

IMMUNOLOGY

Spatial interplay patterns of cancer nuclei and tumor-infiltrating lymphocytes (TILs) predict clinical benefit for immune checkpoint inhibitors

Xiangxue Wang^{1*}, Cristian Barrera¹, Kaustav Bera¹, Vidya Sankar Viswanathan¹, Sepideh Azarianpour-Esfahani¹, Can Koyuncu¹, Priya Velu², Michael D. Feldman³, Michael Yang⁴, Pingfu Fu⁵, Kurt A. Schalper⁶, Haider Mahdi⁷, Cheng Lu¹, Vamsidhar Velcheti⁸, Anant Madabhushi^{1,9*}

Immune checkpoint inhibitors (ICIs) show prominent clinical activity across multiple advanced tumors. However, less than half of patients respond even after molecule-based selection. Thus, improved biomarkers are required. In this study, we use an image analysis to capture morphologic attributes relating to the spatial interaction and architecture of tumor cells and tumor-infiltrating lymphocytes (TILs) from digitized H&E images. We evaluate the association of image features with progression-free (PFS) and overall survival in non–small cell lung cancer (NSCLC) ($N = 187$) and gynecological cancer ($N = 39$) patients treated with ICIs. We demonstrated that the classifier trained with NSCLC alone was associated with PFS in independent NSCLC cohorts and also in gynecological cancer. The classifier was also associated with clinical outcome independent of clinical factors. Moreover, the classifier was associated with PFS even with low PD-L1 expression. These findings suggest that image analysis can be used to predict clinical end points in patients receiving ICI.

INTRODUCTION

Programmed cell death protein 1 (PD-1)/Programmed death-ligand 1 (PD-L1) axis inhibitors have revolutionized the treatment paradigm across multiple cancers, showing remarkable improvement in mortality in melanoma (1), lung (2), head and neck (3), and gynecologic cancers (4), among others. In metastatic non–small cell lung cancer (NSCLC), immune checkpoint inhibitors (ICIs) with or without chemotherapy are now a first-line treatment regimen. ICIs offer a drastic improvement in overall and progression-free survivals (OS and PFS) with minimal treatment-related toxicities (5, 6). In metastatic NSCLC patients with high PD-L1 expression (>50%), where pembrolizumab (a PD-1 inhibitor) is first-line therapy, significant improvements in OS as compared to chemotherapy alone [hazard ratio (HR), 0.63; 95% confidence interval (CI), 0.47 to 0.86; $P = 0.002$] (7, 8) have been reported. Despite robust clinical benefit, one of the caveats is that even when patient selection is driven by PD-L1 expression, treatment response rates ranged from 27 to 45% in the first-line setting (27% in PD-L1 > 1% and 45% in PD-L1 > 50% subgroup of KEYNOTE-024) and 19% in second-line refractory setting (7, 9–11).

In gynecological cancers (including ovarian, cervical, and uterine cancers) (4), preliminary studies with PD-L1 expression–based patient selection have also shown that ICI improved survival and time to progression in patients (12) with metastatic disease, but with similar limited response rates as in NSCLC. These findings suggest

that tissue-derived PD-L1 expression is not an optimal biomarker to identify patients likely to respond to ICI therapy in any of these cancers. This could be due to the lack of an accepted standard in determining PD-L1 expression levels. Other molecular biomarkers, like TMB (tumor mutational burden) or microsatellite instability, have been explored in recent clinical trials to determine the association with clinical outcomes in different tumors (13–16). However, recent studies suggested that high TMB does not always predict response (17, 18), and some studies (19, 20) have emphasized the importance of additional biomarkers that capture the complexity of tumor immune microenvironment. Thus, there is an urgent need to identify novel biomarkers associated with likelihood of response to ICI, helping to identify those patients who are less likely to respond to ICI and hence might benefit from alternative therapeutic strategies (21).

ICI agents rely on augmenting the body's immune response by one or a combination of the following proposed mechanisms: (i) inhibiting co-repressive signals on T lymphocytes during the early immune response (anti–CTLA-4) in lymph nodes, (ii) antagonizing PD-1 receptor signaling in T lymphocytes during the effector phase (anti–PD-1) in peripheral tissues, or (iii) neutralizing PD-L1 (anti–PD-L1) expressed on the surface of tumor cells in peripheral tissues. One of the hallmarks of cellular immune response for solid tumors is the presence of tumor-infiltrating lymphocytes (TILs) in the tumor microenvironment (TME). Studies have shown that TIL density is strongly prognostic of response in a variety of solid tumors, irrespective of the treatment type (22). TIL density has been shown to be a predictor of ICI response as well, with a recent study (23) showing that density, as measured by immunohistochemistry (IHC) of lymphocytes present on the invasive margin of melanomas instead of more central lymphocytic infiltration, can predict anti–PD-1 response. To date, TIL analysis has focused largely on pathologist-defined visual TIL density estimation based on a few high-power fields viewed under the microscope (24–26). This manual TIL estimate suffers from the subjectivity and low consistency between evaluators

¹Center for Computational Imaging and Personalized Diagnostics, Case Western Reserve University, Cleveland, OH, USA. ²Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY, USA. ³Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁴Department of Pathology, University of Colorado School of Medicine, Aurora, CO, USA. ⁵Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA. ⁶Department of Pathology, Yale School of Medicine, New Haven, CT, USA. ⁷Magee-Womens Hospital and Magee-Womens Research Institute, University of Pittsburgh Medical Center, Pittsburgh, PA, USA. ⁸Department of Hematology and Oncology, NYU Langone Health, New York, NY, USA. ⁹Louis Stokes Cleveland Veterans Affairs Medical Center, Cleveland, OH, USA.

*Corresponding author. Email: xxw345@case.edu (X.W.); axm788@case.edu (A.M.)

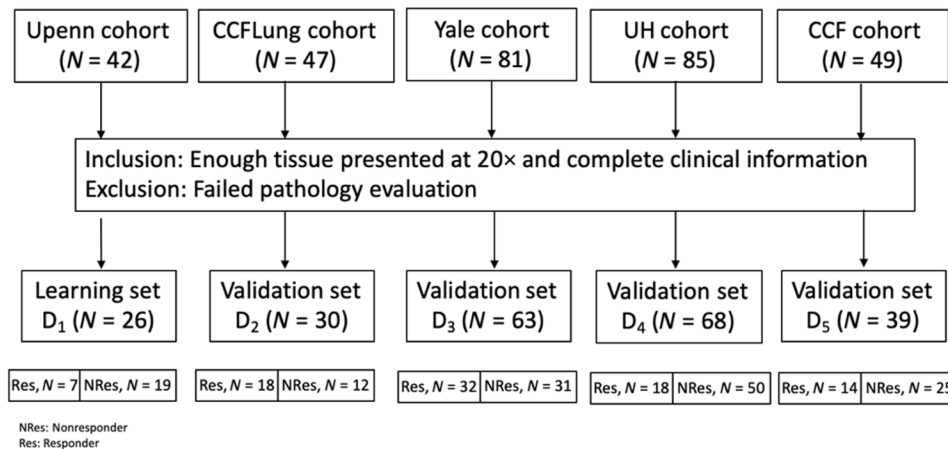


Fig. 1. Consort flow diagram of patient selection. Patient selection from four different NSCLC cohorts (D_{1–4}) and a gynecological cancer cohort (D₅).

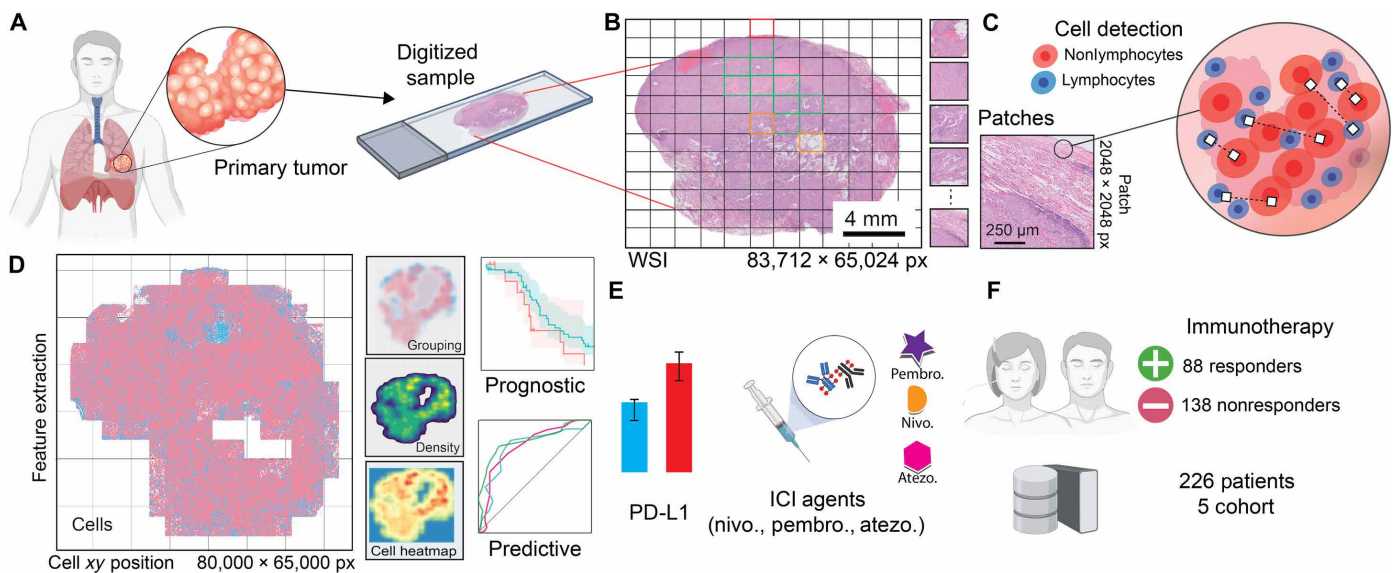


Fig. 2. Flowchart of high-level workflow. Entire workflow comprises (A) tissue preparation and digitization, (B) tile processing from annotation, (C) region and cell segmentation, (D) image analysis and feature extraction, (E) predictive model construction and validation, and (F) cohort overview.

(27). Computer-assisted TIL estimation has been explored based on hematoxylin and eosin (H&E) to decrease the subjectivity (28). Moreover, recent studies have shown that density and arrangement patterns of TILs on H&E images or using quantitative multiplex immunofluorescence are associated with risk of recurrence and outcomes in NSCLC and breast cancer (27, 29–32). While recent studies have reported that neural network–based approaches could predict the treatment response in cancer patients treated with immunotherapy (33, 34), the end-to-end deep learning models tend to lack interpretability, which, in turn, could make it harder to achieve generalizable results on new, unseen, external datasets. A number of findings from our group (27, 35) as well as of others (36, 37) have emphasized the importance of biologically inspired feature related to spatial heterogeneity and complex interactions between TME and immune microenvironment for predicting clinically relevant outcomes in ICI-treated cancer patients.

