Cancer Immunology, Immunotherapy (2024) 73:241
https://doi.org/10.1007/s00262-024-03829-9

RESEARCH

A random survival forest-based pathomics signature classifies immunotherapy prognosis and profiles TIME and genomics in ES-SCLC patients

Yuxin Jiang[1] · Yueying Chen[2] · Qinpei Cheng[2] · Wanjun Lu[2] · Yu Li[3] · Xueying Zuo[2] · Qiuxia Wu[4] · Xiaoxia Wang[5] · Fang Zhang[2,3,6] · Dong Wang[2,3,6] · Qin Wang[2] · Tangfeng Lv[1,2,3,6] · Yong Song[1,2,3,6] · Ping Zhan[1,2,3,6]

Received: 29 July 2024 / Accepted: 6 September 2024 / Published online: 3 October 2024

Abstract

Background Small cell lung cancer (SCLC) is a highly aggressive neuroendocrine tumor with high mortality, and only a limited subset of extensive-stage SCLC (ES-SCLC) patients demonstrate prolonged survival under chemoimmunotherapy, which warrants the exploration of reliable biomarkers. Herein, we built a machine learning-based model using pathomics features extracted from hematoxylin and eosin (H&E)-stained images to classify prognosis and explore its potential association with genomics and TIME.

Methods We retrospectively recruited ES-SCLC patients receiving first-line chemoimmunotherapy at Nanjing Jinling Hospital between April 2020 and August 2023. Digital H&E-stained whole-slide images were acquired, and targeted next-generation sequencing, programmed death ligand-1 staining, and multiplex immunohistochemical staining for immune cells were performed on a subset of patients. A random survival forest (RSF) model encompassing clinical and pathomics features was established to predict overall survival. The function of putative genes was assessed via single-cell RNA sequencing.

Results and conclusion During the median follow-up period of 12.12 months, 118 ES-SCLC patients receiving first-line immunotherapy were recruited. The RSF model utilizing three pathomics features and liver metastases, bone metastases, smoking status, and lactate dehydrogenase, could predict the survival of first-line chemoimmunotherapy in patients with ES-SCLC with favorable discrimination and calibration. Underlyingly, the higher RSF-Score potentially indicated more infiltration of $CD8^+$ T cells in the stroma as well as a greater probability of *MCL-1* amplification and *EP300* mutation. At the single-cell level, MCL-1 was associated with TNFA-NFKB signaling and apoptosis-related processes. Hopefully, this noninvasive model could act as a biomarker for immunotherapy, potentially facilitating precision medicine in the management of ES-SCLC.

Keywords Machine learning · Immunotherapy · Biomarkers · Tumor Microenvironment · Small cell lung cancer

Introduction

Lung cancer is a life-threatening malignancy with the highest mortality, making it a leading cause of cancer-related deaths globally[1]. Small cell lung cancer (SCLC) is a highly aggressive neuroendocrine tumor, featuring strong tendencies of metastases, proliferation, and resistance to definitive therapy[2] and accounting for ~ 15% of all cases of lung cancer. Approximately two-thirds of patients are diagnosed with fatal extensive-stage SCLC (ES-SCLC) at onset, significantly surpassing the prevalence of limited-stage SCLC (LS-SCLC)[3]. Over the past few decades, immunotherapy has grounded its position as the standard first-line treatment for ES-SCLC, achieving an incremental 5-year overall survival (OS) rate of 12%, compared to around 2% for chemotherapy [4]. Despite initial robust responses to immunotherapy, acquired resistance inevitably occurs. Consequently, the exploration of reliable and therapeutic biomarkers is crucial and might facilitate the realm of precision medicine for SCLC.

Histological examination was essential in the diagnostic and therapeutic processes of SCLC. Given the intrinsic heterogeneity of SCLC, molecular subtyping identified via the

Qin Wang, Tangfeng Lv, Yong Song, and Ping Zhan have contributed equally to this work.

Extended author information available on the last page of the article

expression of transcription factors (ASCL1, NEUROD1, and POU2F3) alongside neuroendocrine and non-neuroendocrine classifications, has emerged as a promising framework for SCLC characterization [5]. Nevertheless, the impact of these molecular subtypes on immunotherapeutic efficacy has yielded inconsistent results [4, 6, 7], suggesting that molecular subtyping alone may not suffice as a standalone biomarker for immunotherapy. Recent advancements have highlighted the tumor immune microenvironment (TIME) as a potential biomarker for immunotherapy [8, 9]. Within the context of SCLC, tumor-infiltrating lymphocytes (TILs) have been identified as candidate prognostic biomarkers in SCLC patients undergoing chemotherapy or immunotherapy [7, 10, 11]. However, the challenge of obtaining adequate SCLC tissue samples for comprehensive immune cell marker staining underscores the need for alternative noninvasive methodologies to evaluate the TIME.

Hematoxylin and Eosin (H&E)-stained sections from biopsy samples are routinely employed in the diagnosis of SCLC, which could provide high-resolution images embedded with morphological and textural features of tumor tissues. The extensive array of features extracted from these images offers a wealth of information, contributing to a more comprehensive histological profile. Recently, two studies have applied pathomics to evaluate the prognosis of SCLC patients [12, 13]. However, the sample size of the ES-SCLC cohort was limited, and another study predominantly enrolled LS-SCLC patients who had undergone prior surgical intervention. This warrants the necessity of larger cohorts of ES-SCLC patients in the settings of chemoimmunotherapy. Furthermore, researchers have validated the presumed connection between pathomics and underlying biological functions or gene expression in malignancies [14, 15]. Although a few studies have investigated the relationship between pathomics and prognosis, none have elaborated on the underlying genomic alterations associated with pathomics features in SCLC, highlighting the necessity of our study.

Consequently, we integrated pathomics and clinical features to develop a random survival forest (RSF) model in predicting prognosis in ES-SCLC patients undergoing first-line chemoimmunotherapy. Additionally, we explored the putative correlations between RSF-derived risk-score and genomic alterations, as well as immune cell infiltration within the TIME. This approach may enhance the noninvasive assessment of TIME and guide the selection of appropriate candidates for first-line chemoimmunotherapy in ES-SCLC.

Materials and methods

Study design

This retrospective study was conducted according to the ethics of the Declaration of Helsinki (as revised in 2013) and the Institutional Review Board of Nanjing Jinling Hospital (registration ID. 2024NZKY-050–01). Informed consent was obtained from each patient before sample collection for research purposes. Patients with histologically diagnosed ES-SCLC receiving first-line chemoimmunotherapy between April 2020 and August 2023 were retrospectively recruited. The inclusive criteria were (I) pathologically diagnosed ES-SCLC according to Veterans Administration Lung Study Group staging criteria; (II) receiving lung biopsy at diagnosis at Nanjing Jinling Hospital; (III) receiving first-line programmed cell death 1 (PD-1) or PD-L1 (programmed cell death ligand-1) inhibitors combined with chemotherapy; (IV) with Eastern Cooperative Oncology Group Performance Status (ECOG PS) 0 ~ 3. Patients dying of non-cancer reasons (cardiovascular diseases, infectious diseases, etc.), defined as transformed SCLC, with ECOG PS $\geq$ 4, lost to follow-up within 1 month or receiving biopsy in other medical centers were ruled out from the study. The last follow-up time was in April 2024.

Extraction of clinical variables

The baseline clinicopathological characteristics were collected from electronic medical records, including age, gender, ECOG PS, smoking status, metastatic site (brain, liver, contralateral lung, and bone), chemotherapy regimens, and immunotherapy regimens. Serum parameters including lactate dehydrogenase (LDH), C-reactive protein (CRP), and interleukin-6 (IL-6) were also recorded. Each continuous variable was dichotomized using the "survminer" R package with the OS as the endpoint.

Image acquisition and processing

Archival formalin-fixed paraffin-embedded (FFPE) tumor sections from ES-SCLC patients undergoing needle biopsy of the lung were retrieved from the pathological specimen repository of the Jinling cohort. The digital H&E-stained whole-slide images (WSIs) in MRXS format were captured at 20 × magnification via a digital scanner (Pannoramic MIDI scanner, 3DHISTECH) and were further converted into tiled TIFF using the 3DHISTECH's Slide Converter software. The pathologist (XX. W) with 10 years of clinical experience reviewed and agreed on the image quality for all WSIs. By applying Qupath 0.5.0, all TIFF images were

preprocessed in a pyramidal TIFF format. Next, the tissue regions excluding the white background in the WSIs were identified through the pipeline of "Classify," "Pixel classification," and "Create threshold" (resolution: 17.52 um/pixel; channel: red; prefilter: Gaussian; smoothing sigma: 0; threshold: 240). We further automatically segmented the tissue regions into non-overlapping 1000 × 1000 pixels (0.2738 um/pixel) tiles. Tiles that were fizzy or with tissue folds were excluded. Only the tiles with the tumor region coverage surpassing 60% of the total area were retained, which was further confirmed by an experienced pathologist who was blinded to the clinical responses (XX. W). Finally, a total of 826 tiles were generated from 118 WSIs.

Endpoint events and clinical response

The primary and secondary endpoints were designed as OS and progression-free survival (PFS), which were respectively defined as the time from the baseline to death and disease progression. Clinical response was defined as complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD) according to the modified Response Evaluation Criteria in Solid Tumors version 1.1 for immune-based therapeutics [16].

Extraction of pathomics features

The pipeline of pathomics features extraction was built using CellProfiler[17] (version 4.2.6) according to previously published methods[18–20]. Firstly, the "UnmixColors" module was performed to unmix the selected tiles for tissue foreground identification by a threshold calculated by the Otsu algorithm. Secondly, the pipeline automatically identified the cell nuclei and cytoplasm by using "IdentifyPrimaryObjects" module and "IdentifySecondaryObjects" module with adaptive Otsu thresholds, respectively. Afterward, 401 image features were obtained by multiple modules including "Measure Object Size Shape," "Measure Texture," "Measure Object Intensity," and "Measure Object Intensity Distribution," covering the shapes, sizes, textures, pixel intensity distributions, and proximity relations of the primary and secondary objects. The mean values of each selected feature were calculated for each sample. After eliminating futile features, a total of 345 quantitative image features are listed in Table S1 for further analysis.

