


# A signature for pan-cancer prognosis based on neutrophil extracellular traps

Yi Zhang,<sup>1</sup> Liping Guo,<sup>2</sup> Qichen Dai,<sup>3</sup> Bingqing Shang,<sup>4</sup> Ting Xiao,<sup>1</sup> Xuebing Di,<sup>1</sup> Kaitai Zhang,<sup>1</sup> Lin Feng ,<sup>1</sup> Jianzhong Shou,<sup>4</sup> Yipeng Wang<sup>3</sup>

**To cite:** Zhang Y, Guo L, Dai Q, et al. A signature for pan-cancer prognosis based on neutrophil extracellular traps. *Journal for ImmunoTherapy of Cancer* 2022;**10**:e004210. doi:10.1136/jitc-2021-004210

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jitc-2021-004210>).

YZ, LG and QD contributed equally.

Accepted 13 May 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Lin Feng;  
fenglin@cicams.ac.cn

Dr Jianzhong Shou;  
shoujz@126.com

Dr Yipeng Wang;  
yidoctor99@126.com

## ABSTRACT

**Background** Neutrophil extracellular traps (NETs) were originally thought to be formed by neutrophils to trap invading microorganisms as a defense mechanism. Increasing studies have shown that NETs play a pivotal role in tumor progression and diffusion. In this case, transcriptome analysis provides an opportunity to unearth the association between NETs and clinical outcomes of patients with pan-cancer.

**Methods** The transcriptome sequencing data of The Cancer Genome Atlas pan-cancer primary focus was obtained from UCSC Xena, and a 19-gene NETs score was then constructed using the Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression model based on the expression levels of 69 NETs initial biomarkers we collected from multistudies. In addition, multiple datasets covering multiple cancer types from other databases were collected and used to validate the signature. Gene ontology enrichment analyses were used to annotate the functions of NETs-related pathways. Immunohistochemistry (IHC) was implemented to evaluate the role of NETs-related genes in clinical patients across types of tumors, including lung adenocarcinoma (n=58), colorectal carcinoma (n=93), kidney renal clear cell carcinoma (n=90), and triple-negative breast cancer (n=80).

**Results** The NETs score was calculated based on 19-NETs related genes according to the LASSO Cox model. The NETs score was considered a hazardous factor in most cancer types, with a higher score indicating a more adverse outcome. In addition, we found that NETs were significantly correlated to various malignant biological processes, such as the epithelial to mesenchymal transition (R=0.7444, p<0.0001), angiogenesis (R=0.5369, p<0.0001), and tumor cell proliferation (R=0.3835, p<0.0001). Furthermore, in IHC cohorts of a variety of tumors, myeloperoxidase, a gene involved in the model and a classical delegate of NETs formation, was associated with poor clinical outcomes.

**Conclusions** Collectively, these constitutive and complementary biomarkers represented the ability of NETs formation to predict the development of patients' progression. Integrative transcriptome analyses plus clinical sample validation may facilitate the biomarker discovery and clinical transformation.

## INTRODUCTION

It is known that a constant interplay endures between the tumor and the host immune response, which expedites tumor development and progression.<sup>1 2</sup> The tumor

microenvironment and immune components matter in tumor progression and metastasis. As the most abundant cells in the human immune system, neutrophils play a pivotal role in the tumor microenvironment.<sup>3 4</sup> Neutrophil in peripheral blood has been proposed as a biomarker for risk stratification of patients with cancer. Most clinical evidence in neutrophil to lymphocyte ratios (NLRs) supports the view that neutrophils promote rather than restrain progression.<sup>5</sup> The increase of NLRs reflected the deviation of hematopoiesis to the myeloid system, the decline of lymphoid hematopoiesis, the drop in immune system function, as well as poor prognosis. As the disease progresses and relapses, the neutrophil count, or NLR, ascends, while it descends with the recovery of the disease.<sup>6</sup> Neutrophils within the tumor are often termed tumor-associated neutrophils (TANs). However, the relationship between TANs and clinical outcome is contradictory,<sup>7</sup> suggesting the functional heterogeneity in TANs. In most cases, the existence of TANs was related to a dismal prognosis in most tumors, such as renal cell carcinoma and pancreatic ductal adenocarcinoma.<sup>8 9</sup> TANs reflected a natural cytotoxicity of the host and released plenty of factors exerting mostly pro-tumor functions, including extracellular matrix remodeling, angiogenesis, and promotion of tumor growth and invasion.<sup>10</sup> Inversely, women seemed to have a better prognosis than men in advanced gastric carcinoma, which differences in neutrophil function may explain to some extent.<sup>11</sup>

The function of neutrophil extracellular traps (NETs), a web-like structure expelled from dying neutrophils of decondensed DNA chromatin to extracellular space, was mainly to kill harmful microorganisms. DNA chromatin decondensed through citrullination and then expelled from the cell complexed with citrullinated histones and neutrophilic cytoplasmic contents containing granular enzymes, which was termed 'NETosis'.

Indeed, not all neutrophils release NETs. Only a subset of neutrophils have this capability. For example, the expression of olfactomedin 4 aligned nicely with the percentage of neutrophils that produce NETs<sup>12</sup> when stimulated by lipopolysaccharide, while CD177<sup>−</sup> neutrophils cannot make NETs.<sup>13</sup> As a substantial step of the innate and adaptive immune response, which both infectious and sterile stimuli can trigger, the role of NETosis in the occurrence and development of tumors is still disputable. From the perspective of the antitumor immunity, NETs inhibit tumor growth by activating the immune system. Neutrophils interact with T cells through NETs to reduce the activation threshold and directly stimulate T cells.<sup>14</sup> However, more studies showed that tumors could coach neutrophils to undergo NETosis to promote metastasis. For example, it was reported that the tumor-triggered NETs could protect hepatocellular carcinoma cells from many cytotoxins and assist them in being more invasive.<sup>15</sup>

Myeloperoxidase (MPO), recorded as an essential element for NETs formation, is a heme peroxidase primarily stored in the azurophilic granules of neutrophils and secreted when neutrophils are stimulated.<sup>16–17</sup> A common ELISA detection of MPO-DNA complexes has been used to evaluate the abundance of NETs in human plasma.<sup>18–20</sup> Circulating NETs levels were elevated in patients with advanced esophageal, gastric, and lung cancer compared with local cancers and healthy controls,<sup>21</sup> and the forming of NETs in situ foci rather than peripheral blood may be a more direct prognostic indicator for patients with tumor.

Therefore, to evaluate the relationship between the degree of NETosis in primary tumor lesions and the patients' outcomes, we applied the Least Absolute Shrinkage and Selection Operator (LASSO) Cox model and established a NETs-related score (referred to as NETs score hereafter) for the pan-cancer data sets. According to our findings, the NETs score was associated with poor prognosis in patients with most types of solid tumors.

