

RESEARCH

Open Access



DeepRisk network: an AI-based tool for digital pathology signature and treatment responsiveness of gastric cancer using whole-slide images

Mengxin Tian^{1,2†}, Zhao Yao^{4,5†}, Yufu Zhou^{6†}, Qiangjun Gan^{1,2†}, Leihao Wang^{1,2}, Hongwei Lu^{4,5}, Siyuan Wang^{1,2}, Peng Zhou^{1,2}, Zhiqiang Dai^{7,8}, Sijia Zhang^{1,2}, Yihong Sun^{1,2,3}, Zhaoqing Tang^{1,2,7*}, Jinhua Yu^{4,5*} and Xuefei Wang^{1,2,3,7,8*}

Abstract

Background Digital histopathology provides valuable information for clinical decision-making. We hypothesized that a deep risk network (DeepRisk) based on digital pathology signature (DPS) derived from whole-slide images could improve the prognostic value of the tumor, node, and metastasis (TNM) staging system and offer chemotherapeutic benefits for gastric cancer (GC).

Methods DeepRisk is a multi-scale, attention-based learning model developed on 1120 GCs in the Zhongshan dataset and validated with two external datasets. Then, we assessed its association with prognosis and treatment response. The multi-omics analysis and multiplex Immunohistochemistry were conducted to evaluate the potential pathogenesis and spatial immune contexture underlying DPS.

Results Multivariate analysis indicated that the DPS was an independent prognosticator with a better C-index (0.84 for overall survival and 0.71 for disease-free survival). Patients with low-DPS after neoadjuvant chemotherapy responded favorably to treatment. Spatial analysis indicated that exhausted immune clusters and increased infiltration of CD11b⁺CD11c⁺ immune cells were present at the invasive margin of high-DPS group. Multi-omics data from the Cancer Genome Atlas-Stomach adenocarcinoma (TCGA-STAD) hint at the relevance of DPS to myeloid derived suppressor cells infiltration and immune suppression.

Conclusion DeepRisk network is a reliable tool that enhances prognostic value of TNM staging and aid in precise treatment, providing insights into the underlying pathogenic mechanisms.

[†]Mengxin Tian, Zhao Yao, Yufu Zhou and Qiangjun Gan are contributed equally to this work.

*Correspondence:

Zhaoqing Tang

tang.zhaoqing@zs-hospital.sh.cn

Jinhua Yu

jhyu@fudan.edu.cn

Xuefei Wang

wang.xuefei@zs-hospital.sh.cn

Full list of author information is available at the end of the article



Keywords Deep learning, Gastric cancer, Whole slide image, Artificial intelligence, Suppressive immune microenvironment

Introduction

Gastric cancer (GC) is one of the most lethal malignancies worldwide with poor prognosis [1]. It is extensively heterogeneous at the histological, molecular and genetic levels, making it challenging to treat effectively [2]. Radical gastrectomy remains the mainstay of curative treatment for GCs. Unfortunately, the prognosis after surgery is unsatisfactory due to a high incidence of recurrence and/or distant metastases [3]. Therefore, improvements in the identification of patients with poor clinical outcomes are needed for the optimization of therapeutic strategies and better allocation of adjuvant therapies.

In past decades, growing evidences proved that tumour microenvironment and tumour biology behavior play an important role in gastric cancer progression and prognostic prediction, such as the four-factor immunoscore system [4] and seven-gene signature [5]. However, in clinical practice, these clinical tests are costly and time-consuming, limiting their clinical applications in all patients. Recently, new pathological features such as collagen signature [6], tertiary lymphoid structures [7] and high endothelial venules [8] have been investigated as novel prognostic factors, indicating that pathomics of cancer tissues contain a wealth of information that can provide benefits to the patient beyond diagnosis [9]. Further understanding of the prognostic value of GC pathological features may facilitate prognostic stratification and the establishment of treatment strategies in clinical practice.

Deep learning has made it possible to perform complex analysis of high-resolution histopathology images and to explore the potential correlations between pathological factors and carcinogenesis in various types of cancer, such as colorectal cancer [10] and liver cancer [11, 12]. Accumulating evidence indicates that deep learning algorithms applied to routine histological data can help diagnose origins of cancers [13], detect Epstein–Barr virus status in gastric cancer [14], and predict the therapeutic response [15]. Recently, Wang et al. showed that deep learning could also be used for histological grading based on whole-slide images (WSIs) in breast cancer [16]. Studies have also proven that deep learning is able to identify significant prognostic factors from digital WSIs [11], and to assist in the allocation of adjuvant therapies to patients with a high risk of recurrence in primary melanoma tumours [17]. With respect to GC, a deep learning-based approach based on ResNet18 was applied to hematoxylin and eosin (H&E)-stained histopathological

slides of tissue samples from GC patients and allowed the prediction of microsatellite instability (MSI) [18]. A more recent study applied nomogram to build a machine learning model to predict GC prognosis, which used randomly selected image patches from WSI for modeling [19]. However, the existing methods for WSI analysis faces many limitations. For methods that identify different types of tissue regions, manual annotation is required and therefore time-consuming and subjective, leading to reduced generalization performance [12, 13, 20]. Methods that require manual design of features are often suboptimal because the designed features lack generalization and are not very expressive [17, 19]. In addition, most models lack the fusion of features at different magnification, which is critical for the pathologist's examination [14–16]. Further, most artificial intelligence (AI)-based pathology biomarkers mainly focused on different aspects of prediction, lacking the exploration of underlying biological mechanism.

In the present study, we sought to establish a fully automated histopathological slide analysis pipeline that can use original scanned WSIs as input, automatically locate the tissue area, and output a risk score without any manual annotation. We established a deep risk network (DeepRisk) to obtain a novel digital pathology signature (DPS) as the model output using WSIs from the Zhongshan dataset, and then validated the model performance using 2 external datasets. To explore the underlying pathogenic mechanism and clinical significance of DPS, we also investigated the relationships of the DPS with tumour immune infiltration, spatial immune contexture, as well as treatment response of neoadjuvant treatment.

Materials and methods

Datasets

In our study, we developed a dataset for training and testing the DeepRisk model and determining the DPS values. A total of 1120 patients with a pathologically confirmed diagnosis of GC who underwent surgical resection between April 2006 and November 2019 at Zhongshan Hospital of Fudan University (Shanghai, China) were enrolled in this study. Eligible patients were required to have clinicopathological data and follow-up data, not to have any other malignancies.

The Cancer Genome Atlas (TCGA) and Shanghai Outdo Biotech Company (SOBC) datasets were used as external independent datasets for validation of the

deep learning model. We collected clinical, pathological and prognostic factor data for 268 patients with OS ≥ 1 month in the TCGA–stomach adenocarcinoma (STAD) dataset and downloaded the associated WSIs from the public website (<https://portal.gdc.cancer.gov/projects/TCGA-STAD>) [20]. After excluding WSIs with poor image quality (such as poor staining or tissue overlap), 536 WSIs were available for further analysis. The SOBC dataset, obtained from SOBC with the approval of the Institutional Review Board, only consisted of tissue microarray (TMA) spots. Paired TMAs of tumour and normal tissues from 277 GC patients were scanned using a Hamamatsu (Hamamatsu Photonics, Hamamatsu City, Japan) scanner at 20 \times magnification. All available core samples from each patient were used for analysis as previously described [21]. Due to a lack of tumour recurrence information in the TCGA-STAD and SOBC datasets, OS was used as the primary outcome measure in the survival analysis during model validation.

Tumour characterization and follow-up

Tumour sample collection, postoperative surveillance, adjuvant treatment, and recurrence management were performed as previously described [22]. Tumours were staged according to the 8th edition of the Union for International Cancer Control (UICC)/TNM classification [23]. GC patients were stratified according to tumour differentiation (differentiated or undifferentiated), tumour stage (I-II or III-IV), and adjuvant treatment (received or not received) [24]. The classification of pathological tissue regression after neoadjuvant chemotherapy depends on Ninomiya and Ryan classification systems [25, 26]. The follow-up duration was measured from the time of surgery to the last follow-up date, and patient survival information was collected at the last follow-up. Patients who had not experienced recurrence or were alive at the last follow-up were censored at the last follow-up.

Slide preparation

H&E-stained histopathological slides were prepared as previously described [27]. All slides in the Zhongshan dataset were scanned and digitized using Aperio (Leica Biosystems, Wetzlar, Germany) and Hamamatsu (Hamamatsu Photonics, Hamamatsu City, Japan) digital slide scanners at 40 \times magnification. One or two representative digital slides were selected and analysed for each GC patient.

WSI processing

A threshold-based algorithm was applied for the detection and segmentation of stained tumour tissue on digitized slides [28], which were processed after 8 \times downscaling due to the considerable heterogeneity in resolution and

computational processing. Tissue regions in the approximate contour obtained through segmentation were exhaustively tiled with 256 \times 256 patches at different magnification levels (5 \times , 10 \times and 20 \times). The overlap ratio between adjacent patches was set to 0 at each magnification level during model training, testing and validation, but an overlap ratio of 50% was applied to improve the resolution for feature visualization.

To reduce the computational complexity and extract the information in the image patches, a deep convolutional neural network (CNN) was firstly used to extract low-dimensional features for modelling [29]. Specifically, we used transfer learning strategy and a Resnet50 model pretrained on ImageNet to extract feature maps. As a commonly used backbone model, the pre-trained Resnet50 model is able to extract visual representation from image patches [30]. Average pooling was applied to the output feature maps to obtain 1024-dimensional feature vectors [31]. In this way, each 256 \times 256 patch was converted into a 1024-dimensional feature vector, making it feasible to simultaneously fit thousands of patches for a single slide into GPU memory.

