



OPEN ACCESS

EDITED BY

Jian Song,
University Hospital Münster, Germany

REVIEWED BY

In-Jae Oh,
Chonnam National University Hwasun
Hospital, South Korea
Albino Eccher,
Integrated University Hospital Verona,
Italy

*CORRESPONDENCE

Jie Peng
sank44@sina.com

[†]These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Cancer Immunity
and Immunotherapy,
a section of the journal
Frontiers in Immunology

RECEIVED 03 June 2022

ACCEPTED 10 October 2022

PUBLISHED 07 November 2022

CITATION

Peng J, Zhang J, Zou D, Xiao L, Ma H,
Zhang X, Li Y, Han L and Xie B (2022)
Deep learning to estimate
durable clinical benefit and
prognosis from patients with
non-small cell lung cancer treated
with PD-1/PD-L1 blockade.
Front. Immunol. 13:960459.
doi: 10.3389/fimmu.2022.960459

COPYRIGHT

© 2022 Peng, Zhang, Zou, Xiao, Ma,
Zhang, Li, Han and Xie. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

Deep learning to estimate durable clinical benefit and prognosis from patients with non-small cell lung cancer treated with PD-1/PD-L1 blockade

Jie Peng^{1†*}, Jing Zhang^{2†}, Dan Zou^{1†}, Lushan Xiao^{3†},
Honglian Ma⁴, Xudong Zhang⁵, Ya Li¹, Lijie Han⁶
and Baowen Xie⁷

¹Department of Medical Oncology, The Second Affiliated Hospital, Guizhou Medical University, Kaili, China, ²Department of Radiology, Nanfang Hospital, Southern Medical University, Guangzhou, China, ³Hepatology Unit and Department of Infectious Diseases, Nanfang Hospital, Southern Medical University, Guangzhou, China, ⁴Department of Radiation Oncology, Cancer Hospital of the University of Chinese Academy of Sciences, Hanzhou, China, ⁵The Second Department of Radiation Oncology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, ⁶Department of Hematology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, ⁷Yino Research, Shenzhen Yino Intelligence Technology Development Co., Ltd., Shenzhen, China

Different biomarkers based on genomics variants have been used to predict the response of patients treated with PD-1/programmed death receptor 1 ligand (PD-L1) blockade. We aimed to use deep-learning algorithm to estimate clinical benefit in patients with non-small-cell lung cancer (NSCLC) before immunotherapy. Peripheral blood samples or tumor tissues of 915 patients from three independent centers were profiled by whole-exome sequencing or next-generation sequencing. Based on convolutional neural network (CNN) and three conventional machine learning (cML) methods, we used multi-panels to train the models for predicting the durable clinical benefit (DCB) and combined them to develop a nomogram model for predicting prognosis. In the three cohorts, the CNN achieved the highest area under the curve of predicting DCB among cML, PD-L1 expression, and tumor mutational burden (area under the curve [AUC] = 0.965, 95% confidence interval [CI]: 0.949–0.978, $P < 0.001$; AUC = 0.965, 95% CI: 0.940–0.989, $P < 0.001$; AUC = 0.959, 95% CI: 0.942–0.976, $P < 0.001$, respectively). Patients with CNN-high had longer progression-free survival (PFS) and overall survival (OS) than patients with CNN-low in the three cohorts. Subgroup analysis confirmed the efficient predictive ability of CNN. Combining three cML methods (CNN, SVM, and RF) yielded a robust comprehensive nomogram for predicting PFS and OS in the

three cohorts (each $P < 0.001$). The proposed deep-learning method based on mutational genes revealed the potential value of clinical benefit prediction in patients with NSCLC and provides novel insights for combined machine learning in PD-1/PD-L1 blockade.

KEYWORDS

deep learning, durable clinical benefit, non-small cell lung cancer, PD-1/PD-L1 blockade, prognosis

Introduction

Immune checkpoint blockade (ICB) therapy has been proven to be successful as a treatment for non-small cell lung cancer (NSCLC) (1, 2), and different biomarkers, such as PD-L1 expression (3), tumor mutational burden (TMB) (4, 5), and gene expression profile (GEP) (6), have been recently associated with ICB response. However, the predictive values of these biomarkers are relatively limited because of the low predictive accuracies. Thus, the search for useful and precise biomarkers for predicting ICB response is critical.

Increasing studies have reported that mutated genes carrying single nucleotide variants (SNVs) are significantly related with the ICB response (7, 8). For example, *STK11*, *B2M*, and *EGFR* mutations or *MDM2* amplification were associated with poor responsiveness or even hyper-progressive disease (HPD) (9, 10), whereas *TP53*, *KRAS*, and *POLE* mutations or KP (co-mutations of *KRAS* and *TP53*) molecular sub-type were positively related with ICB response in advanced NSCLC (11, 12). Patients with KL (co-mutations of *KRAS* and *STK11*) showed poor responses (13, 14). Moreover, NSCLC patients with mutations or co-mutations of *DDR* and *Notch* pathways had clinical benefit from ICB (15–17). These findings reveal the potential values of exploiting a novel method for predicting clinical benefit using a mutational database.

Deep learning, especially as convolutional neural network (CNN), has frequently been applied to medical images for the diagnosis, predictive prognosis, and therapy response assessment of patients with cancer (18–20). However, there is unclear whether the CNN based on SNV database could predict the clinical outcome of immunotherapy in the patients with NSCLC.

Therefore, we sought to use CNN algorithm based on a panel of genomic mutations to develop a robust model for selecting patients with advanced NSCLC who are responsive to ICB therapy. CNN based on next-generation sequencing (NGS) and whole-exome sequencing (WES) databases was used to predict ICB benefit in patients with NSCLC from three large cohorts. The CNN model showed a better predictive ability than PD-L1, TMB, and conventional machine learning (cML) models. Moreover, we combined CNN and cML models to

build a nomogram for predicting the prognosis of immunotherapy. A robust comprehensive classification of genomic panels would facilitate the selection of patients who would benefit from ICB.

Materials and methods

Patients treated with immunotherapy

POPLAR/OAK cohort: A total of 287 patients with advanced or metastatic NSCLC were recruited in the POPLAR study (NCT01903993) (21). The exclusion criteria were as follows: second or third-line standard therapy of docetaxel (n=143), treatment with atezolizumab (n=144), and no blood TMB (bTMB) data (n=39). Finally, 105 patients were retained and unchosen for PD-L1 expression status. The OAK study (NCT02008227), a randomized phase III trial, recruited 850 patients with metastatic NSCLC to compare atezolizumab with docetaxel in the primary analysis population (22). A total of 425 patients who received second or third-line standard therapy of docetaxel, 425 patients who were treated with atezolizumab, and 101 patients who had no bTMB data were excluded. Finally, 324 patients were retained and unchosen for PD-L1 expression status. The POPLAR and OAK cohorts had 429 patients in total as a training cohort.

UCMC cohort: Patients with locally advanced or metastatic NSCLC who were treated with only anti-PD-1 or a combination of chemo-immunotherapy treatment in the University of Chicago Medical Center (UCMC) were investigated (23). The patients had undergone tumor NGS test prior to the initiation of ICB therapy. Between 2016 and 2020, of the 426 patients treated with ICB, 139 who undergone tumor NGS test prior to the initiation of ICB therapy were deemed eligible to participate in this study. Two patients without TMB database were excluded. Finally, the tumor samples of 137 patients were evaluated for PD-L1 expression or EGFR/ALK mutational status, and results were validated using deep learning and cML algorithms.

