactions among different immune cell subpopulations with tumor cells in a pCR (left) and a non-pCR (right) patient; Radar plots depicting normalized mixing score (NMS) values for different increasing radii for the interaction of overall immune-

tumor cells (B) and the specific immune cell subpopulations in patients achieving pCR (green) versus non-pCR (red) TN (C). D Entropy gradient slope depicting attraction-like (red lines) and repulsion-like (blue lines) patterns for the interaction between immune and tumor cells in non-pCR (left panel) versus pCR (right panel) patients. Figure 3B was created with *BioRender.com*.

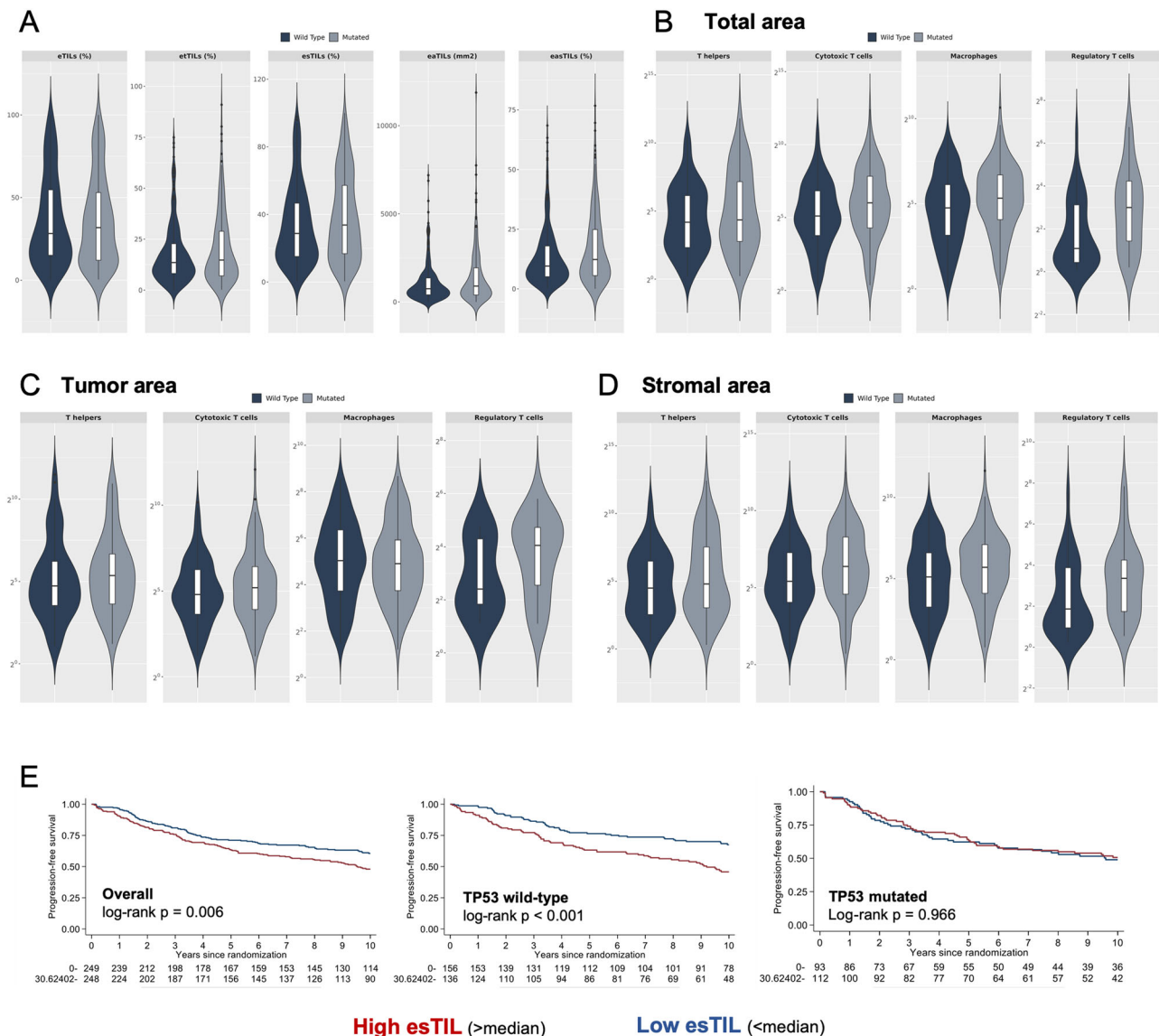


Fig. 5 | Composition of immune landscape according to *TP53* mutational status and prognostic correlates. A Distribution of digital TILs variables in patients with *TP53*-mutated versus *TP53* wild-type tumors; Expression of immune cell subpopulations according to *TP53* mutational status in the total (B), tumor (C) and

stromal (D) areas, cell densities values are log2-transformed; E Kaplan-Meier curves on the prognostic effect of the interaction between esTILs and *TP53* mutational status ($p = 0.028$). Higher TILs were associated with worse PFS in patients with *TP53* wild-type but not in the *TP53*-mutated tumors.

normalized mixing score (NMS) and ii) the entropy gradient. Given that the abundance of immune cells was highest in the triple-negative subtype, we focused on the description of their spatial patterns in this patient subgroup. Higher baseline NMS values for the immune-tumor cell interactions were observed in patients that achieved pCR compared to the non-pCR ones, regardless of radius (Fig. 4A, B), indicating higher degree of cell-cell colocalization for the pCR patients. This effect was maintained when we tested separately the interactions of tumor cells with cytotoxic T-cells or macrophages, but not for T-helpers or regulatory T-cells (Fig. 4C, Supplementary Table 5). Similarly, an “attraction-like” entropy gradient slope pattern for the interaction of immune with tumor cells was observed mostly in patients that achieved pCR compared to the non-pCR patients who had an enriched “repulsion-like pattern” (Chi-squared $p = 0.01$; Fig. 4D). This effect was observed for the interactions of tumor cells with most of the immune cell subpopulations, except for macrophages (Supplementary Fig. 4). Taken together, by using two different ML-based approaches we observed that closer spatial interaction between tumor and immune cells correlated with higher probability of pCR in TNBC subtype.