Feature selection and the establishment of the random survival forest model

To select the candidate features, all pathomics features were standardized using the Z-score method. Subsequently, to prefilter features strongly embedded with prognostic information, only pathomics or clinical features with $P < 0.05$ from univariate Cox regression analysis for OS were retained. Later, the LASSO (least absolute shrinkage and selection operator) method based on Cox regression with tenfold cross-validation was applied to regularize the number of features in proportion to the sample size. The reserved features were subsequently used for model construction.

The Cox regression model (Cox) was developed. The risk-score was calculated based on a linear combination of selected features, weighted by their respective coefficients in the multivariate Cox regression model. The RSF model, a decision tree ensemble learning regression algorithm utilizing feature and sample bragging through random forest, was established using the "randomForestSRC" R package. The parameters were set to 1000 trees, a maximum of 15 levels per decision tree, and a random seed of 123. The estimated cumulative hazard function (CHF) was calculated by integrating the Nelson-Allen estimator within the model [21], serving as the RSF risk-score (RSF-Score). Additionally, variable importance was determined by measuring the decrease in prediction accuracy using out-of-bag data, which were not used in tree construction each time.

Evaluation and comparison of model performances

Multiple methods were employed to evaluate model performance, including model discrimination, calibration and clinical utility [22]. Specifically, the area under the curve (AUC) of the time-dependent receiver operating characteristic (ROC) curve and Harrell's concordance index (C-index) were applied to assess model discrimination. The decision curve analysis (DCA) and calibration plots were utilized to evaluate clinical utility and model calibration, respectively. The bootstrap C-index reported as median [standard error of the mean (SEM)] and ROC curves for the model with 1000 resampling groups were calculated for internal validation. The DeLong test was applied for model comparison at different time points.

Immunohistochemical and multiplex immunohistochemical staining

Consecutive tissue sections, each four micrometers thick, were acquired from FFPE tumor biopsy samples. The expression of PD-L1 was evaluated using the combined positive score (CPS), with sections stained employing the PD-L1 IHC 22C3 pharmDx kit (Agilent, Santa Clara, CA, USA).

The infiltration of immune cells within the TIME was assessed through multiplex immunohistochemistry (mIHC) staining using the PANO Multiplex IHC kit (Panovue, Beijing, China). FFPE sections were stained with antibodies targeting CD8A (Mouse#70,306, CST, Massachusetts, USA), CD56 (Mouse#3576, CST, Massachusetts, USA),

CD68 (Rabbit#76,437, CST, Massachusetts, USA), HLA-DR (Rabbit#ab92511, Abcam, Cambridge, England), and panCK (Mouse#4545S, CST, Massachusetts, USA). Specifically, immune cells were identified as $CD8^+$ T cells (CD8A positive), $CD56^{dim+}$ and $CD56^{bright+}$ natural killer (NK) cells (distinguished by the surface density expression of CD56), M1-like macrophages (double positive for CD68 and HLA-DR), M2-like macrophages (CD68 positive and HLA-DR negative), and tumor cells (panCK positive). Following the application of primary antibodies, tyramide signal amplification and incubation with secondary antibodies conjugated with horseradish peroxidase were performed. All images were analyzed in batches using the InForm automated image analysis software (PerkinElmer, Waltham, MA, USA).

Targeted next-generation sequencing

24 ES-SCLC patients from the Jinling cohort published previously were subjected to a targeted next-generation sequencing (tNGS) panel which included 437 cancer-related genes (Geneseeq PrimeTM, Nanjing Geneseeq Technologies Inc.). Sequencing analysis and sample preparation followed previously published guidelines [23, 24]. In brief, the QIAamp DNA FFPE Tissue Kit (QIAGEN, Dusseldorf, Germany) was used to extract genomic DNA from FFPE tumor tissue samples. Leukocytes were used to isolate genomic DNA for the normal control using the DNeasy Blood and Tissue Kit (QIAGEN, Dusseldorf, Germany). The KAPA Hyper Prep Kit (KAPA Biosystems, Wilmington, MA, USA) was used to prepare the libraries in accordance with the guidelines. Using the GeneseeqPrime™ tNGS panel and the xGen Lock-down Hybridization and Wash Reagents Kit (Integrated DNA Technologies), libraries with different indices were combined for targeted enrichment. Sequencing was performed on a Novaseq6000 platform (Illumina, San Diego, USA).

Bulk and Single-cell RNA sequencing analysis

A public cohort by George et al. [25] accessed via cbioportal (https://www.cbioportal.org/study/summary?id=sclc_ucologne_2015) with 81 limited-stage SCLC patients receiving both whole-genome sequencing and bulk RNA sequencing tests were enrolled for transcriptomics analysis. To evaluate gene expression and function at the single-cell level, lung tissues from untreated SCLC patients with single-cell RNA sequencing were used [26]. Preprocessed single-cell RNA sequencing data, annotated with detailed cell types, were obtained [26] and subjected to analysis using the R package "Seurat." The differentially expressed genes (DEGs) across cell types were identified with "FindMarkers" function through the application of the "MAST" test. Adjusted *P*-values < 0.05 and an absolute average log2 (fold change) larger than 1.5 were used to evaluate significance. Furthermore, Metascape enrichment analysis [27] was carried out to elucidate the molecular processes that underlie the different groups.

Statistical analysis

The mean ± standard deviation (SD) is used to represent continuous variables with a normal distribution, whereas the median (interquartile range) was used to express non-normally distributed data. For continuous and categorical variables, the Fisher's exact test and the Wilcoxon signed-rank test were respectively used. Log-rank tests and Cox proportional hazard models were used to assess survival predictors. The cutoff values of continuous variables were recognized using the "survminer" R package, with OS as the endpoint. Likewise, using the "survminer" R package with OS as the endpoint, patients in the RSF and Cox models were classified into high- and low-risk groups based on their risk scores. The "MICE" R package was utilized to impute missing values through the predictive mean matching method. Two-sided *P*-values less than 0.05 were considered statistically significant. All analyses were conducted using R version 4.3.2 and GraphPad Prism version 9.3.1.

Results

Baseline characteristics

The study design is depicted in Fig. 1. Over a median follow-up period of 12.12 months (interquartile range: 7.61–22.75), 118 patients with ES-SCLC receiving first-line immunotherapy at Nanjing Jinling Hospital were recruited. The cohort predominantly consisted of male patients [101 (85.6%)] with a history of smoking [84 (71.2%)]. A sum of 34 (28.8%) and 23 (19.5%) patients had bone and liver metastases respectively. The majority of patients received etoposide plus cisplatin (EP)-based chemotherapy [42 (35.5%)], while 48 (40.7%) patients were treated with the PD-1 inhibitors as the immunotherapy regimen. In terms of the best treatment response, 65 (55.1%) patients achieved SD, 37 (31.4%) achieved PR, and 16 (13.6%) exhibited PD. Throughout the follow-up period, 77 (65.3%) patients experienced disease progression, and 59 (50.0%) patients died. Detailed information is provided in Table 1.

Analysis of clinical and serum parameters associated with OS

Univariate and multivariate analyses revealed that specific chemotherapy regimens [etoposide plus lobaplatin (EL) vs. etoposide plus cisplatin (EP), Hazard ratio