In this work, we present an automated image classifier that not only computationally enumerates TIL density but also characterizes

the spatial architecture and arrangement of TILs to predict clinical outcomes in ICI-treated cancer patients (Fig. 1). Specifically, we aimed to capture quantitative image biomarkers relating to the spatial interaction between immune and tumor cells in the TME (i.e., epithelium and stroma) and subsequently show the association of these image biomarkers against clinical outcomes for ICI (monotherapy and combination of ICI agents)–treated patients with different cancer types [NSCLC and gynecological cancers (Fig. 2)]. The association of these image biomarkers against clinical outcome in ICI-treated patients was also compared against manual TIL estimation and also evaluated in the context of cancer patients with low PD-L1 expression.

RESULTS

HistoTIL predicts response to ICI

Feature discovery on D₁ NSCLC cohort revealed a total of seven HistoTIL features (intersected area between non-TIL nuclei and TIL

graphs, nuclei shape descriptors in epithelium area, and distance statistics between non-TIL nuclei and TIL graph) that were significantly associated with OS ($P < 0.001$) and PFS ($P = 0.003$). These included four features from the non-TIL nuclei and three from the interplay between non-TIL nuclei and TILs (Fig. 3; detailed feature description and distribution in table S1 and fig. S1). In predicting Response Evaluation Criteria in Solid Tumours (RECIST) as a binary clinical outcome, the linear model achieved 0.71 area under the

curve (AUC) value on NSCLC training set D_1 ; 0.73, 0.62, and 0.73 on independent NSCLC validation cohort D_{2-4} ; and 0.68 on gynecological cohort D_5 , respectively.

Independent prediction of survival

On D_1 , HistoTIL yielded an HR = 5.47 (95% CI, 1.77 to 16.9; $P < 0.001$) and HR = 4.63 (95% CI, 1.38 to 11.8; $P = 0.003$) (Fig. 4) for OS and PFS, respectively. On (D_{2-4}) NSCLC validation set, the

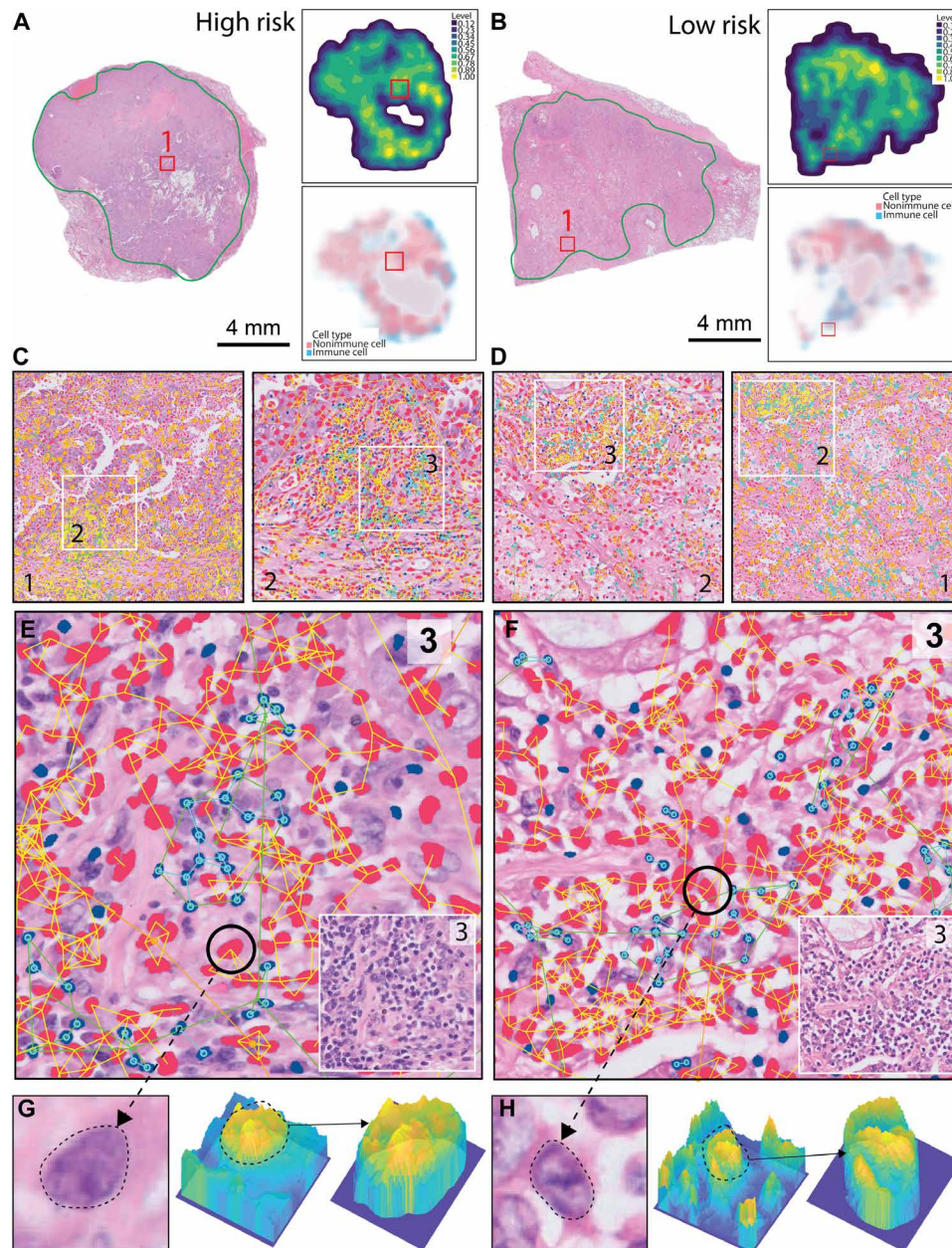


Fig. 3. Visualization of features relating to TIL–non-TIL interactions. A high-risk example is found on the left panel. On the right panel is a low-risk sample. (A and B) The whole-slide image (WSI) is depicted with a green-marked line surrounding the tissue that was indicated as sufficient for processing. (C and D) Two small panels indicate the grouping behavior of the lymphocyte and nonlymphocyte cells by means of a two-dimensional plot with both cell groups intertwined, where the numbers in red indicate the main panel it is coming from. (E and F) A zoomed-in region of interest. The groups of cells form distinctive clusters. The clusters are connected through colored lines, which indicate the distance between lymphocytes (green line), cluster of lymphocytes (cyan line), and cluster of nonlymphocytes (yellow line). Descriptors are calculated that capture general architectural structure and immune-cell interplay. (G and H) Texture-based feature related to the heatmap of the nuclei and surrounding tissue for the low- and high-risk cases. Zernike polynomials and Haralick descriptors are calculated from the shape and texture of the cell.

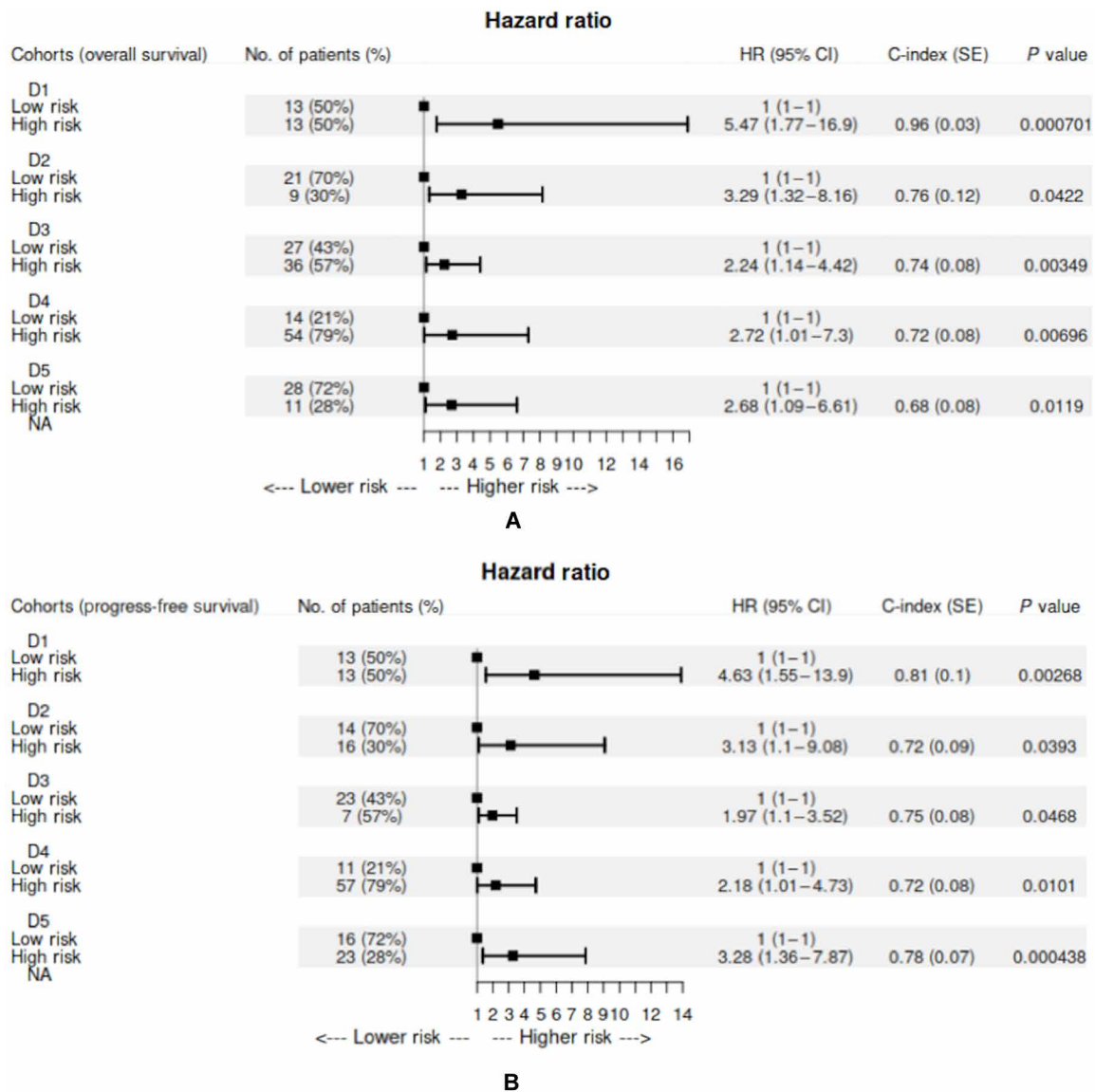


Fig. 4. Forest plot for individual cohort. Survival analysis with OS (A) and PFS (B) between HistoTIL-defined low- and high-risk groups.