Association of *TP53* mutational status with immune infiltrate components

Given that *TP53* mutational status was available for most patients included in EORTC10994/BIG1-00, we aimed to investigate its association with the composition of the immune infiltrate. To this end, we performed correlative analyses in the patient subgroup with known *TP53* status and available dTILs ($n = 497$) or multiplex immunofluorescence ($n = 415$) data. Stroma dTILs metrics were higher in the *TP53*-mutated patient subgroup versus the *TP53* wild-type in the whole population, but not within IHC-subtypes—although numerically higher in the *TP53*-mutated triple-negative group (Fig. 5A, Supplementary Table 6). When evaluating the distribution of immune cell subpopulations per *TP53* mutational status, we observed that cell densities of cytotoxic T-cells (at total, tumor and stroma areas), intra-tumoral T-regulatory cells, and stromal macrophages were significantly higher in the *TP53*-mutated patient subgroup (Fig. 5B–D, Supplementary Table 7). Similarly, PD-L1 expression on T-cells and macrophages was significantly enriched in the *TP53*-mutated patients, regardless of their localization. PD-1 expressing cells were also enriched in the *TP53*-mutated patients, mainly the intra-tumoral T-helpers and stromal cytotoxic T-cells

and macrophages (Supplementary Table 7). Nevertheless, no difference was observed in NMS or entropy gradient score for interactions among tumor and immune cells according to *TP53* mutational status (Supplementary Figs. 5–6, Supplementary Table 8). Lastly, in an exploratory analysis, a significant interaction between esTIL and *TP53* mutational status was noted for PFS ($p = 0.028$), where higher esTIL abundance was prognostic for worse PFS only in patients with *TP53* wild-type tumors but not in mutant ones (Fig. 5E).

Discussion

Breast cancer is a multifaceted ecosystem comprising various cell phenotypes and clones that may provide biological insights and hints towards biomarker development. The best studied immune markers are the assessment of TILs and PD-L1, which although prognostic in the early BC setting, they lack at this time clinical utility for routine use^{1–5,11}. In addition, their static and visual evaluation cannot adequately grasp the complexity of immune infiltrate and the spatial distribution of tumor-host interactions. Toward this end, emerging powerful AI- and ML-based tools and spatial phenotypic quantitative immunohistochemistry/fluorescence technologies could facilitate comprehensive TME profiling^{7,9,12}.

In this study, we investigated the immune cell composition and spatial interactions within the BC TME using ML-based and digital pathology algorithms and demonstrated their feasibility in early BC using small tissue input. By using a digital-assisted classifier on H&E-stained TMA, we showed that easTILs were associated with increased pCR rates, yet we could not demonstrate any significant correlation with long-term survival outcomes. Although this algorithm has been previously tested and validated in adjuvant setting for TNBC patients and demonstrated the prognostic correlation of high dTILs with improved survival¹³, inconsistent results have been reported in the neoadjuvant setting. In previous studies, both in HER2-negative¹⁴ and HER2+ disease¹⁵, easTILs correlated with pCR but not with long-term outcomes. On the other hand, a recent report using another AI-based dTILs algorithm on luminal tumors demonstrated a correlation with worse outcome -however no patient received neoadjuvant treatment¹⁶. Of note, a previous study using gene expression data from the EORTC 10994/BIG 1-00 trial reported that high expression of an 8-gene TILs-related signature correlated with increased pCR rates in ER-negative patients¹⁷. The results of the present study indicate that dTILs were associated with increased pCR in the ER+/HER2- subtype, however low patient numbers of patients in the other subtypes preclude any informative conclusions. Thus, it is unclear if these results could be attributed to biological or technical aspects. Regarding the latter, digital TILs evaluation could be prone to various analytical setbacks including interobserver variability, artifacts, algorithm training and tissue recognition, thus leading to discrepancies with the visual assessment and further impeding clinical utility^{8,18}. The International Immuno-Oncology Biomarker Working Group has previously launched a set of recommendations on the computational TILs assessment, for overcoming the inherent limitations of visual TILs enumeration¹⁹. Furthermore, several versatile interpretable deep-learning algorithms for TILs evaluation have been recently reported i.e., MuTILs²⁰ or the Histomic Prognostic Signature (HiPS)²¹, focusing jointly on tissue region and individual cell nuclei segmentation (compared to our cell nuclei-based classification approach), and outperforming the pathologist-based visual assessment in predicting survival outcomes. Given the increasing number of AI-based models generated for TILs assessment, a recent study compared the validity of ten AI models on TILs scoring, confirming discrepancies and variability in terms of mostly analytical rather than prognostic performance in patients with TNBC²². Therefore, future perspectives towards potential clinical implementation include the refinement of existing or design of next-generation AI models (i.e., foundation models)²³, validation in prospective clinical studies and integration of computational image-based tools to other multi-omic data types for generation of comprehensive prognostic and predictive models of treatment response²⁴.