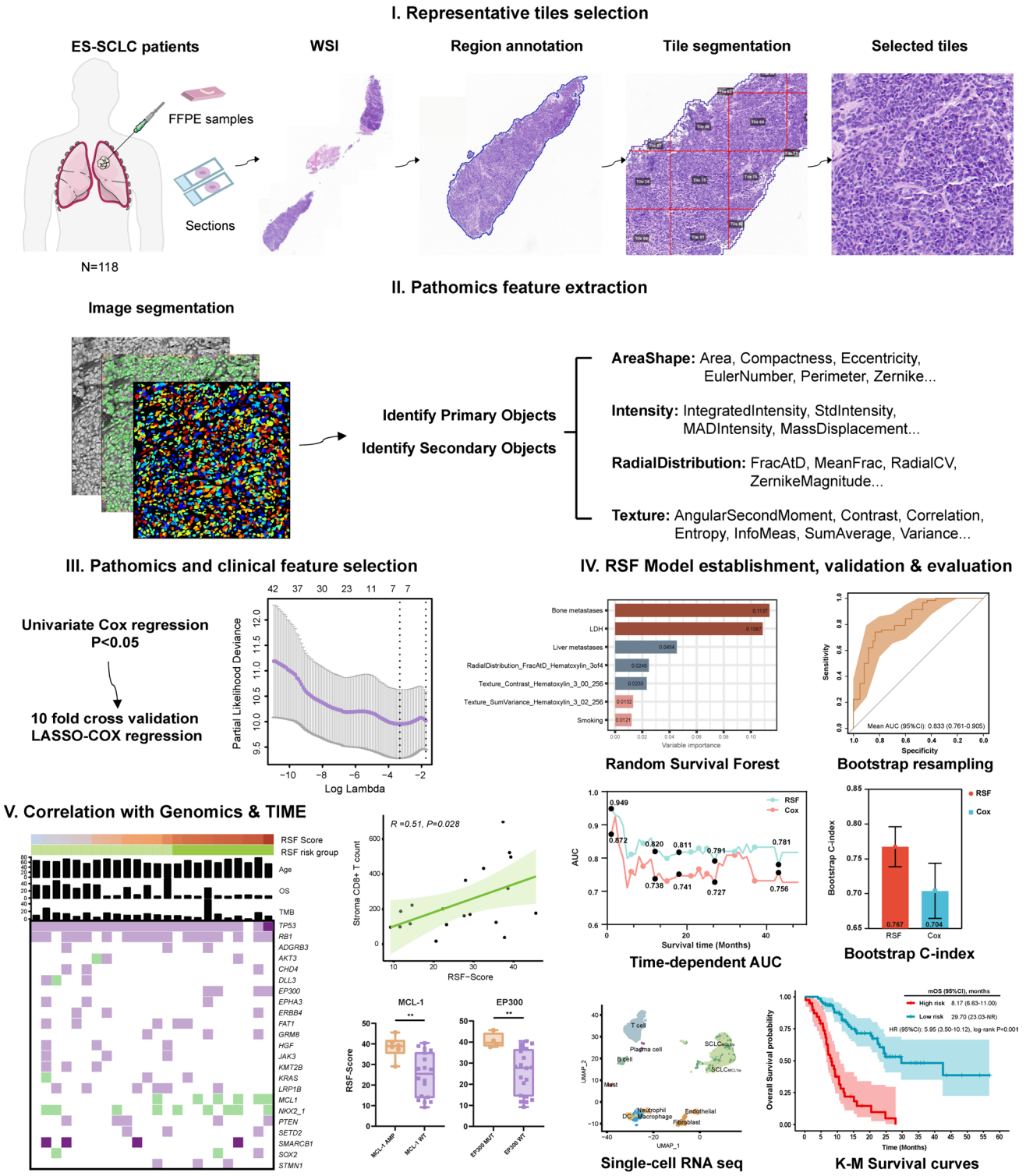

Fig. 1 Flowchart of the study

(HR): 0.68; 95% confidence interval (CI): 0.35–1.31; $P = 0.247$; etoposide plus carboplatin (EC) vs. EP, HR: 0.73; 95% CI: 0.35–1.52; $P = 0.401$] and immunotherapy regimens (HR: 0.72; 95% CI: 0.41–1.28; $P = 0.264$) did not significantly impact survival. Among various serum parameters, elevated levels of LDH were identified as an independent risk factor for OS, whereas IL-6 and CRP were not. Additionally, current or former smokers status

Table 1 Baseline characteristics and Cox regression analysis for predicting overall survival

	Subgroups	Univariate Cox regression		Multivariate Cox regression	
	Mean ± SD/ N (%)	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Age	64 ± 8	1.02 (0.99–1.05)	0.235	1.01 (0.98–1.05)	0.245
ECOG PS					
0	108 (91.5)	Reference			
1–3	10 (8.5)	1.15 (0.41–3.21)	0.784	1.18 (0.40–3.51)	0.766
Gender					
Female	17 (14.4)	0.76 (0.36–1.60)	0.465	1.00 (0.44–2.26)	0.999
Male	101 (85.6)	Reference		Reference	
Stage					
Extensive-stage	118 (100.0)	–		–	
Smoking status					
Current or former	84 (71.2)	1.99 (1.07–3.67)	***0.029***	2.37 (1.19–4.70)	***0.014***
Never	34 (28.8)	Reference		Reference	
Metastatic site					
Brain					
Yes	20 (16.9)	1.22 (0.63–2.35)	0.550	1.07 (0.50–2.28)	0.860
No	98 (83.1)	Reference		Reference	
Contralateral lung					
Yes	14 (11.9)	1.49 (0.76–2.94)	0.249	1.90 (0.87–4.15)	0.109
No	104 (88.1)	Reference		Reference	
Liver					
Yes	23 (19.5)	2.15 (1.19–3.87)	***0.011***	1.96 (1.03–3.72)	***0.040***
No	95 (80.5)	Reference		Reference	
Bone					
Yes	34 (28.8)	1.91 (1.13–3.23)	***0.016***	1.99 (1.07–3.68)	***0.029***
No	84 (71.2)	Reference		Reference	
Chemotherapy regimen					
EP	42 (35.5)	Reference			
EL	48 (40.7)	1.00 (0.56–1.79)	0.999	0.68 (0.35–1.31)	0.247
EC	28 (23.7)	0.98 (0.51–1.89)	0.961	0.73 (0.35–1.52)	0.401
Immunotherapy regimen					
PD-L1 inhibitors	48 (40.7)	0.72 (0.43–1.21)	0.218	0.72 (0.41–1.28)	0.264
PD-1 inhibitors	70 (59.3)	Reference		Reference	
Best response					
Stable disease	65 (55.1)	–		–	
Progressive disease	16 (13.6)	–		–	
Partial response	37 (31.4)	–		–	
LDH					
High (> 259.00)	46 (39.0)	2.20 (1.33–3.63)	***0.002***	1.91 (1.06–3.46)	***0.031***
Low (≤ 259.00)	72 (61.0)	Reference		Reference	
IL-6, U/L					
High (> 12.86)	39 (33.1)	1.29 (0.77–2.17)	0.335	1.22 (0.65–2.27)	0.538
Low (≤ 12.86)	79 (66.9)	Reference		Reference	
CRP, mg/L					
High (> 0.70)	85 (72.0)	1.35 (0.75–2.42)	0.312	0.83 (0.43–1.60)	0.580
Low (≤ 0.70)	33 (28.0)	Reference		Reference	

LDH, lactate dehydrogenase; CRP, C-reactive protein; PD-L1, Programmed cell death ligand-1; PD-1, Programmed cell death protein-1; ECOG PS, Eastern Cooperative Oncology Group Performance Status; EL, etoposide plus lobaplatin; EC, etoposide plus carboplatin; EP, etoposide plus cisplatin; SD, standard deviation; HR, hazard ratio; CI, confidence interval

(HR, 2.37; 95% CI: 1.19–4.70; $P = 0.014$), liver metastases (HR, 1.96; 95% CI: 1.03–3.72; $P = 0.040$), bone metastases (HR, 1.99; 95% CI: 1.07–3.68; $P = 0.029$) and LDH > 259 (HR: 1.91; 95% CI: 1.06–3.46; $P = 0.031$) were independently associated with poorer OS.

Identification of pathomics and clinical features associated with prognosis

In a panel of quantitative characteristics, 42 pathomics and 4 clinical features with $P < 0.05$ via univariate Cox regression analysis for OS were retained for further analysis (Table S2). Finally, seven features were selected using LASSO Cox regression analysis with tenfold cross-validation, including liver metastases, bone metastases, LDH, and three pathomics features (Fig. 2A). Later, we established the Cox and RSF models based on the selected features. According to the Cox model, patients with a predicted OS probability score greater than 1.251 were classified into the high-risk group, while the remaining patients were categorized into the low-risk group. It was observed that patients in the low-risk group had distinctly better OS and PFS compared to those in the high-risk group ($P < 0.001$; $P < 0.001$) (Fig. S1A, B).

Additionally, we developed an ensemble learning random forest classifier for the OS prediction based on the seven features. As illustrated in Fig. 2B-C, the ranking of each variable in the RSF model according to the variable importance is as follows: bone metastases, LDH, liver metastases, factor 1 (Mean_PrimaryObjects_RadialDistribution_FracAtD_Hematoxylin_3of4), factor 2 (Mean_PrimaryObjects_Texture_Variance_Hematoxylin_3_00_256), factor 3 (Mean_PrimaryObjects_Texture_SumVariance_Hematoxylin_3_02_256) and smoking status. Each patient was assigned a score based on the estimated RSF-Score, with an optimized cutoff value of 32.508. It was observed that the patients in the high-risk group had distinctly worse OS (HR: 5.59; 95% CI: 3.35–10.12; $P < 0.001$) and PFS (HR: 3.71; 95% CI: 2.32–5.94; $P < 0.001$) compared to those in the low-risk group. Notably, the median OS in the high-risk group was 8.17 (6.63–11.00) months, compared to 29.70 (23.03-NR) months in the low-risk group (Fig. 2D-E).

We selected representative images of long- and short-term survival patients to visualize the differences in their tumor cell morphology. Both patients presented a central type lung cancer, with patient 1 dead at 8.7 months without liver or bone metastases and smoking history, and patient 2 dead at 20.3 months with smoking history but without liver or bone metastases. However, no distinct morphological differences were visually detected between the two patients (Fig. 3A-B).

Performances of the RSF model

The performances of the Cox and RSF models in risk stratification of OS under chemoimmunotherapy were evaluated. The time-dependent ROC curve estimation of the predicted probability for OS indicated that the time-dependent AUC of the RSF model mainly ranged from 0.781 at 43 months to 0.861 at 11 months, while the Cox model displayed a fluctuating AUC between 0.727 at 21 months and 0.818 at 33 months (Fig. 4A). The RSF model demonstrated superiority over the Cox model according to the Delong test, with improved AUCs at 9 months (84.46 vs. 77.56, $P = 0.018$), 1 year (81.98 vs. 73.76, $P = 0.002$) and 2 years (81.20 vs. 73.70, $P = 0.002$) (Table 2). Furthermore, we assessed the prediction accuracy of the models. The mean C-index (SEM) of the RSF model using bootstrap resampling at 1000 times was 0.767 (0.029), compared to 0.704 (0.040) for the Cox model (Fig. 4B). Bootstrap ROC curves of the RSF model showed a mean AUC of 0.833 (0.761–0.905) (Fig. 4C). The calibration curve demonstrated a satisfactory agreement between the predicted and observed probabilities for OS based on the RSF model at 6 months, 9 months, 1 year, and 2 years (Fig. S2A). The potential clinical utility was further confirmed by DCA (Fig. S2B). To underscore the importance of incorporating pathomics features, we conducted a comparative analysis between the RSF-based combined model integrating pathomics and clinical features and the RSF-based clinical model (Clinical model), which includes bone metastases, liver metastases, LDH, and smoking status. The Delong test indicated that the RSF model outperformed the Clinical model, achieving a superior AUC of 0.812 compared to 0.736 at 2 years ($P = 0.011$) (Table S3).