performance of HistoTIL in predicting OS and PFS was HR = 3.29 (95% CI, 1.32 to 8.16; $P = 0.042$) and HR = 2.45 (95% CI, 1.00 to 6.00; $P = 0.031$) on D₂, HR = 2.03 (95% CI, 1.08 to 3.82; $P = 0.027$) and HR = 2.43 (95% CI, 1.04 to 5.65; $P = 0.018$) on D₃, and HR = 2.58 (95% CI, 1.02 to 6.49; $P = 0.009$) and HR = 3.05 (95% CI, 1.07 to 8.69; $P = 0.001$) on D₄. HistoTIL was also evaluated in terms of association with OS and PFS in gynecological cancers (D₅) with HR = 2.29 (95% CI, 1.15 to 4.58; $P = 0.0037$) and HR = 1.93 (95% CI, 1.10 to 3.38; $P < 0.043$) for OS and PFS, respectively. The differences in the median OS between the high-risk and low-risk groups as determined by HistoTIL were 14.2 and 34.6 months for D₂, 15.7 and 23.3 months for D₃, and 12.7 and 22.6 months for D₄ (all in lung cancer) and 6.5 and 14.2 months for D₅ in gynecological cancers, respectively. Similarly, the differences in the median PFS between the two HistoTIL risk groups were 3.1 and 10.6 months for D₂, 2.7 and 8.5 months for D₃, and 3.4 and 11.2 months for D₄ in NSCLC and 2.5 and 9.1 months for D₅ in gynecological cancers.

Moreover, on multivariable analysis and after adjusting for the effects of clinicopathological variables, HistoTIL was found to be independently prognostic on the NSCLC datasets (D₁: HR = 1.38; 95% CI, 1.04 to 1.68; $P = 0.012$; D₂: HR = 1.69; 95% CI, 1.01 to 3.03; $P = 0.049$; D₃: HR = 1.21; 95% CI, 1.01 to 1.44; $P = 0.038$; D₄: HR = 5.25; 95% CI, 1.65 to 16.78; $P = 0.005$) and for gynecological cancers (D₅ cervical, $n = 10$; ovarian, $n = 14$; and endometrial, $n = 25$; HR = 5.22; 95% CI, 1.24 to 21.94; $P = 0.024$; Table 1). Smoking status was also found to be significant in multivariable analysis (never smoker versus previous/current smoker: HR = 0.271; 95% CI, 0.10 to 0.78; $P = 0.015$) but only in D₃. None of the other clinicopathological covariates were found to be significantly associated with OS.

HistoTIL associated with OS and PFS independent of type of ICI agent

For NSCLC, HistoTIL predicted that low-risk groups had significantly longer survival time for patients who received nivolumab

Table 1. Multivariable analysis of OS with HistoTIL and clinical variables. ADC, adenocarcinoma; SCC, squamous cell carcinoma; values in bold are statistically significant by two-tailed test ($P < 0.05$).

	Characteristics	Hazard ratio	95% CI	P
D ₁ (lung)	Gender			
	Male	0.528	0.096–2.894	0.462
	Female			
	Smoking status			
	Former/current	0.463	0.049–4.396	0.503
	Never			
	Histology subtypes			
	Adenocarcinoma	0.347	0.023–5.187	0.443
	SCC			
	Clinical stage			
	III	3.646	0.292–45.484	0.315
	IV			
	Treatment			
	Mono	0.441	0.020–9.566	0.602
Combo				
TIL counting				
Low	1.385	0.231–8.307	0.721	
High				
Risk scores	1.397	1.071–1.822	0.014	
D ₂ (lung)	Gender			
	Male	0.342	0.056–2.083	0.245
	Female			
	Smoking status			
	Former/current	0.126	0.010–1.531	0.104
	Never			
	Histology subtypes			
	Adenocarcinoma	0.203	0.017–2.488	0.212
	SCC			
	Clinical stage			
	III	3.437	0.179–66.055	0.413
IV				
Risk scores	1.685	1.011–3.029	0.049	

Downloaded from https://www.science.org on May 25, 2023

	Characteristics	Hazard ratio	95% CI	P
D ₃ (lung)	Gender			
	Male	2.334	0.746–7.306	0.145
	Female			
	Smoking status			
	Former/current	0.287	0.840–0.098	0.023
	Never			
	Histology subtypes			
	Adenocarcinoma	0.849	0.19–3.64	0.826
	SCC			
	Clinical stage			
	III	13.023	0–Inf.	0.994
	IV			
	Treatment			
	Mono	0.802	0.246–2.614	0.715
Combo				
PD-L1				
Low	0.593	0.200–1.755	0.345	
High				
Risk scores	1.211	1.013–1.447	0.035	
D ₄ (lung)	Gender			
	Male	1.558	0.744–3.264	0.240
	Female			
	Smoking status			
	Former/current	1.025	0.324–3.241	0.967
	Never			
	Histology subtypes			
	Adenocarcinoma	0.985	0.409–2.373	0.972
	SCC			
	Clinical stage			
	III	2.129	0.553–8.194	0.272
	IV			
	Treatment			
	Mono	1.252	0.311–5.046	0.752
Combo				
Risk scores	5.253	1.645–16.775	0.005	

	Characteristics	Hazard ratio	95% CI	P
D ₅ (gynecology)	Clinical stage			
	III	1.377	0.312–6.078	0.673
	IV			
	Histology subtypes			
	Endometrioid	Reference	Reference	Reference
	Serous	0.382	0.063–2.315	0.295
	SCC	1.980	0–Inf.	1
	Clear cell	3.544	0.817–15.369	0.091
	Primary site			
	Ovarian	Reference	Reference	Reference
	Endometrial	1.031	0.238–4.472	0.967
	Cervix	1.741	0–Inf.	1
	IO type			
	Mono	0.081	0–Inf.	0.983
	Combo			
Risk score	5.217	1.240–21.943	0.024	

(OS: HR = 2.08; 95% CI, 1.27 to 3.41; $P < 0.001$; PFS: HR = 2.04; 95% CI, 1.34 to 3.10; $P < 0.001$) and pembrolizumab (OS: HR = 3.19; 95% CI, 1.06 to 9.61; $P = 0.041$; PFS: HR = 5.34; 95% CI, 1.80 to 15.90; $P < 0.001$) compared to high-risk groups. In the gynecological cohort, patients who received pembrolizumab had significant survival differences (OS: HR = 4.65; 95% CI, 1.08 to 19.9; $P = 0.023$; PFS: HR = 3.55; 95% CI, 1.00 to 12.60; $P = 0.008$) in HistoTIL-predicted risk groups. While patients treated with other ICI agents, which included fewer patients, did not show significant survival difference in low- and high-risk groups (nivolumab, OS: HR = 2.12; 95% CI, 0.67 to 6.75; $P = 0.180$; PFS: HR = 2.55; 95% CI, 0.65 to 10; $P = 0.088$), the effect size (HR) shows a longer survival trend for low-risk groups (fig. S2).

Comparison with PD-L1 expression and TIL estimation

While there was no significant survival difference between low and high HistoTIL-defined risk groups in PD-L1-high group (OS: HR = 2.03; 95% CI, 0.59 to 6.90; $P = 0.166$; PFS: HR = 0.91; 95% CI, 0.25 to 3.40; $P = 0.896$; Fig. 5, A and C), there was a significant survival advantage in HistoTIL-defined low-risk group for PD-L1 low expressed patients for both OS and PFS on D₃ (HR = 2.51; 95% CI, 1.15 to 5.47; $P = 0.017$; HR = 2.59; 95% CI, 1.35 to 4.99; $P = 0.016$; Fig. 5, B and D). However, pathologists' TIL grading failed to stratify patients based on survival in D₁; there was also no correlation between manual TIL grading with HistoTIL-predicted risk groups (Fig. 6).

DISCUSSION

Recent advances in computational power and the advent of whole-slide digitization have led to several studies investigating quantitative approaches to interrogate the TME on high-resolution histology images (38, 39). Research has focused on leveraging artificial intelligence

and digital pathology approaches toward answering clinically relevant questions related to detection, diagnosis, prognosis, and treatment response (40). Broadly speaking, computational pathology-based approaches have considered either handcrafted features based on predefined image patterns such as collagen orientation (41) or spatial architecture of cancer cells (27). Another category of computational pathology approaches involves deep learning (39). These are neural network-based unsupervised feature generation strategies that aim to directly capture patterns from digitized histopathology images without typically requiring specification of predefined image motifs. Deep learning approaches like handcrafted approaches (42) have been shown to be able to predict underlying molecular signatures of the tumor (43) and have also been shown to be associated with disease outcome and treatment response (44). While deep learning has achieved great performances in many medical image tasks recently, its low interpretability could prevent its wide application in the clinical setting. A recent example of the failure of deep learning approaches in real-world clinical settings can be seen in applications for coronavirus disease 2019 (COVID-19) (45). Domain-inspired handcrafted approaches that are human explainable may be an important prerequisite for oncologists in guiding management of ICI-treated patients (40, 46). Abduljabbar *et al.* (36) combined deep learning with RNA sequencing data to reveal the spatial immune landscape of NSCLC, identifying an image signature that was prognostic of recurrence. Beck *et al.* (47) previously reported on stromal morphologic features and identified their prognostic association in breast cancer. These previous studies have suggested the importance of capturing quantitative and interpretable domain-inspired attributes of the disease; however, these studies were primarily in the context of disease prognosis and outcome prediction. It seems logical, therefore, that these interpretable approaches will play an even more significant role in helping clinicians make treatment decisions and in assessing therapeutic benefit.