## MATERIAL AND METHODS

### NETs-initial biomarkers

When NETosis occurred, the dead neutrophils secreted decondensed nuclear chromatin and mixed cytoplasmic and granule components. We looked for neutrophils and NETosis-related gene sets as NETs-initial biomarkers to reproduce this process. The two parts of genes were input into the training set for modeling. Specifically, the gene set of neutrophils was from Şenbabaoglu *et al.*<sup>22</sup> (online supplemental table S1). The NETosis-related gene set was a summary of the research progress of NETs in immunity and various diseases,<sup>23</sup> of which mainly cover the ligands and receptors that stimulate the formation of NETs, downstream-related signals, and the molecules identified to adhere to the framework of NETs (online supplemental table S2). To summarize, we converged a total of 69 genes as the NETs-initial biomarkers for signature training.

### Patients and data sets

This study collected pan-cancer data from diverse data sets consisting of four individual platforms: The Cancer Genome Atlas program (TCGA, n=8739), Chinese Glioma Genome Atlas (CGGA, n=651), Molecular Taxonomy of Breast Cancer International Consortium (METABRIC, n=1868), and Gene Expression Omnibus (GEO, n=2459). The TCGA data sets obtained from USCS Xena (<http://xenabrowser.net/hub>), used here for independent observations and validations, involved 8739 cases of pan-cancer from 32 kinds of tumors, without samples of acute myeloid leukemia. We randomly assorted 70% of the samples as the training set (6117/8739) and the remaining 30% as the test set (2622/8739). Online supplemental table S3 contains the details of samples in training and test sets involving specific tumors in TCGA. From CGGA, a database specific to brain tumors, we obtained a total of 651 samples of patients with glioma for the study, including data sets of messenger RNA sequencing (mRNAseq\_693 and mRNAseq\_325, with 422 and 229 patients, respectively). The expression profile and clinical information of 1868 patients with breast cancer from METABRIC were downloaded from cbiportal (<https://www.cbiportal.org>). In addition, we retrieved and obtained the data for non-small cell lung cancer (NSCLC) and breast cancer from the GEO database, including GSE12276 (n=204), GSE17705 (n=298), GSE19615 (n=115), GSE21653 (n=266), GSE2990 (n=189), GSE6532\_U133A (n=741), GSE7390 (n=198), GSE8894 (n=138), GSE43767 (n=69), GSE19188 (n=91) and GSE67061 (n=150).

### Construction of prognostic NETs-related signature

We used the LASSO algorithm for the NETs-related genes and ensured the simplicity of the model and minimized the overfitting in the process of model training. We set the regularization of LASSO regression as a one-time SE for the most concise model. We subsequently implemented the Cox proportional hazard regression model to assess the association of each gene expression of NETs signature with the disease-specific survival (DSS) of patients in the TCGA pan-cancer training series. Based on the Cox proportional hazard regression model of NETs-related genes, if the coefficient  $\beta$  of each NETs-related gene were obtained, then each patient would get a NETs score as follows:

$$\text{NETs score} = \sum_{i=1}^n G_i^* \beta_i$$

Where n was the number of NETs signature genes,  $G_i$  was the normalized expression value of gene i, and  $\beta_i$  was the regression coefficient of gene i in the univariate Cox regression analysis. The median NETs score as a cut-off value divided tumor samples into the high-risk and low-risk groups. Likewise, we investigated the NETs score in predicting patients' survival on external data sets.

### The z-score evaluation of the biological process

The z-score was an algorithm proposed by Lee *et al.*<sup>24</sup> to mirror the activity of given pathways by integrating feature gene expressions. Gene sets containing the genes referring to angiogenesis,<sup>22</sup> epithelial to mesenchymal transition (EMT, obtained from Gene Ontology (GO): 0001837), cell cycle,<sup>25</sup> as well as the NETs initial biomarkers (online supplemental table S1, S2 and S4), were subjected to the z-score algorithm implemented in the R package GSVA. The value of each gene set was enumerated as angiogenesis z-score, EMT z-score, Cell Cycle z-score, and NETs z-score, respectively.

### Validation using tissue microarray analysis and immunohistochemical staining

Tissue samples from multiple tumor tissue microarrays were used for clinical validation, containing the tumor and adjacent normal tissues from lung adenocarcinoma (LUAD), triple-negative breast cancer (TNBC), colon adenocarcinoma (COAD) and kidney renal clear cell carcinoma (KIRC). In particular, we collected 58 pairs of LUAD tissues matched to their adjacent normal tissue samples in the Cancer Hospital of the Chinese Academy of Medical Sciences (Beijing, China). The use of LUAD specimens in the present study has been approved by the Ethics Committee of the Cancer Hospital, Chinese Academy of Medical Sciences. Informed consents were signed after patients had been fully aware of the purpose of the research. The COAD tissue array, which contains 93 tumor samples with 87 corresponding adjacent normal tissues, and the KIRC tissue array, containing 90 paired tumor-normal tissues, were purchased from Outdo Biotech (Shanghai, China). The tissue array with 80 pairs of TNBC samples was purchased from Super Biotek (Shanghai, China). The clinical information for all patients covered was listed in online supplemental table S5. Immunohistochemistry (IHC) staining was implemented with MPO antibody (#14569, Cell Signaling Technology), interleukin 17A (IL-17A) antibody (ab79056, Abcam), and histone hypercitrullination (H3Cit) antibody (ab5103, Abcam), according to their manufacturers' protocols. MPO is specifically expressed on stromal cells, and each sample was assigned a score based on the proportion of MPO<sup>+</sup> cells (0%=0, 1%–9%=1, 10–25%=2, 26%–100%=3). IL-17 is expressed in both tumor cells and stromal cells, so we scored tumor cells and stromal cells separately by the extent of staining (percentage of positive cells with the intensity).<sup>26,27</sup> The score was independently judged by two experienced pathologists with unknown clinicopathological information during this process.

### Functional annotation of differently expressed genes between patients with high and low NETs scores

The R package 'clusterProfiler' package was used to conduct Gene Ontology (GO) enrichment analysis based on the differentially expressed genes (DEGs) ( $|\log_2FC| \geq 1$ , q-value (FDR) < 0.05) between the high-risk and low-risk groups.

### Statistical analyses

All statistical analyses were performed using the R V.3.5.2 (<http://www.r-project.org>). The training and test sets were standardized by zero-mean normalization. We used the Kaplan-Meier analysis to assess the probability of survival outcomes in patients within low-risk and high-risk groups. The log-rank test was accustomed to determining the statistical significance of differential groups. By integrating multiforecast indicators, nomogram construction and validation allowed the multivariable Cox regression analysis to be visualized and readable. Decision curve analysis (DCA), first proposed by Vickers and Elkin,<sup>28</sup> was used to show the clinical benefit provided by each model. The time-dependent receiver operating characteristic curve was used to measure the prognosis performance by the areas under the curves (AUC) of the NETs score and the nomogram model. For all, a p value < 0.05 was considered statistically significant.