The underlying association between genomics and the RSF-Score

The correlation between the pathomics signature and genomic alterations was investigated based on 24 ES-SCLC patients from the Jinling cohort who underwent additional tNGS testing, which has been published by our group beforehand (Fig. 5A). The cohort was divided into high-risk (10 patients) and low-risk (14 patients) subgroups according to the RSF-Score. We further analyzed genes with an alteration frequency greater than three. It was found that the *MCL-1* amplification and *EP300* mutation were more prevalent in the high-risk group. Specifically, the RFS-Score was significantly higher in patients with *MCL-1* amplification (AMP) (*P*) and *EP300* mutant (MUT) ($P = 0.008$) types compared to their counterparts (Fig. 5B, Table S4). Additionally, no significant differences were observed between the risk groups in terms of tumor mutation burden (TMB) ($P = 0.105$) (Fig. S3A) or PD-L1 CPS ($P = 0.338$) (Figure S3B).

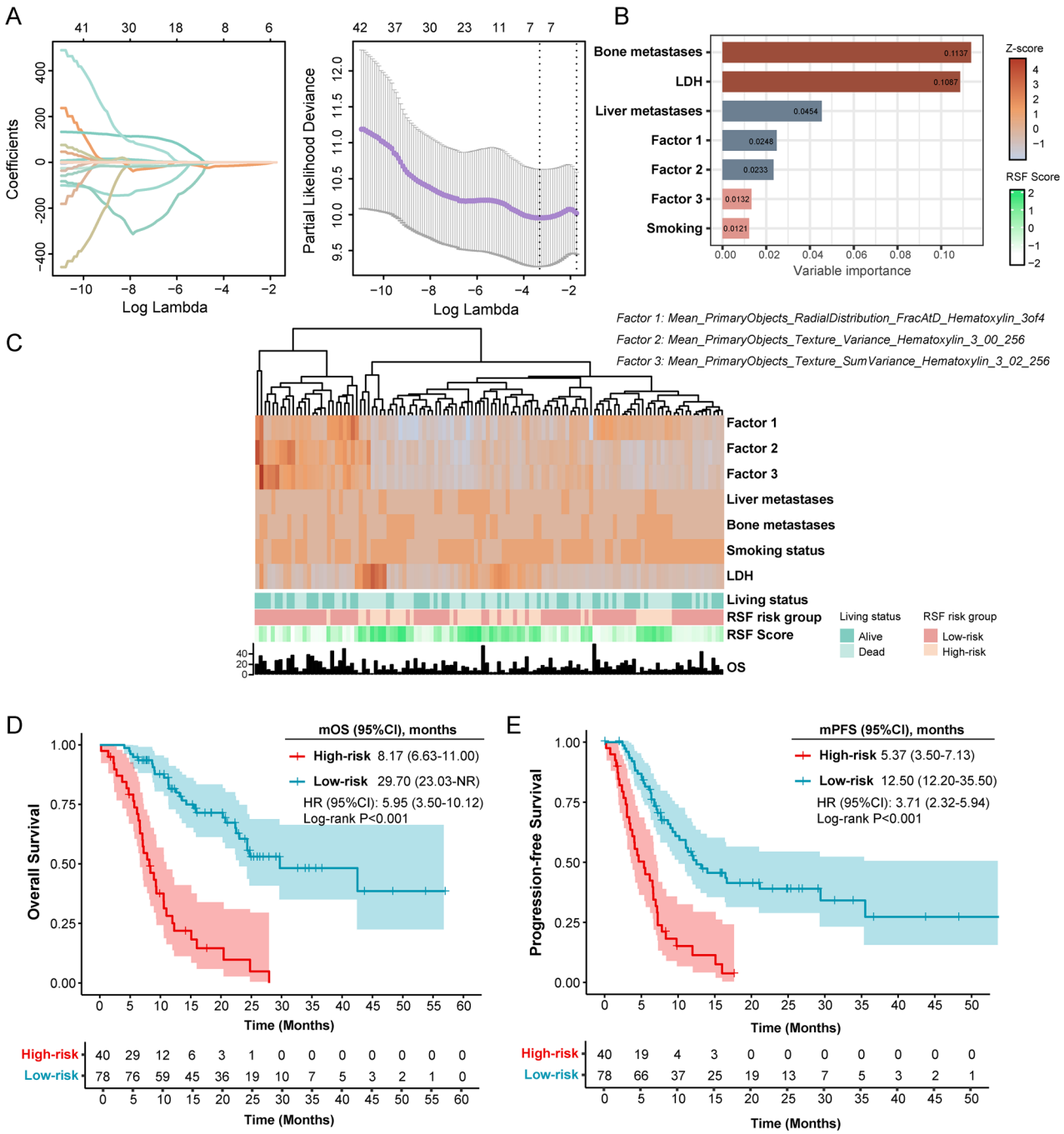

Fig. 2 Establishment of the random survival forest model indicative of survival in ES-SCLC patients receiving first-line chemoimmunotherapy. (**A**) Seven parameters were selected via LASSO-Cox regression analysis based on overall survival. (**B**) Variable importance of reserved parameters in the RSF model. (**C**) The heatmap of the RSF model. (**D**–**E**) The K-M survival curves of OS and PFS between patients in the high-risk (RSF-Score>32.508) and low-risk (RSF-Score≤32.508) groups

In a public cohort by George et al. with 81 limited-stage SCLC patients receiving both whole-genome sequencing and bulk RNA sequencing tests, we found that 9 (11.1%) patients harbored *EP300* mutation and none of them exhibited *MCL-1* amplification. Furthermore, we compared the *EP300* mRNA levels among the MUT and wild-type (WT) tumors. However, no significant disparities in *EP300* expression were found between *EP300* MUT and WT patients (Fig. S4). Furthermore, we probed into the specific mutation sites of *EP300*. In our cohort, a total of four patients harbored *EP300* mutations with varying survival outcomes and mutation sites. Specifically, three patients (exon 26: c.4195G > T,

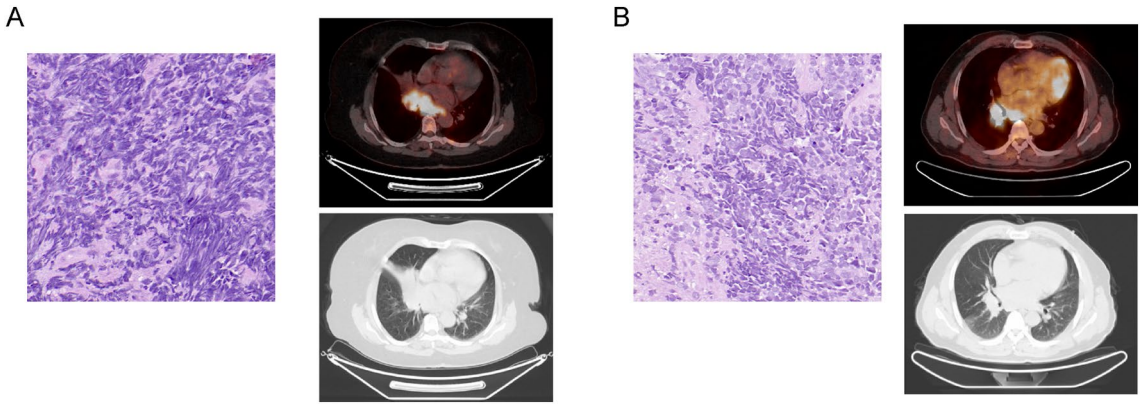

Fig. 3 Representative histopathological images and transverse section PET/CT images. (**A**) Representative images of a short-term survival patient. (**B**) Representative images of a long-term survival patient

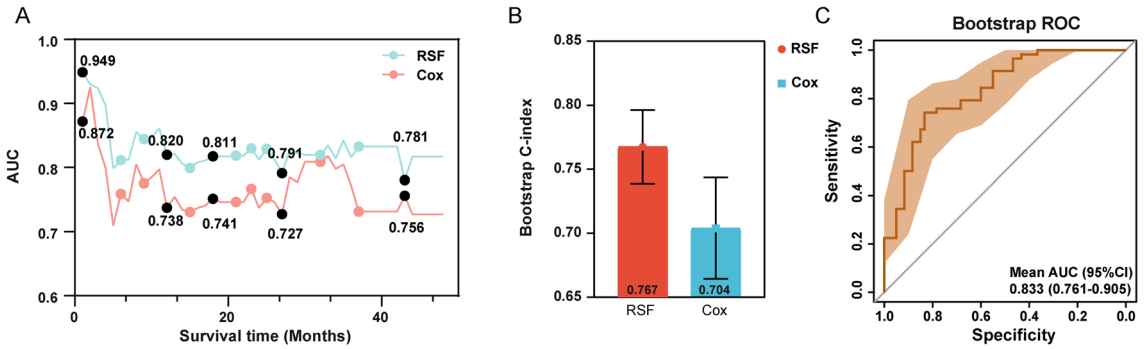

Fig. 4 The performance of the RSF and Cox models. (**A**) Time-dependent area under the curve of the RSF and Cox models. (**B**) Bootstrap C-index for the Cox and RSF model. The heights of the columns represent the average of 1000 resampling groups, and vertical lines represent the SEM of 1000 resampling groups. (**C**) Bootstrap ROC curves of the RFS model at 1000 resampling times

Table 2 The performances of the RSF model and Cox regression model for predicting overall survival

Models	Time of prediction	tAUC (95% CI)	*P*-value*
RSF model	9 mo	84.46 (76.40–92.52)	Reference
Cox model	9 mo	77.56 (66.45–88.66)	**0.018**
RSF model	1 yr	81.98 (73.43–90.53)	Reference
Cox model	1 yr	73.76 (63.04–84.48)	**0.002**
RSF model	2 yr	81.20 (71.52–90.88)	Reference
Cox model	2 yr	73.70 (62.43–84.97)	**0.022**
RSF model	3 yr	81.60 (65.59–97.61)	Reference
Cox model	3 yr	77.25 (57.23–96.08)	0.565

* Delong test comparing the RSF model to the LASSO model at different time points

LASSO, least absolute shrinkage and selection operator; tAUC, time-dependent area under the curve; RSF, random survival forest; mo, month(s); yr, year(s)

exon 13: c.2299C > G, and exon 26: c.4199G > A) succumbed within seven months, while one patient (exon 11: c.2084G > A) survived for more than 27 months.