MSKCC cohort: The Memorial Sloan Kettering Cancer Center (MSKCC) cohort of patients with advanced NSCLC

receiving anti-PD-1 treatment were derived from three clinical studies. In the first cohort, 75 patients with stage IV NSCLC were treated with a combination of nivolumab and ipilimumab in the CheckMate-012 clinical trial (NCT01454102) between February 2013 and March 2015 (24). In the second cohort, 34 patients with metastatic NSCLC treated with anti-PD-1 treatment were derived from the MSKCC (n=29) and University of California at Los Angeles (n=5) (NCT01295827) studies (25). In the third cohort, the data of 240 patients treated with only anti-PD-1 or a combination of anti-CTLA-4 and anti-PD-1 between April 2011 and January 2017 were retrospectively collected (10). A total of 349 patients were considered as another validation cohort.

This study was approved by the institutional review board of the Second Affiliated Hospital of Guizhou Medical University and was conducted in accordance with the tenets of the Declaration of Helsinki.

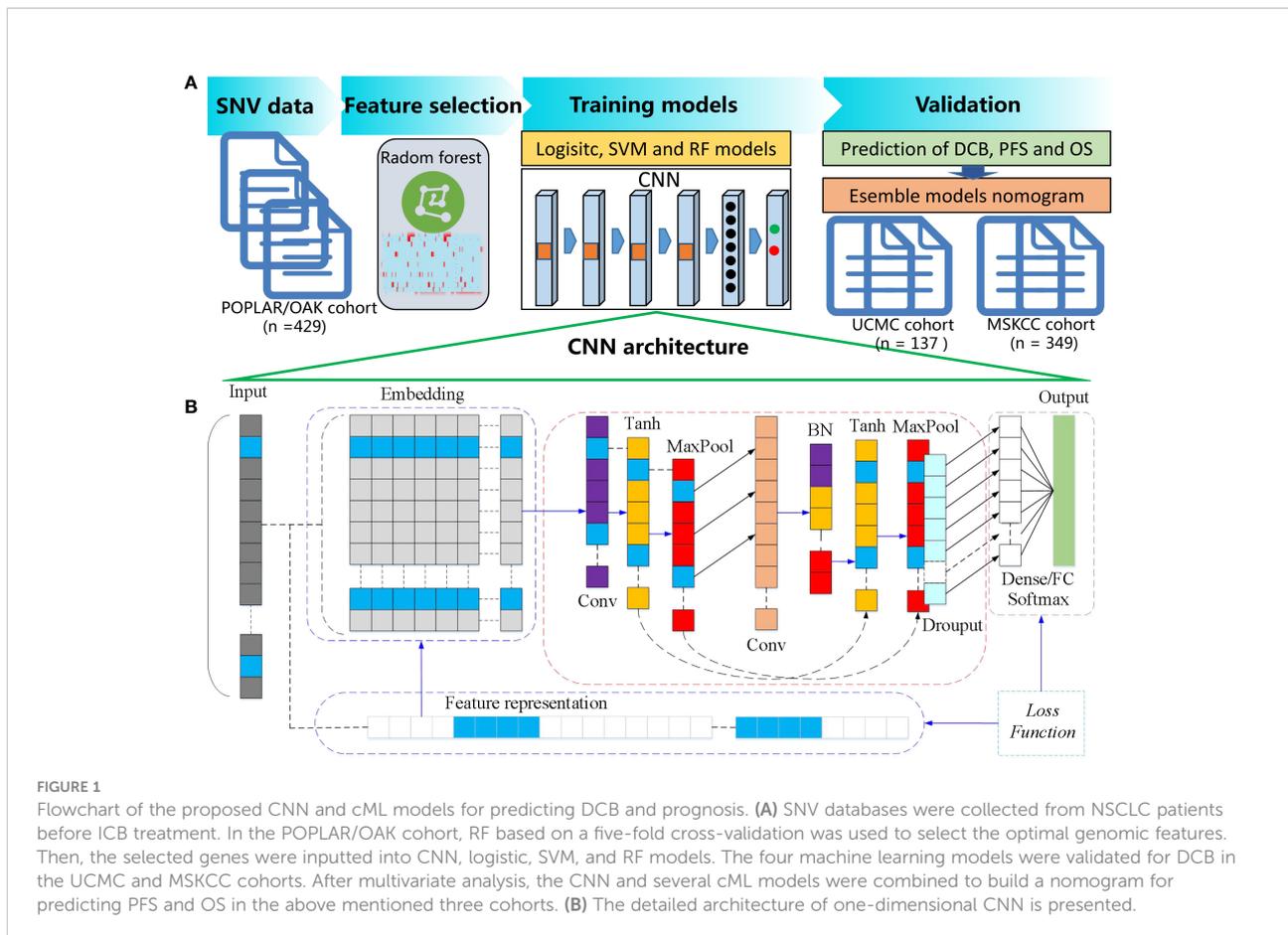
Study design

The flowchart of the proposed CNN and cML models for predicting DCB and prognosis is shown in Figure 1A. The SNV databases of sequencing results of tumor or blood samples from

the patients with NSCLC before ICB treatment were collected. In the POPLAR/OAK cohort, the optimal genomic features were selected by RF algorithm based on a five-fold cross-validation as previously described (26). The selected genes were input into CNN, logistic, support vector machine (SVM), and random forest (RF) models and were used to train for DCB prediction. After adjusting the parameters, the four machine learning models were validated for DCB in the UCMC and MSKCC cohorts. The associations between the predictive scores of four models and prognosis were analyzed. CNN model was further trained on the data of the clinical subgroups. After multivariate analysis, the CNN and cML models were combined to build a nomogram for predicting PFS and OS in the above mentioned three cohorts.

DCB, PFS, and OS

The primary outcome measures of this study were durable clinical benefit (DCB), progression-free survival (PFS), and overall survival (OS). The Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 were used to evaluate complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). We defined DCB as CR, PR,



or SD lasting ≥ 6 months, whereas no durable benefit (NDB) was defined as either SD lasting < 6 months or PD. We defined PFS as the time from the start of PD-1/PD-L1 blockade therapy to death or the first confirmation of PD based on RECIST version 1.1. OS was defined as the time from the start of PD-1/PD-L1 blockade therapy until the last contact or death.

WES, targeted NGS, TMB, and PD-L1 analysis

WES and targeted NGS of tumor and blood samples were performed before ICB. DNA and circulating tumor DNA (ctDNA) were separately extracted from formalin-fixed paraffin-embedded (FFPE) tumor masses and peripheral blood samples from the patients. The different sequencing assays, including WES and targeted NGS (MSK-IMPACT and OncoPlus), were performed as described in [Supplementary Methods](#).

Based on the results of WES and targeted NGS profiling, we defined a high TMB as ≥ 20 /Mb or total somatic nonsynonymous as ≥ 200 , and low TMB as < 20 /Mb or total nonsynonymous mutations as < 200 . The E1L3N (Cell Signaling, Danvers, MA, USA), 22C3 (DAKO), and 28-8 (DAKO) were used to determine the PD-L1 expression of tumor cells. Positive PD-L1 expression was defined as $> 1\%$ staining.

