ABSTRACT

Objective Fast progression (FP) represents a desperate situation for advanced non-small cell lung cancer (NSCLC) patients undergoing immune checkpoint inhibitor therapy. We aimed to develop a predictive framework based on machine learning (ML) methods to identify FP in advanced NSCLC patients using blood test biomarkers.

Methods and analysis We extracted data of 1546 atezolizumab-treated patients from four multicentre clinical trials. In this study, patients from the OAK trial were taken for model training, whereas patients from the other trials were used for independent validations. The FP prediction model was developed using 21 pretreatment blood test variables in seven ML approaches. Prediction performance was evaluated by the receiver operating characteristic (ROC) curve.

Results The prevalence of FP was 7.6% (118 of 1546) in all atezolizumab-treated patients. The most important variables for the prediction model were: C reactive protein, neutrophil count, lactate dehydrogenase and alanine transaminase. The Support Vector Machine (SVM) algorithm applied to these four blood test parameters demonstrated good performance: the area under the ROC curve obtained from the training cohort (OAK), validation cohort 1 (BIRCH) and cohort 2 (merged POPLAR and FIR) were 0.908, 0.666 and 0.776, respectively. In addition, the absolute difference in median survival between the SVM-predicted FP and non-FP groups was significant in both progression-free survival and overall survival (p<0.001).

Conclusion SVM trained using a 4-biomarker panel has good performance in predicting the occurrence of FP regardless of programmed cell death ligand 1 expression, hence providing evidence for decision-making in single-agent atezolizumab immunotherapy for patients with advanced NSCLC.

INTRODUCTION

Immune checkpoint inhibitors (ICIs) have provided promising therapy strategies and remarkably improved overall survival (OS) for advanced non-small cell lung cancer (NSCLC) patients.1-5 However, some patients suffer from accelerated disease progression or early death after ICI administration, leading to shorter OS. Several concepts have been proposed to describe this phenomenon, such as fast progression (FP),6 hyperprogression disease (HPD),7 early progression8 or early death.9 Despite these nomenclatures, achieving a universally satisfactory and precise definition for this phenomenon has proven elusive.
Champiat et al. initially defined HPD, which used two pretreatment CT scan results to calculate tumour growth rate (TGR). However, the availability of two pretreatment CT scans, especially in the context of first-line immunotherapy, is often limited. Consequently, studies using the HPD or TGR definitions have faced challenges when it comes to validation and practical application in clinical trials. To overcome those disadvantages, Gandara et al. proposed a concept named FP which facilitates the analysis of the phenomenon, especially when only one pretreatment CT scan is available.

Recent research has identified potential biomarkers associated with FP, shedding light on the underlying mechanisms of this phenomenon. These biomarkers, often detected in routine blood tests, hold promise for early identification of patients at risk of FP. In the current study, we aim to use machine learning (ML) approaches to develop and validate prediction models based on the routine clinical laboratory parameters to identify FP patients before atezolizumab initiation.

**MATERIAL AND METHODS**

**Study cohort and patient-level data extraction**

This study included four clinical trials of advanced NSCLC patients treated with atezolizumab: BIRCH (NCT02031458), FIR (NCT01846416), POPLAR (NCT01903993) and OAK (NCT02008227). Specifically, BIRCH and FIR were single-arm studies, and patients received 1200 mg atezolizumab intravenously every 3 weeks. In contrast, OAK and POPLAR were randomised trials of atezolizumab (1200 mg intravenously every 3 weeks) versus docetaxel (75 mg/m² intravenously every 3 weeks) in platinum-containing treated advanced NSCLC. Only the atezolizumab treated patients with or without previous treatment with platinum-based chemotherapy from the above clinical trials are analysed in the study.