In order to further investigate the immune cell composition at a spatial resolution, we performed multiplex immunofluorescence for T-cell and

macrophage markers and for the expression of PD-1/PD-L1 checkpoints. Higher immune infiltration was seen in the TNBC subtype and similarly to dTILs analysis, CD8+ and CD4+ T-cells were correlated with increased pCR rates but not with long-term outcomes. Conflicting data have been reported regarding the association of single immune markers with pCR in early BC, highlighting thus the challenges of multiplex assays^{9,25}. However, these multidimensional methods could provide further directions beyond the immune cell abundance, hence a growing body of literature is focusing on the spatial interactions among the different cells within the TME⁶. In the ARTEMIS trial on neoadjuvant chemotherapy for TNBC, a closer spatial proximity of T cells to cancer cells was associated with increased pCR rates²⁵. Similar results on spatial immune infiltrate patterns have been recently reported in association with benefit to immunotherapy in TNBC^{26,27} and other tumor types^{28,29}, further indicating that close interaction of tumor with activated immune cells is a major predictor of treatment response. In line with the aforementioned reports, we demonstrated here a higher and closer proximity of immune cells to cancer cells in TNBC patients that achieved pCR compared to those that did not respond to neoadjuvant chemotherapy treatment. Therefore, coupling deep-learning features with the spatial TME distribution in common predictive models could lead to improved patient stratification and treatment outcomes³⁰.

Considering the previous reports on the link of the genetic determinants with antitumor immunity and that investigating the predictive role of *TP53* mutational status was a primary objective of the trial, we explored the potential effect of *TP53* mutation on TME. *TP53* has been shown to regulate the PD-1/PD-L1 axis in lung cancer studies, with IHC p53 positivity being associated with higher PD-L1 tumor cell expression^{31,32}—a finding which was not confirmed in the case of breast cancer and *TP53* mutations³³. On the other hand, it was recently shown that *TP53*-mutated BC exhibits significantly higher expression of chronic inflammation markers (i.e. interferon signaling, CD8 + T-cell infiltration) and immune checkpoints, indicating that *TP53*mut-targeting compounds could restore effective immune surveillance³⁴. In accordance with previous reports, we demonstrated an increased immune cell abundance (and respective checkpoint expression) in the *TP53*-mutated group of patients and especially in the TNBC subtype, possibly indicating that the presence of the mutation contributes to a more inflammatory phenotype. Moreover, in an exploratory analysis, we demonstrated a significant interaction between *TP53* mutational status and esTILs, with higher esTIL abundance being prognostic for worse PFS only in patients with *TP53* wild-type tumors. It is unclear if this prognostic effect could solely rely on the *TP53* biological effect or it reflects the association of increased TILs with more aggressive luminal B tumors^{4,35}.

Although the presented results stem from a randomized phase III trial with long-term follow-up and highlight the advantages of automated and ML-based approaches, our study has several limitations needed to be addressed. A major limitation is the use of TMA, with cores of very small diameter (0.6 mm) which could have affected cellularity, immune cell abundance, tumor heterogeneity aspects and spatial tissue morphology analyses compared to whole tissue section³⁶. The selection of representative tumor-rich areas for TMA construction and the inherent limitations of the used classifier could not distinguish between stromal and intratumoral TILs, thus affecting variable definitions and outcome correlations. The multicentric nature of the study could have also contributed to the tissue quality and TMA handling heterogeneity. Moreover, tumor phenotyping was performed on TMA and by using $\geq 1\%$ cut-off for ER/PR positivity^{37,38}; however, ER/PR cut-off $\geq 10\%$ is used for positivity in some centers, based on previous studies demonstrating that patients with tumors with ER 1–9% and HER2 negativity have mostly similar outcomes to those with ER $\leq 1\%$ ^{39–41}. Although we showed that immune infiltrate correlated with chemosensitivity in the whole cohort, the relatively low number of patients within subgroups precluded definite conclusions in the exploratory analyses. Moreover, due to the exploratory nature of this study, no formal adjustment for multiplicity was performed. Lastly, at the time the trial was conducted, neoadjuvant chemotherapy alone was standard treatment. Therefore, any prognostic or predictive value of the immune infiltrate for the

effect of newer neoadjuvant therapies (i.e., immunotherapy, anti-HER2) could not be evaluated.

In conclusion, our data indicate that ML-based algorithms could be used for characterizing the immune infiltrate in situ in early BC, which in turn contains prognostic information. Further studies are warranted in order to provide insights into TME cell phenotype features, interactions and underlying molecular determinants with the overall aim to improve prognostication and select candidates for future therapies.

Methods

Study design and patient cohort description

The present analysis is a translational substudy of the multicenter, international, randomized phase III EORTC 10994/BIG 1-00 clinical trial (ClinicalTrials.gov identifier NCT00017095). The study enrolled women (<71 years) with histologically-proven locally advanced, inflammatory, or large operable primary invasive BC who were candidates for neoadjuvant chemotherapy. Patients were randomly assigned in a 1:1 ratio to receive six cycles of taxane-based (docetaxel for three cycles followed by three cycles of epirubicin + docetaxel, T-ET; 928 patients) or non-taxane-based (5-fluorouracil, epirubicin, cyclophosphamide, FEC100 or tailored FEC; 928 patients) chemotherapy, prior to surgery. Post-surgery therapy included hormonal treatment and radiotherapy, according to the protocol guidelines. None of the patients received neoadjuvant trastuzumab or radiotherapy before surgery¹⁰. Patients who had signed informed consent, with available tissue microarray (TMA) and with IHC-based subtype defined per TMA were included in this translational study. The analyses are reported according to the Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) guidelines⁴² (Supplementary Table 9). The EORTC 10994/BIG 1-00 randomized phase III trial was registered at ClinicalTrials.gov (NCT00017095) and approved by the Ethics Committees in all participating centers (Supplementary Table 10). The proposed translational study has been reviewed and approved by the EORTC Translational Research Advisory Committee and Headquarters and the tissue material has been handled according to the signed material transfer agreement by the two organizations (Karolinska Institutet and EORTC). Only the tissue from the patients who have previously provided/signed informed consent for additional biologic research on their samples was included and analyzed. All patients have signed informed consent -prior registration- for inclusion in the trial and the assessment of p53 mutational status. The present study has been performed according to the Declaration of Helsinki, principles of Good Clinical Practice and was also approved by the Swedish Medical Product Agency and the Regional Ethical Committee in Stockholm (Dnr 01-387, Dnr 2005/472-32, Dnr 2005/738-32, Dnr 2008/897-32, Dnr 2013/954-32, Dnr 2018/540-32, Dnr 2021-05174, Dnr 2022-00657-02).