Selection of genomics features and construction of conventional machine learning models

The genomics features of the SNVs were selected by RF function and five-fold cross-validation sampling. Three cML algorithms, including SVM, logistic, and RF, were employed. Based on the selected genomics features, three cML models were built by R packages (“randomForest”, “caret”, and “e1071”).

CNN architecture

As shown in the [Figure 1B](#), the architecture of the one-dimensional CNN included a one-dimensional convolution layer, with a convolution kernel of 16, a spatial domain of convolution kernel of 128, and a stride of 1. Firstly, the input information was processed as embedded. Secondly, we used the tanh activation function, followed by the Maxpooling method to reduce the dimension. After the first dimensionality reduction, one-dimensional convolution calculation was carried out for the vector, with a convolution kernel of 32 and a spatial domain of convolution kernel of 3. Then, batch normalization (BN) was carried out. Similarly, tanh activation function was used, and dimensionality was reduced by Maxpooling. Adam was used as deep neural network optimization, the gradient descent method

was SGD, and the learning rate was 0.01. On the basis of the above, the fully-connected feed forward network (FCN) of dense and the output result of the Softmax activation function were used as previously described ([27](#)).

$$Z^{l+1}(i, j) = [Z^l \otimes w^{l+1}](i, j) + b$$

$$= \sum_{k=1}^{K_i} \sum_{x=1}^f \sum_{y=1}^f [Z_k^l(s_0 i + x, s_0 j + y) w_k^{l+1}(x, y)] + b$$

$$(i, j) \in \{0, 1, \dots, L_{l+1}\} \quad L_{l+1} = \frac{L_l + 2p - f}{s_0} + 1$$

The summation part in the above formula is equivalent to solving a cross correlation, where b is the deviation; Z^l and Z^{l+1} represent the convolution input and output of layers $l+1$, respectively, also known as feature map; L_{l+1} is the dimension of Z_{l+1} ; K is the number of channels; and f , s_0 , and p are the convolution layer parameters corresponding to the size of convolution kernel, convolution stride, and the number of padding layers, respectively.

The tanh activation function is shown below:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

The Softmax activation function is expressed as follows:

$$S_i = \frac{e^{v_i}}{\sum_c e^{v_i}}$$

where V_i is the output of the output unit of the front stage of the classifier, i represents the category index, C is the total number of categories, and S_i represents the ratio of the index of the current element to the sum of the indexes of all elements.

Cross Entropy Loss was calculated as below:

$$loss = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})]$$

where y is the real value, and \hat{y} is the predicted value.

CNN implementation

The selected 55 genomics features were input to the CNN model. To ensure adequate performance of training process, the maximum number of epochs was set to 600. Implementation of deep learning was based on the TensorFlow-1.14 in Python (<https://www.python.org/>). The experiment was performed in a Windows environment with a 3.7 GHz Intel i7-12700KF CPU, NVIDIA GeForce RTX 3090, and 32 GB of RAM.

Statistical analysis

The ROC curves were plotted and evaluated for accuracy using the “pROC” package. The area under the curve (AUC) and the corresponding 95% confidence interval (CI) were calculated

in the three cohorts. PFS and OS curves were analyzed by Kaplan–Meier method and plotted by the survminer package. Multivariate Cox regression of the CNN, logistic, SVM, and RF models was analyzed, and significant variables ($P < 0.01$) were used to build the nomogram using “rms” package. The HRs for PFS and OS in the CNN-low and CNN-high subgroups were analyzed and visualized by the “Forestplot” package. KEGG (Kyoto Encyclopedia of Genes and Genomes) and Gene Ontology (GO) were analyzed in DAVID (<https://david.ncifcrf.gov/home.jsp>). All statistical analyses were conducted in R version 3.5.1 (<https://www.r-project.org/>) and GraphPad Prism 7.01 (<https://www.graphpad.com/>). $P < 0.05$ was defined as statistically significant.

Results

Characteristics of patients who received ICB therapy

The basic clinical characteristics of the patients with NSCLC in the POPLAR/OAK, UCMC, and MSKCC cohorts are presented in [Supplementary Table 1](#). A total of 275 (64.10%), 61 (44.53%), and 171 (49.00%) patients were men in the POPLAR/OAK, UCMC, and MSKCC cohorts, respectively. In the three cohorts, there were 265 (61.77%), 92 (67.15%), and 222 (63.61%) patients aged >60 years old. Majority (82.05%, 87.59%, and 80.51%) of the patients were current or ever smokers. Patients with no-squamous NSCLC comprised 70.86% and 94.26% of the POPLAR/OAK and MSKCC cohorts. We found that 15 (3.50%), 20 (14.60%), and 71 (20.34%) patients had high TMB (≥ 200 or >20 /Mb) in the three cohorts and that the TMB status is a stratifying variable in the different populations. In the POPLAR/OAK, UCMC, and MSKCC cohorts, 59 (13.75%), 45 (32.84%), and 43 (12.33%) patients had positive PD-L1 expression ($>1\%$), respectively. There were 286 (31.27%) patients in the three cohorts who were not tested for PD-L1 expression. In the POPLAR/OAK, UCMC, and MSKCC cohorts, 134 (31.24%), 57 (41.61%), and 131 (37.54%) patients achieved DCB, respectively.