Attention-based model building and risk score prediction

The DeepRisk was optimized using the features of all patches of a single case as model inputs and using survival time and survival status as labels. A multiple instance learning (MIL) scheme is used to aggregate the features of all patches of the WSIs for a single case and predict the bag label [32]. The model integrates N patch features of a WSI by using an attention pooling module; by predicting N corresponding attention scores, the attention module enables the model to unambiguously learn which morphological features should be considered prognostic indicators of high risk. Specifically, given a patient with a single slide or multiple slides, we denote instance-level features embedding in the bag with B instances as $\{f_i, i = 1, 2, \dots, B\}$. Following the attention module, the attention score of instance k is calculated as follows:

$$a_k = \frac{\exp\{w_k^T (\tanh(Vf_k^T) \odot \text{sigmoid}(Uf_k^T))\}}{\sum_{j=1}^K \exp\{w_j^T (\tanh(Vf_j^T) \odot \text{sigmoid}(Uf_j^T))\}}$$

where U and V are learnable parameters, \odot is element-wise multiplication, and $\tanh(\cdot)$ and $\text{sigmoid}(\cdot)$ are nonlinearities.

The case-level presentation aggregating all instance embedding features is given by:

$$h_{case} = \sum_{k=1}^B a_k f_k$$

Then the prediction results are given via the fully connection layer by $p_{case} = W_c h_{case}$, where W_c is the trainable parameters in fully connection layer. The training process is supervised using the cross entropy-based Cox proportional loss function following for survival analysis (detailed in Additional file 1: Methods S1).

Integration of multi-scale features and clinical features

Analogous to the diagnostic process used by pathologists to obtain pathological image information in the clinic, DeepRisk contains a hierarchical network to extract WSI features under different magnification scales. Specifically, the WSIs were preprocessed at different magnifications (5×, 10× and 20×) to extract image patches and corresponding feature embeddings. Then, the attention module corresponding to each magnification aggregated image patches into case-level features. Finally, we concatenated the case-level features at different magnifications and input them to the fully connection (FC) layer for generating DPS. Specifically, the survival time of patients is subdivided into n intervals $[0, a_1), [a_1, a_2), \dots, [a_{n-1}, \infty)$, which results in a discrete scale $t = 1, 2, \dots, n$. And each time point $t \in \{1, \dots, n\}$ is represented by a neuron in the FC layer. Then the survival function can be calculated by

$$S = \prod 1 - p_{case}$$

where p_{case} is the discrete hazard of FC layer output. The risk related DPS can be expressed as

$$PRS_{GC} = - \sum S$$

Which is positively correlated with case risk (detailed in Additional file 1: Methods S1). For the fusion of clinical information, we concatenated the gender and normalized age of each case with the corresponding WSI features.

Visualization of prognosis-related features

To visually interpret the clues that DeepRisk used to predict prognosis from the WSIs, we analysed the clues hierarchically from three perspectives. Firstly, deep features were extracted from the penultimate hidden layer of the model, and t-distributed stochastic neighbor embedding (t-SNE) was used to reduce the dimension of the feature vector to visualize the distribution of each case in the feature space. Second, the attention scores across all patches for each WSI were computed and were converted to percentile scores between 0.0 (low attention) to 1.0 (high attention) (Additional file 1: Fig. S1). Top-scored patches (100 patches for each case) were extracted and their feature embedding was reduced to two dimensions with t-SNE and plotted using hexagonal heat map. Finally, we used QuPath

software (version 0.2.3) to segment cells in the patches and trained a support vector machine (SVM) model to identify different cell types. We determined the proportions of different cell types in GC tumour samples from patients with different risk prognoses (for more details, refer to Additional file 1: Methods S1).

Tumour immune infiltration, differential gene expression and pathway enrichment analysis.

To explore the potential correlation between the DPS and immune cell infiltration, the TCGA-STAD dataset was divided into 2 subgroups: low- and high-DPS groups. As we previously described [33], 22 human immune phenotypes of TCGA-STAD dataset were analyzed based on the CIBERSORT algorithm and the LM22 gene signature. We used the DESeq2 package (version 1.38.0) from Bioconductor in R environment (version 3.6.1, <https://www.rproject.org/>) to perform differential gene expression analysis, and set $P_{adj} < 0.05$ and a $|\log_2FC| \geq 1.5$ as screening criteria [34]. We next performed the pathway enrichment analysis of those differentially expressed genes by using clusterProfiler V3.19.0 [35] (enricher function, KEGG gene sets, or cancer hallmark gene sets from msigdb). $P=0.05$ was set as the cut-off value.

Multiplex Immunohistochemistry (mIHC) and data analysis

To investigate the composition of different immune cells in GC, 12 tissue samples were enrolled and conducted with mIHC analysis, every GC tissue contains normal region, tumor region and invasive margin (IM, 500 μ m width on each side of the intra- and peri-tumor interface) [36]. Among these samples, 6 were from high-DPS GCs and another 6 from low-DPS GCs. Consecutive slides from these samples were stained by mIHC and H&E. The mIHC was performed on consecutive slides by using the Opal 7-color kit (Akoya Biosciences) as previously described [37]. Three mIHC panels were employed to characterize different subsets of tumour-infiltrating immune cells (TIICs), including panel 1 (lymphocytes): CD3, CD4, CD8, CD16, CD56 and Foxp3; panel 2 (myeloid cells and B cells): CD11b, CD11c, CD20, CD45RO, CD68, and MPO; panel 3 (immune inhibitory molecules): PD-1, PD-L1, LAG3, TIM3, CTLA4, and IDO. Detailed information of primary antibodies is provided in the Additional file 1: Table S1. The slides were scanned and imaged using a Vectra 3.0 Quantitative Pathology Imaging System (Perkin Elmer, Waltham, MA). Image analysis software, including InForm 2.3 (Perkin Elmer) and HALO (Indica Labs), was used for spectral unmixing, cell segmentation, and the identification and quantification of cellular subsets [38]. The fraction of each lineage of cells

was normalized to the number of tumour cells in each analysed field [39].

Statistical analysis

Percentages or median values are presented as descriptive summary statistics. Pearson's χ^2 test or Fisher's exact test was employed to compare categorical variables. The Wilcoxon rank sum test or Student's *t* test was used to evaluate continuous variables. OS and disease-free survival (DFS) were estimated by the Kaplan–Meier method and compared using the log-rank test. Cox's proportional hazards regression model was used to analyze independent prognostic factors. Model performance was assessed with Harrell's concordance index (C-index). The predictive accuracy of DPS was evaluated by the integrated area under the receiver operating characteristic (ROC) curve (iAUC) with 1000×bootstrap resampling [40]. Correlation coefficients were computed by Spearman and distance correlation analyses. All correlation heatmaps were generated using the pheatmap function (<https://github.com/raivokolde/pheatmap>). $p < 0.05$ was considered to indicate statistical significance.

Results

Clinicopathological characteristics of patients

Demographics and clinical characteristics of GC patients, including tumour location, tumour differentiation, Lauren type and TNM stage, are summarized in Additional file 1: Table S2. In Zhongshan dataset, death was reported in 215 patients, 37.6% of whom had tumour recurrence, while the recurrence rate was 25.0% among the other 905 patients who survived. The median follow-up duration was 110.8 months (range: 1.5–147.0 months). The 1-, 3- and 5-year overall survival (OS) rates were 94.2%, 77.0% and 70.0%, respectively, and the 1-, 3- and 5-year recurrence rates were 20.0%, 42.5% and 53.9%, respectively.

Model construction of DeepRisk network

The model was built based on the Zhongshan dataset ($n=1120$) and further validated in the TCGA-STAD ($n=268$) and SOBC datasets ($n=277$). We performed threefold cross-validation, which split cases randomly (1:1:1) in Zhongshan cohort into training set, validation set and testing set, and this was repeated 3 times to obtain all cases results. All WSIs were first preprocessed to extract image patches at different magnifications, and a total of about 18 million $256 \times 256 \times 20 \times$ image patches were extracted for further analysis (Fig. 1A). The training process was supervised by patient overall survival time and status. All patch features are integrated by attention scores in a weighted summation to get the case features, and the case risk is predicted by a FC layer. The DPS is the case risk accumulation. Once the training process

is finished, in the inference stage the model outputs the DPS based on the input WSIs. C-index via the DPS was used as evaluation metric for the performance of different models. Then, we examined the predictive performance of DeepRisk using the Zhongshan dataset. The DeepRisk model exhibited C-index of 0.781, 0.812 and 0.821 at $5 \times$, $10 \times$ and $20 \times$ magnification, respectively, based on WSI features alone (Fig. 1B). The combination of the three magnification scales ($5 \times$, $10 \times$ and $20 \times$) showed consistently better performance (C-index=0.828) than each single scale. Furthermore, the incorporation of clinical factors (age and sex) with WSI features significantly improved model performance (C-index) at every magnification scale compared with WSI features only ($p < 0.001$). The combination of the three scales plus clinical factors achieved a C-index as high as 0.856. These results indicate that DeepRisk effectively captures multiscale patch information, and its performance is further enhanced by adding clinical demographic features.

Prognostic value of the DPS in Zhongshan dataset

With the DPS as the model output, the patients were dichotomized into high- and low-DPS groups based on various DPS cut-off values. Our model demonstrated significant separation in OS and DFS outcomes across a wide range of DPS cut-off values (ranging from 5 to 95%), with hazard ratios (HRs) > 5 for OS and > 1 for DFS (Fig. 1C, D). Given a relatively even distribution of patients across the continuum of DPS values, 50% was chosen as the cut-off to maximize the number of patients in each risk group. Patients with high-DPS had significantly worse outcomes with a lower survival rate and higher recurrence rate than those with low-DPS (Fig. 1E, F).