## RESULTS

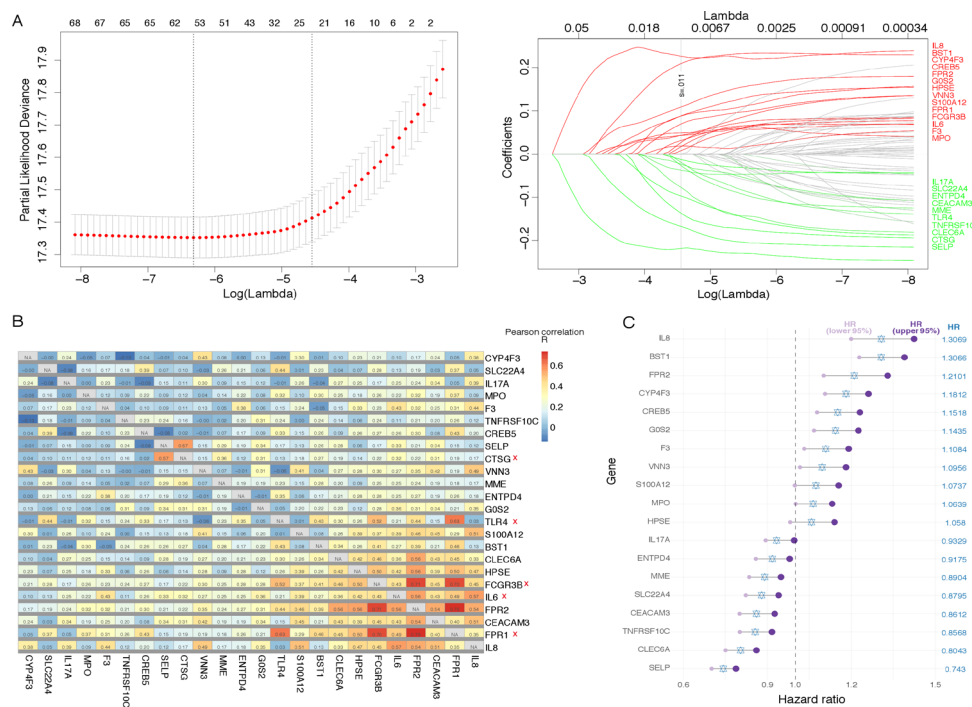
### Identification of a 19-gene NETs-related signature for pan-cancer

The workflow for constructing the NETs-related signature of pan-cancer was depicted in online supplemental figure S1. First, 69 NETs-initial biomarkers were input into LASSO regression analysis in the TCGA pan-cancer training set (figure 1A), and 24 NETs-related genes with non-zero coefficients were obtained. Subsequently, pairwise correlation analysis was performed on these genes, and we found that some of the genes had similar expression patterns (figure 1B). The correlation coefficients between CTSG and SELP, IL-6 and IL-8, IL-6 and FRP2, TLR4 and FPR1, FCGR3B and FPR1, FCGR3B and FPR2, FPR1 and FPR2 were higher than the others. Therefore, we removed CTSG, IL-6, TLR4, FPR1, and FCGR3B and used the remaining 19 NETs-related genes for further model construction.

The Cox regression analysis was applied to the 19 NETs-related genes, and a score (the NETs score) based on their normalized expression levels was derived (online supplemental table S6). In addition, the hazard ratio (HR) of every single gene for the prognosis was observed (figure 1C). Not surprisingly, MPO, as an important representative of the tendency of NETosis, was involved in the model.

### Panorama of NETs score in cancers

The NETs score showed an obvious organ specificity, as the tumors originating from the brain and gastrointestinal tract generally had greater NETs scores, while the tumors from some secreting glands, for example, prostate adenocarcinoma, thyroid carcinoma, and breast invasive carcinoma (BRCA), usually had lower ones (figure 2A). Patients in the TCGA training cohort were assigned to the high-risk or low-risk group according to whether their NETs scores were higher than the median of the population. Compared with those in the low-risk



**Figure 1** Construction of the neutrophil extracellular traps-characteristic signature for pan-cancer.

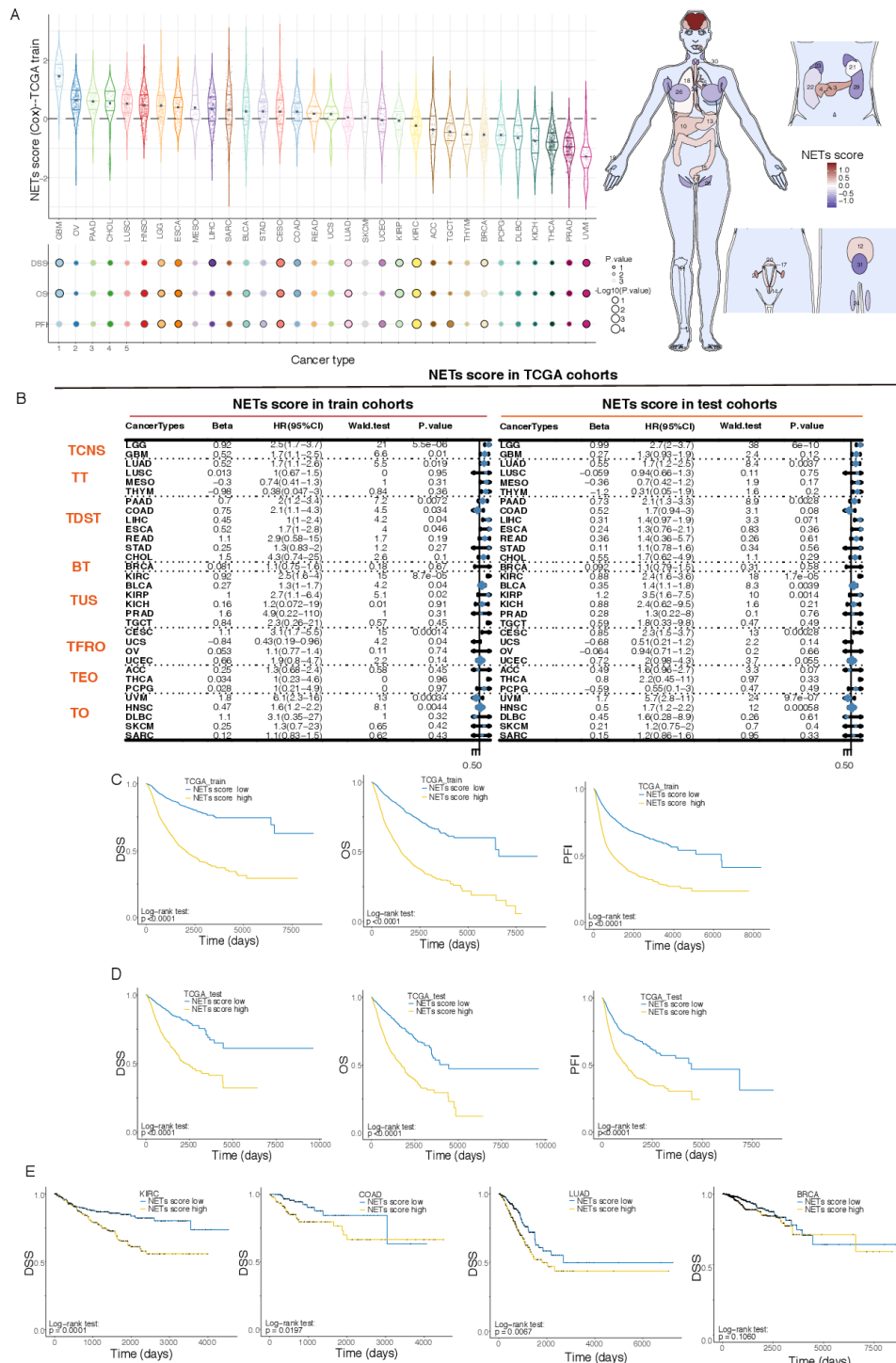
group, patients with pan-cancer with higher NETs scores were associated with various adverse survival indicators, including DSS, overall survival (OS), and progression-free interval (PFI) (figure 2C). The Kaplan-Meier survival curves referring to DSS, OS, and PFI of each type of tumor in the training set were presented in online supplemental figures S2-S4. The result of the univariate Cox analyses used to characterize the effect of NETs score on various predicting prognoses showed that NETs score was related to a dismal prognosis in most types of cancers (figure 2B).