Landscape of selection genomics and building of cML for DCB

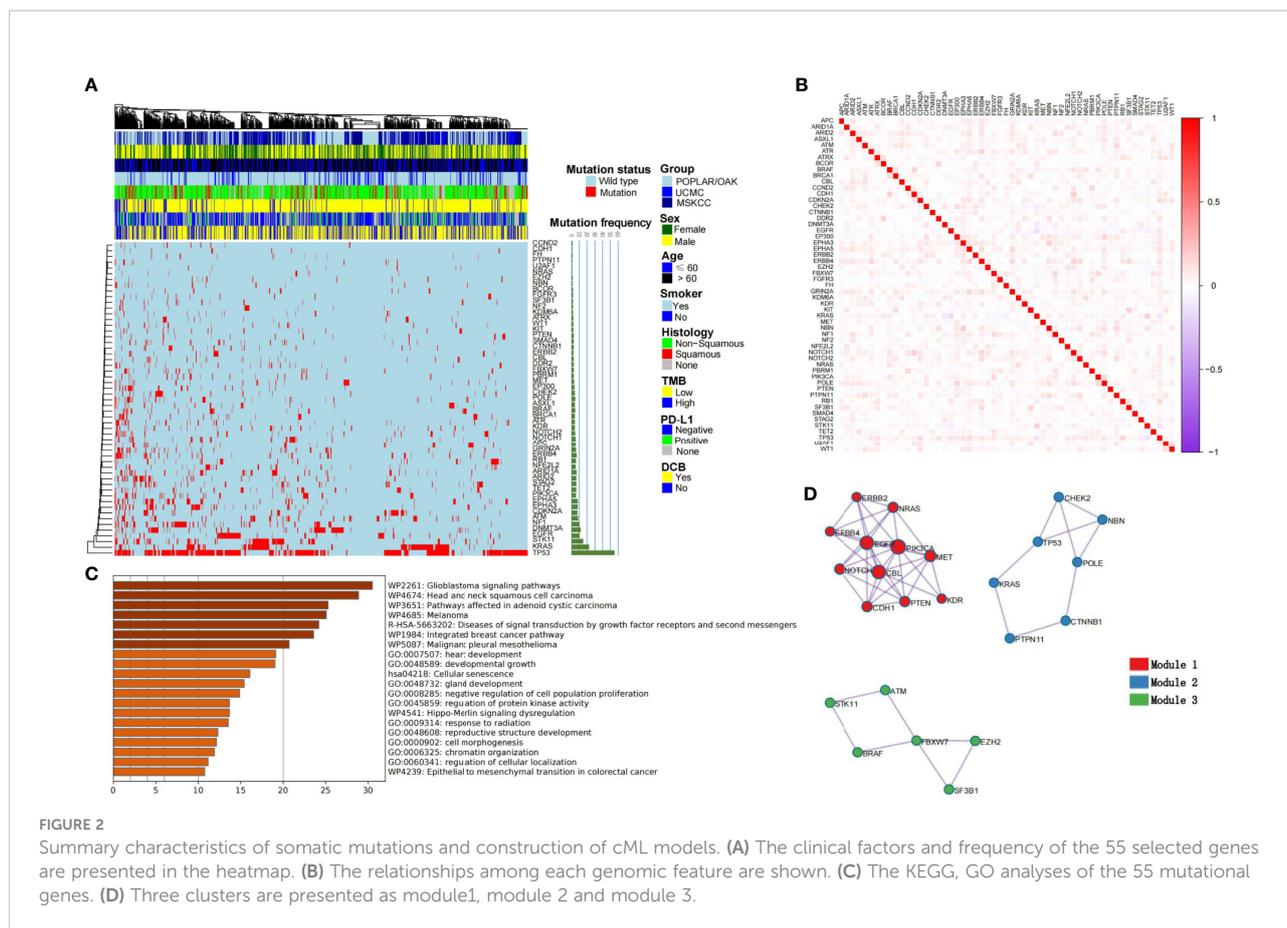
Based on the five-fold cross-validation, the RF was used to select the optimal mutational genomics from the POPLAR/OAK cohort, and 55 somatic mutations were finally chosen. The importance and Gini coefficients of the top 30 somatic mutations are shown in [Supplementary Figure 1](#). The ARID1A and ERBB4 showed the important roles of the 55 genomics features. We summarized the clinical and somatic mutations in the three patient cohorts with NSCLC ([Figure 2A](#)). Seven mutational subtypes (nonsense mutation, missense mutation, frame shift del,

frame shift ins, splice site, inframe del, and multi hit) were detected in the three cohorts, and the frequencies of the 55 selected genes in each case are shown as a heatmap. *TP53*, *KRAS*, and *STK11* showed high mutational frequency (55.30%, 22.62%, and 15.19%) in the total cohort ($n = 915$). The correlations among 55 somatic mutations are presented in [Figure 2B](#). *Notch1* and *POLE* had a positive correlation ($r = 0.133$, $P < 0.001$), whereas *KRAS* and *EGFR* showed a negative correlation ($r = -0.150$, $P < 0.001$).

KEGG revealed 55 mutational genes that are associated with different cancer pathways, such as the glioblastoma signaling pathway, head and neck squamous cell carcinoma, and melanoma (FDR $P < 0.001$; [Figure 2C](#)). GO analysis showed that these genomics are related with development growth, regulation of protein kinase activity (PKA), and response to radiation. Modular analysis revealed 55 genes that could be classified into three groups. Model1 consisted of *ERBB2*, *ERBB4*, *NRAS*, *EGFR*, *PIK3CA*, *NOTCH1*, *CBL*, *MET*, *CDH1*, *PTEN*, and *KDR* ([Figure 2D](#)). Based on the three cML algorithms, we used the panel of 55 mutational genes to train the model to predict DCB. The training process of RF is shown in [Supplementary Figure 2](#), with the number of trees set at 100. The RF model showed a better and more stable prediction than the logistic and SVM methods in the POPLAR/OAK, UCMC, and MSKCC cohorts ([Supplementary Table 2](#)).

Convolutional neural network was trained and tested for DCB in the three ICB cohorts

The deep-learning model of CNN was implemented in the TensorFlow platform as the set parameters. The POPLAR/OAK cohort was trained, and the UCMC cohort was validated in the process of 600 epochs ([Figure 3A](#)). The curves of training accuracy and loss had consistent trend with the validating curves. Then, the MSKCC cohort was tested using the trained CNN model. We found significant associations between the three cML (logistic, SVM, and RF) models and the CNN model (each $P < 0.001$) ([Figure 3B](#)), with the strongest correlation between the CNN and SVM models ($r = 0.905$, 95% CI: 0.886–0.920, $P < 0.001$). The CNN model had the highest AUCs in the POPLAR/OAK, UCMC, and MSKCC cohorts (AUC = 0.965, 95% CI: 0.949–0.978, $P < 0.001$; AUC = 0.965, 95% CI: 0.940–0.989, $P < 0.001$; AUC = 0.959, 95% CI: 0.942–0.976, $P < 0.001$, respectively; [Figure 3C](#)). Comparing the predictive accuracy of TMB, PD-L1, logistic, SVM, RF, and CNN, the CNN had the highest AUCs in the three cohorts, and we found that TMB and PD-L1 had similar AUCs ([Figure 3D](#)). The CNN model also had significantly higher sensitivity and specificity than TMB and PD-L1 in the POPLAR/OAK cohort (sensitivity = 97.01, 95% CI: 92.53–99.18, $P < 0.001$; specificity = 76.95, 95% CI: 71.72–81.63, $P < 0.001$), the UCMC cohort (sensitivity = 94.74, 95% CI: 85.38–98.90, $P < 0.001$; specificity = 85.00, 95% CI: 75.26–92.00, $P < 0.001$), and the MSKCC cohort (sensitivity = 83.97, 95% CI: 76.55–89.79, $P < 0.001$; specificity



= 90.37, 95% CI: 85.65–93.94, $P < 0.001$; [Supplementary Table S3](#)). A case with a positive PD-L1 expression and a low TMB showed a significant response. The somatic mutations included *TP53*, *KRAS*, *EPHA5*, *ARID1A*, and *POLE*. The CNN and the three cML models showed different scores in this patient ([Figure 3E](#)).