Multivariate Cox regression analysis revealed that a high-DPS was significantly associated with unfavourable OS (HR: 5.41, 95% CI 3.55–8.24, $p < 0.0001$) and DFS (HR: 1.97, 95% CI 1.59–2.44, $p < 0.0001$), independent of other clinicopathological features (Table 1). Notably, subanalyses stratifying patients into subgroups based on different clinicopathological characteristics consistently demonstrated superior survival outcomes (OS and DFS) for high-DPS patients (Figs. 1G, H, 2 and Additional file 1: Fig. S2), indicating that the DPS strongly correlates with patient prognosis and is nonredundant with other prognosis-related baseline variables.

To evaluate the prognostic accuracy of the DPS, iAUC analysis with 1000×bootstrap resampling was conducted. The C-index of DPS in predicting OS (0.84, 95% CI: 0.81–0.87) and DFS (0.71, 95% CI: 0.68–0.75) was superior to that of existing clinicopathological parameters ($p < 0.0001$), including TNM stage, tumour differentiation, tumour location, and Lauren type

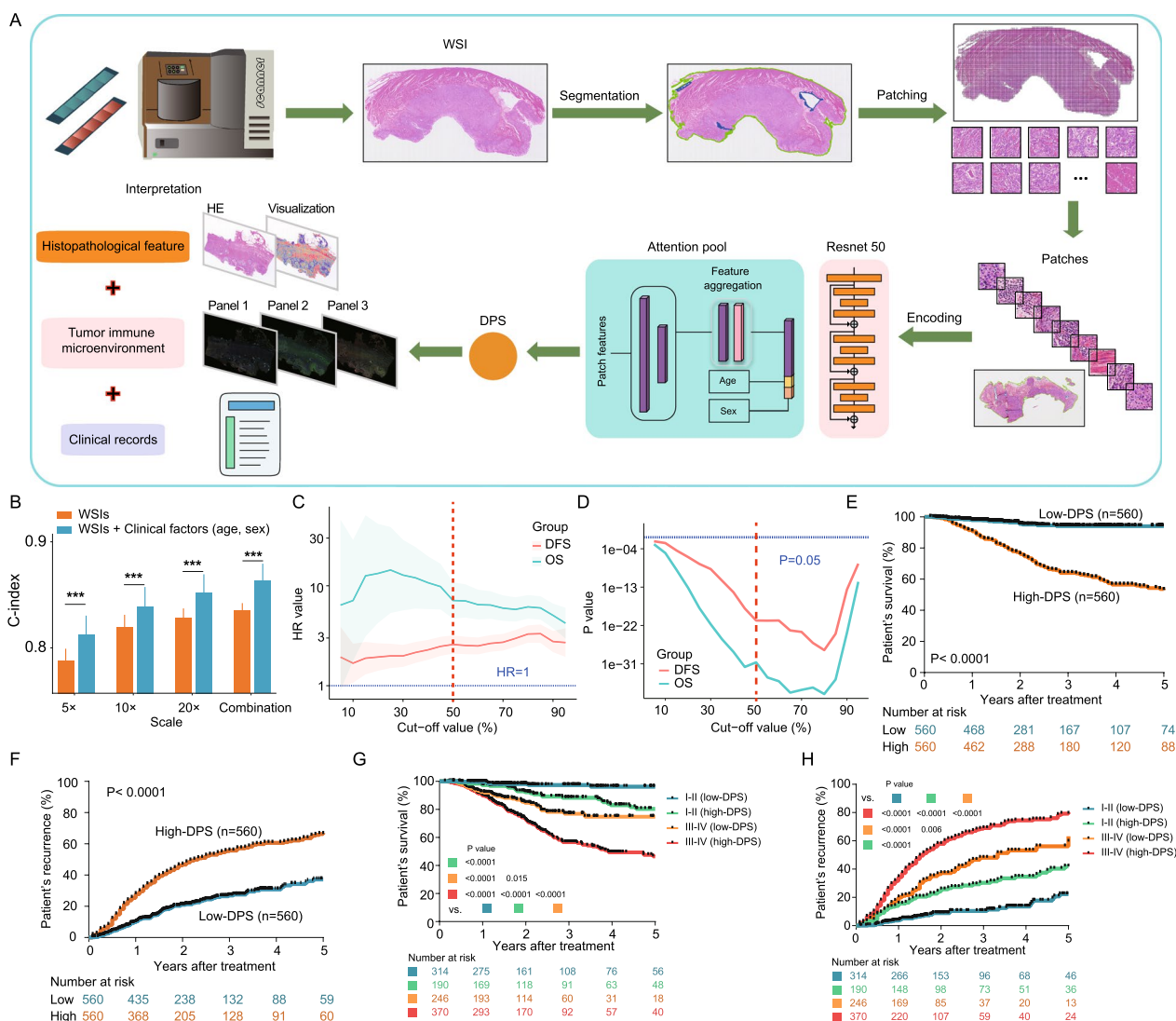


Fig. 1 Workflow for DeepRisk model building and evaluation of DPS performance. **A** Underwent segmentation and patching process from WSIs, all the patches were encoded using a deep CNN model into descriptive feature representations. A pre-trained ResNet50 model was used to extract feature maps, to which average pooling was applied to obtain feature vectors. Attention-based MIL was used to aggregate all patch features of a single case and deliver an output label. Then, we built the DeepRisk network without annotation and further validated in 2 external cohorts (TCGA-STAD and SOBC). Histopathological features, immune contexture, transcriptomics and clinical information were used to investigate the correlations between model output (DPS) and underlying GC features. **B** C-indices based on DPS under different magnification scales, with and without incorporation of demographic factors (age and gender). **C, D** Performance of DPS in predicting patient survival in the Zhongshan dataset, using a series of DPS cut-off values for patient dichotomising. **E, F** Kaplan–Meier curves of OS and DFS for high-DPS (>50%) and low-DPS (≤50%). **G, H** Kaplan–Meier curves for survival and recurrence of DPS at different TNM stages. ***p < 0.001. CNN, convolutional neural network; GC, gastric cancer; WSI, whole-slide image; DPS, digital pathology signature; MIL, multi-instance learning; TCGA, The Cancer Genome Atlas

(Fig. 3A, B). Furthermore, by calculating the continuous net reclassification improvement (NRI) for 5-year postoperative mortality, we combined the TNM stage and the DPS into a single model and compared its predictive ability to TNM stage alone. The combined

model containing the DPS showed a continuous NRI of 34.5% (95%CI: 0.209–0.455; p = 0.002), suggesting that the novel DPS may serve as a potential complement to the TNM staging system.

Table 1 Univariate and multivariate cox analysis of disease-free survival and overall survival in the Zhongshan dataset

Variables	Overall survival						Disease-free survival					
	Univariate			Multivariate			Univariate			Multivariate		
	HR	95% CI	P	HR	95% CI	P	HR	95% CI	P	HR	95% CI	P
Age (>60 vs ≤60 years)	1.32	1.01–1.73	0.043	1.18	0.89–1.55	0.245	1.11	0.91–1.34	0.306			
Sex (Female vs male)	0.79	0.59–1.08	0.136				0.80	0.65–1.00	0.045	0.94	0.76–1.17	0.60
Tumor location	1.05	0.88–1.25	0.582				0.90	0.80–1.02	0.099			
Tumor size (>4 vs ≤4)	2.74	2.09–3.59	<0.0001	2.11	1.49–2.99	<0.0001	2.11	1.74–2.56	<0.0001	1.33	1.09–1.63	0.005
Lauren type	1.18	1.00–1.40	0.047	1.05	0.88–1.25	0.61	1.06	0.94–1.20	0.303			
LVI (positive vs negative)	2.31	1.74–3.06	<0.0001	1.39	1.03–1.87	0.030	1.92	1.57–2.33	<0.0001	1.26	1.02–1.55	0.029
PNI (positive vs negative)	2.45	1.80–3.33	<0.0001	1.44	1.03–2.00	0.031	2.13	1.72–2.64	<0.0001	1.28	1.01–1.61	0.040
TNM stage (III–IV vs I–II)	5.09	3.57–7.26	<0.0001	2.72	1.83–4.05	<0.0001	3.91	3.11–4.91	<0.0001	2.50	1.93–3.23	<0.0001
Tumor differentiation (Undifferentiated vs differentiated)	1.03	0.79–1.35	0.83				0.80	0.66–0.97	0.024	0.69	0.57–0.84	0.0002
DPS (high vs low)	7.12	4.69–10.81	<0.0001	5.41	3.55–8.24	<0.0001	2.58	2.09–3.18	<0.0001	1.97	1.59–2.44	<0.0001

LVI, Lymphovascular invasion; PNI, Peripheral nerve invasion; DPS, digital pathology signature; TNM, tumor, node, and metastasis