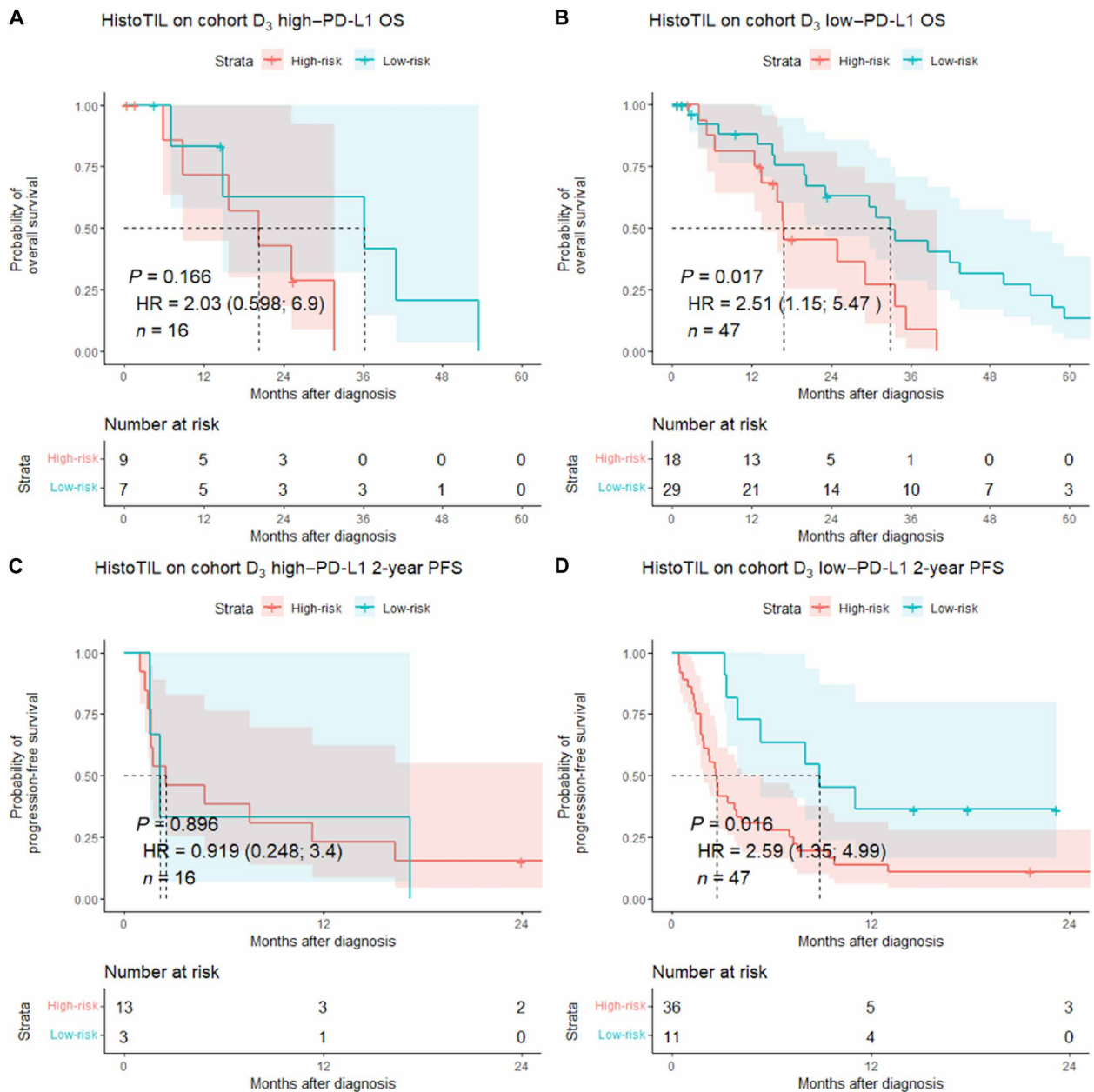


Fig. 5. Survival analysis of PD-L1 expression. The Kaplan-Meier survival analysis of OS in high-PD-L1 (A) and low-PD-L1 (B) expression group and PFS in high-PD-L1 (C) and low-PD-L1 (D) expression group on D₃.

In this work, we presented HistoTIL, a computational-based assay that spatially characterizes TILs in tumor specimens to predict response to ICI in patients with NSCLC and gynecological cancer. HistoTIL is composed of a set of quantitative image descriptors that characterize local TIL–non-TIL interactions within the separate stromal and epithelial compartments of the TME. This is different from measures of TIL density or manual TIL grading that result in a single number that express the average abundance of lymphocytes over an entire slide. In high-risk patients who showed poor response to ICI agents (as determined by HistoTIL), the cancer nuclei clusters dominated the TME and appeared more chaotic, with only limited immune cell infiltration. However, in the ICI responders

identified by HistoTIL, there was a preponderance of TIL clusters that appeared to be encasing the cancer nuclei (Fig. 3), possibly hinting at how the ICI agents require the recruitment of the body’s immune defenses to fight against cancer (9). This is further strengthened by having four completely blinded validation sets from two different cancer types, as HistoTIL was only exposed to a single institution dataset for training. This is also advantageous because while TILs have recently been shown to be prognostic in various cancer types (24, 25), their prognostic significance is plagued by the wide variability in pathologist or manual estimation of TILs between different pathologists (48). Moreover, in comparison with current RECIST-based response evaluation, HistoTIL could also predict

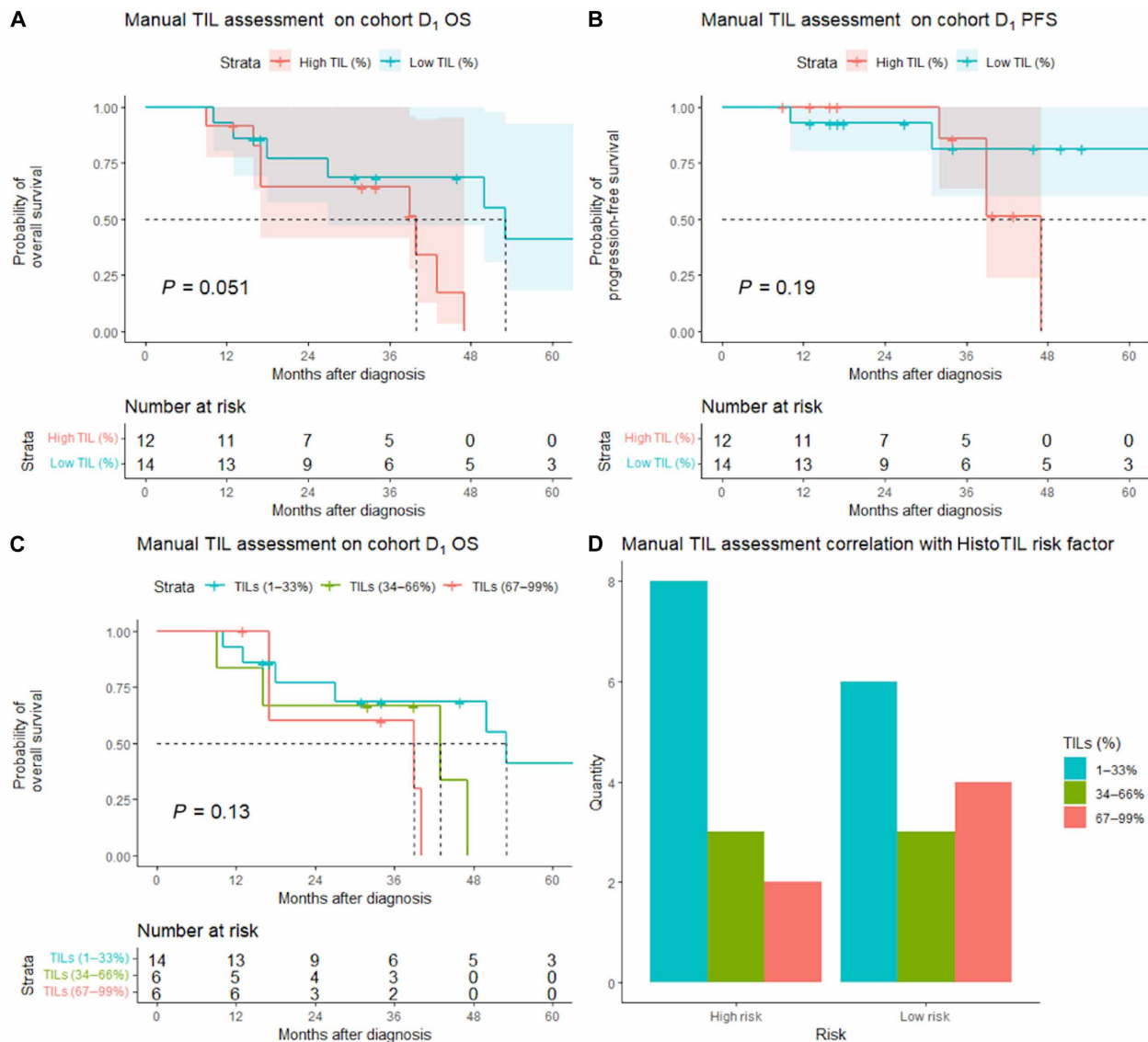


Fig. 6. Survival analysis of manual TIL grading and comparing with HistoTIL. Kaplan-Meier survival analysis of (A) manual TIL grading with OS, (B) manual TIL grading with PFS, and (C) manual TIL grading with OS, three stratified groups. (D) Correlation between manual grading and HistoTIL prediction.

short-term clinical outcomes across different datasets (Fig. 7). HistoTIL was shown to be associated with survival for NSCLC patients treated with nivolumab and pembrolizumab and gynecological cancer treated with pembrolizumab (Fig. 8 and fig. S2). HistoTIL was also associated with OS in the individual cancer types under gynecological cancer (cervical, $n = 10$; ovarian, $n = 14$; and endometrial, $n = 25$). HistoTIL was associated with clinical outcome independent of specific ICI agent, likely because features relating to immune cell and cancer cell interplay on H&E images are a reflection of tumor biology and disease aggressiveness. Independent of the type of ICI agent used or disease indication, HistoTIL-predicted low-risk patients had a relatively lower hazard of disease progression or death compared to high-risk patients.

HistoTIL could also significantly stratify two risk groups within the PD-L1 low expressed cohorts for both OS and PFS (HR = 2.2; 95% CI, 0.97 to 4.95; $P = 0.01$; HR = 2.59; 95% CI, 1.35 to 4.99;

$P = 0.01$; Fig. 5, B and D). This finding could potentially suggest that HistoTIL-identified low-risk patients with PD-L1 low expression, typically recommended for ICI-chemo combination therapy currently, might be eligible for ICI monotherapy directly. Manual TIL grading of the slides by pathologists ($N = 26$) did not result in significant stratification of patients into different survival groups in D₁; the high-TIL groups even presented with longer survival trend after ICI (Fig. 6). This might be due to the lack of standard agreement of manual grading and intra-observer variation, which has been previously reported (27).