The function of *MCL-1* and *EP300* in SCLC at single-cell level

We further elaborated on the expression of *MCL-1* and *EP300* at the single-cell level (Fig. 6A). Utilizing a dataset comprising 14,283 cells from three treatment-naïve SCLC patients, we observed robust *MCL-1* expression across all SCLC subtypes (SCLC-A, SCLC-N, and SCLC-P). Notably, elevated *MCL-1* expression was observed across various immune cell types, including mast cells, B cells, monocytes, natural killer cells, neutrophils, dendritic cells, plasma cells, and multiple T cell subtypes (Fig. 6B, C). Metascape enrichment analysis further revealed that SCLC cells exhibiting

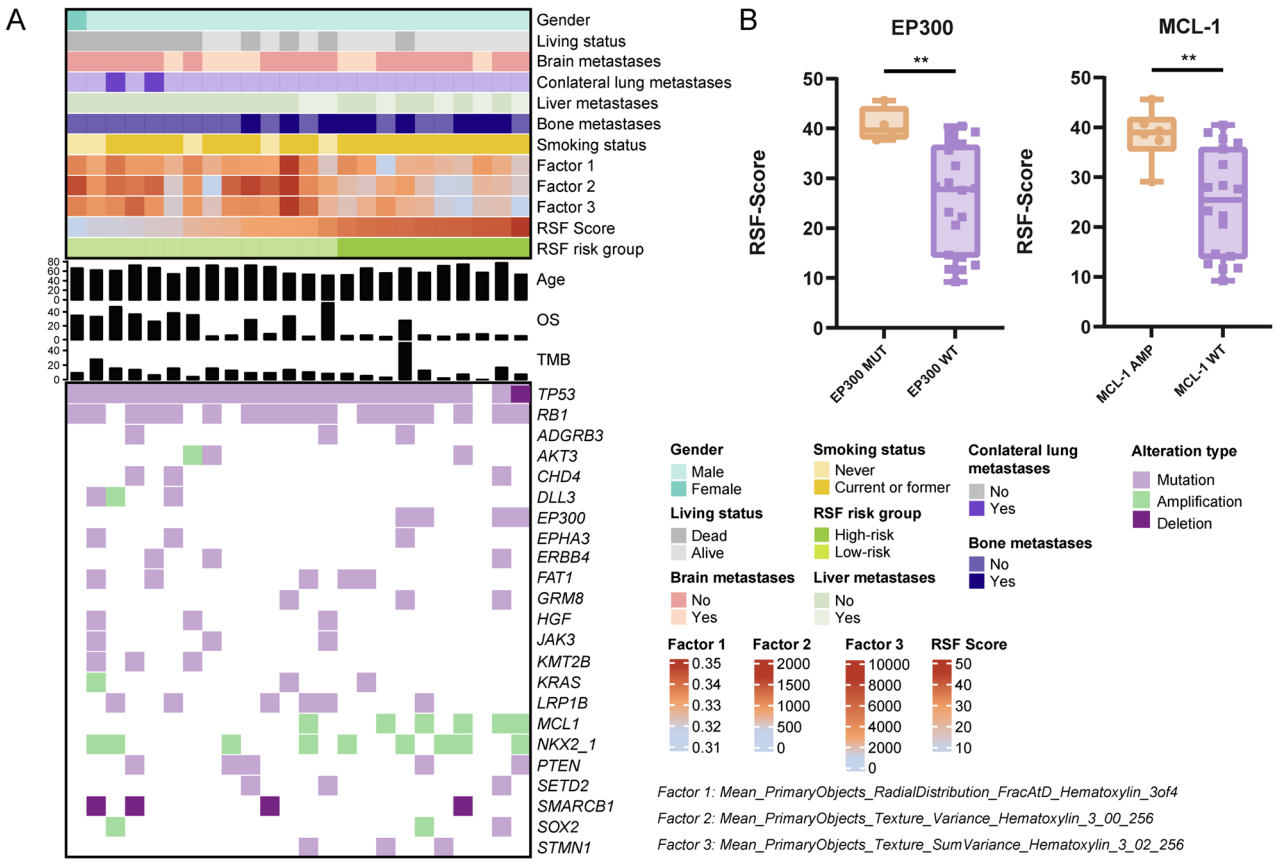

Fig. 5 The correlation between genomics and RFS-Score. (**A**) Heatmap of the genomics and clinical information patients with different RFS-Score. (**B**) The associations between RSF-Score and alterations in EP300 and *MCL-1*

increased *MCL-1* levels were concomitantly enriched in TNFA-NFKB signaling, apoptosis-related processes, and hypoxia pathways (Fig. 6D). However, *EP300* expression remained relatively limited across all subtypes, with only moderate expression observed in plasma cells (Fig. S5).

The correlation between the RSF-Score and TIME

We further investigated whether the RSF-Score could reflect the TIME by analyzing the correlation between the RSF-Score and immune cell infiltration using mIHC. Only 19 patients had accessible data for both tNGS and mIHC staining (Fig. 7A, B). Among the four classical immune cell types ($CD8^+$ T cells, $CD56^{dim+}$ and $CD56^{bright+}$ NK cells, M1-like and M2-like macrophages) identified in the parenchyma, stroma and total areas, the RSF-Score was positively correlated with the count and percentage of $CD8^+$ T cells in the stroma, both by density ($R = 0.51$, $P = 0.028$) and percentage ($R = 0.48$, $P = 0.041$) (Fig. 7C, D). However, no significant association was observed between the infiltration of $CD8^+$ T cells in the parenchyma and the RSF-Score ($R = 0.06$, $P = 0.81$; $R = 0.11$, $P = 0.65$) (Fig. S6). Additionally, we investigated the potential impact of gene alterations on immune cell population infiltration. Our analysis revealed that a sum of three patients harbored an *EP300* mutation and five patients exhibited *MCL-1* amplification. However, no significant correlation was identified for the immune cells between the wild-type and mutant-type (amplification/mutation) groups (Table S5).

Discussion

SCLC is a highly aggressive and heterogeneous neuroendocrine lung cancer with a poor prognosis and limited treatment options. The challenge of obtaining samples has significantly impeded the progress of biomarker research in SCLC, which warrants the discovery of easily accessible prognostic tools. Recently, pathomics, which involves the extraction of features from H&E-stained digital images, has emerged as a promising field in various malignancies[29, 30]. However, studies on the application of pathomics in SCLC remain scarce. In 2024, Shibaki et al. [12] pioneered the use of pathomics features derived from TIME images in SCLC. They established a deep-learning (DL) pathomics model based on tumor and immune cell features to predict

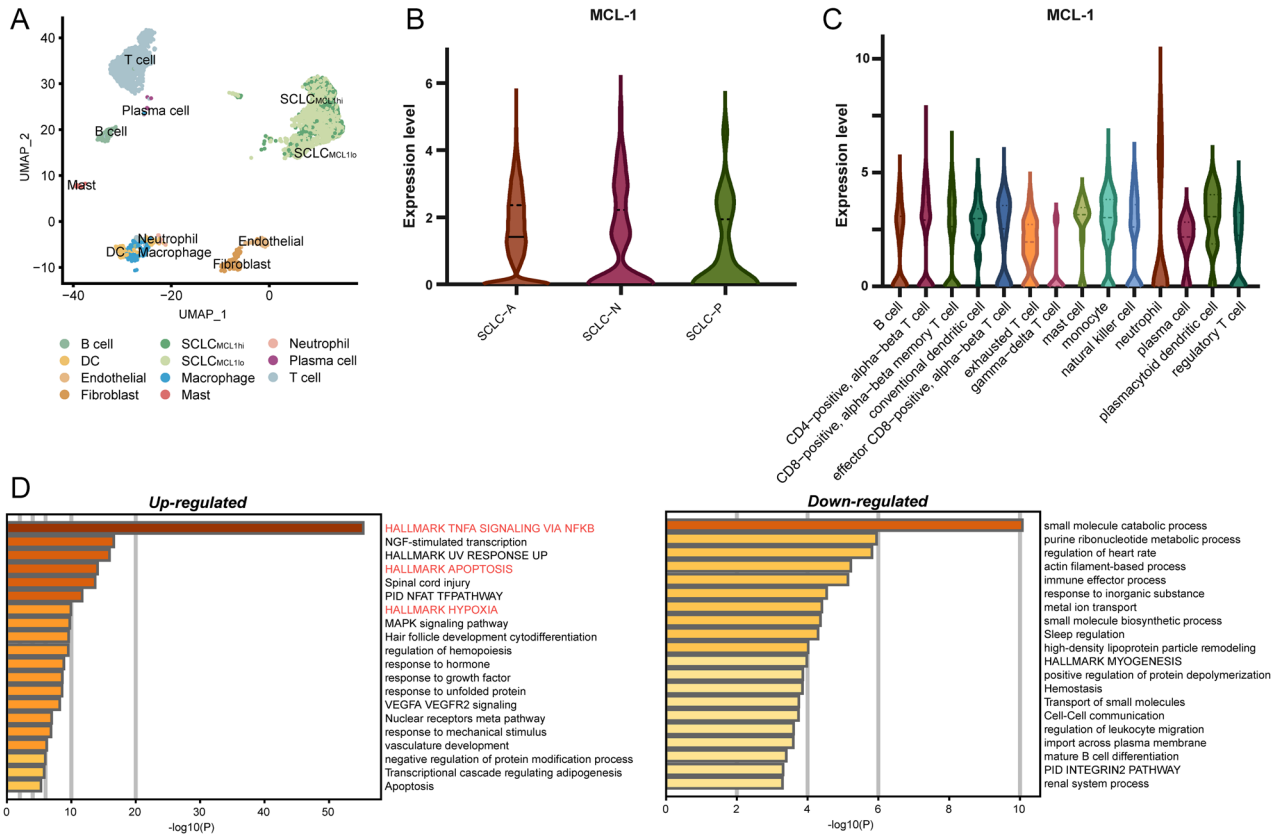

Fig. 6 The function of *MCL-1* in SCLC at single-cell level. (**A**) The UMAP plot of *MCL-1*^{hi} and *MCL-1*^{lo} SCLC cells and other cell types in the tumor microenvironment. (**B**-**C**) The expression of *MCL-1* among different SCLC molecular subtypes and immune cells. (**D**) The enrichment analysis between *MCL-1*^{hi} and *MCL-1*^{lo} SCLC cells

1-year PFS of immunotherapy in a 78-patient ES-SCLC cohort[12]. Subsequently, Zhang et al.[13] enrolled SCLC patients undergoing resection and identified 50 intricate histomorphological phenotypes for prognosis classification. However, these studies were either limited to LS-SCLC patients with surgery or bearing a small sample size with 1-year PFS as the endpoint, thus constraining the optimization of DL models. Consequently, these findings may not apply to the broader population of extensive-stage patients, and PFS may not fully represent OS, necessitating further investigations in larger cohorts. Additionally, no existing studies have explored the underlying genomic alterations beneath the pathomics signatures.