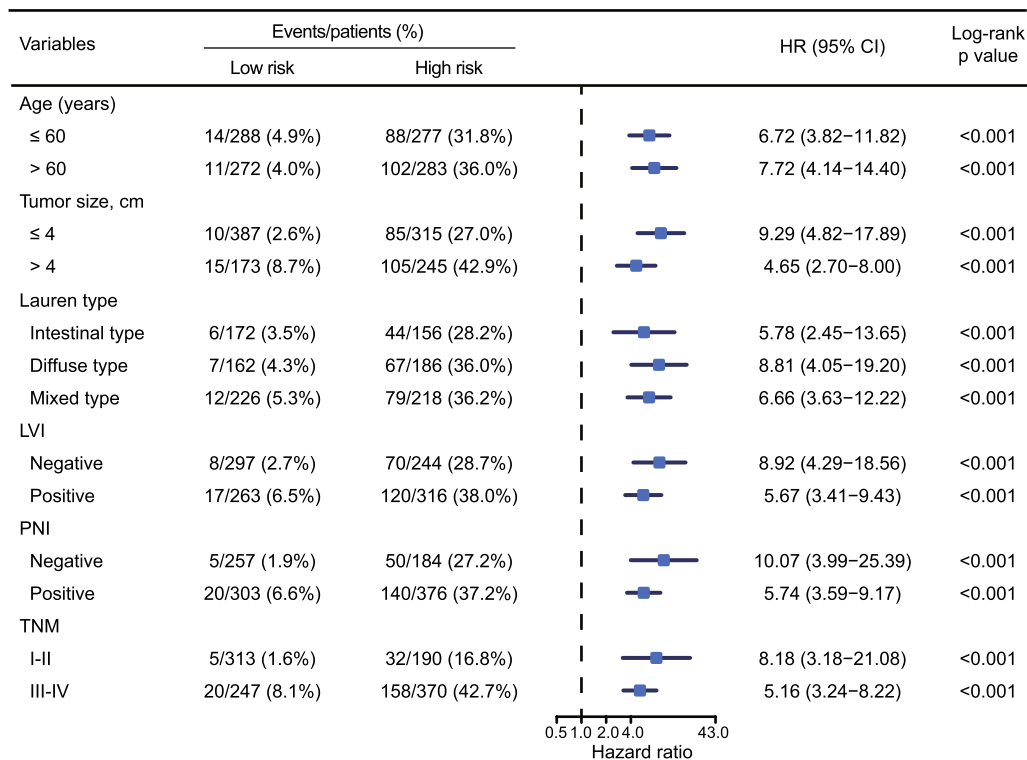


Fig. 2 Forest plot of DPS for the Zhongshan dataset in overall survival analysis

Validation of the DPS

Next, the DPS was validated in the TCGA-STAD dataset (268 patients with 583 WSIs). After dividing the TCGA-STAD dataset using DPS 50th percentile as a cut-off value, Kaplan–Meier analysis indicated that both groups were well stratified (p=0.006, Fig. 1F). The C-index of the

DPS in predicting OS was 0.74 (95% CI 0.67–0.80). Multivariable Cox analysis remained an independent predictor of OS in the TCGA-STAD dataset (HR: 1.76, 95% CI 1.04–2.98, p=0.036, Additional file 1: Table S3). Moreover, the predictive power of the DPS was statistically superior to that of age, sex, tumour differentiation and

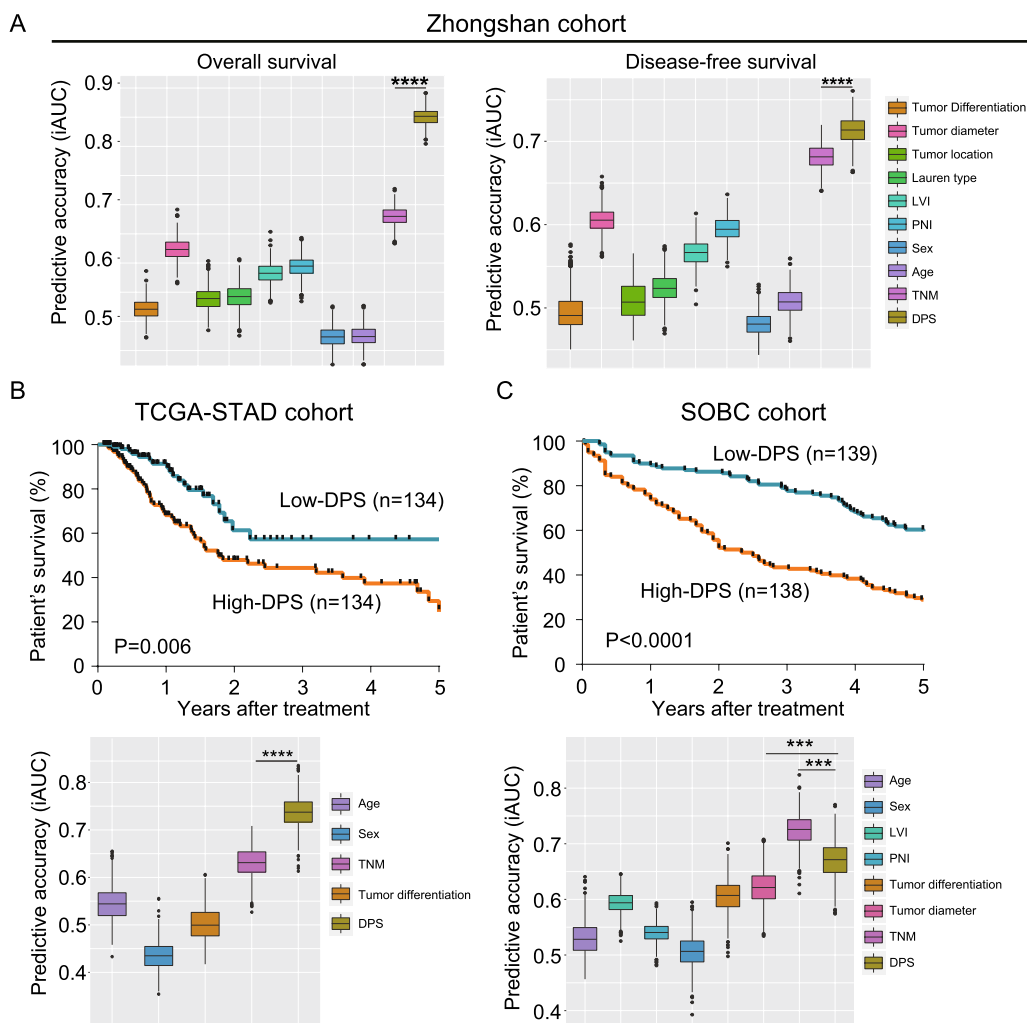


Fig. 3 Performance of DPS compared with clinicopathological parameters in three independent datasets. **A** Clinical performance of tumour-related parameters and DPS. The predictive accuracy for survival and recurrence based on the iAUC with 1000×bootstrap resampling for each parameter is shown in a box plot. **B, C** Kaplan–Meier curves and predictive accuracy for overall survival in TCGA-STAD and SOBC cohorts

TNM stage ($p < 0.0001$, Fig. 3C). Although the majority of patients in the TCGA-STAD dataset had a short follow-up duration, TNM stage + DPS had significantly better predictive performance for the 1-year postoperative mortality rate, with an NRI of 0.368 (95% CI 0.074–0.518, $p = 0.032$). These findings suggest that the DPS could add prognostic value to the TNM staging system. To investigate the prognostic performance of DPS in tissue microarray (TMA) with limited pathological information, the SOBC dataset (277 patients with 552 TMA spots) was also enrolled. Similar trends were observed in the survival stratification ($p < 0.0001$, Fig. 1G) and prognostic prediction (C-index, 0.67, 95% CI 0.61–0.74; Fig. 3D and Additional file 1: Table S3).

To investigate whether the DPS is superior to existing gene or immune signatures, we compared the DPS with previous prognostic models using the mRNA expression

profile of TCGA-STAD [5, 41, 42]. The DPS showed stronger prognostic power for OS ($p < 0.001$, Additional file 1: Fig. S3) compared to the seven-gene signature, collagen score, and T-effector signature. As expected, our novel AI-based pathological predictor demonstrated better prognostic performance for OS.

DPS and neoadjuvant chemotherapy

To investigate whether GCs with different DPS level show differential responses to neoadjuvant chemotherapy (NAC), 82 GCs received NAC was enrolled in our study. As shown in Fig. 4A, the DPS level of GCs after receiving NAC was significantly decreased, in contrast to those not receiving NAC treatment. According to the Ninomiya and Ryan classification systems [25, 26], we observed that the DPS level after NAC treatment was negatively associated with tumor regression grade (TRG)

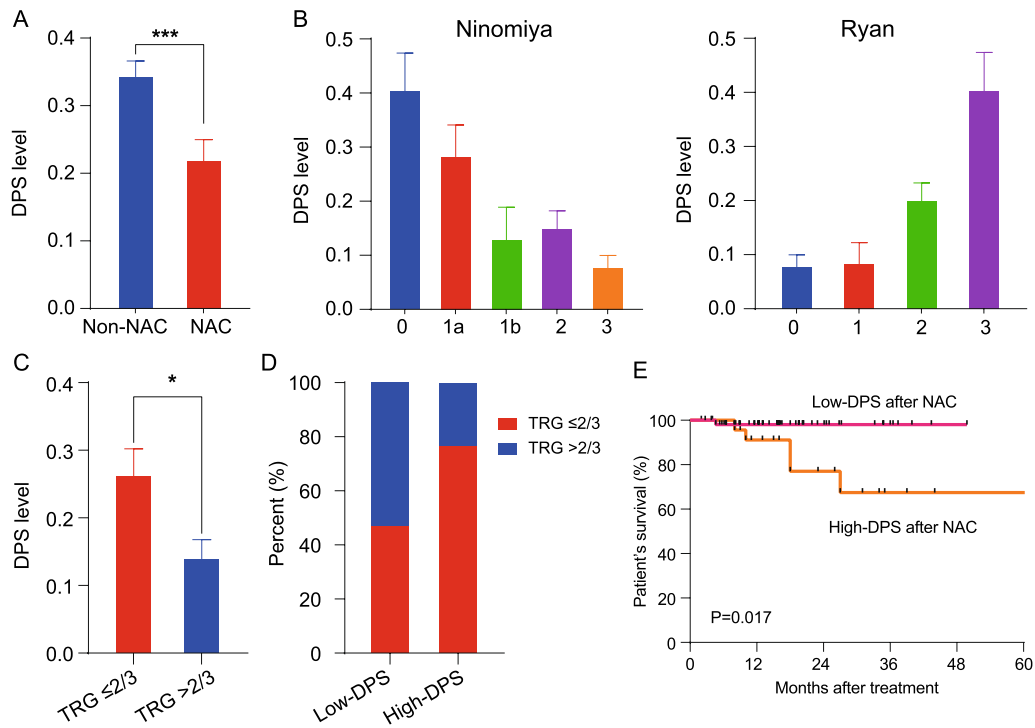


Fig. 4 DPS levels was negatively correlated with the treatment response of NAC. **A** Comparison of the DPS levels between Non-NAC and NAC treatment groups; **B–D** The relationship between treatment response and DPS levels; **E** Kaplan–Meier survival analysis between patients with low-DPS after NAC and those with high-DPS after NAC treatment. * $p < 0.05$, **** $p < 0.0001$. TRG, tissue regression grade; NAC, neoadjuvant chemotherapy

(Fig. 4B). Our analysis of GCs with low-DPS after NAC treatment revealed that 53.8% of GCs presented a better treatment response (TRG >2/3) (Fig. 4C, D). Furthermore, a significant difference was observed between low- and high-DPS patients after receiving NAC treatment (Fig. 4E). Consequently, these findings suggest that DPS after NAC treatment provides potential evaluation efficacy and prognostic value for GC patients, facilitating the subsequent optimization of treatment strategy.