Tissue sample preparation and histopathology assessment

Formalin-fixed paraffin-embedded (FFPE) tumor tissue blocks from the initial diagnostic biopsies (before initiation of neoadjuvant chemotherapy) have been collected from all patients. Hematoxylin & eosin (H&E)-stained slides were reviewed by a certified pathologist (G.M.) and representative tumor-rich areas were selected for TMA construction, performed at Institut Bergonié, Bordeaux, France. For each patient, up to three cores (0.6 mm in diameter each) were used for the TMA, as previously described⁴³. TP53 mutational status analysis (assessed via a yeast functional assay and further validated through Sanger and next-generation sequencing) has been previously performed on baseline frozen tumor biopsies from all patients included in the study, with most (80%) having evaluable TP53 mutational status^{10,43,44}. Tumor IHC-based phenotyping and biomarker interpretation were performed on TMA as follows: Estrogen receptor (ER) and progesterone receptor (PR) were defined as positive in case of ≥1% tumor cell expression; high Ki67 expression was defined as ≥14% of positive cells; human Epidermal Growth Factor Receptor 2 (HER2) was evaluated according to the ASCO/CAP 2013 recommendations and was considered positive if immunohistochemistry (IHC) 3+ or if IHC 2+ and ≥6 HER2

gene copies using in situ hybridization⁴⁵. Tumors were classified into three subtypes as such: hormone receptor-positive/HER2-negative (HR+/HER2-negative), HER2-positive and triple negative (TN). Histological grade, type as well as pathological response (after neoadjuvant chemotherapy) were locally evaluated at each participating center.

Digital evaluation of tumor-infiltrating lymphocytes

One FFPE section (4 μm thickness) was obtained from each TMA tissue block and stained with hematoxylin and eosin (H&E). H&E slides were digitized using the Nanozoomer 2.0-HT (Hamamatsu Photonics K.K.) platform at 20× magnification. TILs were subsequently enumerated using the digitally-assisted, image-based automated scoring CNN11 algorithm^{13,45}. This previously trained and optimized classifier -which can be used with the open-source QuPath software- was used for detecting lymphocytes, tumor cells, stromal cells (e.g., fibroblasts) and “other” cells and for defining annotation and accumulative areas of each cell type (measured in mm²). Quality control of the tissue morphology (artifacts, necrosis) and cell segmentation was performed and subsequently five different TILs variables were derived from the different annotations as follows: i) eTILs (%) = 100 * (TILs/sum of tumor cells and TILs); ii) etTILs (%) = 100 * (TILs/All detected cells (i.e., tumor cells + fibroblasts + other cells)); iii) esTILs (%) = 100 * (TILs/Stromal cells (i.e. TILs + fibroblasts + other cells)); iv) eaTILs (mm²) = TILs/tumor region areas analyzed; v) easTILs (%) = 100 * (sum of TILs area (mm²)/stroma area (mm²)), mirroring the definition provided by the International TILs Working Group guidelines¹. Intratumoral TILs were included in the respective calculations of all the aforementioned metrics.

Multiplex immunofluorescence staining, image processing, and analysis

Staining. We performed multiplex immunofluorescence staining on FFPE TMA tumor tissue sections (thickness: 4 μm), using the OpalTM 7-color Solid Tumor Immunology Kit (Akoya Biosciences, Marlborough, MA, USA), according to the manufacturer's instructions. The automated Leica Bond RX[™] system (Leica Biosystems, Buffalo Grove, IL, USA) was used for sequential staining with a panel of lymphocytic and macrophage markers using antibodies against CD4, CD8, PD-L1, PD-1, FoxP3, CD68. For the detection of epithelial tissue, a combination of antibodies against cytokeratin and E-cadherin was used, as previously described^{36,46,47}. The list of antibodies, reagents, and experimental conditions used for the multiplex immunofluorescence staining are listed in Supplementary Table 11.

Imaging, image analysis, and quality control. The stained TMA were imaged using the Vectra[®] PolarisTM Automated Quantitative Pathology Imaging System (Akoya Biosciences, Marlborough, MA, USA) in multispectral mode at a resolution of 0.496 μm/pixel, which resulted in a total of 2350 multispectral images of size 0.93 mm × 0.7 mm. Each of the images was manually reviewed by three investigators (I.Z., A.M., C.B.) and curated to exclude artifacts, staining defects, and accumulation of immune cells in necrotic areas and intraglandular structures. Any discrepancies were resolved by a pathologist (A.Mez.). The number of images used for further analysis was 1414. The vendor-provided machine learning algorithm, implemented in the inForm[®] image analysis software, was trained to split tissue into three categories: tumor compartment, stromal compartment, or blank areas. The training was performed on a selection of representative cores by providing a set of samples that was manually annotated. Cell segmentation was performed using DAPI nuclear staining. The perinuclear region at 7 pixels from the nuclear border was considered as the cytoplasm area.