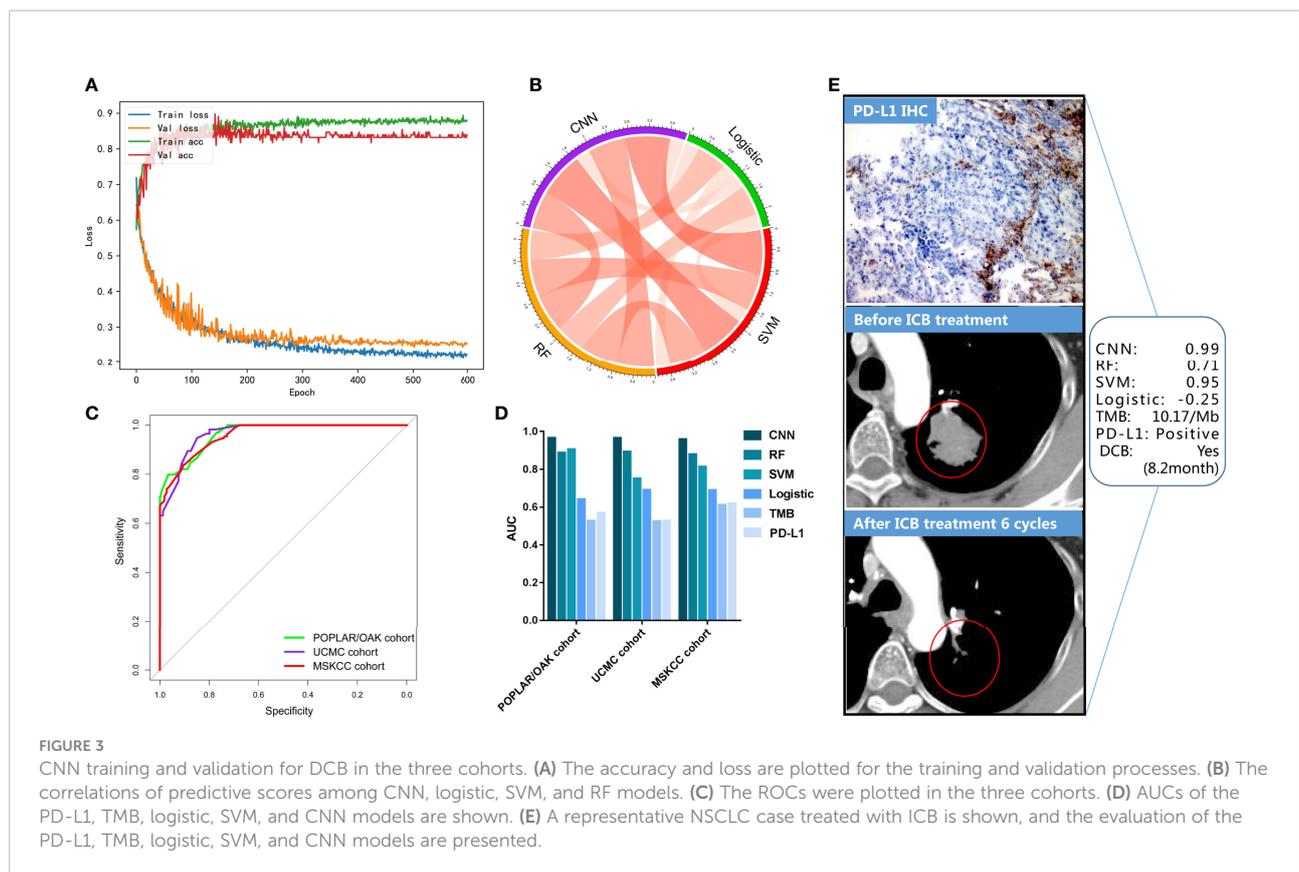
Convolutional neural network and conventional machine learning predict PFS and OS in NSCLC patients with ICB

According to the cut-off value (0.31) of CNN scores as the biggest Youden Index (YI), the patients with NSCLC treated with ICB were stratified into CNN-high (> 0.31) or CNN-low (≤ 0.31) groups. The CNN-high group had a longer median PFS than that of the CNN-low group (mPFS: 1.41 vs 9.29 months) (HR = 3.67 [2.94–.57], $P < 0.001$; [Figure 4A](#)) in the POPLAR/ OAK cohort with anti-PD-1 therapy. In the UCMC and MSKCC cohorts, the CNN-high group also showed better PFS than that by the CNN-low group (both $P < 0.001$; [Figures 4B, C](#)). We then validated the predictive OS of the CNN model in the POPLAR/OAK cohort and found that the CNN-high group had a longer median OS (mOS: 6.70 vs 22.04 months) (HR = 3.20 [2.52–4.06], $P < 0.001$) than that of the CNN-low group ([Figure 4D](#)). The CNN-high group had better OS than

that of the CNN-low group in both the UCMC and MSKCC cohorts (both $P < 0.001$; [Figures 4E, F](#)). PD-L1 expression was a significant predictor of PFS in the MSKCC cohort ($P < 0.001$), and TMB was a significant biomarker for PFS in the POPLAR/OAK and MSKCC cohorts ($P = 0.004$ and $P < 0.001$; [Supplementary Table 4](#)). PD-L1 expression was a significant predictor of OS in the POPLAR/OAK and MSKCC cohorts ($P = 0.005$ and $P = 0.022$), and TMB was a significant predictor of OS in the MSKCC cohort ($P = 0.049$, [Supplementary Table 5](#)). The three cML models accurately predicted the OS and PFS in the three cohorts ([Supplementary Tables 4, 5](#)). We analyzed the prediction of CNN in the clinical subgroups across different variables. Combining the three cohorts, the CNN model showed a good prediction as a biomarker of PFS and OS ([Figures 4G–H](#)).

Nomogram of combining CNN, SVM, and RF predicts PFS and OS in patients with ICB

In the POPLAR/OAK cohort, we used six variables (including logistic, SVM, RF, CNN, TMB, and PD-L1) for the multivariate analysis of PFS and OS. The SVM, RF, and CNN models were identified as significant independent factors for PFS ($P < 0.001$,



0.014, and < 0.001) and SVM, CNN, and PD-L1 ($P < 0.001$, < 0.001 , and 0.005) were significant independent factors for OS (Supplementary Table 6).

Considering the low predictive ability of PD-L1 expression for prognosis, we used the SVM, RF, and CNN models to develop a comprehensive nomogram for predicting PFS (herein, EMN; Figure 5A). According to the two cut-off values (0 and 0.8), we stratified the three cohorts into EMN-low (≤ 0), EMN-intermediate ($0 < \text{and} \leq 0.8$), and EMN-high (> 0.8) groups. The EMN-low had a longer median PFS and OS than the EMN-intermediate and EMN-high groups (mPFS: 13.37 vs 2.72 vs 1.41 months, $P < 0.00$; mOS: not reached [NR] vs 12.41 vs 6.66 months, $P < 0.001$) (Figures 5B, E) in the POPLAR/OAK cohort. Similarly, the patients with EMN-low had better PFS and OS than those with EMN-intermediate and EMN-high in the UCMC and MSKCC cohorts (mPFS: 20.41 vs 8.94 vs 2.79 months, $P < 0.001$; 18.90 vs 3.60 vs 2.10 months, $P < 0.001$; mOS: NR vs 20.42 vs 7.36 months, $P < 0.001$; NR vs 18.00 vs 8.00 months, $P < 0.001$; Figures 5C–E, G).

Discussion

In this study, we used deep learning and cML methods based on the NGS or WES data to develop predictive models for

DCB in 915 patients with NSCLC treated with ICB from three independent cohorts. To our knowledge, this is the largest study conducted to predict ICB response based on the sequencing data of patients with NSCLC. To avoid overfitting of the training model, the RF algorithm was first used to reduce the genomics features, and 55 somatic mutations were finally selected. After tuning the parameters, the CNN and cML models showed high prediction accuracies for DCB in the POPLAR/OAK, UCMC, and MSKCC cohorts. In the four models, namely CNN, SVM, RF, and logistic, a significant association with PFS and OS in the above three cohorts was noted. Subgroup analysis of CNN revealed that deep learning had robust prediction in different clinical variables. After multivariate analysis, we used SVM, RF, and CNN to build a nomogram for predicting PFS and found that this nomogram could stratify patients into three groups. The three groups of patients had different PFS and OS in the POPLAR/OAK, UCMC, and MSKCC cohorts.