In order to examine the effectiveness and generality of the signature, we first verified the prognostic effect of the NETs score in the TCGA test series. Consistent with the training set results, survival analyses of the patients in the test cohort showed that the higher NETs score corresponded to the worse progression (figure 2D). The multivariable Cox proportional hazards model comprising the NETs score and cancer types illustrated that the NETs score was an independent predictor of the OS ( $p=0.0070$ ,  $HR=1.22$ ) of the patients in the TCGA test cohort. We observed that there was a quite powerful predictive ability of the NETs score on prognosis in a variety of tumors (figure 2E), including KIRC (log-rank test:  $p=0.0001$ ), COAD (log-rank test:  $p=0.0197$ ), and LUAD (log-rank test:  $p=0.0067$ ). However, for BRCA with relatively low NETs score, the prognostic predictive value of NETs score could not be reproduced in the test cohort (log-rank test:  $p=0.1060$ ). As the most frequently diagnosed cancer in women, there was a huge repertoire of BRCA sample data in TCGA, which gave rise to the fact that there was high statistical efficiency and overfitting in the model of the training cohort. Interestingly, we

found that NETs score had different effects on cancers from the same organ, as it was significantly correlated with the unfavorable prognosis of patients with LUAD, but not on patients with lung squamous cell carcinoma (LUSC) even through the latter ones had a higher NETs score than the former ones (figure 2B), which indicated the roles of NETosis in tumor progression are not only affected by its hosting organs but also related to the nature of the tumor.

### Validation of NETs-related signature in other independent data sets

To further confirm the prognostic value of the NETs score, we applied it to multiple external validation cohorts and obtained similar results. We got 651 patients with glioma mRNA matrix and clinical information from the CGGA database and proved that the patients with glioma with higher NETs scores were accompanied by worse prognoses, especially for those with high-grade disease (grade III or IV, figure 3A). External validations suggested that the NETs score was not a reliable prognosis predictor for breast cancer as it was not significant in seven out of the eight validation data sets (figure 3B), which was in concert with the validation result of the TCGA cohort. Extra NSCLC data sets were collected to confirm the distinct prognostic effects of NETs scores on patients with LUAD and LUSC. Similarly, we found that the NETs score was significantly related to the prognosis of patients with LUAD ( $p=0.0087$  in GSE8894, and  $p=0.0023$  in GSE43767) (figure 3C), but not to those with LUSC ( $p=0.5970$  in GSE19188, and  $p=0.4970$  in GSE67061) (figure 3D).

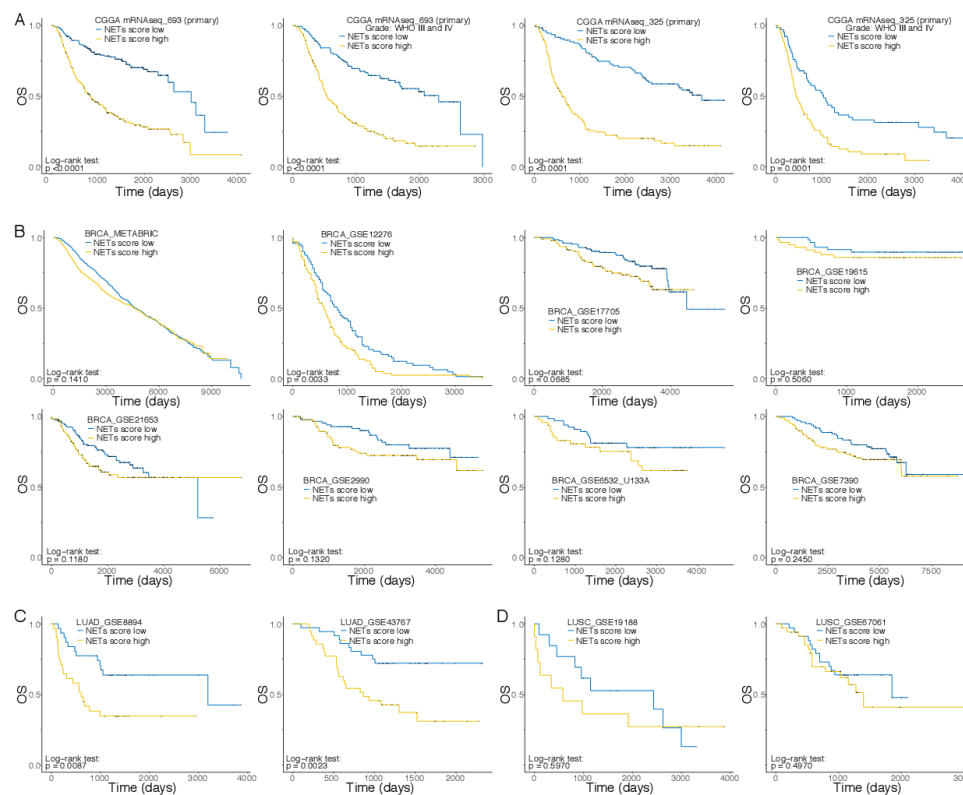


**Figure 2** Prognostic performance of the 19-gene NETs score. BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; DSS, disease-specific survival; KIRC, kidney renal clear cell carcinoma; LUAD, lung adenocarcinoma; OS, overall survival; PFI, progression-free interval; TCGA, The Cancer Genome Atlas. TCNS, Tumors of Central Nervous System; TT, Thoracic Tumors; TDST, Tumors of Digestive System; BT, Breast Tumor; TUS, Tumors of Urinary System and Male Genital Organs; TFRO, Tumors of Female Reproductive Organs; TEO, Tumors of Endocrine Organs; TO, Tumors of others.