Cell phenotyping. The cell phenotyping function of the inForm[®] image analysis software was used to manually define a representative subset of cells positive to expression of each of the markers and a subset of cells negative to all markers (also considering the rare immune phenotypes).

The marker expression was evaluated in the cytoplasm of the segmented cells as the number of photons, normalized to exposure time. Exceptions were made for i) FoxP3 (expressed in nuclei and thus its signal was analyzed in the nuclear region of segmented cells) and ii) CD4 marker, which demonstrated more diffuse staining (probably due to epitope instability) often covering also nuclei regions, and thus, the average marker expression level in total cell region (nucleus + cytoplasm) was used⁴⁸. Intensity cutoffs for the markers were determined in the R programming environment (version 3.6.0). The marker-specific cutoffs were defined by analyzing the distributions of the positive intensities for the marker in the manually-annotated cells and by controlling for the background levels visually and in the negatively annotated cells. The empirical cumulative distribution function was used for visual control. The minimal cutoff was selected as minimal recorded signal level (among the manually annotated cells) if the empirical cumulative distribution demonstrated steep monotonous tendency. In case of individual low-level outliers, they were considered as negative. In case of prominent “kneepoint” in the graph the expression of the marker was back-controlled by reviewing the annotated images and controlling for the marker expression. Additionally, for each of the markers the images were screened for the areas of non-specific staining which were considered as background. Due to the high fraction of CK-positive cells in the entire cohort, a different approach was used to define its cutoff. The complete dataset, containing 1,790,165 cells, was analyzed to evaluate the distribution of CK expression, measured as marker mean expression in cell cytoplasm. The distribution had two peaks, reflecting non-CK cells (low expression) and CK-positive cells (high expression) and Otsu algorithm⁴⁹ was used to find minima between two groups.

Cell classification. Using the established cutoff levels, every cell was characterized as positive or negative for each marker in the panel. These data were used to classify the cell and define its immune subtype as illustrated in Supplementary Table 12. Multiple TMA cores from the same tumor/patient were merged and cell abundance was normalized to the total tissue area. Cell subclasses were quantified and normalized to tissue area thus resulting in final metric called “cell density” (units per mm²), which was calculated in total tissue area, stroma, and tumor regions for each cell subclass. For illustration purposes, the cell density values were log2- transformed. The multispectral imaging/analysis workflow is depicted in Supplementary Fig. 7.

Spatial image analysis and topographical interactions

The previously described Spatial Image Analysis of Tissues (SPIAT) toolkit⁵⁰ was utilized to facilitate the detection of cellular interactions among immune cells (including regulatory T-cells, cytotoxic T-cells, macrophages, and T-helpers and the respective expression of PD-L1/PD-1 checkpoints as reference cells) with tumor cells (target cells) and spatial cell localization. The applied metrics included the normalized mixing score and the entropy gradient which both evaluated the co-localization of immune and tumor cells. The analysis was performed on single reconstructed images from multiple patient-specific TMA stitched with a proper distance between them to avoid potential overlaps. Normalized mixing score (NMS) and the “aggregated entropy” were calculated as a function of different radii around the reference cells (immune cells here), varying from 50 up (close-distance interactions) to 600 μm to cover the whole TMA area with a step of 50 μm. The NMS uses a normalization scale factor -accounting for cell number- and receives a numerical value according to the following formula, described in SPIAT: $NMS = \frac{\text{number of interactions between reference and target cells} \times (\text{total number of reference cells} - 1)}{2 \times \text{number of interactions between reference cell types} \times \text{total number of target cells}}$. The higher the score, the higher the degree of co-localization between immune and tumor cells. Given that cell number and types in TMAs are fixed, the definition of entropy entails mainly the variability in cell number rather than degree of randomness. Immune cells were selected as reference, target cells were identified for each selected radius around the reference cells and entropy

scores were subsequently calculated for each radius (“aggregated entropy” as described in the SPIAT toolkit⁵⁰). Higher entropy scores correspond to high cell colocalization between reference and target cells and vice versa for low entropy scores (unbalanced cell types – low colocalization). Next, patient-specific gradient entropy curves were generated from the “aggregated entropy” values along the different radii (for illustration purposes all curves were graphically displayed following a 10 μm radius step) and classified either as “Attraction-like” or “Repulsion-like” patterns according to the slope of the curve. If the slope was negative close to the lowest radius, the pattern of interaction was defined as “Attraction”, indicating higher aggregated entropy values close to reference cells, thus high colocalization. If the slope was positive close to the lowest radius, the pattern of interaction is defined as “Repulsion”, indicating higher aggregated entropy values far from reference cells, thus low colocalization. In case of no difference on entropy values across the radii, there was a zero slope⁵⁰. The aforementioned analyses were performed in the R programming environment.

Outcomes definition and statistical analysis

Pathologic complete response (pCR) was defined as in the original publication of the primary efficacy analysis of the trial as the absence of residual invasive cancer (or very few scattered cancer cells left), with or without residual DCIS and with negative axillary lymph nodes (ypT0/is ypN0). Non-pCR patients included those who presented with tumor progression on neoadjuvant chemotherapy or who did not undergo surgery⁴³. Patients with missing information were excluded from the analysis. Progression-free survival (PFS), which was the primary endpoint of the trial¹⁰, was defined as the time from randomization to progression on neoadjuvant therapy, locoregional relapse (invasive cancer), first distant metastasis, invasive contralateral breast cancer or death from any cause, whichever occurred first. Second primary invasive non-breast cancers, DCIS or LCIS (ipsilateral or contralateral) were not considered as events.