Feature visualization and interpretation

To visualize and interpret the importance of each region in the WSI, whole-slide-level heatmaps were generated by the attention scores as described previously [28]. As

shown in Fig. 5A, clear separation of patient-level features between high- and low-DPS patients was observed in the feature space. Pseudocolours on the heatmap show the clustering of high-risk features (purple) and low-risk features (yellow) in the upper left and lower right corners of the heatmap, respectively (Fig. 5B). All the top-ranked patches were retrieved from the high- and low-DPS groups, and reviewed by 2 pathologists independently. Possible annotations of the featured patches were provided. We observed that top patches from the high-DPS cluster mainly contained tumour cells, lymphocytes and dense tumor stroma, while those from the low-DPS cluster mainly contained normal mucosa and loose stroma (Fig. 5C). To identify and determine the

(See figure on next page.)

Fig. 5 High-DPS associated with suppressive tumor immune microenvironment at the invasive margin. **A** Separability of high- and low-risk area on feature space. **B–D** Identification of DPS high-risk and low-risk areas on heatmap. **E** The cell fractions of lymphocytes and tumor cells between low- and high-DPS groups. **F** IHC panels showing spatial distribution patterns of TIICs on high- and low-DPS specimens. **G** Correlation matrix followed by unsupervised hierarchical clustering of 54 immune features in high- and low-DPS groups. **H** Differences in the density of immune cell types at the invasive margin between high- and low-DPS groups. **I** IHC expression pattern of CD11b⁺CD11c⁺ immune cells and B cells at the invasive margin. **J** Spatial density of immune cells (memory T cells, B cells and Macs) within a 50 μm radius of CD11b⁺CD11c⁺ immune cells at the invasive margin. Macs, macrophages. TIIC, tumor-infiltrating immune cells. * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$. Data are represented as mean ± SEM. ** $p < 0.01$, **** $p < 0.0001$. P, peritumor tissues; IM, invasive margin; T, tumor core

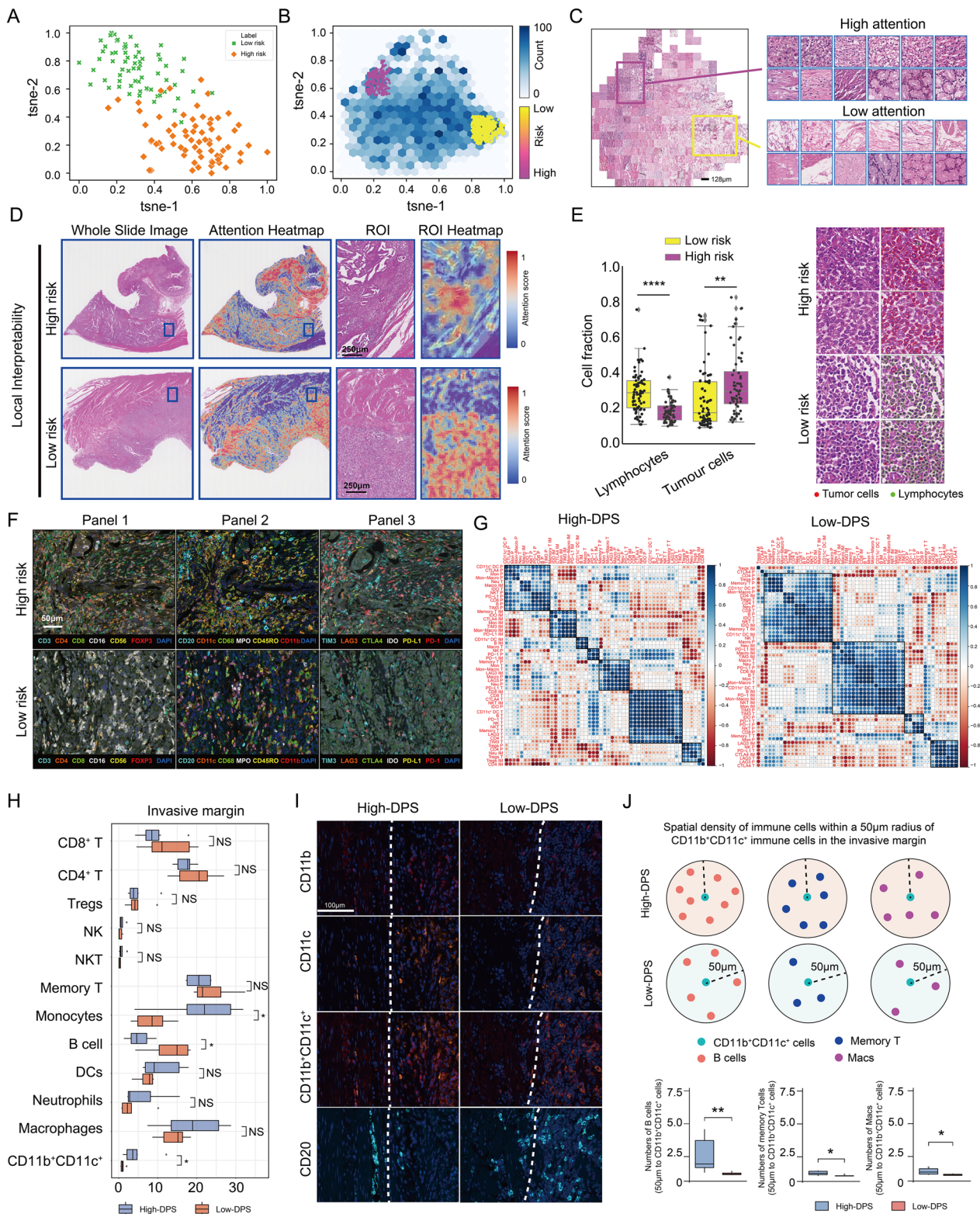


Fig. 5 (See legend on previous page.)

proportions of different cell types in GC, the QuPath software was used as previously described [15]. Interestingly, our results revealed higher proportion of tumour cells and lower proportion of lymphocytes in the high-DPS group (Fig. 5D, E), suggesting that the imbalance of tumor immune microenvironment may facilitate tumor progression and metastasis in GC.

The imbalance of local tumor immune microenvironment associated with high-DPS at the invasive margin

The in situ immune microenvironment plays a critical role in carcinogenesis and distant dissemination [43]. In our study, we performed mIHC with 3 immune panels on the high-DPS (n=6) and low-DPS (n=6) specimens (Fig. 5F). The density of 54 immune features in 3 regions of GC was evaluated: normal region, invasive margin and tumor region. Based on correlation analysis, we identified 7 major clusters in the high-DPS group. In the high-DPS group, Fig. 5G illustrated that one cluster in the invasive margin was characterized by myeloid cells and immune inhibitory markers (including memory T, Mon, Mon-Macro, CD11c⁺ DC, TIM3, IDO and CTLA4). Additionally, another one cluster in tumor region was characterized by lymphocytes and immune inhibitory markers (including CD4, CD8, NK, NKT, B cells, memory T, LAG3, TIM3, CTLA4, IDO and PD-1). These findings suggested an exhausted immune state within the high-DPS group. Then, we compared the immune subsets between high- and low-DPS groups in tumor region and invasive margin, respectively. Interestingly, we observed increased infiltration of CD11b⁺CD11c⁺ immune cells (a phenotype of myeloid derived suppressor cells, MDSCs) [44] and monocytes in tumor region and invasive margin, and a decreased infiltration of B cells at the invasive margin of high-DPS group (Fig. 5H, I and Additional file 1: Fig S4). To explore the relationship between immune cells and tissue architecture, we focused on invasive margin (tumor-immune boundary), which has been previously implicated in playing a prognostic role in tumor progression [36]. Spatial analysis showed a significant increase in B cells, memory T cells and macrophages around CD11b⁺CD11c⁺ immune cells (within 50 μ m radius) at the invasive margin of high-DPS group (Fig. 5J). Consistent with previous studies, the potential interaction between CD11b⁺CD11c⁺ cells and other 3 immune cells may promote the imbalance of local tumor immune microenvironment [45].

To investigate the underlying relevance of pathomics features to molecular mechanism and tumor immune microenvironment, the mRNA expression profile of TCGA-STAD was used in our study. The differentially expressed genes identified between the low- and high-DPS groups were shown in Fig. 6A, B. By using

CIBERSORT method with TCGA-STAD data [33], we observed that high-DPS group was characterized by a higher abundance of MDSCs (Fig. 6C). Furthermore, we investigated the expression levels of immune co-inhibitors and found that BTLA and BTN3A1 had higher expression (Fig. 6D), delineating that high-DPS group might escape immune surveillance by MDSCs infiltration and immune co-inhibitors expression. Additionally, we performed Gene Set Enrichment Analysis (GSEA) to identify gene expression signatures that might be correlated with DPS status. Figure 6E revealed the activation of the chemokine signaling pathway, cytokine-cytokine receptor interaction, Th17 cell differentiation, JAK-STAT signaling pathway, NF- κ B signaling pathway, and PI3K-AKT signaling pathway, implicated in immune regulation [46], cancer progression [47] and metastasis [48].