### Establishment of a nomogram based on the NETs score for clinical predicting the survival of pan-cancer

To provide a readable and quantitative measurement for the NETs signature to clinically predict the probability of adverse events, we constructed a comprehensive nomogram that combined the NETs score and several

clinicopathological characteristics, including the patients' age and cancer types (figure 4A). The calibration curve of the 3-year DSS, was plotted in figure 4B, almost coincided with the standard one, a straight line with a slope of 1 through a coordinate axis dot, indicating the nomogram signature was very close to the actual survival probability.



**Figure 3** Prognostic performance of the NETs score in the external cohorts. BRCA, breast invasive carcinoma; CGGA, Chinese Glioma Genome Atlas; GSE, GEO series; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; OS, overall survival.

As shown in [figure 4C](#), we found that the AUC of nomogram prediction was better than NETs score alone in the training set and the test cohort (TCGA\_train: 0.700 vs 0.726, TCGA\_test: 0.661 vs 0.718), suggesting the prediction efficiency of nomogram signature would be better after considering a variety of clinical factors. Furthermore, we found the net benefits (NB) of both NETs score alone and comprehensive nomogram were higher than 0, with the maximum NB greater than 15% in DCA clinical evaluation ([figure 4D,E](#)), indicating the importance of the NETs score coordinated with other clinical traits for prognostic prediction.

### The NETs signature and malignant features of the tumor

In the process of normal cells dissimulating to a malignant state, rapid proliferation, active EMT, and angiogenesis were acquired, which were the hallmarks of malignant tumor capabilities.<sup>29</sup> It was reported that NETs could advance the progression of the neoplasm by stimulating angiogenesis<sup>30</sup> and transforming tumor cells from dormant to the proliferative phase.<sup>31</sup> To investigate the connection between the NETosis and these malignant features, we quantified the ability of tumor in NETs promotion, angiogenesis, EMT, and cell cycle by z-score algorithm (see Methods for details). We found that there were pronounced positive correlations between NETs z-score with Angiogenesis z-score ( $R=0.5369$ ,  $p<0.0001$ ), EMT z-score ( $R=0.7444$ ,  $p<0.0001$ ), and Cell Cycle z-score ( $R=0.3835$ ,  $p<0.0001$ ), respectively, in the overall TCGA

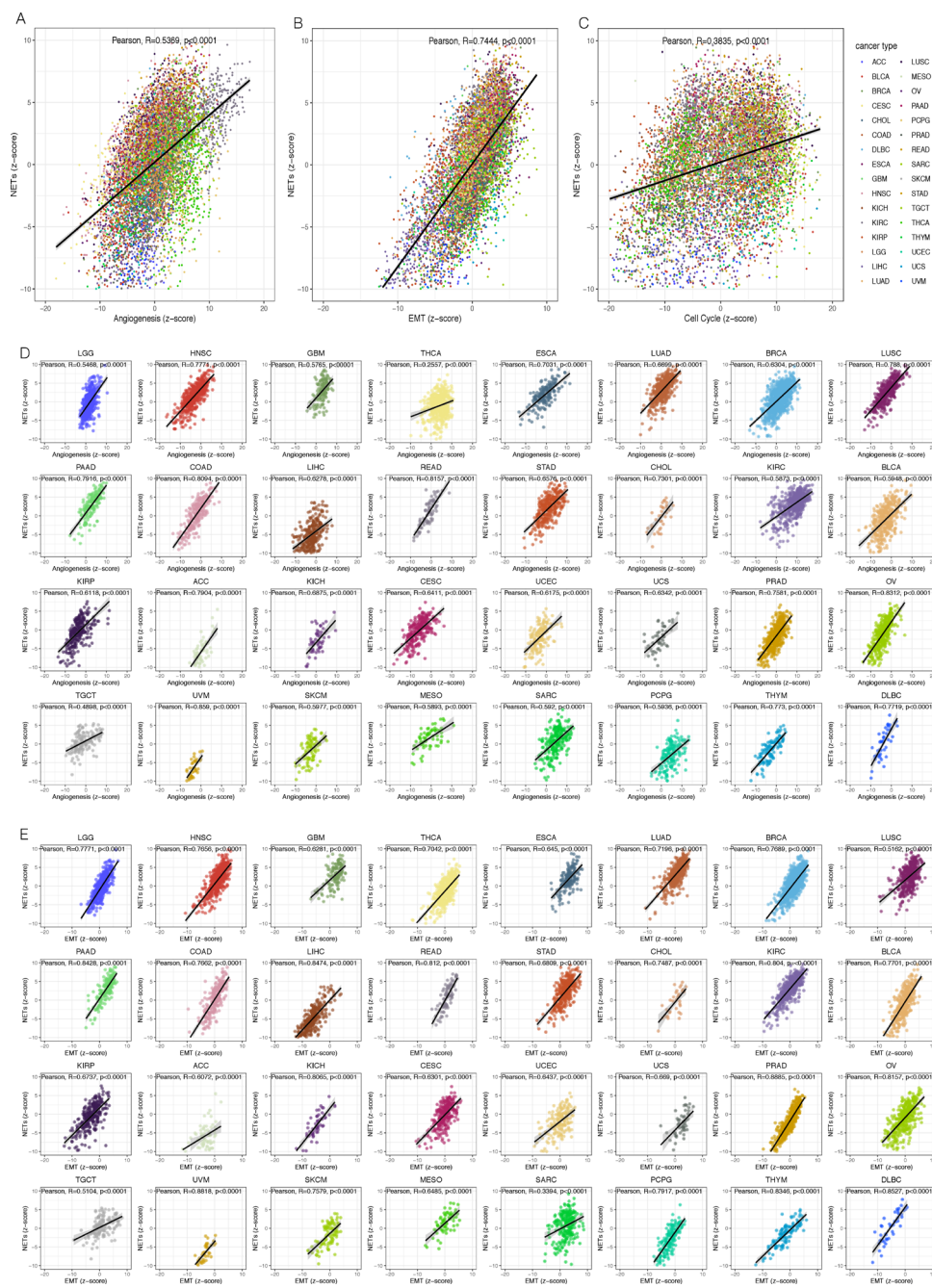
pan-cancer cohort ([figure 5A–C](#)), or in most of the tumor types ([figure 5D,E](#), and online supplemental figure S5). In other words, the tumor with strong potential to promote NETosis was generally accompanied by more active angiogenesis in microenvironments, and more aggressive tumor cells.

In conclusion, these results suggested that the more actively the NETosis occurred in the tumor microenvironment, the more aggressively the malignant cells acted.