Continuous outcomes were tested using the Wilcoxon rank-sum or Kruskal–Wallis test while binary outcomes with Pearson’s chi-squared or Fisher’s exact test. The Spearman’s rank correlation coefficient was used for correlations between continuous variables. The different immune cell subpopulations and digital metrics derived by dTILs and multiplex immunofluorescence were correlated to pCR using univariate and multivariable logistic regression models. Median follow-up was calculated using the reverse Kaplan–Meier method. Survival (PFS) was calculated using the Kaplan–Meier method and differences in PFS were calculated with the log-rank test. Time to failure was modeled using proportional hazard regression. Logistic regression and Cox regression models are presented as odds ratios and hazard ratios, respectively, with the respective 95% confidence intervals and Wald p-values. Multivariable models were adjusted for tumor size, nodal status and treatment, and stratified for subtype since subtype violated proportionality of hazards. All p-values were two-sided with 5% as the level of significance. Adjustment for multiplicity was not performed. All analyses were performed the Stata software (v. 17, StataCorp, College Station, TX, USA).

Data availability

Data shall be shared according to the EORTC data release policy (<http://www.eortc.org/datasharing/>).

Code availability

The code for the multiplex immunofluorescence-based data analysis could be found in the following link: https://github.com/gmanikis/EORTC_NPJ_BC/tree/main.

Received: 23 June 2024; Accepted: 1 February 2025;

Published online: 07 March 2025

References

- 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 (2015).

2. 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 (2019).
3. 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 (2013).
4. 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 (2018).
5. Denkert, C. et al. Tumor-infiltrating lymphocytes and response to neoadjuvant chemotherapy with or without carboplatin in human epidermal growth factor receptor 2-positive and triple-negative primary breast cancers. *J. Clin. Oncol.* **33**, 983–991 (2015).
6. Page, D. B. et al. Spatial analyses of immune cell infiltration in cancer: current methods and future directions: a report of the International Immuno-Oncology Biomarker Working Group on Breast Cancer. *J. Pathol.* **260**, 514–532 (2023).
7. Acs, B., Rantalainen, M. & Hartman, J. Artificial intelligence as the next step towards precision pathology. *J. Intern. Med.* **288**, 62–81 (2020).
8. Thagaard, J. et al. Pitfalls in machine learning-based assessment of tumor-infiltrating lymphocytes in breast cancer: a report of the International Immuno-Oncology Biomarker Working Group on Breast Cancer. *J. Pathol.* **260**, 498–513 (2023).
9. Tzoras, E. et al. Dissecting tumor-immune microenvironment in breast cancer at a spatial and multiplex resolution. *Cancers* **14**, 1999 (2022).
10. Bonnefoi, H. et al. TP53 status for prediction of sensitivity to taxane versus non-taxane neoadjuvant chemotherapy in breast cancer (EORTC 10994/BIG 1-00): a randomised phase 3 trial. *Lancet Oncol.* **12**, 527–539 (2011).
11. Matikas, A. et al. Prognostic implications of PD-L1 expression in breast cancer: systematic review and meta-analysis of immunohistochemistry and pooled analysis of transcriptomic data. *Clin. Cancer Res.* **25**, 5717–5726 (2019).
12. Harms, P. W. et al. Multiplex immunohistochemistry and immunofluorescence: a practical update for pathologists. *Mod. Pathol.* **36**, 100197 (2023).
13. Bai, Y. et al. An open-source, automated tumor-infiltrating lymphocyte algorithm for prognosis in triple-negative breast cancer. *Clin. Cancer Res.* **27**, 5557–5565 (2021).
14. Fanucci, K. A. et al. Image analysis-based tumor-infiltrating lymphocytes measurement predicts breast cancer pathologic complete response in SWOG S0800 neoadjuvant chemotherapy trial. *NPJ Breast Cancer* **9**, 38 (2023).
15. Matikas, A. et al. Survival outcomes, digital TILs, and on-treatment PET/CT during neoadjuvant therapy for HER2-positive breast cancer: results from the randomized PREDIX HER2 Trial. *Clin. Cancer Res.* **29**, 532–540 (2023).
16. Makhoul, S. et al. Evaluation of tumour infiltrating lymphocytes in luminal breast cancer using artificial intelligence. *Br J Cancer* (2023).
17. West, N. R. et al. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. *Breast Cancer Res* **13**, R126 (2011).
18. Kos, Z. et al. Pitfalls in assessing stromal tumor-infiltrating lymphocytes (sTILs) in breast cancer. *NPJ Breast Cancer* **6**, 17 (2020).
19. 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 (2020).
20. Liu, S. et al. A panoptic segmentation dataset and deep-learning approach for explainable scoring of tumor-infiltrating lymphocytes. *NPJ Breast Cancer* **10**, 52 (2024).
21. Amgad, M. et al. A population-level digital histologic biomarker for enhanced prognosis of invasive breast cancer. *Nat. Med.* **30**, 85–97 (2024).
22. Vidal Martinez, J. et al. The analytical and clinical validity of AI algorithms to score TILs in TNBC: can we use different machine learning models interchangeably? *eClinicalMedicine* **78**, 102928 29 (2024).
23. Chen, R. J. et al. Towards a general-purpose foundation model for computational pathology. *Nat. Med.* **30**, 850–862 (2024).
24. Sammut, S. J. et al. Multi-omic machine learning predictor of breast cancer therapy response. *Nature* **601**, 623–629 (2022).
25. Yam, C. et al. Immune phenotype and response to neoadjuvant therapy in triple-negative breast cancer. *Clin. Cancer Res.* **27**, 5365–5375 (2021).
26. Wang, X. Q. et al. Spatial predictors of immunotherapy response in triple-negative breast cancer. *Nature* **621**, 868–876 (2023).
27. Hammerl, D. et al. Spatial immunophenotypes predict response to anti-PD1 treatment and capture distinct paths of T cell evasion in triple-negative breast cancer. *Nat. Commun.* **12**, 5668 (2021).
28. Park, S. et al. Artificial intelligence-powered spatial analysis of tumor-infiltrating lymphocytes as complementary biomarker for immune checkpoint inhibition in non-small-cell lung cancer. *J. Clin. Oncol.* **40**, 1916–1928 (2022).
29. Berry, S. et al. Analysis of multispectral imaging with the AstroPath platform informs efficacy of PD-1 blockade. *Science*. **372**, eaba2609 (2021).
30. Huang, Z. et al. Artificial intelligence reveals features associated with breast cancer neoadjuvant chemotherapy responses from multi-stain histopathologic images. *NPJ Precis. Oncol.* **7**, 14 (2023).
31. Cortez, M. A. et al. PDL1 regulation by p53 via miR-34. *J. Natl. Cancer Inst.* **108**, djv303 (2016).
32. Yu, X. Y. et al. Correlation and prognostic significance of PD-L1 and P53 expression in resected primary pulmonary lymphoepithelioma-like carcinoma. *J. Thorac. Dis.* **10**, 1891–1902 (2018).
33. Ali, H. R. et al. PD-L1 protein expression in breast cancer is rare, enriched in basal-like tumours and associated with infiltrating lymphocytes. *Ann. Oncol.* **26**, 1488–1493 (2015).
34. Zhou, X. et al. Mutant p53 gain of function mediates cancer immune escape that is counteracted by APR-246. *Br. J. Cancer* **127**, 2060–2071 (2022).
35. Dumay, A. et al. Distinct tumor protein p53 mutants in breast cancer subgroups. *Int J. Cancer* **132**, 1227–1231 (2013).
36. Zerdes, I. et al. Interplay between copy number alterations and immune profiles in the early breast cancer Scandinavian Breast Group 2004-1 randomized phase II trial: results from a feasibility study. *NPJ Breast Cancer* **7**, 144 (2021).
37. Cheang, M. C. et al. Defining breast cancer intrinsic subtypes by quantitative receptor expression. *Oncologist* **20**, 474–482 (2015).
38. Benefield, H. C. et al. Borderline estrogen receptor-positive breast cancers in Black and White women. *J. Natl. Cancer Inst.* **112**, 728–736 (2020).
39. Villegas, S. L. et al. Therapy response and prognosis of patients with early breast cancer with low positivity for hormone receptors—an analysis of 2765 patients from neoadjuvant clinical trials. *Eur. J. Cancer* **148**, 159–170 (2021).
40. Schrodri, S. et al. Outcome of breast cancer patients with low hormone receptor positivity: analysis of a 15-year population-based cohort. *Ann. Oncol.* **32**, 1410–1424 (2021).
41. Fredriksson, I. et al. 241MO Patient characteristics and real-world outcomes in HER2 negative/ER zero and ER low patients treated as triple-negative breast cancer in Sweden 2008–2020. *Ann. Oncol.* **34**, S279 (2023).
42. McShane, L. M. et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *J. Natl. Cancer Inst.* **97**, 1180–1184 (2005).
43. Bonnefoi, H. et al. Molecular apocrine tumours in EORTC 10994/BIG 1-00 phase III study: pathological response after neoadjuvant chemotherapy and clinical outcomes. *Br. J. Cancer* **120**, 913–921 (2019).