Although WES, WGS, and NGS databases from blood or tumor tissue samples have been increasingly and extensively used in cancer research, most studies have frequently focused on several gene panels or sole driver mutational gene, resulting in the inefficient use of large sequencing data, especially in somatic mutations (28–30). In this study, we mainly sought to use a relatively small panel of mutational genes to develop a robust

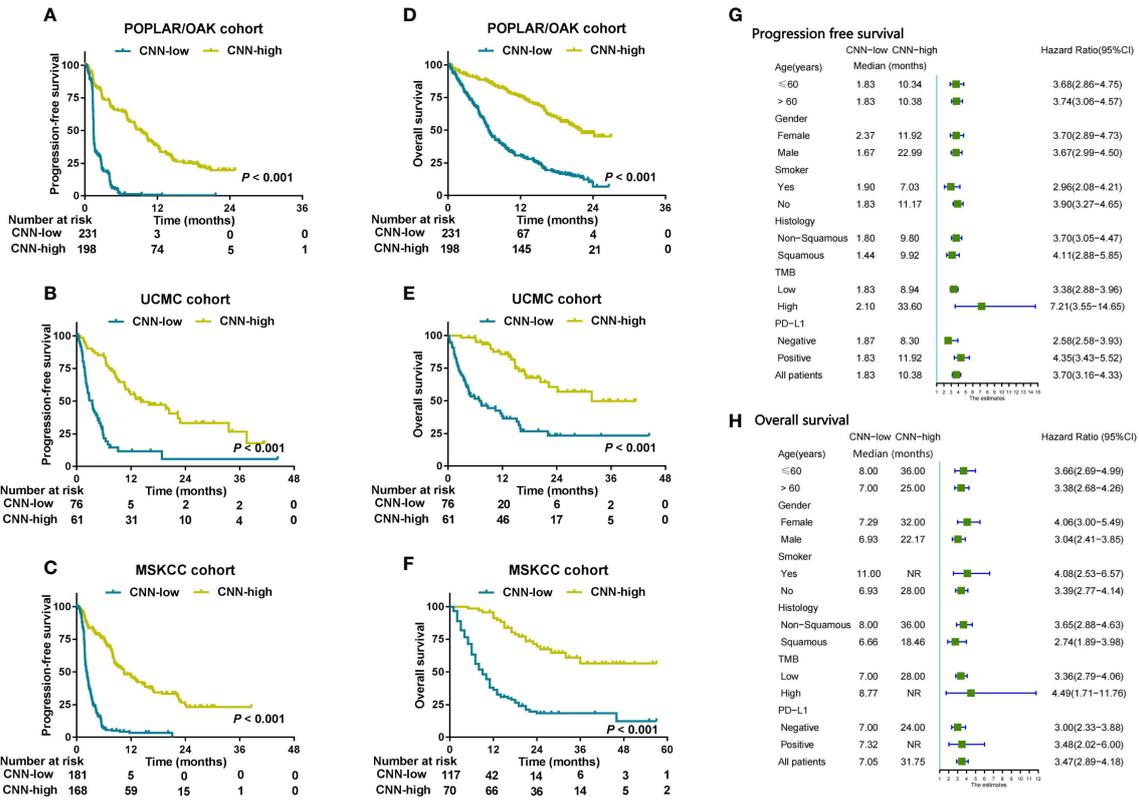


FIGURE 4 CNN was used to predict prognosis and analyzed in clinical subgroups. (A–C) Kaplan–Meier survival curves showing PFS between the CNN-low and CNN-high groups in patients from the three cohorts. (D–F) Kaplan–Meier survival curves showing OS between the CNN-low and CNN-high groups in three cohorts. (G, H) Subgroup analysis of CNN for PFS and OS from the combination cohorts (n = 915) according to basic clinical variables.

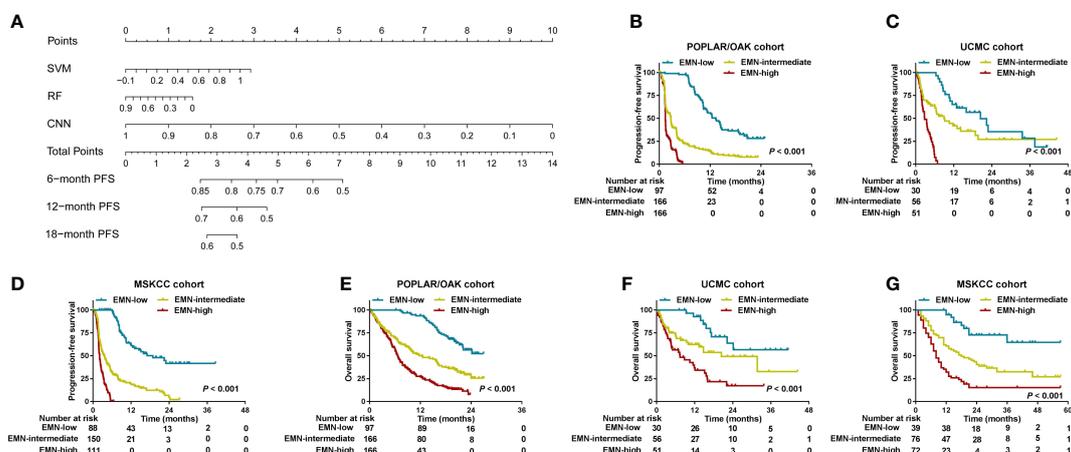


FIGURE 5 Development and analysis of a nomogram for PFS and OS in the three cohorts. (A) A nomogram was constructed using SVM, RF, and CNN methods in the POPLAR/OAK cohort. (B–D) PFS curves were compared among the EMN-low, EMN-intermediate, and EMN-high subgroups of patients from the three cohorts. (E–G) OS curves were compared among the EMN-low, EMN-intermediate, and EMN-high subgroups of patients from the three cohorts.

predicting model for ICB response. Comparing the different cML models (SVM, RF, and logistic), we found that the CNN model had the highest precision for DCB prediction. This was the first study that used CNN to train somatic mutations but not routine images. More research should be conducted to determine whether CNN can potentially perform on classification of predicting DCB from large WES, NGS, or WGS data but not use a simple TMB. We also found CNN had a higher and more stable ability for predicting DCB than TMB and PD-L1 expression status. Because our genomics sequencing data were collected from circulating tumor DNA (ctDNA) and analysis in the training model (POPLAR/OAK) were validated from the other two cohorts (UCMC and MSKCC) with tumor tissue sequencing, these findings reveal that our CNN model could be potentially used as a non-invasive method to predict ICB response of patients with NSCLC.