Recently, Hu *et al.* (49) showed a deep learning model to predict anti-PD-1 response in melanoma ($N = 54$) and lung cancer ($N = 55$) (49). While Hu *et al.* (49) also found that the model trained on one tumor type could potentially apply to another tumor type for predicting RECIST-defined response, our results were validated with larger set ($N = 226$) with more consistent performance (i.e., prognostic

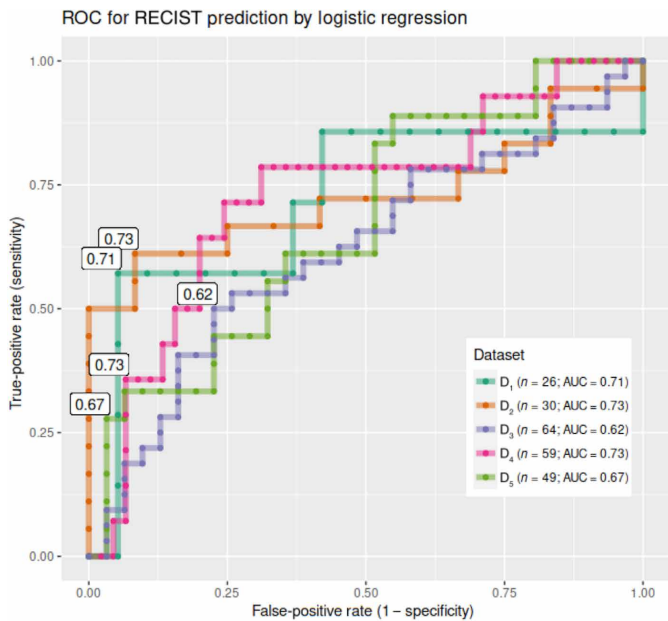


Fig. 7. Prediction of clinical response by HistoTIL. Receiver operating characteristic (ROC) analysis for HistoTIL in predicting clinical response by RECIST among five independent cohorts.

of survival and predictive of RECIST) among different cancers treated with different ICI agents. Moreover, post hoc analysis of deep learning approaches usually requires the dividing of whole-slide images (WSIs) into smaller image tiles (256 × 256) (49), which can dilute and cause loss of information related to spatial heterogeneity within TME (29, 50). Moreover, the deep learning model was trained on ImageNet, and hidden features were extracted with further principal components analysis (PCA) projection (49) for response prediction, which further added another layer of opacity for explaining the result. In contrast, HistoTIL achieved higher and more robust accuracy among different tumor types using human explainable features. A recent surge of applying deep learning on histopathology images ranged from object (i.e., tumor and cell) detection and classification (51) to prognosticating disease-related survival and response to different treatment options (49, 52). However, the nature of deep learning-based classifiers makes it difficult to fully interpret the basis of the predictions. It has also been suggested that the opacity associated with the deep learning-based predictions could be an impediment from deploying them in a clinical setting for treatment management-based decision-making (40, 46). Moreover, studies have indicated that biologically inspired image features could provide more transparent and interpretable options in the clinical setting (53). Hence, our HistoTIL approach leverages deep learning for object detection and segmentation tasks but subsequently uses interpretable

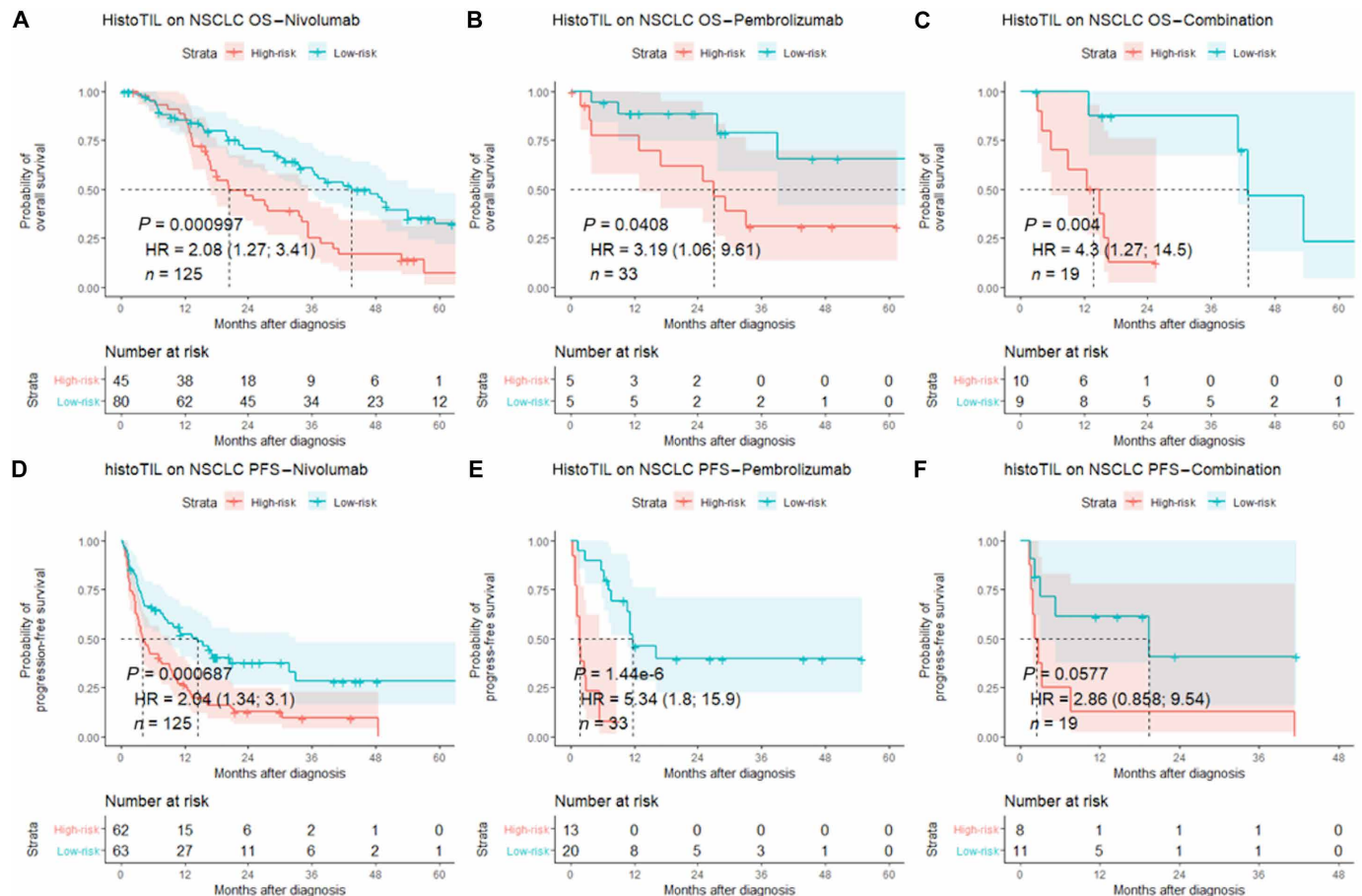


Fig. 8. Survival analysis of NSCLC among ICI agents. Kaplan-Meier OS assessment among ICI agents (nivolumab, pembrolizumab, and combination therapy) for patients with NSCLC with OS (A to C) and PFS (D to F).

and engineered features to capture and describe the tumor-immune interaction.

However, we do acknowledge some limitations of our work. This work was limited by the retrospective nature of data collection. Future work will involve validating HistoTIL on completed clinical trial datasets of cancer patients treated with ICI and also evaluating HistoTIL in a prospective setting. Such a prospective clinical trial validation would also allow for head-to-head comparisons with currently used patient selection biomarkers including PD-L1 and TMB. We also acknowledge that the approach invoking five large image patches for analysis could bias the approach when only smaller samples are available for interrogation.

In conclusion, we presented a computational method to prognosticate survival in advanced NSCLC and gynecological patients treated with multiple ICI. Our approach, HistoTIL, was trained on a dataset from a single institution and independently validated on datasets from two different institutions for predicting both OS and PFS just based on H&E WSIs. In addition to being robust to external multisite validation, HistoTIL also uncovered the prognostic significance of the spatial relationships and structural interplay between TILs and tumor cell nuclei in determining response to ICI agents that modulate the body's immune system. Last, HistoTIL is tissue nondestructive, based on H&E alone, and thus much less time consuming and inexpensive, as it eliminates the need for tissue to be processed and/or analyzed with expensive molecular methods.

MATERIALS AND METHODS

Dataset and image acquisition

H&E-stained pretreatment tumor specimens of patients with advanced NSCLC and gynecological cancers who received ICIs were collected from four independent institutions for this study. The ICI treatment response was evaluated independently by response criteria used for solid tumors (RECIST 1.1) (54). All retrospective analyses were conducted under a protocol approved by the institutional review board (IRB), and the requirement of patient informed consent was waived by IRB. From the University of Pennsylvania Hospital, 42 patients diagnosed with NSCLC between 2008 and 2016 with follow-up information were collected. Tissue quality evaluation was performed at the pathology laboratory to exclude tissue with major artifacts (i.e., tissue folding, blur, staining issues, and anthracosis), and only samples with sufficient tissue (accommodating at least five image patches of size 2000×2000 pixels) and complete treatment and follow-up information were included in this study. This process rendered 26 patients to construct the learning set (D_1). In addition, we collected 47 (chart review between 2005 and 2016), 81 (chart review between 2009 and 2017), and 85 (chart review between 2011 and 2017) NSCLC patients from the Cleveland Clinic, Yale University, and University Hospital Cleveland Medical Center, respectively. Forty-nine patients with gynecological cancer (cervical, $n = 10$; ovarian, $n = 14$; and endometrial, $n = 25$) were collected from the Cleveland Clinic. After applying the same inclusion and exclusion criteria, we obtained independent validation datasets D_2 ($N = 30$), D_3 ($N = 63$), D_4 ($N = 68$), and D_5 ($N = 39$), respectively (Fig. 1). Detailed demographics including the histology subtypes and ICI agents are available in Table 2.

H&E slides from D_{1-2} were scanned with Philips IntelliSite Pathology Solution (software version: Philips DP v1.0) at $\times 40$ magnification. An Aperio CS2 tissue scanner (software version: Aperio

Image Library v10.2.41) was used at D_{3-5} to scan slides at $\times 20$ magnification. To compensate for the variations during the tissue imaging process among different scanners and institutions, an open source software, HistoQC (55), was applied to identify digitized tissue slide images eligible for downstream analysis. Expert pathologists manually annotated the tumor region on digitized H&E slides. All the subsequent image analyses were conducted at $\times 20$ magnification (images were down-sampled from $\times 40$ to $\times 20$) from the annotated regions (see Fig. 2 for flowchart illustrating the methodology).

Image analysis: Nuclei detection and region separation

For segmentation of the nuclei, each WSI was divided into a set of adjacent 2000×2000 pixel tiles. A U-Net style deep convolutional neural network (CNN) at $\times 20$ magnification was used for automated nuclei segmentation (56). Following nuclear segmentation, a previously reported lymphocyte detection approach (27) was used to identify TILs and cancer cell nuclei. Another U-Net style CNN model was used to specifically differentiate and identify epithelial and stromal regions at $\times 5$ (see the Supplementary Materials for more details).