44. Iggo, R. et al. Validation of a yeast functional assay for p53 mutations using clonal sequencing. *J. Pathol.* **231**, 441–448 (2013).
45. Acs, B. et al. An open-source automated tumor-infiltrating lymphocyte algorithm for prognosis in melanoma. *Nat. Commun.* **10**, 5440 (2019).
46. Zeldes, I. et al. Discordance of PD-L1 expression at the protein and RNA levels in early breast cancer. *Cancers* **13**, 4655 (2021).
47. Mezheyski, A. et al. Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J. Pathol.* **244**, 421–431 (2018).
48. Backman, M. et al. Spatial immunophenotyping of the tumour microenvironment in non-small cell lung cancer. *Eur. J. Cancer* **185**, 40–52 (2023).
49. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst., Man Cybern.* **9**, 62–66 (1979).
50. Feng, Y. et al. Spatial analysis with SPIAT and spaSim to characterize and simulate tissue microenvironments. *Nat. Commun.* **14**, 2697 (2023).

Acknowledgements

Ioannis Zeldes is supported by the Region Stockholm (clinical postdoctoral appointment, FoUI-977295), the Swedish Society of Oncology Postdoctoral Award, the Swedish Society of Medicine and Iris, Stig och Gerry Castenbäcks Foundation. Alexios Matikas is supported by the Swedish Cancer Society (Cancerfonden), the Iris, Stig and Gerry Castenbäcks Foundation, The Swedish Breast Cancer Association (Bröstcancerförbundet), and The Research Funds at Radiumhemmet. Georgios C. Manikis is a recipient of a postdoctoral scholarship from the Wenner-Gren Foundations (www.swgc.org). Balazs Acs is supported by The Swedish Society for Medical Research (Svenska Sällskapet för Medicinsk Forskning) Postdoctoral grant and the Swedish Breast Cancer Association (Bröstcancerförbundet). Data analysis for the scope of this translational study was performed at Karolinska Institute, Stockholm, Sweden independently from all funding bodies. This study was supported by grants from Region Stockholm, Karolinska Institutet including Cancer Research KI, the Swedish Research Council (Vetenskapsrådet), the Swedish Cancer Society (Cancerfonden), the Research Funds at Radiumhemmet, the Swedish Breast Cancer Association (Bröstcancerförbundet) and Percy Falks Stiftelse. The funding bodies had no role in the data collection, analysis, interpretation, writing or submission of the manuscript. We thank HistoOne AB, Uppsala, Sweden (<https://histo.one>), Susanne Agartz, and Evangelos Tzoras, MD for assistance with image processing and multiplex data analysis used in this study. We thank also Ioannis Tsarantanis for his contribution to data management. An earlier version of the results was presented in poster form at the European Society of Medical Oncology (ESMO) Congress 2023.