Previous studies have reported that patients with various cancers who are responsive to ICB treatment frequently have a better prognosis than nonresponsive patients (31, 32). The CNN, SVM, logistic, and RF models for predicting DCB were significantly associated with PFS and OS. The patients with CNN-high had significantly better PFS and OS than those with CNN-low in the POPLAR/OAK, UCMC, and MSKCC cohorts. This finding revealed that the CNN model for predicting DCB could effectively assess the clinical prognosis of ICB therapy in patients with NSCLC. In contrast, TMB and PD-L1 expression showed unsatisfactory results for predicting PFS and OS in the three cohorts. We speculated that the different cut-off values or detection platform for TMB could have contributed to the uncertain predictive effect. New method of tumor mutational burden on cytological samples from a pilot study could be feasible and the application was unclear (33). PD-L1 is becoming a fundamental data for patient management and the role of pathologist and immunohistochemical assessment should be emphasized (34). The testing results of the PD-L1 assay varied across various reagents from several manufacturers (35). New evidences regarding clones, platforms, reporting system are issues (36). Additionally, PD-L1 expression from different areas of the tumor may also differ (37). Therefore, using PD-L1 as a prognostic and predictive marker of response to therapy was insufficient. This suggests that our CNN model is a feasible tool and could supplement the limitations of PD-L1 expression and TMB for predicting the PFS and OS in patients with NSCLC who were treated with ICB.

In the combination of three cohorts, subgroup analysis of patients with CNN-high also had significantly longer PFS and OS than those with CNN-low. These results support that the CNN model had a good prognostic prediction and was not affected by clinical variables. Combining deep learning and cML methods based on radiology images was suggested in previous studies and using this combination could improve the predictive

abilities of various models (38–40). Interestingly, this is the first study to integrate CNN and cML methods based on WES and NGS databases to construct a nomogram for the stratification of patients treated with ICB. We found that patients in the EMN-low, EMN-intermediate, and EMN-high groups had different PFS and OS. This finding indicated that ensemble models, such as CNN, SVM, and RF, are promising and precise tools for predicting clinical benefit (for example, DCB, PFS, and OS) in patients undergoing immunotherapy.

However, this study had some limitations. First, although the number of patients was relatively large, the POPLAR/OAK, UCMC, and MSKCC cohorts were obtained from the American population. Thus, the molecular characteristics of these NSCLC patients may differ from those of East Asian descent, thus affecting clinical treatment outcomes (41). The CNN model should be trained with large international multi-center datasets to further improve prediction performance. Second, a panel of genomic variants based on the WES or NGS data was used in this study; however, DNA methylation, mRNA expression, radiology, and pathology were not used to predict the DCB. A multi-omics model, in addition to a genomics model, should be studied. Third, although extraction of the ctDNA of peripheral blood samples is a non-invasive method, the predictive ability of the model requires further investigation.

In summary, CNN classification based on a panel of 55 mutational genes serves as a novel and robust model for predicting DCB from ICB therapy in patients with NSCLC. A combination model that integrated CNN, SVM, and RF algorithms could better predict the PFS and OS. These findings may contribute to the discovery of a new strategy for patients with NSCLC treated with PD-1/PD-L1 blockade. Our method based on NGS and WES databases provides new insights for predicting clinical outcome in pan-cancer immunotherapy.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. The data has been resubmitted to Dryad (<https://doi.org/10.5061/dryad.hdr7sqvmf>).

Author contributions

(I) Conception and design: JP; (II) Administrative support: JP; (III) Provision of study materials or patients: JP and DZ; (IV) Collection and assembly of data: JP; (V) Data analysis and interpretation: JP; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Funding

This work was supported by National Nature Science Foundation of China (grant number: 82060327); the Science and Technology Foundation of Guizhou Province (grant numbers: Qian ke he ji chu-ZK 2021 and yi ban 454); and the Qian Dong Nan Science and Technology Program (grant number: qdnkhjz2020-013).

Conflict of interest

Author BX is employed by Yino Research.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Socinski MA, Jotte RM, Cappuzzo F, Orlandi F, Stroyakovskiy D, Nogami N, et al. Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. *N Engl J Med* (2018) 378:2288–301. doi: 10.1056/NEJMoa1716948
- Lemery S, Keegan P, Pazdur R. First FDA approval agnostic of cancer site – when a biomarker defines the indication. *N Engl J Med* (2017) 377:1409–12. doi: 10.1056/NEJMp1709968
- Brahmer JR, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Health-related quality-of-life results for pembrolizumab versus chemotherapy in advanced, PD-L1-positive NSCLC (KEYNOTE-024): a multicentre, international, randomised, open-label phase 3 trial. *Lancet Oncol* (2017) 18:1600–9. doi: 10.1016/S1470-2045(17)30690-3
- Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* (2019) 51:202–6. doi: 10.1038/s41588-018-0312-8
- Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther* (2017) 16:2598–608. doi: 10.1158/1535-7163.MCT-17-0386
- Cristescu R, Mogg R, Ayers M, Albricht A, Murphy E, Yearley J, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* (2018) 362:eaar3593. doi: 10.1126/science.aar3593
- Miao D, Margolis CA, Vokes NI, Liu D, Taylor-Weiner A, Wankowicz SM, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat Genet* (2018) 50:1271–81. doi: 10.1038/s41588-018-0200-2
- Ricciuti B, Arbour KC, Lin JJ, Vajdi A, Vokes N, Hong L, et al. Diminished efficacy of programmed death-1 inhibition in STK11- and KEAP1-mutant lung adenocarcinoma is affected by kras mutation status. *J Thorac Oncol* (2022) 17:399–410. doi: 10.1016/j.jtho.2021.10.013
- Kato S, Goodman A, Walavalkar V, Barkauskas DA, Sharabi A, Kurzrock R. Hyperprogressors after immunotherapy: analysis of genomic alterations associated with accelerated growth rate. *Clin Cancer Res* (2017) 23:4242–50. doi: 10.1158/1078-0432.CCR-16-3133
- Rizvi H, Sanchez-Vega F, La K, Chatila W, Jonsson P, Halpenny D, et al. Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing. *J Clin Oncol* (2018) 36:633–41. doi: 10.1200/JCO.2017.75.3384
- Dong ZY, Zhong WZ, Zhang XC, Su J, Xie Z, Liu SY, et al. Potential predictive value of TP53 and KRAS mutation status for response to PD-1 blockade immunotherapy in lung adenocarcinoma. *Clin Cancer Res* (2017) 23:3012–24. doi: 10.1158/1078-0432.CCR-16-2554
- Vauchier C, Pluvy J, Theou-Anton N, Soussi G, Poté N, Brosseau S, et al. Poor performance status patient with long-lasting major response to pembrolizumab in advanced non-small-cell lung cancer with coexisting POLE

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.960459/full#supplementary-material>