Discussion

Though cancer prognostication and precise treatment are increasingly driven by biomarkers in histology, clinical data and genomics, pathomics remains a cornerstone of oncology. In this study, we developed a weakly supervised DeepRisk model for GC based on WSIs. Model development included adopting the process of expanding magnification scales, which mimics the way in which pathologists examine histopathological images, and utilizing an attention-based mechanism to focus on image regions important for patient prognosis, without any manual annotation. Our findings demonstrate that this model is a robust tool for GC prognostic evaluation and is applicable to various types of datasets (WSIs and TMAs) obtained from different cohorts and processed following different protocols. As the model output, a novel artificial intelligence (AI)-based pathomic score, the DPS, was developed. Our results suggest that the DPS is independent of and stronger than that of a variety of clinicopathological features (including TNM stage system), and may be a potential supplement to the TNM staging system for GC prognosis prediction and patient stratification.

Because of the 'black box' nature of neural networks [49], the lack of interpretability has become a major limitation that hinders the clinical application of deep learning models. In this study, we interpreted the model hierarchically from case-, patch- and cell-level, and visualized the feature embeddings of cases with different prognostic risks via t-SNE dimensionality reduction. The good separability of high- and low-risk labels in the t-SNE space illustrated the informativeness of feature embeddings converted from WSI patches using our model. Similarly, the top-scoring patches of cases with different prognostic risks were extracted, and the separability of these patches in the t-SNE space indicated that different

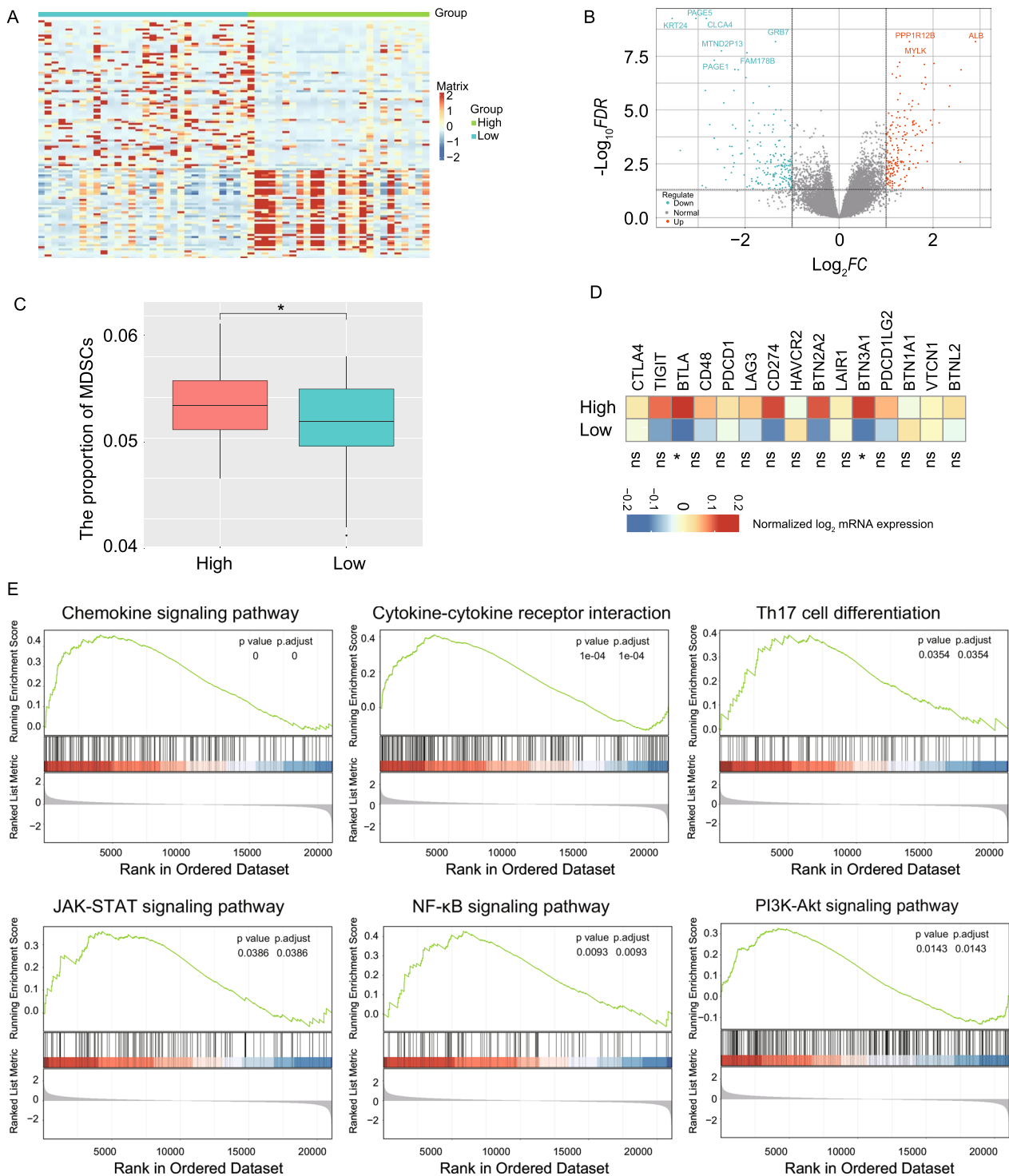


Fig. 6 Transcriptomic analysis of tumor suppressive immune microenvironment associated with high-DPS. **A** Heatmap shows the expression levels of differentially expressed genes between high- and low-DPS groups with TCGA-STAD profile. **B** Volcano plot shows the differentially expressed genes in the high-DPS group compared with the low-DPS group. **C** Comparison of the abundance of MDSCs between high- and low-DPS groups; **D** Normalized mRNA expression levels of immune co-inhibitor between low- and high-DPS groups. **E** GSEA results showing the enrichment of six representative pathways in the high-DPS group. MDSCs, myeloid-derived suppressor cells; GSEA, gene set enrichment analysis

patterns of patches dominated the model prediction. These visualizations demonstrate that the model is able to automatically locate regions crucial for prognosis prediction from WSIs and extract informative features from these regions. Our model eliminates the need for manual annotation, which reduces human intervention and therefore generalizes better. The multi-scale approach used in our model is more in line with the pathologist's diagnostic process. According to the comparison results of our model with several advanced models in Table S4, our model achieved the better performance [9, 11, 12, 50]. Therefore, our method could automate many time-consuming and repetitive tasks, easing the growing workload of clinicians and pathologists.

Recently, Huang et al. developed two deep learning algorithms that were able to identify tumor regions, extract the most suspicious tiles, and generate risk scores for GC diagnosis and prognosis based on digital histopathology derived from TMAs [51]. However, our study has several advantages, including multi-scale sampling, spatial immune contexture analysis, and multi-omics analysis. In our study, we developed a DeepRisk model based on WSIs from GC tissue specimens, which provide more abundant information than TMAs (1–1.5 mm in diameter for each spot). Furthermore, we enrolled a larger sample size in our study, and our DPS achieved excellent performance in the Zhongshan and TCGA-STAD datasets, surpassing any other clinicopathological risk parameters. Our findings are consistent with previous studies showing depleted B cells and increased CD11b⁺CD11c⁺ immune cells at the invasive margin across high-DPS patients, suggesting that the invasive margin is a unique site of immune inhibition with altered expression profiles of immune cells [44, 52]. Previous studies reported that CD11b⁺CD11c⁺ myeloid cells could impair the T cell-stimulating ability by arginase-1, lysosomal protease and COX-2 [53], promote the distant metastasis via CCL2/CCR2 in colorectal cancer [54] and suppress the tumor immune microenvironment and tumor growth through CSF1R/PD-L1 upregulation [55]. Wang et al. also reported that MDSCs suppress B cell proliferation in vitro in an arginase-dependent manner and promote tumor escape from immune surveillance [45]. Interestingly, we also observed the activation of chemokine signaling pathway, cytokine-cytokine receptor interaction, JAK-STAT signaling pathway in the high-DPS group. The above results suggested that the imbalance of the tumor immune microenvironment may influence the immune status under different DPS states.

Neoadjuvant chemotherapy is a common treatment for patients with GC [56]. However, a considerable proportion of GCs do not respond well to NAC [57]. In clinical practice, it is challenging for clinicians to extract

comprehensive prognosticators or treatment-related features directly from pathological images. Herein, we proposed a clinical applicable method to simplify this process and aid clinicians in treatment decision-making. Our results indicated a significant decrease of DPS levels in patients responding to NAC, and GCs with low-DPS after NAC could obtain better survival benefit, implying that DPS might be a dynamic parameter for evaluating the efficacy of NAC. Further, our proposed DPS could reflect the intrinsic characteristics of the tumor biology, immune contexture and aggressiveness, which may be associated with response and outcomes to neoadjuvant chemotherapy. The underlying mechanism between pathological features and chemotherapy have not been thoroughly elucidated, and further investigation may provide insight and strategies for treatment [11, 19].