Quantitative histomorphometric features

A total of 753 quantitative histomorphometric (QH) features of two main categories were extracted: nuclei-based features and TIL-related features, from regions of stroma, epithelium, and the entire annotated tumor (table S2).

Nuclear features were extracted from within the pixels of nuclei regions. The four types of features relating to nuclei were as follows: shape descriptors, orientation entropy, image texture, and local contextual. Shape features included statistical measurements of the nuclei including nuclear area, perimeter, and eccentricity to mathematical operator like Zernike shape descriptors (57). Orientation entropy measures the randomness of cell directionality. Nuclear texture quantitatively characterizes the local pixel distribution pattern in the nuclei and is prognostic among different cancer types. Last, local contextual features capture the local neighboring and topological information of individual nuclei.

TIL-related features were extracted from the different tissue regions (i.e., epithelium and stroma) and included enumeration of TILs and determination of the architecture and spatial interaction of TILs and cancer cells. Previous work (27) has shown that a densely packed cluster of these two cell types has a poorer prognosis than a sparse arrangement in early-stage NSCLC. The extent of TIL infiltration was determined through the counts and density of TILs by automated detection (27). Aspects of spatial arrangement were measured by first assigning TILs and non-TIL cells into neighboring clusters (see the "Graph construction" section in the Supplementary Materials) and then characterizing relative abundance and spatial closeness of TILs and non-TIL cells within the cluster as quantitative image features. The variations in cell composition (i.e., ratio of number of TILs to non-TIL nuclei) among each neighboring cluster were also captured to reflect the heterogeneity within TME. These patterns quantitatively describe the histomorphometry (i.e., shape, staining, and local density) among neighboring cells and, in turn, reflect the complex interaction within the tumor immune microenvironment. Details of each feature description are included in table S2.

HistoTIL: An image-based digital risk score

The top-performing QH features were selected from a total of 753 features by applying an elastic net regularized Cox proportional

Table 2. Demographic and characteristics of the patients.

	D ₁	D ₂	D ₃	D ₄	P	D ₅		
Lung	Gender						Primary site	
	Male	9	19	36	32	0.128	Ovarian	12
	Female	17	11	27	36		Endometrial	20
	Histology						Histology subtypes	
	ADC	19	20	47	52	<0.001	Endometrioid	12
	SCC	7	7	12	13		Serous	14
	Others	0	3	4	3		SCC	7
	Smoking status							Clear cell
	Former/current	21	25	51	61	0.552	Clinical stage	
	Never smoker	5	5	12	7		III	27
	Clinical stage						Gynecology	
	III	7	5	2	11	0.062	IV	12
	IV	19	25	61	57		Treatment	
	Treatment						Nivo	16
	Nivo	13	30	43	39	<0.001	Pembro	20
	Pembro	7	0	4	22		Atezo	2
	Atezo	4	0	6	0		Combo	1
	Combo	2	0	10	7		Age (years)	
	Age (years)						Mean ± SD	67.6 ± 12.1
	Mean ± SD	65.2 ± 10.1	N.A.	68.8 ± 10.5	63.3 ± 12.2	0.015	Systematic treatment before ICI	
Systematic treatment before ICI						Yes	35	
Yes	22	N.A.	53	58	0.99	No	4	
No	4		10	10				

hazard model on D₁. Tenfold cross-validation was used to select the regularization parameter (lambda) by grid search, and the best penalty hyperparameter (alpha) controlling the ratio between L₁ and L₂ was found by changing the value from 0 to 1 with step size of 0.1. The selected features were used to build another Cox proportional hazard model to learn the coefficient for each of the selected features on D₁. Survival indices (risk score) were then calculated by a linear combination of selected features and learned coefficients. The median index from the training set D₁ was picked up as the threshold to divide the patients into low and high risk based on OS and PFS. D₁ to D₄ were entirely composed of metastatic NSCLC patients. To evaluate the robustness of HistoTIL, another independently collected set (D₅) of 39 gynecological cancer patients was used for validation using the same cutoff value derived from D₁. We evaluated the association of HistoTIL against PFS and OS on independent sets D₂, D₃, D₄, and D₅.

Statistical analysis

OS was measured from the date of initiation of ICI to the date of death and censored at the date of last follow-up for survivors. PFS was calculated from the date of initiation of ICI to the date of recurrence or death, whichever occurred earlier, and censored at the date of last follow-up for those still alive without recurrence. The HistoTIL classifier assigned every patient to either a low- or high-risk survival group based on the median risk score learned from training set D₁. Kaplan-Meier survival curves were obtained to visualize the differences in sets D₂, D₃, D₄, and D₅. The log-rank test was then used to evaluate the prognostic performance of HistoTIL in predicting PFS and OS on validation sets. HR between survival groups with 95% CIs and P values were calculated. Multivariable Cox analysis was conducted to assess the relationships between the various covariates and survival while adjusting for baseline factors (i.e., tumor subtypes and pathological stage). The Fisher exact test was used to test the

difference of demographics in NSCLC patients across the training and validation datasets. All statistical tests were two sided with significance level set at 0.05. In addition, we also built a logistic regression model in conjunction with the same HistoTIL features identified from D₁ to predict patients' response versus nonresponse of ICI. While the response category included both complete response (CR) and partial response (PR), the nonresponse category included progressive disease (PD) and stable disease (SD) defined by RECIST (54). The response prediction was evaluated by AUC of the receiver operating characteristics (ROCs).

ICI agent separated analysis

To evaluate whether HistoTIL could stratify ICI-treated patients into low- and high-risk categories independent of specific agents (i.e., nivolumab, $N = 141$; pembrolizumab, $N = 53$; atezolizumab, $N = 12$; and combination of agents, $N = 20$), we further conducted the Kaplan-Meier survival analysis for each subset of patients treated with different ICI agents. The same cutoff value of HistoTIL learned from D₁ was used to identify patients into low- and high-risk survival groups.

Comparison with PD-L1 expression and manual TIL estimation

HistoTIL was also compared with the current PD-L1 expression-based patient selection strategy, especially to evaluate the performance of HistoTIL in the low-PD-L1 expression group. D₃ with available IHC PD-L1 expression data was used to validate the HistoTIL prediction of clinical outcome after ICI in PD-L1 low and high groups separately (PD-L1 cutoff was empirically determined at 1% expression). Expert thoracic pathologists with experience in TIL grading were involved in deciding the level of lymphocyte infiltration for each WSI in D₁. Four available categories were provided to the pathologists for TIL grading: no infiltration (visual absence of TILs from the entire slide), low TILs (1 to 33% TILs were found from tumor area), moderate (34 to 66% tumor region were found with TILs), and high (more than 66% of tumor bed was infiltrated with lymphocytes). The same Kaplan-Meier survival analysis was conducted to compare the survival difference between manually defined low- and high-TIL groups.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abn3966>

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

- J. J. Luke, K. T. Flaherty, A. Ribas, G. V. Long, Targeted agents and immunotherapies: Optimizing outcomes in melanoma. *Nat. Rev. Clin. Oncol.* **14**, 463–482 (2017).
- J. Malhotra, S. K. Jabbar, J. Aisner, Current state of immunotherapy for non-small cell lung cancer. *Transl. Lung Cancer Res.* **6**, 196–211 (2017).
- R. Mandal, Y. Şenbabaoğlu, A. Desrichard, J. J. Havel, M. G. Dalin, N. Riaz, K.-W. Lee, I. Ganly, A. A. Hakimi, T. A. Chan, L. G. Morris, The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight* **1**, e89829 (2016).
- O. Kooshkaki, A. Derakhshani, H. Safarpour, S. Najafi, P. Vahedi, O. Brunetti, M. Torabi, P. Lotfinejad, A. V. Paradiso, V. Racaneli, N. Silvestris, B. Baradaran, The latest findings of PD-1/PD-L1 inhibitor application in gynecologic cancers. *Int. J. Mol. Sci.* **21**, 5034 (2020).
- S. J. Antonia, J. F. Vansteenkiste, E. Moon, Immunotherapy: Beyond anti-PD-1 and anti-PD-L1 therapies. *Am. Soc. Clin. Oncol. Educ. Book* **35**, e450-458 (2016).
- D. H. Owen, L. Wei, M. A. Villalona-Calero, E. M. Bertino, K. He, P. G. Shields, D. P. Carbone, G. A. Otterson, Impact of immune-related adverse events (irAE) on overall survival (OS) in patients treated with immunotherapy for non-small cell lung cancer (NSCLC). *J. Clin. Oncol.* **35** (suppl. 15), 9080–9080 (2017).