### The characteristics comparison of the high-risk and low-risk patients

To further characterize and compare the tumor biological behaviors and the immune microenvironments of the high-risk and low-risk patients, we extracted a total of 1599 patients with LUAD, COAD, KIRC, and TNBC from TCGA data set. Based on their NETs scores, the patients with each type of tumor were divided into high-risk and low-risk groups according to the optimal cut-off value. The DEGs between the high-risk and low-risk groups were applied to GO enrichment analyses. As expected, DEGs were characterized consistently by the features of the NETs formation process, such as neutrophil activation, neutrophil degranulation, and neutrophil-mediated immunity, as well as the steps referred to the tumor metastasis, for example, positive regulation of cell adhesion, extracellular organization, and the extracellular matrix degradation (online supplemental figure S6).





**Figure 5** The neutrophil extracellular traps score was highly correlated with many malignant features of the tumor.

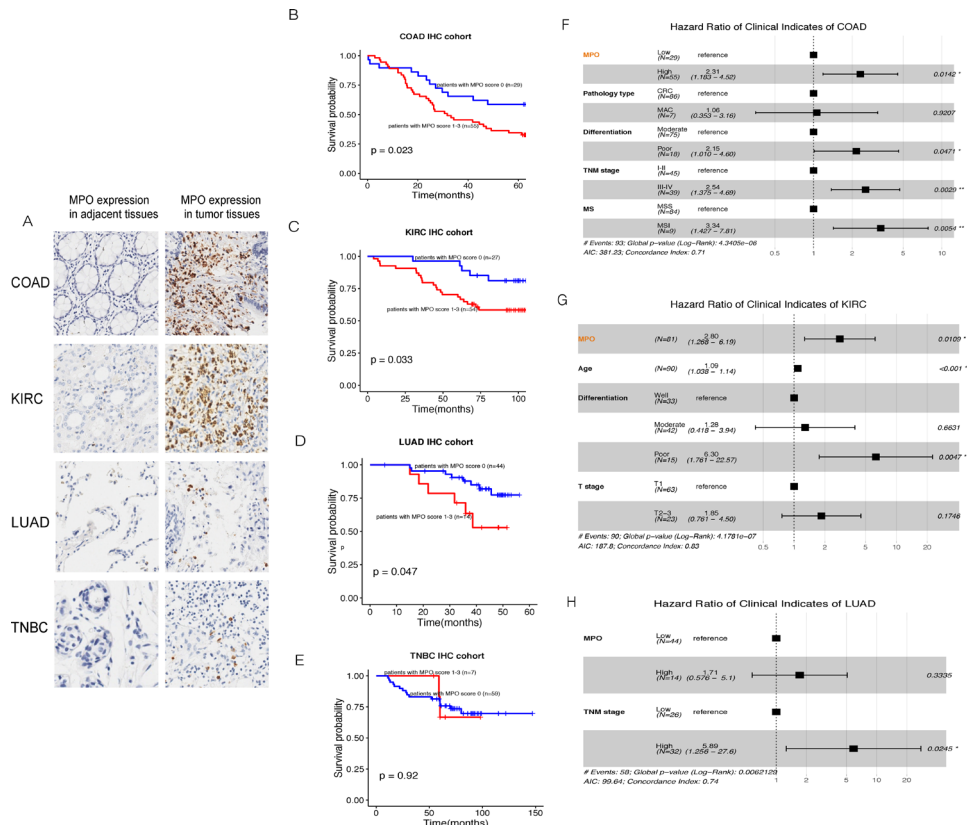
that it was not an independent prognostic factor in the multivariable Cox model (figure 6H).

H3Cit is known as a substantial step in NETosis and considered an essential and specific element of NETs.<sup>32</sup> Maybe due to its high specificity for indicating the formed NETs, the positive rate of H3Cit in our IHC cohort was too low to be evaluated (only eight patients were stained positively). Therefore, we could not evaluate the prognostic value of NETs formation tendency based on H3Cit.

Since the NETs score model was an evaluating system composed of multigenes, the expression levels of any one of them alone might not directly map the prediction of the model. For example, IL-17, encoded by IL-17A, has been known as a neutrophil recruiter and NETs

promoter,<sup>33,34</sup> and consequently was originally included in the model construction as a component contributing to NETs formation. However, the coefficient of IL-17A turned to be negative in the final model (online supplemental table S6). We speculated that it may be related to the wide variety of sources of IL-17 in tumor tissues, which complicated the biological effects of this gene. Our IHC results indicated that IL-17 was expressed by various types of cells, including tumor cells, in concert with previous reports.<sup>35–37</sup> The overall levels of IL-17 in tumor tissue were not correlated with the patients' clinical outcomes in our cohort (data not shown). However, the proportion score of IL-17<sup>+</sup> cells in tumor stroma (referred to as IL-17 score, hereafter) was significantly correlated to





**Figure 6** IHC validation of MPO in our clinical cohorts. COAD, colon adenocarcinoma; IHC, immunohistochemistry; KIRC, kidney renal clear cell carcinoma; LUAD, lung adenocarcinoma; MPO, myeloperoxidase; TNBC, triple-negative breast cancer; TNM, tumor, node, metastases. MS, microsatellite; MSI, microsatellite instability; MSS, microsatellite stable.

the MPO score (Kruskal-Wallis test,  $p=0.0411$ ). Likely aligned with the MPO score, the IL-17 score was also an independent prognostic factor for the patients with COAD and KIRC (online supplemental figure S8A,B), as a higher IL-17 score was accompanied by a shorter OS (online supplemental figure S8C,D). However, the situation was not optimistic in patients with LUAD, which might be explained by the MPO score was not being an independent prognostic factor in our LUAD cohort as mentioned. The constitutive results of the IL-17 and MPO scores suggested that it was the expression of IL-17 by stromal cells rather than other cells that may have an important role in promoting NETosis. Accordingly, the overall expression level of a single gene like IL-17A in tumor tissues may be insufficient for representing the results of the NETs score model.

## DISCUSSION

The death of patients with tumor is attributed more frequently to metastases, rather than the primary tumors. About 90% of patients died of organ failure, cancer-associated thrombosis, or other complications caused by tumor metastasis.<sup>38</sup> Increasingly, more evidence has demonstrated that heterogeneity exists across tumor types and even among tumors arising in the same tissue. Therefore, it is hard to construct a prediction system that affects the prognosis of pan-cancer from the tumor essence.