There are several limitations to our study. Firstly, due to its retrospective nature, further prospective validation is necessary before our model and the DPS can be routinely applied in clinical settings. Secondly, the study cohorts (Zhongshan and SOBC cohorts) mainly consisted of Chinese population, while the TCGA-STAD cohort included populations from Caucasian (167, 62.3%), African (11, 4.1%) and Asian (42, 15.7%). Thus, further validation with various population is needed. Thirdly, our attention-based case-level feature aggregation method discarded the spatial information of the patches, which may reduce model performance, although our attention-based feature aggregation method still achieved better C-indexes than two previously reported WSI-based networks at multiple magnification scales (Additional file 1: Table S5). Finally, the lack of mechanistic insights is a major limitation. The biological mechanisms and immune contexture underlying DPS remain to be further investigated in order to understand the roles of CD11b⁺CD11c⁺ cells in tumor progression and the interactions between CD11b⁺CD11c⁺ cells and other immune cells (such as B cells and macrophages) in high-DPS GCs.

Conclusion

In conclusion, our study developed a DeepRisk model and DPS score based on WSIs for GC prognostic evaluation and adjuvant treatment. Our findings suggest that the DPS is a novel AI-based pathomics signature that may supplement the TNM stage system and lead to potential individualized treatment in clinical practice.

Abbreviations

GC	Gastric cancer
DPS	Digital pathology signature
WSI	Whole-slide image
H&E	Haematoxylin and eosin
MSI	Microsatellite instability
TCGA	The Cancer Genome Atlas

SOBC	Shanghai Outdo Biotech Company
OS	Overall survival
DFS	Disease-free survival
STAD	Stomach adenocarcinoma
TMA	Tissue microarray
CNN	Convolutional neural network
MIL	Multiple instance learning
t-SNE	T-distributed stochastic neighbour embedding
SVM	Support vector machine
TIIC	Tumour-infiltrating immune cells

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-023-04838-5>.

Additional file 1: Methods S1. Table S1. Antibody sources and staining conditions. **Table S2.** Demographic, clinical, and tumor characteristics of patients with gastric cancer in the Zhongshan, TCGA-STAD and SOBC dataset. **Table S3.** Univariate and multivariate Cox analysis of overall survival in the TCGA-STAD and SOBC datasets. **Table S4.** Comparison of C-index of different models on different datasets. **Table S5.** C-index performance comparison of different feature aggregation methods on Zhongshan cohort. **Figure S1.** Heatmaps generated based on WSIs from the Zhongshan dataset. Distribution features of intratumoral fibroblast and extracellular matrix were investigated through α -SMA staining and Masson staining, respectively. **Figure S2.** Forest plot of DPS in the Zhongshan dataset in disease-free survival analysis. **Figure S3.** ROC curves of overall survival (OS) for the DPS and other 3 prognostic models with TCGA-STAD mRNA profile. **Figure S4.** Differences in the density of immune cell types at the tumor region between high-and low-DPS groups.

Acknowledgements

We thank Chunyan Wang for the bioinformatic analysis, Zunsong Hu for the statistical analysis, Meng Gao and Yu Chen for the mlHC analysis.

Author contributions

Study and design: MXT, ZY, YFZ, QJG, YHS, ZQT, JHY and XFW. Data acquisition: LHW, SYW, PZ, ZQD and SJZ. Data analysis and interpretation: MXT, ZY, YFZ, QJG, YHS, ZQT, JHY and XFW. Development of the artificial intelligence: ZY and JHY. Drafting of the manuscript: all authors. All the authors approved the final version of the manuscript.

Funding

This work was supported by the grants from the National Natural Science Foundation of China (No. 81902963, 82273303, 82072679 and 81972228), the Major research plan of the National Natural Science Foundation (No. 91959127), National Key R&D Programmes of China (2021YFF1201001), Natural Science Foundation of Shanghai (23ZR1410700 and 22ZR1461600), Natural Science Foundation of Fujian Province (Grant No.2023J011692), Science and Technology Benefiting People Program of Xiamen, Fujian Province (3502Z20184010).

Availability of data and materials

The data from Zhongshan Hospital supporting the findings are available upon reasonable request from the corresponding author (XFW) due to patient privacy concerns. The external validation of TCGA data set is publicly available at the TCGA portal (<https://portal.gdc.cancer.gov>). We have uploaded our source code to the Github repository with site (<https://github.com/yyzzzhao/DeepRisk>). The TCGA-STAD dataset can be downloaded from (<https://portal.gdc.cancer.gov/projects/TCGA-STAD>).

Declarations

Ethics approval and consent to participate

The study was approved by the Institutional Review Board of Zhongshan Hospital (No. B2021-785R), and complied with the standards of the Declaration of Helsinki and current ethical guidelines. Written informed consent was obtained from each patient at the time of surgery.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Gastrointestinal Surgery, Zhongshan Hospital, Fudan University, 180 Fenglin Road, Shanghai 200032, China. ²Gastric Cancer Center, Zhongshan Hospital, Fudan University, Shanghai, China. ³Cancer Center, Zhongshan Hospital, Fudan University, Shanghai, China. ⁴Biomedical Engineering Center, School of Information Science and Technology, Fudan University, Shanghai 200433, China. ⁵The Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai, China. ⁶Department of Immunology and Pathogenic Biology, School of Basic Medical Sciences, Shanghai University of Traditional Chinese Medicine, Shanghai, People's Republic of China. ⁷Department of General Surgery, Zhongshan Hospital (Xiamen), Fudan University, Xiamen, China. ⁸Xiamen Clinical Research Center for Cancer Therapy, Zhongshan Hospital (Xiamen), Fudan University, Xiamen, China.

Received: 7 May 2023 Accepted: 26 December 2023

Published online: 19 February 2024

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7.
- Wang R, Dang M, Harada K, Han G, Wang F, Pool Pizzi M, Zhao M, Tatlonghari G, Zhang S, Hao D, et al. Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nat Med.* 2021;27(1):141–51.
- Lee JH, Chang KK, Yoon C, Tang LH, Strong VE, Yoon SS. Lauren histologic type is the most important factor associated with pattern of recurrence following resection of gastric adenocarcinoma. *Ann Surg.* 2018;267(1):105–13.
- Wen T, Wang Z, Li Y, Li Z, Che X, Fan Y, Wang S, Qu J, Yang X, Hou K, et al. A four-factor immunoscore system that predicts clinical outcome for stage II/III gastric cancer. *Cancer Immunol Res.* 2017;5(7):524–34.
- Smyth E, Nyamundanda G, Cunningham D, Fontana E, Ragulan C, Tan I, Lin S, Wotherspoon A, Nankivell M, Fassan M. A seven-Gene Signature assay improves prognostic risk stratification of perioperative chemotherapy treated gastroesophageal cancer patients from the MAGIC trial. *Ann Oncol.* 2018;29(12):2356–62.
- Chen D, Liu Z, Liu W, Fu M, Jiang W, Xu S, Wang G, Chen F, Lu J, Chen H, et al. Predicting postoperative peritoneal metastasis in gastric cancer with serosal invasion using a collagen nomogram. *Nat Commun.* 2021;12(1):179.
- Mori T, Tanaka H, Suzuki S, Deguchi S, Yamakoshi Y, Yoshii M, Miki Y, Tamura T, Toyokawa T, Lee S, et al. Tertiary lymphoid structures show infiltration of effective tumor-resident T cells in gastric cancer. *Cancer Sci.* 2021;112(5):1746–57.
- Park HS, Kim YM, Kim S, Lee WS, Kong SJ, Yang H, Kang B, Cheon J, Shin S-J, Kim C, et al. High endothelial venule is a surrogate biomarker for T-cell inflamed tumor microenvironment and prognosis in gastric cancer. *J Immunother Cancer.* 2021;9(10):e003353.
- Song Z, Zou S, Zhou W, Huang Y, Shao L, Yuan J, Gou X, Jin W, Wang Z, Chen X, et al. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun.* 2020;11(1):4294.
- Kather JN, Krisam J, Charoentong P, Luedde T, Herpel E, Weis CA, Gaiser T, Marx A, Valous NA, Ferber D, et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. *PLoS Med.* 2019;16(1):e1002730.
- Shi J-Y, Wang X, Ding G-Y, Dong Z, Han J, Guan Z, Ma L-J, Zheng Y, Zhang L, Yu G-Z, et al. Exploring prognostic indicators in the pathological images of hepatocellular carcinoma based on deep learning. *Gut.* 2020. <https://doi.org/10.1136/gutjnl-2020-320930>.
- Saillard C, Schmauch B, Laifa O, Moarii M, Toldo S, Zaslavskiy M, Pronier E, Laurent A, Amaddeo G, Regnault H, et al. Predicting survival after