- mutation and deficient mismatch repair pathway. *Lung Cancer* (2021) 160:28–31. doi: 10.1016/j.lungcan.2021.07.016
- Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor JF, et al. STK11/LKB1 mutations and PD-1 inhibitor resistance in KRAS-mutant lung adenocarcinoma. *Cancer Discov* (2018) 8:822–35. doi: 10.1158/2159-8290.CD-18-0099
- Biton J, Mansuet-Lupo A, Pécuchet N, Alifano M, Ouakrim H, Arrondeau J, et al. TP53, STK11, and EGFR mutations predict tumor immune profile and the response to anti-PD-1 in lung adenocarcinoma. *Clin Cancer Res* (2018) 24:5710–23. doi: 10.1158/1078-0432.CCR-18-0163
- Wang Z, Zhao J, Wang G, Zhang F, Zhang Z, Zhang F, et al. Comutations in DNA damage response pathways serve as potential biomarkers for immune checkpoint blockade. *Cancer Res* (2018) 78:6486–96. doi: 10.1158/0008-5472.CAN-18-1814
- Li X, Wang Y, Li X, Feng G, Hu S, Bai Y. The impact of NOTCH pathway alteration on tumor microenvironment and clinical survival of immune checkpoint inhibitors in NSCLC. *Front Immunol* (2021) 12:638763. doi: 10.3389/fimmu.2021.638763
- Zhang Z, Gu Y, Su X, Bai J, Guan W, Ma J, et al. Co-Occurring alteration of NOTCH and DDR pathways serves as novel predictor to efficacious immunotherapy in NSCLC. *Front Oncol* (2021) 11:659321. doi: 10.3389/fonc.2021.659321
- Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* (2018) 172:1122–1131.e9. doi: 10.1016/j.cell.2018.02.010
- Jiang Y, Zhang Z, Yuan Q, Wang W, Wang H, Li T, et al. Predicting peritoneal recurrence and disease-free survival from CT images in gastric cancer with multitask deep learning: a retrospective study. *Lancet Digit Health* (2022) 4:e340–50. doi: 10.1016/S2589-7500(22)00040-1
- Peng J, Huang J, Huang G, Zhang J. Predicting the initial treatment response to transarterial chemoembolization in intermediate-stage hepatocellular carcinoma by the integration of radiomics and deep learning. *Front Oncol* (2021) 11:730282. doi: 10.3389/fonc.2021.730282
- Fehrenbacher L, Spira A, Ballinger M, Kowanzet M, Vansteenkiste J, Mazieres J, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (Poplar): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet* (2016) 387:1837–46. doi: 10.1016/S0140-6736(16)00587-0
- Rittmeyer A, Barlesi F, Waterkamp D, Park K, Ciardiello F, von Pawel J, et al. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. *Lancet* (2017) 389:255–65. doi: 10.1016/S0140-6736(16)32517-X
- Gutierrez SI, Turchan WT, Spurr LF, Rouhani SJ, Chervin CS, Steinhardt G, et al. CDKN2A loss-of-function predicts immunotherapy resistance in non-small cell lung cancer. *Sci Rep* (2021) 11:20059. doi: 10.1038/s41598-021-99524-1
- Hellmann MD, Nathanson T, Rizvi H, Creelan BC, Sanchez-Vega F, Ahuja A, et al. Genomic features of response to combination immunotherapy in patients

with advanced non-small-cell lung cancer. *Cancer Cell* (2018) 33:843–52.e4. doi: 10.1016/j.ccell.2018.03.018

25. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology: mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* (2015) 348:124–8. doi: 10.1126/science.aaa1348
26. Peng J, Zou D, Han L, Yin Z, Hu X. A support vector machine based on liquid immune profiling predicts major pathological response to chemotherapy plus anti-PD-1/PD-L1 as a neoadjuvant treatment for patients with resectable non-small cell lung cancer. *Front Immunol* (2021) 12:778276. doi: 10.3389/fimmu.2021.778276
27. Peng J, Kang S, Ning Z, Deng H, Shen J, Xu Y, et al. Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from CT imaging. *Eur Radiol* (2020) 30:413–24. doi: 10.1007/s00330-019-06318-1
28. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the evolution of non-small-cell lung cancer. *N Engl J Med* (2017) 376:2109–21. doi: 10.1056/NEJMoa1616288
29. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* (2016) 48:607–16. doi: 10.1038/ng.3564
30. Johannet P, Coudray N, Donnelly DM, Jour G, Illa-Bochaca I, Xia Y, et al. Using machine learning algorithms to predict immunotherapy response in patients with advanced melanoma. *Clin Cancer Res* (2021) 27:131–40. doi: 10.1158/1078-0432.CCR-20-2415
31. Harding JJ, Nandakumar S, Armenia J, Khalil DN, Albano M, Ly M, et al. Prospective genotyping of hepatocellular carcinoma: clinical implications of next-generation sequencing for matching patients to targeted and immune therapies. *Clin Cancer Res* (2019) 25:2116–26. doi: 10.1158/1078-0432.CCR-18-2293
32. Arbour KC, Luu AT, Luo J, Rizvi H, Plodkowski AJ, Sakhi M, et al. Deep learning to estimate RECIST in patients with NSCLC treated with PD-1 blockade. *Cancer Discov* (2021) 11:59–67. doi: 10.1158/2159-8290.CD-20-0419
33. Pepe F, Pisapia P, Gristina V, Rocco D, Micheli M, Micheli P, et al. Tumor mutational burden on cytological samples: A pilot study. *Cancer Cytopathol* (2021) 129(6):460–7. doi: 10.1002/cncy.22400
34. Munari E, Mariotti FR, Quatrini L, Bertoglio P, Tumino N, Vacca P, et al. PD-1/PD-L1 in cancer: Pathophysiological, diagnostic and therapeutic aspects. *Int J Mol Sci* (2021) 22(10):5123. doi: 10.3390/ijms22105123
35. Rimm DL, Han G, Taube JM, Yi ES, Bridge JA, Flieder DB, et al. A prospective, multi-institutional, pathologist-based assessment of 4 immunohistochemistry assays for PD-L1 expression in non-small cell lung cancer. *JAMA Oncol* (2017) 3:1051–8. doi: 10.1001/jamaoncol.2017.0013
36. Marletta S, Fusco N, Munari E, Luchini C, Cimadamore A, Brunelli M, et al. Atlas of PD-L1 for pathologists: Indications, scores, diagnostic platforms and reporting systems. *J Pers Med* (2022) 12(7):1073. doi: 10.3390/jpm12071073
37. Ilie M, Long-Mira E, Bence C, Butori C, Lassalle S, Bouhrel L, et al. Comparative study of the PD-L1 status between surgically resected specimens and matched biopsies of NSCLC patients reveal major discordances: a potential issue for anti-PD-L1 therapeutic strategies. *Ann Oncol* (2016) 27:147–53. doi: 10.1093/annonc/mdv489
38. Wei J, Jiang H, Zeng M, Wang M, Niu M, Gu D, et al. Prediction of microvascular invasion in hepatocellular carcinoma via deep learning: a multi-center and prospective validation study. *Cancers (Basel)* (2021) 13:2368. doi: 10.3390/cancers13102368
39. Ning Z, Luo J, Li Y, Han S, Feng Q, Xu Y, et al. Pattern classification for gastrointestinal stromal tumors by integration of radiomics and deep convolutional features. *IEEE J BioMed Health Inform* (2019) 23:1181–91. doi: 10.1109/JBHI.2018.2841992
40. Ran J, Cao R, Cai J, Yu T, Zhao D, Wang Z. Development and validation of a nomogram for preoperative prediction of lymph node metastasis in lung adenocarcinoma based on radiomics signature and deep learning signature. *Front Oncol* (2021) 11:585942. doi: 10.3389/fonc.2021.585942
41. Chen J, Yang H, Teo ASM, Amer LB, Sherbaf FG, Tan CQ, et al. Genomic landscape of lung adenocarcinoma in East Asians. *Nat Genet* (2020) 52:177–86. doi: 10.1038/s41588-019-0569-6