- M. Reck, D. Rodríguez-Abreu, A. G. Robinson, R. Hui, T. Csőszi, A. Fülöp, M. Gottfried, N. Peled, A. Tafreshi, S. Cuffe, M. O'Brien, S. Rao, K. Hotta, K. Vandormael, A. Riccio, J. Yang, M. C. Pietanza, J. R. Brahmer, Updated analysis of KEYNOTE-024: Pembrolizumab versus platinum-based chemotherapy for advanced non-small-cell lung cancer with PD-L1 tumor proportion score of 50% or greater. *J. Clin. Oncol.* **37**, 537–546 (2019).
- J. R. Brahmer, D. Rodríguez-Abreu, A. G. Robinson, R. Hui, T. Csőszi, A. Fülöp, M. Gottfried, N. Peled, A. Tafreshi, S. Cuffe, M. O'Brien, S. Rao, K. Hotta, J. Zhang, G. M. Lubiniecki, A. C. Deitz, R. Rangwala, M. Reck, Health-related quality-of-life results for pembrolizumab versus chemotherapy in advanced, PD-L1-positive NSCLC (KEYNOTE-024): A multicentre, international, randomised, open-label phase 3 trial. *Lancet Oncol.* **18**, 1600–1609 (2017).
- J. Dine, R.-A. Gordon, Y. Shames, M. K. Kasler, M. Barton-Burke, Immune checkpoint inhibitors: An innovation in immunotherapy for the treatment and management of patients with cancer. *Asia Pac. J. Oncol. Nurs.* **4**, 127–135 (2017).
- C. Teixidó, N. Vilariño, R. Reyes, N. Reguart, PD-L1 expression testing in non-small cell lung cancer. *Ther. Adv. Med. Oncol.* **10**, 1758835918763493 (2018).
- L. Paz-Ares, A. Luft, D. Vicente, A. Tafreshi, M. Güümüş, J. Mazières, B. Hermes, F. Çay Şenler, T. Csőszi, A. Fülöp, J. Rodríguez-Cid, J. Wilson, S. Sugawara, T. Kato, K. H. Lee, Y. Cheng, S. Novello, B. Halmos, X. Li, G. M. Lubiniecki, B. Piperdi, D. M. Kowalski, Pembrolizumab plus chemotherapy for squamous non-small-cell lung cancer. *N. Engl. J. Med.* **379**, 2040–2051 (2018).
- A. Varga, S. Piha-Paul, P. A. Ott, J. M. Mehnert, D. Berton-Rigaud, A. Morosky, P. Yang, J. Ruman, D. Matei, Pembrolizumab in patients with programmed death ligand 1-positive advanced ovarian cancer: Analysis of KEYNOTE-028. *Gynecol. Oncol.* **152**, 243–250 (2019).
- T. A. Chan, M. Yarchoan, E. Jaffee, C. Swanton, S. A. Quezada, A. Stenzinger, S. Peters, Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. *Ann. Oncol.* **30**, 44–56 (2019).
- M. D. Hellmann, T.-E. Ciuleanu, A. Pluzanski, J. S. Lee, G. A. Otterson, C. Audigier-Valette, E. Minenza, H. Linardou, S. Burgers, P. Salman, H. Borghaei, S. S. Ramalingam, J. Brahmer, M. Reck, K. J. O'Byrne, W. J. Geese, G. Green, H. Chang, J. Szustakowski, P. Bhagavatheswaran, D. Healey, Y. Fu, F. Nathan, L. Paz-Ares, Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N. Engl. J. Med.* **378**, 2093–2104 (2018).
- T. R. Hodges, M. Ott, J. Xiu, Z. Gatalica, J. Swensen, S. Zhou, J. T. Huse, J. de Groot, S. Li, W. W. Overwijk, D. Spetzler, A. B. Heimmerger, Mutational burden, immune checkpoint expression, and mismatch repair in glioma: Implications for immune checkpoint immunotherapy. *Neuro Oncol.* **19**, 1047–1057 (2017).
- A. Marabelle, M. Fakih, J. Lopez, M. Shah, R. Shapira-Frommer, K. Nakagawa, H. C. Chung, H. L. Kindler, J. A. Lopez-Martin, W. H. Miller Jr., A. Italiano, S. Kao, S. A. Piha-Paul, J.-P. Delord, R. R. McWilliams, D. A. Fabrizio, D. Aurora-Garg, L. Xu, F. Jin, K. Norwood, Y.-J. Bang, Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: Prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol.* **21**, 1353–1365 (2020).
- T. Powles, M. Kockx, A. Rodriguez-Vida, I. Duran, S. J. Crabb, M. S. van der Heijden, B. Szabados, A. F. Pous, G. Gravis, U. A. Herranz, A. Protheroe, A. Ravaud, D. Maillet, M. J. Mendez, C. Suarez, M. Linch, A. Prendergast, P.-J. van Dam, D. Stanoeva, S. Daelemans, S. Mariathasan, J. S. Tea, K. Mousa, R. Banchereau, D. Castellano, Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the ABACUS trial. *Nat. Med.* **25**, 1706–1714 (2019).
- D. J. McGrail, P. G. Pilié, N. U. Rashid, L. Voorwerk, M. Slagter, M. Kok, E. Jonasch, M. Khasraw, A. B. Heimmerger, B. Lim, N. T. Ueno, J. K. Litton, R. Ferrarotto, J. T. Chang, S. L. Moulder, S.-Y. Lin, High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types. *Ann. Oncol.* **32**, 661–672 (2021).
- J. H. Strickler, B. A. Hanks, M. Khasraw, Tumor mutational burden as a predictor of immunotherapy response: Is more always better? *Clin. Cancer Res.* **27**, 1236–1241 (2021).
- Y. R. Murciano-Goroff, A. B. Warner, J. D. Wolchok, The future of cancer immunotherapy: Microenvironment-targeting combinations. *Cell Res.* **30**, 507–519 (2020).
- V. Verma, T. Sprave, W. Haque, C. B. Simone, J. Y. Chang, J. W. Welsh, C. R. Thomas, A systematic review of the cost and cost-effectiveness studies of immune checkpoint inhibitors. *J. Immunother. Cancer* **6**, 128 (2018).
- A. Uryvaev, M. Passhak, D. Hershkovits, E. Sabo, G. Bar-Sela, The role of tumor-infiltrating lymphocytes (TILs) as a predictive biomarker of response to anti-PD1 therapy in patients with metastatic non-small cell lung cancer or metastatic melanoma. *Med. Oncol.* **35**, 25 (2018).
- P. C. Tumeh, C. L. Harview, J. H. Yearley, I. P. Shintaku, E. J. M. Taylor, L. Robert, B. Chmielowski, M. Spasic, G. Henry, V. Ciobanu, A. N. West, M. Carmona, C. Kivork, E. Seja, G. Cherry, A. J. Gutierrez, T. R. Grogan, C. Mateus, G. Tomicic, J. A. Glaspy, R. O. Emerson, H. Robins, R. H. Pierce, D. A. Elashoff, C. Robert, A. Ribas, PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).
- S. Adams, R. J. Gray, S. Demaria, L. Goldstein, E. A. Perez, L. N. Shulman, S. Martino, M. Wang, V. E. Jones, T. J. Saphner, A. C. Wolff, W. C. Wood, N. E. Davidson, G. W. Sledge,