To overcome these constraints, we enrolled a larger cohort of ES-SCLC patients to develop a RSF-based combined model that integrates pathomics and clinical parameters, to predict OS under first-line chemoimmunotherapy. Unlike previous studies, we employed a more accessible and automated pipeline of tile delineation and feature selection using an open-source software. The RSF model, which encompassed pathomics features, liver metastases, bone metastases, LDH, and smoking status, demonstrated robust performance with a 2-year tAUC of 0.81. Moreover, our study was the first to suggest that the proposed RFS-Score was higher in patients with *MCL-1* AMP and *EP300* MUT. At the single-cell level, MCL-1 was associated with TNFA-NFKB signaling and apoptosis-related processes. The RSF-Score was also positively correlated with the infiltration of $CD8^+$ T cells in stromal areas. Promisingly, our findings might prompt the application of routinely performed H&E staining sections in the assessment of chemoimmunotherapy prognosis, thereby facilitating precision medicine in the management of ES-SCLC.

In this study, we identified three pathomics and four clinical features—liver metastases, bone metastases, LDH, and smoking status—as prognostic factors. It was reported that the primary and secondary objects identified by the Cellprofiler software respectively represented the cell nuclei and cell cytoplasm [18]. The selected image factors, including PrimaryObjects Radial Distribution FracAtD, Texture Variance Hematoxylin, and Texture SumVariance Hematoxylin, were all related to the shape and staining depth of the cell nucleus [31]. Specifically, Radial Distribution FracAtD represented the intensity distribution of the cell

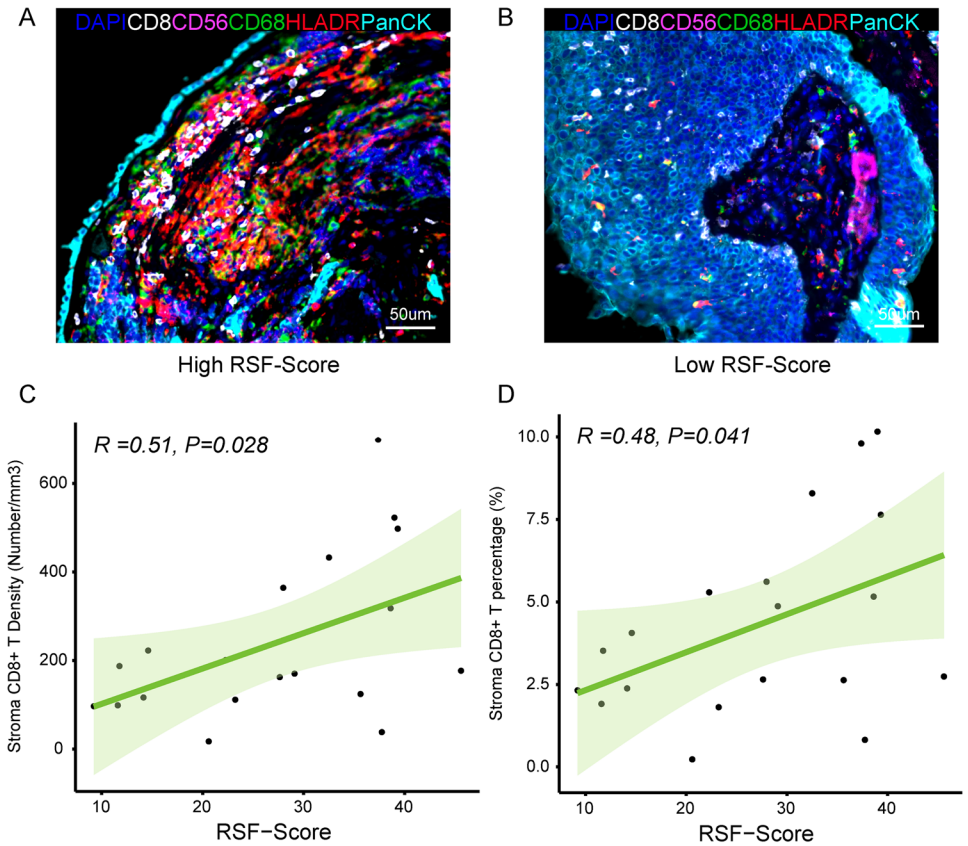

Fig. 7 The correlation between the RSF-Score and the infiltration of $CD8^+$ T cells in the stroma. (**A**–**B**) Representative multiple immunohistochemistry images of the immune cells ($CD8^+$, $CD56^+$, $CD68^+HLADR^-$, and $CD68^+HLA\text{-}DR^-$ cells) of two patients with different RSF-Scores. (**C**) The correlation between the RSF-Score and the density of $CD8^+$ T cells in the stroma. (**D**) The correlation between the RSF-Score and the percentage of $CD8^+$ T cells in the stroma

nucleus at a pixel position, while texture features provided statistical properties of variance of the gray levels of the points, descriptive of the spatial regularity or smoothness/coarseness of the cell nucleus. Pathologically, SCLC tumors are characterized by small size, scant cytoplasm, ill-defined borders, finely granular nuclear chromatin, and inconspicuous nucleoli [28]. This may explain why all the pathomics features retained in our model were associated with tumor cell nuclei. Furthermore, the subtle differences in pathological images, which are not easily discernible by visual inspection, highlight the importance of quantitative feature measurement for survival classification.

A prior DL-based study suggested that pathological images were more significant than clinical information in evaluating the efficacy of immunotherapy for SCLC [12]. This conclusion was contradictory to our findings, which underscored the importance of metastatic sites and LDH. The discrepancy may stem from different feature extraction methodologies. It has been proved that the presence of liver metastases and elevated LDH levels are associated with worse clinical responses in treatment-naïve or relapsed/refractory SCLC patients undergoing immunotherapy [32, 33]. Fundamentally, PD-1/PD-L1 inhibitors function by activating cytotoxic T lymphocytes, hindering immune escape, and sensitizing antitumor activities [34]. Consequently, markers related to the systemic immune or inflammation status, including LDH, are naturally correlated with immunotherapy response. Therefore, incorporating these clinical parameters in prognosis prediction is essential.

TIME has been identified as a potential biomarker in SCLC [11]. It has been inferred that TILs, macrophages, and their ratio are associated with survival in LS-SCLC patients treated with chemotherapy [10, 11, 35]. Recently, the infiltration of $CD8^+$ T cells was found to correlate with survival in ES-SCLC patients treated with immuno-monotherapy or doublets in CheckMate 032 [7]. Our previous investigations [36] indicated that the count of $CD8^+$T cells in the stroma and M1-like macrophages at any site could predict the survival of ES-SCLC patients with chemoimmunotherapy. We hypothesized that this might be attributed to the inability of $CD8^+$T cells to migrate into tumor parenchyma to exert their antitumoral effects. Notably, Shibaki et al. [12] confirmed the superior performance of a machine learning-based pathomics model in delineating tumor and immune cells compared to the manual measurement of individual immune cells in assessing the 1-year PFS of ES-SCLC receiving chemoimmunotherapy. The AUC of their combined model was 0.789, compared to 0.681, 0.626, and 0.567 for $CD8^+$, $FoxP3^+$ TILs, and PD-L1 score respectively. We further validated that our RSF risk group was correlated with the infiltration of $CD8^+$T cells in stromal areas, which could facilitate the assessment of TIME in a noninvasive manner.

Moreover, intrinsic alterations within cancer cells have been shown to correlate with an immunosuppressive "cold" TIME and poor responses to immunotherapy in lung cancer [37]. The alterations in several potent signaling pathways have been detected in SCLC [38], including those involving epigenetic regulators, metabolism, and DNA damage repair pathways [39]. Given the high cost of DNA sequencing techniques such as NGS, the development of a noninvasive tool to assess gene alterations might be beneficial. Promisingly, using the genomics data from the Jinling cohort published by our group [36], we demonstrated that the higher estimated RSF-Score was associated with *MCL-1* amplification and *EP300* mutation. MCL-1, an antiapoptotic member of the BCL-2 protein family, has been observed to be amplified in various cancers, including breast and pancreatic cancers [40, 41]. MCL-1 could confer resistance to chemotherapy or targeted therapy [41]. Previous studies have indicated that *MCL-1* amplification occurs in approximately 4% of all SCLC cases [42]. Although the relationship between *MCL-1* amplification and prognosis in SCLC has not been extensively studied, it is known that MCL-1 is highly expressed in SCLC cell lines and promotes tumor growth [43]. Inhibition of MCL-1 has been shown to restore antitumoral effects in SCLC cells with elevated MCL-1 levels and overcome venetoclax resistance in SCLC patients [44–46]. However, the exploration of these studies was mainly implemented in chemotherapy settings, and our exploratory analysis illustrated that its amplification could not reflect TIME. More investigations should be implemented on it in the settings of immunotherapy. Furthermore, our analysis via the public cohort revealed high expression levels of *MCL-1* in immune cells. Underlyingly, it has been reported that distinct subsets of immune cells rely on specific antiapoptotic BCL-2 family proteins for survival, with MCL-1 being a critical molecule in B cells, germinal center B cells, plasma cells, naïve T cells, memory T cells, regulatory T cells, natural killer cells, plasmacytoid dendritic cells, and neutrophils [47]. MCL-1 functions as an antiapoptotic "guardian," promoting cell survival by preventing mitochondrial outer membrane permeabilization via the downstream proapoptotic effectors BAX and BAK, which further induce apoptosis.