Author contributions

Conceptualization: I.Z., A.M., J.B., H.B., T.F.; methodology: I.Z., A.M., A.Mez., G.M., B.A., C.Boy, G.M.G., Y.B., D.R.; data extraction/curator: I.Z., A.M., A.Mez., G.M., B.A., C.Boy, C.Bom; formal analysis: I.Z., A.M., A.Mez., G.M., B.A., H.J.; data interpretation: I.Z., A.M., A.Mez., G.M., H.B., T.F.; writing—original draft preparation: I.Z., A.M., T.F.; writing—review and editing: all authors; visualization: I.Z., A.M., A.Mez., G.M., H.J., B.A.; supervision, J.B., H.B., T.F.; resources and funding acquisition: I.Z., A.M., J.B., T.F. All authors have read and agreed to the published version of the manuscript. I.Z. and A.M. have contributed equally to this work as co-first authors. A.Mez. and G.M. are listed as joint co-second authors.

Funding

Open access funding provided by Karolinska Institute.

Competing interests

I. Zeldes: Financial Interests, Institutional, Invited Speaker: Novartis. Financial Interests (Institutional), Research Grant: Gilead; A. Matikas:

Financial Interests, Institutional, Coordinating PI, International co-PI of academic trial ARIADNE (EU CT: 2022-501504-95-00): AstraZeneca, Novartis, Veracyte; Non-Financial Interests, Advisory Role: Veracyte. Invited Speaker: Roche, Seagen. M. Ignatiadis: Financial Interests, Personal, Invited Speaker: Novartis; Financial Interests, Personal, Other, Independent Monitoring Committee: Seattle Genetics; Financial Interests, Institutional, Coordinating PI: Pfizer, Roche, Natera, Inivata, Rejuveron; Non-Financial Interests, Officer: EORTC; Financial Interests, Personal, Other, Travel Grants: Roche, Gilead. D. Rimm: Financial Interests, Personal, Advisory Board: AstraZeneca, Cell Signaling Technology, Cepheid, Danaher, Daiichi Sankyo, Konica/Minolta, Lilly, Merck, Nanostring, PAIGE.AI, Regeneron, Roche; Financial Interests, Personal, Advisory Role: NextCure, Regeneron, Sanofi, Verily; Financial Interests, Institutional, Other: Cepheid, Navigate Biopharma, NextCure, Konica/Minolta; Financial Interests, Personal, Royalties: Rarecyte. D.A. Cameron: Financial Interests, Institutional, Advisory Board: Roche, Pfizer, AstraZeneca, Daiichi Sankyo, SeaGen, Synthon, Zymeworks; Financial Interests, Institutional, Other, Has done advisory boards, spoken at a webinar and involved in a manuscript with Lilly - all recompense to my institution: Lilly; Financial Interests, Institutional, Other, Have done advisory boards and been involved in a health economic analysis and publication. All recompense to my institution: Novartis; Financial Interests, Institutional, Coordinating PI, funding and drug for UK participation in a French-led Novartis-funded study. Institutional funding received for consultancy work: Novartis; Non-Financial Interests, Principal Investigator, one of three PIs for the Ameera-6 trial—which has now closed early as the company has stopped all development of the drug: Sanofi; Other, Chair of the board of this small Scottish charity for secondary breast cancer: Make Seconds Count; Other, Chair of the Board of B.I.G.—a group of international breast cancer research groups: Breast International Group; Other, Chair of the European Breast Cancer Council which organizes the bi-annual EBCC meetings: EBCC. J. Bergh: Research grants from Amgen, AstraZeneca, Bayer, Merck, Pfizer, Roche, and Sanofi-Aventis to Karolinska Institutet and/or University Hospital. No personal payments. Co-author on a chapter on “Prognostic and Predictive factors in early, non-metastatic breast cancer” in UpToDate. Honoraria to Asklepios Medicin HB. Stocks in Stratipath AB, a company involved in AI-based diagnostics for breast cancer. Chairperson for Coronis and Asklepios Cancer Research HB. Honoraria from Roche and AstraZeneca for chairmanship and lectures at scientific meetings and consultations for Stratipath AB. H. Bonnefoi: Financial Interests, Personal, Advisory Board: AstraZeneca/ Daiichi Sankyo, Pfizer; Financial Interests, Personal, Advisory role: Stratipath; Financial Interests, Institutional, Coordinating PI, UCBG3-06 prospective study: Bayer. T. Foukakis: Financial Interests, Institutional, Invited Speaker: Roche, AstraZeneca, Gilead Sciences; Financial Interests, Personal, Advisory Board: Novartis, Veracyte, Exact Sciences, Affibody; Financial Interests, Personal, Invited Speaker: Pfizer; Financial Interests, Personal, Royalties, Authorship of two chapters in UpToDate: Wolters Kluwer; Financial Interests, Institutional, Coordinating PI, Clinical trial support (research grant and study drug): Pfizer; Financial Interests, Institutional, Sponsor and Coordinating PI, International co-PI of academic trial ARIADNE (EU CT: 2022-501504-95-00): AstraZeneca, Novartis, Veracyte. All other authors have declared no conflicts of interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41523-025-00730-1>.

Correspondence and requests for materials should be addressed to Ioannis Zeldes.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025