- hepatocellular carcinoma resection using deep learning on histological slides. *Hepatology* (Baltimore, MD). 2020;72(6):2000–13.
13. Lu MY, Chen TY, Williamson DFK, Zhao M, Shady M, Lipkova J, Mahmood F. AI-based pathology predicts origins for cancers of unknown primary. *Nature*. 2021;594(7861):106–10.
 14. Muti HS, Heij LR, Keller G, Kohlruss M, Langer R, Dislich B, Cheong JH, Kim YW, Kim H, Kook MC, et al. Development and validation of deep learning classifiers to detect Epstein–Barr virus and microsatellite instability status in gastric cancer: a retrospective multicentre cohort study. *Lancet Digit Health*. 2021;3(10):e654–64.
 15. Reichling C, Taieb J, Derangere V, Klopfenstein Q, Le Malicot K, Gornet JM, Becheur H, Fein F, Cojocarasu O, Kaminsky MC, et al. Artificial intelligence-guided tissue analysis combined with immune infiltrate assessment predicts stage III colon cancer outcomes in PETACC08 study. *Gut*. 2020;69(4):681–90.
 16. Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wählby C, Hartman J, Rantalainen M. Improved breast cancer histological grading using deep learning. *Ann Oncol*. 2022;33(1):89–98.
 17. Kulkarni PM, Robinson EJ, Sarin Pradhan J, Gartrell-Corrado RD, Rohr BR, Trager MH, Geskin LJ, Kluger HM, Wong PF, Acs B, et al. Deep learning based on standard H&E images of primary melanoma tumors identifies patients at risk for visceral recurrence and death. *Clin Cancer Res*. 2020;26(5):1126–34.
 18. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, Marx A, Boor P, Tacke F, Neumann UP, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med*. 2019;25(7):1054–6.
 19. Chen D, Fu M, Chi L, Lin L, Cheng J, Xue W, Long C, Jiang W, Dong X, Sui J, et al. Prognostic and predictive value of a pathomics signature in gastric cancer. *Nat Commun*. 2022;13(1):6903.
 20. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, Liu J, Yue YG, Wang J, Yu K, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med*. 2015;21(5):449–56.
 21. Zhang Q, Cheng L, Qin Y, Kong L, Shi X, Hu J, Li L, Ding Z, Wang T, Shen J, et al. SLAMF8 expression predicts the efficacy of anti-PD1 immunotherapy in gastrointestinal cancers. *Clin Transl Immunology*. 2021;10(10): e1347.
 22. Liu F, Huang C, Xu Z, Su X, Zhao G, Ye J, Du X, Huang H, Hu J, Li G, et al. Morbidity and mortality of laparoscopic vs open total gastrectomy for clinical stage I gastric cancer: the CLASS02 multicenter randomized clinical trial. *JAMA Oncol*. 2020;6(10):1590–7.
 23. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, Meyer L, Gress DM, Byrd DR, Winchester DP. The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging. *CA Cancer J Clin*. 2017;67(2):93–9.
 24. Han DS, Suh YS, Kong SH, Lee HJ, Choi Y, Aikou S, Sano T, Park BJ, Kim WH, Yang HK. Nomogram predicting long-term survival after d2 gastrectomy for gastric cancer. *J Clin Oncol*. 2012;30(31):3834–40.
 25. Ryan R, Gibbons D, Hyland JM, Treanor D, White A, Mulcahy HE, O’Donoghue DP, Moriarty M, Fennelly D, Sheahan K. Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. *Histopathology*. 2005;47(2):141–6.
 26. Ninomiya Y, Yanagisawa A, Kato Y, Kitagawa T, Ishihara S, Nakajima T. Histological indications of a favorable prognosis with far-advanced gastric carcinomas after preoperative chemotherapy. *J Cancer Res Clin Oncol*. 1999;125(12):699–706.
 27. Tian M, Li Y, Liu W, Jin L, Jiang X, Wang X, Ding Z, Peng Y, Zhou J, Fan J, et al. The nanomechanical signature of liver cancer tissues and its molecular origin. *Nanoscale*. 2015;7(30):12998–3010.
 28. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng*. 2021;5(6):555–70.
 29. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Process Syst*. 2012: 25.
 30. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR): 27–30 June 2016. 2016; 2016: 770–778.
 31. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein MS, et al. ImageNet large scale visual recognition challenge. *Int J Comput Vision*. 2015;115:211–52.
 32. Ilse M, Tomczak J, Welling M. Attention-based deep multiple instance learning. In: Jennifer D, Andreas K, editors. International conference on machine learning, vol. 80. Westminister: PMLR; 2018. p. 2127–36.
 33. Tian MX, Liu WR, Wang H, Zhou YF, Jin L, Jiang XF, Tao CY, Tang Z, Zhou PY, Fang Y, et al. Tissue-infiltrating lymphocytes signature predicts survival in patients with early/intermediate stage hepatocellular carcinoma. *BMC Med*. 2019;17(1):106.
 34. Saliari M, Jalal R, Javadmanesh A. Differential expression analysis of genes and long non-coding RNAs associated with KRAS mutation in colorectal cancer cells. *Sci Rep*. 2022;12(1):7965.
 35. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, Cheng Y, Huang S, Liu Y, Jiang S, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discov*. 2021. <https://doi.org/10.1158/2159-8290.CD-21-0316>.
 36. Halama N, Michel S, Kloor M, Zoernig I, Benner A, Spille A, Pommerenke T, von Knebel DM, Folprecht G, Lubber B, et al. Localization and density of immune cells in the invasive margin of human colorectal cancer liver metastases are prognostic for response to chemotherapy. *Cancer Res*. 2011;71(17):5670–7. <https://doi.org/10.1158/0008-5472.CAN-11-0268>.
 37. Ding GY, Ma JQ, Yun JP, Chen X, Ling Y, Zhang S, Shi JY, Chang YQ, Ji Y, Wang XY, et al. Distribution and density of tertiary lymphoid structures predict clinical outcome in intrahepatic cholangiocarcinoma. *J Hepatol*. 2021. <https://doi.org/10.1016/j.jhep.2021.10.030>.
 38. Chan JY, Lim JQ, Yeong J, Ravi V, Guan P, Boot A, Tay TKY, Selvarajan S, Md Nasir ND, Loh JH, et al. Multiomic analysis and immunoprofiling reveal distinct subtypes of human angiosarcoma. *J Clin Invest*. 2020;130(11):5833–46.
 39. Giraldo NA, Nguyen P, Engle EL, Kaunitz GJ, Cottrell TR, Berry S, Green B, Soni A, Cuda JD, Stein JE, et al. Multidimensional, quantitative assessment of PD-1/PD-L1 expression in patients with Merkel cell carcinoma and association with response to pembrolizumab. *J Immunother Cancer*. 2018;6(1):99.
 40. Pages F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, Lugli A, Zlobec I, Rau TT, Berger MD, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet*. 2018;391(10135):2128–39.
 41. Rosenberg JE, Hoffman-Censits J, Powles T, Van Der Heijden MS, Balar AV, Necchi A, Dawson N, O’Donnell PH, Balmanoukian A, Loriot Y. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *The Lancet*. 2016;387(10031):1909–20.
 42. Wang J, Liu Z, Lin L, Wu Z, Gao X, Cai X, Chang L, Xia X, Zhang H, Chen G. Collagen-related gene expression level predicts the prognosis and immune therapy response. *Gastric Cancer*. 2023;26(6):891–903.
 43. Kim R, An M, Lee H, Mehta A, Heo YJ, Kim KM, Lee SY, Moon J, Kim ST, Min BH, et al. Early tumor-immune microenvironmental remodeling and response to first-line fluoropyrimidine and platinum chemotherapy in advanced gastric cancer. *Cancer Discov*. 2022;12(4):984–1001.
 44. Keren L, Bosse M, Marquez D, Angoshtari R, Jain S, Varma S, Yang SR, Kurian A, Van Valen D, West R, et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell*. 2018;174(6):1373–1387 e1319.
 45. Wang Y, Schafer CC, Hough KP, Tousif S, Duncan SR, Kearney JF, Ponnazhagan S, Hsu HC, Deshane JS. Myeloid-derived suppressor cells impair B cell responses in lung cancer through IL-7 and STAT5. *J Immunol*. 2018;201(1):278–95.
 46. Gabrilovich DI, Nagaraj S. Myeloid-derived suppressor cells as regulators of the immune system. *Nat Rev Immunol*. 2009;9(3):162–74.
 47. Jeong HY, Ham IH, Lee SH, Ryu D, Son SY, Han SU, Kim TM, Hur H. Spatially distinct reprogramming of the tumor microenvironment based on tumor invasion in diffuse-type gastric cancers. *Clin Cancer Res*. 2021;27(23):6529–42.
 48. Condamine T, Ramachandran I, Youn JI, Gabrilovich DI. Regulation of tumor metastasis by myeloid-derived suppressor cells. *Annu Rev Med*. 2015;66:97–110.

49. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol*. 2019;16(11):703–15.
50. Campanella G, Hanna MG, Geneslaw L, Mirafior A, Werneck Krauss Silva V, Busam KJ, Brogi E, Reuter VE, Klimstra DS, Fuchs TJ. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25(8):1301–9.
51. Huang B, Tian S, Zhan N, Ma J, Huang Z, Zhang C, Zhang H, Ming F, Liao F, Ji M, et al. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. *EBioMedicine*. 2021;73: 103631.
52. Li K, Shi H, Zhang B, Ou X, Ma Q, Chen Y, Shu P, Li D, Wang Y. Myeloid-derived suppressor cells as immunosuppressive regulators and therapeutic targets in cancer. *Signal Transduct Target Ther*. 2021;6(1):362.
53. Kitamura H, Ohno Y, Toyoshima Y, Ohtake J, Homma S, Kawamura H, Takahashi N, Taketomi A. Interleukin-6/STAT3 signaling as a promising target to improve the efficacy of cancer immunotherapy. *Cancer Sci*. 2017;108(10):1947–52.
54. Zhao L, Lim SY, Gordon-Weeks AN, Tapmeier TT, Im JH, Cao Y, Beech J, Allen D, Smart S, Muschel RJ. Recruitment of a myeloid cell subset (CD11b/Gr1 mid) via CCL2/CCR2 promotes the development of colorectal cancer liver metastasis. *Hepatology* (Baltimore, MD). 2013;57(2):829–39.
55. Zhao T, Liu S, Ding X, Johnson EM, Hanna NH, Singh K, Sen CK, Wan J, Du H, Yan C. Lysosomal acid lipase, CSF1R, and PD-L1 determine functions of CD11c+ myeloid-derived suppressor cells. *JCI Insight*. 2022;7(17): e156623.
56. Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L, Ye G, Deng H, Mou T, Cai S, et al. ImmunoScore signature: a prognostic and predictive tool in gastric cancer. *Ann Surg*. 2018;267(3):504–13.
57. Li Z, Gao X, Peng X, May Chen MJ, Li Z, Wei B, Wen X, Wei B, Dong Y, Bu Z, et al. Multi-omics characterization of molecular features of gastric cancer correlated with response to neoadjuvant chemotherapy. *Sci Adv*. 2020;6(9): eaay4211.

Publisher's Note

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