Importantly, as the soil of tumor development, the tumor immune microenvironment of different types of tumors has something in common. Neutrophils could be the link between the tumor parenchyma and the immune microenvironment. NETs, a unique derivate of neutrophils, could prevent tumor cells from being attacked by the immune system and promote invasion and metastasis. Meanwhile, the mechanism of NETs promoting malignant metastasis is gradually emerging. For example, some studies have proved that NETs enhanced diffuse large B-cell lymphoma (DLBCL) migration by TLR9 signaling in human DLBCL cell lines.<sup>39</sup> Not alone, another study suggested that the formation of NETs boosted breast cancer to lung dissemination through cathepsin C.<sup>40</sup>

Neutrophil accumulation varies in different tumor types. Neutrophils with certain molecular markers may have a variety of biological functions. The subset that produces NETs is important in terms of the patient's prognosis. Our findings shed light on a constitutive NETs-related signature that contributed to the progression of patients with pan-cancer. The NETs score classified all patients of pan-cancer into high-risk and low-risk groups. Most tumor progressions occurred in high-risk patients while low-risk patients had longer survival.

Clinical validation and application are the core practical significance of our model construction. For selecting the delegate gene, MPO is one of the most abundant

proteins in neutrophils, stored in the azurophilic granules and released when neutrophils are stimulated. MPO contributed to DNA decondensation, binding to DNA, and catalyzing oxidative reactions, promoting the relocation of ELANE from the cytoplasm to the nucleus.<sup>41</sup> Furthermore, it was recorded that MPO was necessary for making NETs and acted cell-autonomously to promote NETs formation<sup>42</sup> and was considered as a representative marker of NETosis. In the clinical validation phase, we found that the MPO score in tumors was significantly higher than in adjacent normal tissues and was associated with weaker survivals of patients with COAD, KIRC, and LUAD. Consistent with the relatively low NETs scores of the BRCA data set among TCGA cohorts, we found the MPO score was statistically lower in patients with TNBC than that of patients with COAD, KIRC and LUAD in our IHC cohorts, and did not have any prognostic value. We conjectured that MPO staining in tumor stroma cells reflected the infiltration and function of TAN and could act as a representer of the tendency of NETs formation in tumor tissues, as well as the NETs score model did, based on evaluating multiexpression.

NETs are competent to damage vascular endothelial cells to activate inflammatory responses as well as promote angiogenesis. When stimulated by angiopoietin 1/2, the amounts of NETs produced by neutrophils were about 2.5 times higher than that in the natural state, which further promoted angiogenesis.<sup>43</sup> By evaluating the activity of tumor-related biological pathways, we found that the synchronization between NETs formation and vascularization. In pancreatic ductal adenocarcinoma, a very malignant and one of the worst prognostic tumors has been confirmed that many TANs promote angiogenesis in several studies,<sup>44 45</sup> which was likely part of contributed by the NETs. Besides allowing patients in a hypercoagulable state, NETs also boost vascular endothelial cells proliferating and further shaping into tubular structures,<sup>31</sup> which could be explained by the coherence between NETs z-score and cell-cycle z-score in our study.

There are several limitations in our study. First of all, the number of evaluated NETs-representing genes is small. Also, only a few cancer types were validated in this study, and the number of samples in each tumor group is limited. Last but not least, although MPO is a reliable marker of TANs, it is not sufficient for a comprehensive assessment of NETosis and should be implemented in combination with other markers, such as H3Cit. In light of the unsuccessful test of H3Cit in IHC cohorts, we will explore and validate other reliable markers of NETosis as a supplementation in the future.

## CONCLUSIONS

In summary, the NETs formation potential of tumor tissues was highly coordinated with many malignant features and was associated with the clinical outcomes of patients with pan-cancer. Evaluating and targeting the

NETs formation in tumor lesions may facilitate personalized treatment for cancer.

## Author affiliations

<sup>1</sup>State Key Laboratory of Molecular Oncology, Department of Etiology and Carcinogenesis, National Cancer Center/ National Clinical Research Center for Cancer/ Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

<sup>2</sup>Peking University People's Hospital, Peking University Institute of Hematology, National Clinical Research Center for Hematologic Disease, Beijing Key Laboratory of Hematopoietic Stem Cell Transplantation, Beijing, China

<sup>3</sup>Department of Breast Surgery, National Cancer Center/ National Clinical Research Center for Cancer/ Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

<sup>4</sup>Department of Urology, National Cancer Center/ National Clinical Research Center for Cancer/ Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

**Acknowledgements** We sincerely thank the researchers who collected, managed, and maintained TCGA, CGGA, METABRIC and GEO data. Their high-quality work and efforts provide great help for our research.

**Contributors** YW, LF and JS developed the concept and design the study. LG, YZ and QD carried out the data analysis and statistical analysis. YZ and TX conducted the laboratory work. BS, KZ and XD managed the resources and interpretation. YZ wrote the manuscript. All authors contributed to the critical review of the manuscript. YW is responsible for the overall content as the guarantor.

**Funding** This work has been supported by the CAMS Innovation Fund for Medical Sciences (CIFMS, ID Number: 2021-I2M-1-014), National Natural Science Foundation of China (ID Number: 81972318, 82072837, 81972317), and Natural Science Foundation of Beijing Municipality (ID Number: 7212083).

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by Ethics Committee of the Cancer Hospital Chinese Academy of Medical Sciences (ID: NCC2016YJC-01). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Lin Feng <http://orcid.org/0000-0003-0951-713X>

## REFERENCES

- Janssen LME, Ramsay EE, Logsdon CD, *et al*. The immune system in cancer metastasis: friend or foe? *J Immunother Cancer* 2017;5:79.
- Bruni D, Angell HK, Galon J. The immune contexture and immunoscore in cancer prognosis and therapeutic efficacy. *Nat Rev Cancer* 2020;20:662–80.
- Giese MA, Hind LE, Huttenlocher A. Neutrophil plasticity in the tumor microenvironment. *Blood* 2019;133:2159–67.