EP300 encodes the E1A binding protein P300, which plays a crucial role in epigenetic regulation, including transcription, cell proliferation, and cell differentiation. *EP300* mutations have been reported to occur at a frequency of 13% in SCLC. Consistent with our findings, the co-occurrence of *TP53* missense mutations and *CREBBP/EP300* mutations has been found to be associated with shorter relapse periods in SCLC patients undergoing chemotherapy [48]. Preclinical studies have demonstrated that SCLC cells with *EP300* mutations are susceptible to histone deacetylase (HDAC) inhibitors, which have been shown to upregulate both PD-L1 and MHC-I, potentially enhancing the efficacy of combined immunotherapy. Consequently, a phase I study combining HDAC inhibitors and chemoimmunotherapy in treatment-naïve ES-SCLC patients is currently underway [49]. Additionally, *EP300* mutations have demonstrated prognostic significance in various malignancies. For instance, *EP300* mutations are associated with poor survival outcomes in patients with classic Hodgkin lymphoma [50] and cervical cancer [51]. Conversely, T cell acute lymphoblastic leukemia patients harboring *EP300* mutations, in the absence of alterations in *N/K-RAS*, the PI3K pathway (*PTEN*, *PIK3CA*, and *PIK3R1*), *TP53*, *DNMT3A*, *IDH1/2*, and *IKZF*, have shown favorable prognoses [52].

TMB is an established biomarker of immunotherapy [53]. Some studies have reported that higher TMB was associated with greater immunogenicity and a favorable prognosis with immunotherapy in SCLC, which warrants further evaluation [54, 55]. Therefore, we hypothesized that our RSF risk group, which incorporated prognostic information, might reflect different levels of TMB. Although the results did not reach statistical significance, it was observed that the patients in the high-risk group exhibited lower levels of TMB compared to those in the low-risk group. The lack of statistical significance may be attributed to the small proportion of patients who underwent TMB testing; thus, larger cohorts are warranted to validate this potential connection.

To our knowledge, this study represents the largest cohort of ES-SCLC patients treated with first-line chemoimmunotherapy and evaluates immunotherapy prognosis with OS as the endpoint through RSF-based models integrating pathomics and clinical parameters. Nonetheless, certain limitations warrant further elaboration. Firstly, the retrospective and single-center nature of the study might mitigate the generalizability of our RSF model. Secondly, tile segmentation and feature extraction were not conducted using deep-learning algorithms. However, the processing procedure was carried out using a standardized pipeline with open-access software, potentially making it more accessible to physicians without a background in artificial intelligence. Additionally, the WSIs of H&E sections from biopsy samples represented only a small proportion of the tumors, which could introduce bias due to tumor heterogeneity. Lastly, only a small subset of patients underwent additional tNGS testing and TIME staining, diminishing the reliability of examining the association between our RSF-Score and genomics or TIME. Therefore, larger cohorts with multi-omics analyses are warranted to validate our findings prospectively.

Conclusion

In summary, the establishment of a RSF model combining clinical and pathomics features could prompt the prognosis assessment of chemoimmunotherapy in patients with

ES-SCLC. Higher RSF-Score putatively indicated more infiltration of $CD8^+$ T cells in tumor stroma as well as a greater probability of *MCL-1* amplification and *EP300* mutation. Hopefully, this noninvasive model holds promise as a biomarker for immunotherapy, potentially facilitating precision medicine in the management of SCLC.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00262-024-03829-9.

Acknowledgements The authors thank Geneseeq Technology Inc., Nanjing, Jiangsu, China, for their technical support and sponsorship in this research.

Author contributions P.Z., Y.S., T.F.L, and W.Q. conceived, designed, and supervised the study; Y.X.J., Q.P.C., Y.Y.C., W.J.L., X.Y.Z, Y.L., Q.X.W., and X.X.W. collected and assembled the data; Y.X.J., Q.P.C., Y.Y.C., W.J.L., and X.X.W were responsible for data analysis and interpretation; Y.X.J., Y.Y.C., W.J.L., F.Z., D.W., and Q.P.C wrote the manuscript. All athors reviewed the manuscript.

Funding This work was supported by the 16th batch "Summit of the Six Top Talents" Program of Jiangsu Province (Grant number WSN-154), China Postdoctoral Science Foundation 12th batch Special fund (Postdoctoral number: 45786), Jiangsu Provincial Postdoctoral Science Foundation (Grant number 2018K049A), Natural Science Foundation of Jiangsu Province (Grant number BK20180139), Jiangsu Provincial Health Committee Medical projects (Grant number M2022110), Natural Science Foundation of Jiangsu Province (Grant number BK20210146), and National Natural Science Foundation of China (82100095).

Data availability The raw data were accessible from the corresponding author under reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Consent for publication Informed consent was obtained from each patient before sample collection for research purposes, and all research process was performed under the relevant guidelines and regulations.

Ethics approval This retrospective study was conducted according to the ethics of the Declaration of Helsinki (as revised in 2013) and the Institutional Review Board of Nanjing Jinling Hospital (registration ID. 2024NZKY-050–01). Informed consent was obtained from each patient before sample collection for research purposes.

References

1. Bray F, Laversanne M, Sung H et al (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. https://doi.org/10.3322/caac.21834
2. Megyesfalvi Z, Gay CM, Popper H et al (2023) Clinical insights into small cell lung cancer: tumor heterogeneity, diagnosis, therapy, and future directions. CA A Cancer J Clinicians 73:620–652. https://doi.org/10.3322/caac.21785
3. Yu Y, Chen K, Fan Y (2023) Extensive-stage small-cell lung cancer: current management and future directions. Int J Cancer 152:2243–2256. https://doi.org/10.1002/ijc.34346
4. Liu SV, Dziadziuszko R, Sugawara S et al (2023) OA01.04 five-year survival in patients with ES-SCLC treated with atezolizumab in IMpower133: imbrella a extension study results. J Thorac Oncol 18:S44–S45. https://doi.org/10.1016/j.jtho.2023.09.025
5. Lissa D, Takahashi N, Desai P et al (2022) Heterogeneity of neuroendocrine transcriptional states in metastatic small cell lung cancers and patient-derived models. Nat Commun 13:2023. https://doi.org/10.1038/s41467-022-29517-9
6. Xie M, Chugh P, Broadhurst H, et al (2022) Abstract CT024: Durvalumab (D) + platinum-etoposide (EP) in 1L extensive-stage small-cell lung cancer (ES-SCLC): Exploratory analysis of SCLC molecular subtypes in CASPIAN. Cancer Research 82:CT024. https://doi.org/10.1158/1538-7445.AM2022-CT024
7. Rudin CM, Balli D, Lai WV et al (2023) Clinical benefit from immunotherapy in patients with SCLC is associated with tumor capacity for antigen presentation. J Thorac Oncol 18:1222–1232. https://doi.org/10.1016/j.jtho.2023.05.008
8. Bindea G, Mlecnik B, Tosolini M et al (2013) Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 39:782–795. https://doi.org/10.1016/j.immuni.2013.10.003
9. Angelova M, Mlecnik B, Vasaturo A et al (2018) Evolution of metastases in space and time under immune selection. Cell 175:751-765.e16. https://doi.org/10.1016/j.cell.2018.09.018
10. Longo V, Catino A, Montrone M et al (2021) What are the biomarkers for immunotherapy in SCLC? Int J Mol Sci 22:11123. https://doi.org/10.3390/ijms222011123
11. Li T, Qiao T (2022) Unraveling tumor microenvironment of small-cell lung cancer: implications for immunotherapy. Semin Cancer Biol 86:117–125. https://doi.org/10.1016/j.semcancer.2022.09.005
12. Shibaki R, Fujimoto D, Nozawa T et al (2024) Machine learning analysis of pathological images to predict 1-year progression-free survival of immunotherapy in patients with small-cell lung cancer. J Immunother Cancer 12:e007987. https://doi.org/10.1136/jitc-2023-007987
13. Zhang Y, Yang Z, Chen R, Zhu Y, Liu L, Dong J, Zhou M (2024) Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer. NPJ digital medicine 7(1):15
14. Wang Y, Li X, Gang Q et al (2024) Pathomics and single-cell analysis of papillary thyroid carcinoma reveal the pro-metastatic influence of cancer-associated fibroblasts. BMC Cancer 24:710. https://doi.org/10.1186/s12885-024-12459-4
15. Wang Y, Zhou C, Li T, Luo J (2024) Prognostic value of CDKN2A in head and neck squamous cell carcinoma via pathomics and machine learning. J Cell Mol Med 28:e18394. https://doi.org/10.1111/jcmm.18394
16. Seymour L, Bogaerts J, Perrone A et al (2017) iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol 18:e143–e152. https://doi.org/10.1016/S1470-2045(17)30074-8