- J. A. Sparano, S. S. Badve, 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.* **32**, 2959–2966 (2014).
25. R. D. Bense, C. Sotiriou, M. J. Piccart-Gebhart, J. B. A. G. Haanen, M. A. T. M. van Vugt, E. G. E. de Vries, C. P. Schröder, R. S. N. Fehrmann, Relevance of tumor-infiltrating immune cell composition and functionality for disease outcome in breast cancer. *J. Natl. Cancer Inst.* **109**, djw192 (2017).
 26. A. S. Berghoff, E. Fuchs, G. Ricken, B. Mlecnik, G. Bindea, T. Spanberger, M. Hackl, G. Widhalm, K. Dieckmann, D. Prayer, A. Bilocq, H. Heinzl, C. Zielinski, R. Bartsch, P. Birner, J. Galon, M. Preusser, Density of tumor-infiltrating lymphocytes correlates with extent of brain edema and overall survival time in patients with brain metastases. *Oncotargets Ther.* **5**, e1057388 (2016).
 27. G. Corredor, X. Wang, Y. Zhou, C. Lu, P. Fu, K. Syrigos, D. L. Rimm, M. Yang, E. Romero, K. A. Schalper, V. Velcheti, A. Madabhushi, Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell lung cancer. *Clin. Cancer Res.* **25**, 1526–1534 (2019).
 28. J. Saltz, R. Gupta, L. Hou, T. Kurc, P. Singh, V. Nguyen, D. Samaras, K. R. Shroyer, T. Zhao, R. Batiste, J. Van Arnam; Cancer Genome Atlas Research Network, I. Shmulevich, A. U. K. Rao, A. J. Lazar, A. Sharma, V. Thorsson, Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep.* **23**, 181–193.e7 (2018).
 29. A. Heindl, I. Sestak, K. Naidoo, J. Cuzick, M. Dowsett, Y. Yuan, Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. *J. Natl. Cancer Inst.* **110**, 10.1093/jnci/djx137, (2018).
 30. K. A. Schalper, V. Velcheti, D. Carvajal, H. Wimmerly, J. Brown, L. Pusztai, D. L. Rimm, In situ tumor PD-L1 mRNA expression is associated with increased TILs and better outcome in breast carcinomas. *Clin. Cancer Res.* **20**, 2773–2782 (2014).
 31. K. A. Schalper, J. Brown, D. Carvajal-Hausdorf, J. M. Laughlin, V. Velcheti, K. N. Syrigos, R. S. Herbst, D. L. Rimm, Objective measurement and clinical significance of TILs in non-small cell lung cancer. *J. Natl. Cancer Inst.* **107**, dju435 (2015).
 32. H. Wimmerly, J. R. Brown, K. Schalper, H. Haack, M. R. Silver, C. Nixon, V. Bossuyt, L. Pusztai, D. R. Lannin, D. L. Rimm, PD-L1 expression correlates with tumor-infiltrating lymphocytes and response to neoadjuvant chemotherapy in breast cancer. *Cancer Immunol. Res.* **3**, 326–332 (2015).
 33. P. Johannet, N. Coudray, D. M. Donnelly, G. Jour, I. Illa-Bochaca, Y. Xia, D. B. Johnson, L. Wheless, J. R. Patrinely, S. Nomikou, D. L. Rimm, A. C. Pavlick, J. S. Weber, J. Zhong, A. Tsigirgos, I. Osman, Using machine learning algorithms to predict immunotherapy response in patients with advanced melanoma. *Clin. Cancer Res.* **27**, 131–140 (2021).
 34. L. A. Hildebrand, C. J. Pierce, M. Dennis, M. Paracha, A. Maoz, Artificial intelligence for histology-based detection of microsatellite instability and prediction of response to immunotherapy in colorectal cancer. *Cancer* **13**, 391 (2021).
 35. C. Lu, K. Bera, X. Wang, P. Prasanna, J. Xu, A. Janowczyk, N. Beig, M. Yang, P. Fu, J. Lewis, H. Choi, R. A. Schmid, S. Berezowska, K. Schalper, D. Rimm, V. Velcheti, A. Madabhushi, A prognostic model for overall survival of patients with early-stage non-small cell lung cancer: A multicentre, retrospective study. *Lancet Digit Health* **2**, e594–e606 (2020).
 36. K. AbdulJabbar, S. E. Ahmed Raza, R. Rosenthal, M. Jamal-Hanjani, S. Veeriah, A. Akarca, T. Lund, D. A. Moore, R. Salgado, M. A. Bakir, L. Zapata, C. T. Hiley, L. Officer, M. Sereno, C. R. Smith, S. Loi, A. Hackshaw, T. Marafioti, S. A. Quezada, N. M. Granahan, J. Le Quesne; TRACERx Consortium, C. Swanton, Y. Yuan, Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. *Nat. Med.* **26**, 1054–1062 (2020).
 37. A. Heindl, S. Nawaz, Y. Yuan, Mapping spatial heterogeneity in the tumor microenvironment: A new era for digital pathology. *Lab. Invest.* **95**, 377–384 (2015).
 38. N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyo, A. L. Moreira, N. Razavian, A. Tsigirgos, Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* **24**, 1559–1567 (2018).
 39. P. Courtiol, C. Maussion, M. Moari, E. Pronier, S. Pilcer, M. Sefta, P. Manceron, S. Toldo, M. Zaslavskiy, N. le Stang, N. Girard, O. Elemento, A. G. Nicholson, J.-Y. Blay, F. Galateau-Sallé, G. Wainrib, T. Clozel, Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat. Med.* **25**, 1519–1525 (2019).
 40. K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti, A. Madabhushi, Artificial intelligence in digital pathology—New tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* **16**, 703–715 (2019).
 41. H. Li, K. Bera, H. Gilmore, N. E. Davidson, L. J. Goldstein, A. Madabhushi, Abstract P5-06-16: Histomorphometric measure of disorder of collagen fiber orientation is associated with risk of recurrence in ER+ breast cancers in ECOG-ACRIN E2197 and TCGA-BRCA. *Poster Session Abstracts* (American Association for Cancer Research, 2020) [cited 21 January 2021], p. P5-06-16-P5-06-16; <http://cancerres.aacrjournals.org/lookup/doi/10.1158/1538-7445.SABCS19-P5-06-16>; doi:10.1158/1538-7445.SABCS19-P5-06-16.
 42. J. Cheng, Y. Liu, W. Huang, W. Hong, L. Wang, X. Zhan, Z. Han, D. Ni, K. Huang, J. Zhang, Computational image analysis identifies histopathological image features associated with somatic mutations and patient survival in gastric adenocarcinoma. *Front. Oncol.* **11**, 623382 (2021).
 43. F. Xie, J. Zhang, J. Wang, A. Reuben, W. Xu, X. Yi, F. S. Varn, Y. Ye, J. Cheng, M. Yu, Y. Wang, Y. Liu, M. Xie, P. du, K. Ma, X. Ma, P. Zhou, S. Yang, Y. Chen, G. Wang, X. Xia, Z. Liao, J. V. Heymach, I. I. Wistuba, P. A. Futreal, K. Ye, C. Cheng, T. Xia, Multifactorial deep learning reveals pan-cancer genomic tumor clusters with distinct immunogenomic landscape and response to immunotherapy. *Clin. Cancer Res.* **26**, 2908–2920 (2020).
 44. P. Johannet, N. Coudray, D. M. Donnelly, G. Jour, I. Illa-Bochaca, Y. Xia, D. B. Johnson, L. Wheless, J. R. Patrinely, A. C. Pavlick, J. S. Weber, J. Zhong, A. Tsigirgos, I. Osman, Using machine learning to predict immunotherapy response in advanced melanoma. *J. Clin. Oncol.* **38**(suppl. 15), 10010–10010 (2020).
 45. M. Roberts, D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, A. I. Aviles-Rivero, C. Etmann, C. M. Cague, L. Beer, J. R. Weir-McCall, Z. Teng, E. Gkrania-Klotsas; AIX-COVNET, J. H. F. Rudd, E. Sala, C.-B. Schönlieb, Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat. Mach. Intell.* **3**, 199–217 (2021).
 46. K. Bera, N. Braman, A. Gupta, V. Velcheti, A. Madabhushi, Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat. Rev. Clin. Oncol.* **19**, 132–146 (2021).
 47. A. H. Beck, A. R. Sangoi, S. Leung, R. J. Marinelli, T. O. Nielsen, M. J. van de Vijver, R. B. West, M. van der Rijn, D. Koller, Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci. Transl. Med.* **3**, 108ra113 (2011).
 48. E. Brambilla, G. le Teuff, S. Marguet, S. Lantuejoul, A. Dunant, S. Graziano, R. Pirker, J.-Y. Douillard, T. le Chevalier, M. Filipits, R. Rosell, R. Kratzke, H. Popper, J.-C. Soria, F. A. Shepherd, L. Seymour, M. S. Tsao, Prognostic effect of tumor lymphocytic infiltration in resectable non-small-cell lung cancer. *J. Clin. Oncol.* **34**, 1223–1230 (2016).
 49. J. Hu, C. Cui, W. Yang, L. Huang, R. Yu, S. Liu, Y. Kong, Using deep learning to predict anti-PD-1 response in melanoma and lung cancer patients from histopathology images. *Transl. Oncol.* **14**, 100921 (2021).
 50. C. Rudin, Please stop explaining black box models for high stakes decisions. arXiv:1811.10154 [stat.ML] (25 November 2018).
 51. J. van der Laak, G. Litjens, F. Ciompi, Deep learning in histopathology: The path to the clinic. *Nat. Med.* **27**, 775–784 (2021).
 52. A. Hosny, C. Parmar, T. P. Coroller, P. Grossmann, R. Zeleznik, A. Kumar, J. Bussink, R. J. Gillies, R. H. Mak, H. J. W. L. Aerts, Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLoS Med.* **15**, e1002711 (2018).
 53. E. G. Peyster, S. Arabyarmohammadi, A. Janowczyk, S. Azarianpour-Esfahani, M. Sekulic, C. Cassol, L. Blower, A. Parwani, P. Lal, M. D. Feldman, K. B. Margulies, A. Madabhushi, An automated computational image analysis pipeline for histological grading of cardiac allograft rejection. *Eur. Heart J.* **42**, 2356–2369 (2021).
 54. E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, J. Verweij, New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
 55. A. Janowczyk, R. Zuo, H. Gilmore, M. Feldman, A. Madabhushi, HistoQC: An open-source quality control tool for digital pathology slides. *J. Clin. Cancer Inform.* **3**, 1–7 (2019).
 56. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi, Eds. (Springer International Publishing, 2015) [accessed 24 February 2020], pp. 234–241; http://link.springer.com/10.1007/978-3-319-24574-4_28; doi:10.1007/978-3-319-24574-4_28.
 57. W.-Y. Kim, Y.-S. Kim, A region-based shape descriptor using Zernike moments. *Signal Process. Image Commun.* **16**, 95–102 (2000).

Acknowledgments

Funding: Research reported in this publication was supported by the National Cancer Institute under award numbers 1U24CA199374-01, R01CA249992-01A1, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, R01CA257612-01A1, 1U01CA239055-01, 1U01CA248226-01, and 1U54CA254566-01; National Heart, Lung and Blood Institute 1R01HL15127701A1; National Institute of Biomedical Imaging and Bioengineering 1R43EB028736-01; National Center for Research Resources under award number 1 C06 RR12463-01; VA Merit Review Award IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service; the Office of the Assistant Secretary of Defense for Health Affairs, through the Breast Cancer Research Program (W81XWH-19-1-0668); the Prostate Cancer Research Program (W81XWH-15-1-0558 and W81XWH-20-1-0851); the Lung Cancer Research Program (W81XWH-18-1-0440 and W81XWH-20-1-0595); the Peer Reviewed Cancer Research Program (W81XWH-18-1-0404); the Kidney Precision Medicine Project (KPMP) Glue Grant; the Ohio Third Frontier Technology Validation Fund; the Clinical and Translational Science Collaborative of Cleveland (UL1TR0002548) from the National Center for Advancing Translational Sciences (NCATS) component of the NIH and NIH Roadmap for Medical Research; The Wallace H. Coulter Foundation Program in the Department of

Biomedical Engineering at Case Western Reserve University; and sponsored research agreements from Bristol Myers Squibb, Boehringer Ingelheim, and AstraZeneca. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, the U.S. Department of Veterans Affairs, the Department of Defense, or the U.S. government. **Author contributions:** X.W., C.B., K.B., V.S.V., C.L., P.V., P.F., and A.M. designed and conducted the experiments. X.W., K.B., C.B., and A.M. wrote the manuscripts. M.Y., S.A.-E., C.K., M.D.F., H.M., K.A.S., and V.V. acquired and provided the image and clinical data, as well as the medical writing guidance. All authors read and approved the final version of the manuscript and have verified the underlying data. **Competing interests:** X.W., C.B., K.B., V.S.V., S.A.-E., C.K., P.V., M.D.F., M.Y., P.F., H.M., and C.L. disclosed no conflicts of interest. K.A.S. has received research funding from Navigate Biopharma, Tesaro/GSK, Moderna Inc., Takeda, Surface Oncology, Pierre-Fabre Research Institute, Merck Sharp & Dohme, Bristol Myers Squibb, AstraZeneca, Ribon Therapeutics, Akoya Biosciences, Boehringer Ingelheim, and Eli Lilly and honoraria for consultancy, advisory, or speaker roles from Shattuck Labs, Pierre-Fabre, EMD Serono, Moderna Inc., Clinica Alemana de Santiago, Genmab, Takeda, Merck Sharp & Dohme, Bristol Myers Squibb, AstraZeneca, Agenus, Repertoire Therapeutics, OnCusp, and Ariagen. For activities related to the present article, V.V. disclosed no relevant relationships. For activities not related to the present article, V.V. is a consultant for BMS, Genentech, AstraZeneca, Celgene, Foundation Medicine, Takeda, Merck, Alkermes, and Nektar Therapeutics; institution has grants or grants pending with AstraZeneca, Merck, BMS, Genentech, and Alkermes; is on

the speakers bureaus of Novartis, BMS, Celgene, and Foundation Medicine; and has received payment for the development of educational presentations from BMS and Foundation Medicine. For other relationships, V.V. disclosed no relevant relationships. For activities related to the present article, A.M. disclosed no relevant relationships. For activities not related to the present article, A.M. is on the board of and is a consultant to Inspirata; institution has three NCI RO1 grants with Inspirata and one ongoing U24 with Pathcore; Elucid Bioimaging and Inspirata has licensed some of institution patents (both Case Western Reserve University and Rutgers University); receives royalties from Elucid Bioimaging and Inspirata; and holds stock equity in Elucid Bioimaging and Inspirata. A.M., X.W., C.B., and V.V. are inventors on two patents related to this work filed by Case Western Reserve University (U.S. Patent No. 11055844, filed on 21 February 2019 and issued on 6 July 2021, and U.S. Patent No. 10902256, filed on 15 February 2019 and issued on 26 January 2021). The authors declare no other competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 23 November 2021

Accepted 14 April 2022

Published 1 June 2022

10.1126/sciadv.abn3966

Spatial interplay patterns of cancer nuclei and tumor-infiltrating lymphocytes (TILs) predict clinical benefit for immune checkpoint inhibitors

Xiangxue Wang, Cristian Barrera, Kaustav Bera, Vidya Sankar Viswanathan, Sepideh Azarianpour-Esfahani, Can Koyuncu, Priya Velu, Michael D. Feldman, Michael Yang, Pingfu Fu, Kurt A. Schalper, Haider Mahdi, Cheng Lu, Vamsidhar Velcheti, and Anant Madabhushi

Sci. Adv., **8** (22), eabn3966.
DOI: 10.1126/sciadv.abn3966

View the article online

<https://www.science.org/doi/10.1126/sciadv.abn3966>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.
Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).