- 4 Liang W, Ferrara N. The complex role of neutrophils in tumor angiogenesis and metastasis. *Cancer Immunol Res* 2016;4:83–91.
- 5 Templeton AJ, McNamara MG, Šeruga B, et al. Prognostic role of neutrophil-to-lymphocyte ratio in solid tumors: a systematic review and meta-analysis. *J Natl Cancer Inst* 2014;106:dju124.
- 6 Coffelt SB, Wellenstein MD, de Visser KE. Neutrophils in cancer: neutral no more. *Nat Rev Cancer* 2016;16:431–46.
- 7 Ocana A, Nieto-Jiménez C, Pandiella A, et al. Neutrophils in cancer: prognostic role and therapeutic strategies. *Mol Cancer* 2017;16:137.
- 8 Reid MD, Basturk O, Thirabanasak D, et al. Tumor-Infiltrating neutrophils in pancreatic neoplasia. *Mod Pathol* 2011;24:1612–9.
- 9 Jensen HK, Donskov F, Marcussen N, et al. Presence of intratumoral neutrophils is an independent prognostic factor in localized renal cell carcinoma. *J Clin Oncol* 2009;27:4709–17.
- 10 Powell DR, Huttenlocher A. Neutrophils in the tumor microenvironment. *Trends Immunol* 2016;37:41–52.
- 11 Caruso RA, Bellocco R, Pagano M, et al. Prognostic value of intratumoral neutrophils in advanced gastric carcinoma in a high-risk area in northern Italy. *Mod Pathol* 2002;15:831–7.
- 12 Clémensen SN, Bohr CT, Rørvig S, et al. Olfactomedin 4 defines a subset of human neutrophils. *J Leukoc Biol* 2012;91:495–500.
- 13 Zhou G, Yu L, Fang L, et al. CD177<sup>+</sup> neutrophils as functionally activated neutrophils negatively regulate IBD. *Gut* 2018;67:1052–63.
- 14 Tillack K, Breiden P, Martin R, et al. T lymphocyte priming by neutrophil extracellular traps links innate and adaptive immune responses. *J Immunol* 2012;188:3150–9.
- 15 Yang L-Y, Luo Q, Lu L, et al. Increased neutrophil extracellular traps promote metastasis potential of hepatocellular carcinoma via provoking tumorous inflammatory response. *J Hematol Oncol* 2020;13:3.
- 16 Ramachandra CJA, Ja KPMM, Chua J, et al. Myeloperoxidase as a multifaceted target for cardiovascular protection. *Antioxid Redox Signal* 2020;32:1135–49.
- 17 Klebanoff SJ. Myeloperoxidase. *Proc Assoc Am Physicians* 1999;111:383–9.
- 18 Yoo D-goon, Floyd M, Winn M, et al. Net formation induced by *Pseudomonas aeruginosa* cystic fibrosis isolates measured as release of myeloperoxidase-DNA and neutrophil elastase-DNA complexes. *Immunol Lett* 2014;160:186–94.
- 19 Le Joncour A, Martos R, Loyau S, et al. Critical role of neutrophil extracellular traps (nets) in patients with Behcet's disease. *Ann Rheum Dis* 2019;78:1274–82.
- 20 Miller-Ocuin JL, Liang X, Boone BA, et al. Dna released from neutrophil extracellular traps (nets) activates pancreatic stellate cells and enhances pancreatic tumor growth. *Oncimmunology* 2019;8:e1605822.
- 21 Rayes RF, Mouhanna JG, Nicolau I, et al. Primary tumors induce neutrophil extracellular traps with targetable metastasis promoting effects. *JCI Insight* 2019;5:e128008.
- 22 Şenbabaoğlu Y, Gejman RS, Winer AG, et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol* 2016;17:231.
- 23 Papayannopoulos V. Neutrophil extracellular traps in immunity and disease. *Nat Rev Immunol* 2018;18:134–47.
- 24 Lee E, Chuang H-Y, Kim J-W, et al. Inferring pathway activity toward precise disease classification. *PLoS Comput Biol* 2008;4:e1000217.
- 25 Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016;352:189–96.
- 26 Amicarella F, Muraro MG, Hirt C, et al. Dual role of tumour-infiltrating T helper 17 cells in human colorectal cancer. *Gut* 2017;66:692–704.
- 27 Wang D, Yu W, Lian J, et al. Th17 cells inhibit CD8<sup>+</sup> T cell migration by systematically downregulating CXCR3 expression via IL-17A/STAT3 in advanced-stage colorectal cancer patients. *J Hematol Oncol* 2020;13:68.
- 28 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74.
- 29 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–74.
- 30 Aldabbous L, Abdul-Salam V, McKinnon T, et al. Neutrophil extracellular traps promote angiogenesis: evidence from vascular pathology in pulmonary hypertension. *Arterioscler Thromb Vasc Biol* 2016;36:2078–87.
- 31 Li B, Liu Y, Hu T, et al. Neutrophil extracellular traps enhance procoagulant activity in patients with oral squamous cell carcinoma. *J Cancer Res Clin Oncol* 2019;145:1695–707.
- 32 Kuczia P, Zuk J, Iwaniec T, et al. Citrullinated histone H3, a marker of extracellular trap formation, is increased in blood of stable asthma patients. *Clin Transl Allergy* 2020;10:31.
- 33 Chuammitri P, Wongsawan K, Pringproa K, et al. Interleukin 17 (IL-17) manipulates mouse bone marrow-derived neutrophils in response to acute lung inflammation. *Comp Immunol Microbiol Infect Dis* 2019;67:101356.
- 34 Frangou E, Chrysanthopoulou A, Mitsios A, et al. REDD1/autophagy pathway promotes thromboinflammation and fibrosis in human systemic lupus erythematosus (SLE) through nets decorated with tissue factor (TF) and interleukin-17A (IL-17A). *Ann Rheum Dis* 2019;78:238–48.
- 35 Liu J, Duan Y, Cheng X, et al. IL-17 is associated with poor prognosis and promotes angiogenesis via stimulating VEGF production of cancer cells in colorectal carcinoma. *Biochem Biophys Res Commun* 2011;407:348–54.
- 36 Vasilescu F, Arsene D, Cionca F, et al. Foxp3 and IL17 expression in tumor infiltrating lymphocytes (TIL) and tumor cells - correlated or independent factors? *Rom J Morphol Embryol* 2013;54:43–9.
- 37 Chen X, Wan J, Liu J, et al. Increased IL-17-producing cells correlate with poor survival and lymphangiogenesis in NSCLC patients. *Lung Cancer* 2010;69:348–54.
- 38 Cedervall J, Zhang Y, Olsson A-K. Tumor-induced NETosis as a risk factor for metastasis and organ failure. *Cancer Res* 2016;76:4311–5.
- 39 Nie M, Yang L, Bi X, et al. Neutrophil extracellular traps induced by IL8 promote diffuse large B-cell lymphoma progression via the TLR9 signaling. *Clin Cancer Res* 2019;25:1867–79.
- 40 Xiao Y, Cong M, Li J, et al. Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation. *Cancer Cell* 2021;39:423–37.
- 41 Papayannopoulos V, Metzler KD, Hakkim A, et al. Neutrophil elastase and myeloperoxidase regulate the formation of neutrophil extracellular traps. *J Cell Biol* 2010;191:677–91.
- 42 Metzler KD, Fuchs TA, Nauseef WM, et al. Myeloperoxidase is required for neutrophil extracellular trap formation: implications for innate immunity. *Blood* 2011;117:953–9.
- 43 Lavoie SS, Dumas E, Vulesevic B, et al. Synthesis of human neutrophil extracellular traps contributes to angiotensin-mediated in vitro proinflammatory and proangiogenic activities. *J Immunol* 2018;200:3801–13.
- 44 Nozawa H, Chiu C, Hanahan D. Infiltrating neutrophils mediate the initial angiogenic switch in a mouse model of multistage carcinogenesis. *Proc Natl Acad Sci U S A* 2006;103:12493–8.
- 45 Bausch D, Pausch T, Krauss T, et al. Neutrophil granulocyte derived MMP-9 is a VEGF independent functional component of the angiogenic switch in pancreatic ductal adenocarcinoma. *Angiogenesis* 2011;14:235–43.