17. Carpenter AE, Jones TR, Lamprecht MR et al (2006) Cell Profiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7:R100. https://doi.org/10.1186/gb-2006-7-10-r100
18. Yu K-H, Zhang C, Berry GJ et al (2016) Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 7:12474. https://doi.org/10.1038/ncomms12474
19. Chen S, Zhang N, Jiang L et al (2021) Clinical use of a machine learning histopathological image signature in diagnosis and survival prediction of clear cell renal cell carcinoma. Int J Cancer 148:780–790. https://doi.org/10.1002/ijc.33288
20. Chen S, Jiang L, Zheng X et al (2021) Clinical use of machine learning-based pathomics signature for diagnosis and survival prediction of bladder cancer. Cancer Sci 112:2905–2914. https://doi.org/10.1111/cas.14927
21. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS (2008) Random survival forests. The Annals of Applied Statistics 2:841–860. https://doi.org/10.1214/08-AOAS169
22. Steyerberg EW, Vickers AJ, Cook NR et al (2010) Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 21:128–138. https://doi.org/10.1097/EDE.0b013e3181c30fb2
23. Yang Z, Yang N, Ou Q et al (2018) Investigating novel resistance mechanisms to third-generation EGFR tyrosine kinase inhibitor osimertinib in non-small cell lung cancer patients. Clin Cancer Res 24:3097–3107. https://doi.org/10.1158/1078-0432.CCR-17-2310
24. Shu Y, Wu X, Tong X et al (2017) Circulating tumor DNA mutation profiling by targeted next generation sequencing provides guidance for personalized treatments in multiple cancer types. Sci Rep 7:583. https://doi.org/10.1038/s41598-017-00520-1
25. George J, Lim JS, Jang SJ et al (2015) Comprehensive genomic profiles of small cell lung cancer. Nature 524:47–53. https://doi.org/10.1038/nature14664
26. Chan JM, Quintanal-Villalonga Á, Gao VR et al (2021) Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. Cancer Cell 39:1479-1496.e18. https://doi.org/10.1016/j.ccell.2021.09.008
27. Zhou Y, Zhou B, Pache L et al (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10:1523. https://doi.org/10.1038/s41467-019-09234-6
28. Petty WJ, Paz-Ares L (2023) Emerging strategies for the treatment of small cell lung cancer: a review. JAMA Oncol 9:419–429. https://doi.org/10.1001/jamaoncol.2022.5631
29. Amgad M, Hodge JM, Elsebaie MAT et al (2024) A population-level digital histologic biomarker for enhanced prognosis of invasive breast cancer. Nat Med 30:85–97. https://doi.org/10.1038/s41591-023-02643-7
30. Wang Y, Pan X, Lin H et al (2022) Multi-scale pathology image texture signature is a prognostic factor for resectable lung adenocarcinoma: a multi-center, retrospective study. J Transl Med 20:595. https://doi.org/10.1186/s12967-022-03777-x
31. Haralick R, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3(6):610–621
32. Stratmann JA, Timalsina R, Atmaca A et al (2022) Clinical predictors of survival in patients with relapsed/refractory small-cell lung cancer treated with checkpoint inhibitors: a German multicentric real-world analysis. Ther Adv Med Oncol 14:17588359221097192. https://doi.org/10.1177/17588359221097191
33. Xie J, Xu K, Cai Z et al (2024) Efficacy and safety of first-line PD-L1/PD-1 inhibitors in limited-stage small cell lung cancer: a multicenter propensity score matched retrospective study. Transl Lung Cancer Res. 13(3):526–539
34. He X, Xu C (2020) Immune checkpoint signaling and cancer immunotherapy. Cell Res 30:660–669. https://doi.org/10.1038/s41422-020-0343-4
35. Muppa P, Parrilha Terra SBS, Sharma A et al (2019) Immune cell infiltration may be a key determinant of long-term survival in small cell lung cancer. J Thorac Oncol 14:1286–1295. https://doi.org/10.1016/j.jtho.2019.03.028
36. Jiang Y, Xie J, Cheng Q et al (2024) Comprehensive genomic and spatial immune infiltration analysis of survival outliers in extensive-stage small cell lung cancer receiving first-line chemo-immunotherapy. Int Immunopharmacol 141:112901. https://doi.org/10.1016/j.intimp.2024.112901
37. Otegui N, Houry M, Arozarena I et al (2023) cancer cell-intrinsic alterations associated with an immunosuppressive tumor microenvironment and resistance to immunotherapy in lung cancer. Cancers (Basel) 15:3076. https://doi.org/10.3390/cancers15123076
38. Yuan M, Zhao Y, Arkenau H-T et al (2022) Signal pathways and precision therapy of small-cell lung cancer. Signal Transduct Target Ther 7:187. https://doi.org/10.1038/s41392-022-01013-y
39. Li W, Ye L, Huang Y et al (2022) Characteristics of Notch signaling pathway and its correlation with immune microenvironment in SCLC. Lung Cancer 167:25–33. https://doi.org/10.1016/j.lungcan.2022.03.019
40. Yue H, Huang R, Shan Y, Xing D (2020) Delivery of cas13a/crRNA by self-degradable black phosphorus nanosheets to specifically inhibit Mcl-1 for breast cancer therapy. J Mater Chem B 8:11096–11106. https://doi.org/10.1039/d0tb01914c
41. Merino D, Whittle JR, Vaillant F et al (2017) Synergistic action of the MCL-1 inhibitor S63845 with current therapies in preclinical models of triple-negative and HER2-amplified breast cancer. Sci Transl Med. https://doi.org/10.1126/scitranslmed.aam7049
42. Ross JS, Wang K, Elkadi OR et al (2014) Next-generation sequencing reveals frequent consistent genomic alterations in small cell undifferentiated lung cancer. J Clin Pathol 67:772–776. https://doi.org/10.1136/jclinpath-2014-202447
43. Inoue-Yamauchi A, Jeng PS, Kim K et al (2017) Targeting the differential addiction to anti-apoptotic BCL-2 family for cancer therapy. Nat Commun 8:16078. https://doi.org/10.1038/ncomms16078
44. Yasuda Y, Ozasa H, Kim YH et al (2020) MCL1 inhibition is effective against a subset of small-cell lung cancer with high MCL1 and low BCL-XL expression. Cell Death Dis 11:177. https://doi.org/10.1038/s41419-020-2379-2
45. Valko Z, Megyesfalvi Z, Schwendenwein A et al (2023) Dual targeting of BCL-2 and MCL-1 in the presence of BAX breaks venetoclax resistance in human small cell lung cancer. Br J Cancer 128:1850–1861. https://doi.org/10.1038/s41416-023-02219-9
46. Khan S, Kellish P, Connis N et al (2023) Co-targeting BCL-XL and MCL-1 with DT2216 and AZD8055 synergistically inhibit small-cell lung cancer growth without causing on-target toxicities in mice. Cell Death Discov 9:1. https://doi.org/10.1038/s41420-022-01296-8
47. Carrington EM, Zhan Y, Brady JL et al (2017) Anti-apoptotic proteins BCL-2, MCL-1 and A1 summate collectively to maintain survival of immune cell populations both in vitro and in vivo. Cell Death Differ 24:878–888. https://doi.org/10.1038/cdd.2017.30
48. George J, Maas L, Abedpour N et al (2024) Evolutionary trajectories of small cell lung cancer under therapy. Nature 627:880–889. https://doi.org/10.1038/s41586-024-07177-7
49. Gentzler RD, Villaruz LC, Rhee JC et al (2023) Phase i study of entinostat, atezolizumab, carboplatin, and etoposide in previously untreated extensive-stage small cell lung cancer, ETCTN 10399. Oncologist 28:1007-e1107. https://doi.org/10.1093/oncolo/oyad221
50. Santisteban-Espejo A, Bernal-Florindo I, Montero-Pavon P et al (2024) Pathogenic variants associated with epigenetic control and

the NOTCH pathway are frequent in classic hodgkin lymphoma. Int J Mol Sci 25:2457. https://doi.org/10.3390/ijms25052457
51. He H, He M, Zhou Q et al (2024) Genetic analysis of cervical cancer with lymph node metastasis. J Gynecol Oncol. https://doi.org/10.3802/jgo.2024.35.e102
52. Simonin M, Vasseur L, Lengliné E et al (2024) NGS-based stratification refines the risk stratification in T-ALL and identifies a very high-risk subgroup of patients. Blood blood. https://doi.org/10.1182/blood.2023023754
53. Ettinger DS, Wood DE, Aisner DL et al (2023) NCCN guidelines® insights: non-small cell lung cancer, version 2.2023. J Natl Compr Canc Netw 21:340–350. https://doi.org/10.6004/jnccn.2023.0020
54. Ricciuti B, Kravets S, Dahlberg SE et al (2019) Use of targeted next generation sequencing to characterize tumor mutational burden and efficacy of immune checkpoint inhibition in small cell lung cancer. J Immunother Cancer 7:87. https://doi.org/10.1186/s40425-019-0572-6
55. Hellmann MD, Callahan MK, Awad MM et al (2019) Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with Ipilimumab in small-cell lung cancer. Cancer Cell 35:329. https://doi.org/10.1016/j.ccell.2019.01.011

Authors and Affiliations

Yuxin Jiang[1] · Yueying Chen[2] · Qinpei Cheng[2] · Wanjun Lu[2] · Yu Li[3] · Xueying Zuo[2] · Qiuxia Wu[4] · Xiaoxia Wang[5] · Fang Zhang[2,3,6] · Dong Wang[2,3,6] · Qin Wang[2] · Tangfeng Lv[1,2,3,6] · Yong Song[1,2,3,6] · Ping Zhan[1,2,3,6]

✉ Qin Wang
wq.026@163.com

✉ Tangfeng Lv
bairoushui@163.com

✉ Yong Song
yong.song@nju.edu.cn

✉ Ping Zhan
zhanping207@163.com

[1] School of Medicine, Southeast University, Nanjing 210000, China

[2] Department of Respiratory and Critical Care Medicine, Affiliated Hospital of Medical School, Jinling Hospital, Nanjing University, Nanjing 210002, China

[3] Department of Respiratory and Critical Care Medicine, Jinling Hospital, Affiliated Hospital of Nanjing Medical School, Nanjing 210002, China

[4] Jinling Clinical Medical College, Nanjing University of Chinese Medicine, Nanjing 210002, China

[5] Department of Pathology, Affiliated Hospital of Medical School, Jinling Hospital, Nanjing University, Nanjing 210002, China

[6] Department of Respiratory and Critical Care Medicine, School of Medicine, Jinling Hospital, Southeast University, 305 Zhongshan East Road, Nanjing 210002, China