Baseline characteristics included in the analysis are age, sex, race, smoking history, Eastern Cooperative Oncology Group Performance Status (ECOG-PS), the sum of longest diameters (SLD) and metastasis information. The definition of programmed cell death ligand 1 (PD-L1) positive in OAK and POPLAR was Combined Positive Score (CPS) ≥1%, but in BIRCH and FIR was CPS ≥5%. Laboratory variables such as white cell count, neutrophil/lymphocyte ratio (NLR), C reactive protein (CRP), albumin levels and other variables were also extracted for model development. Key demographic and clinicopathological characteristics of the training and validation cohorts were summarised in Table 1. Two hundred and twenty-seven patients were filtered with 40% biomarker missed. By finding out the common laboratory test variables in four cohorts, we obtained 68 variables, then after filtering out the laboratory variables with missing data greater than 5%, there are 27 laboratory test values that were preliminarily retained. Six tests (‘magnesium (mmol/L)’, ‘phosphate (mg/dL)’, ‘calcium (mg/dL)’, ‘chloride (mmol/L)’, ‘sodium (mmol/L)’ and ‘pH scale’), without clinical correlation of FP were removed based on clinical insight. The missing data have been imputed with multiple imputation methods with mice package (V.3.15.0). Finally, we have 1319 patients with full 21 blood test results from whole cohorts (n=1546) in the model development (Figure 1).

**Labelling of FP patients**

The definition of FP based on Gandara et al. included two situations: one situation was SLD of target lesions had an increase of at least more than 50% from baseline scans to the first assessment at 6 weeks (±7 days); another situation was patient died within 12 weeks due to disease progression (evaluated by the investigators), but without any post-treatment CT evaluation. Importantly, for patients who had a post-treatment scan but also died within 12 weeks, FP was evaluated based only on the scan results and not on the death event. All patients included were labelled as FP or non-FP with the above definition.

**ML algorithms and model development**

The OAK and POPLAR trials specifically included participants who had progressed on first-line platinum-based chemotherapy. However, participants with newly diagnosed, chemotherapy-naive advanced NSCLC were included in the BIRCH and FIR trials. Considering the sample size, the atezolizumab arm from OAK was set as model development cohort, BIRCH study was selected as independent validation cohort 1 which provided a contrast to the OAK trial population. To enhance the robustness of our findings and further validate our model, we integrated the atezolizumab arms of the POPLAR and FIR studies as our second independent validation cohort. All the ML algorithms were performed in the vivil platform by R language with corresponding packages: e1071 (V.1.7.9) for Support Vector Machine (SVM), randomForest (V.4.6.14) for Random Forest (RF), rpart (V.4.1.15) for Decision Tree (DT), gbm (V.2.1.8) for Gradient-Boosted Machine (GBM), xgboost (V.0.4.2) for XGBoost, glmnet (V.4.1.3) for Generalised Linear Models and Least Absolute Shrinkage and Selection Operator. Model performance was then assessed in the training and two independent validation cohorts using area under the receiver operating characteristic curve (AUC).

Initially, all 21 laboratory variables were included in the ML methods to develop the primary prediction models using the training cohort. Next, we added important clinical variables to develop the combined models and compared their performance with the primary models. Finally, we screened the variables to obtain the optimised models using fewer variables. We define an optimal panel the one with a minimum of variables without a significant decrease in the AUC value. This was done by counting the frequencies of each variable appearing in the high-ranking importance scores of each ML approach, followed by a selection of variables appearing more than three times in the top-ranking variables of all seven ML.
approaches and applying them to the FP prediction models to test the output AUC values; all combinations of the most frequent, top-ranking variables in terms of importance scores were tested. These variables were eventually reduced one by one to obtain the optimal model consisting of the minimal number of variables with output AUCs comparable to the primary and combined models. We applied this approach to screen for the best combination of laboratory and clinical variables.

### Statistics
Categorical variables were presented as percentages and continuous variables were presented as median and IQRs. We used the $\chi^2$ test for categorical data and the Wilcoxon rank-sum test for continuous variables between FP and non-FP patients. Kaplan-Meier curves, HR and CIs based on stratified Cox models are shown along with log-rank p values, and statistical tests were two-sided. All statistical analyses were performed in R on the *vivli* platform. The significance level was set at 0.05.

### RESULTS

#### Patient population and prevalence of FP

One thousand five hundred and forty-eight advanced NSCLC patients treated with atezolizumab were included in this study. The prevalence of FP was 7.6% (118 of 1546) in all atezolizumab treated patients. Leaving out the records with pre-specified missing data criterion, we included 1319 patients in training and validation cohorts as shown in figure 1. In OAK study, the prevalence of FP was 9.5% (53 of 558 atezolizumab treated patients) and the prevalence of FP was 6.3% and 10.4% in validation cohorts 1 and 2, respectively. The lower incidence of FP in the BIRCH may be related to its enrolled criteria:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Overall, N=1319*</th>
<th>FP, N=109*</th>
<th>Non-FP, N=1210*</th>
<th>P value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort name</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training cohort</td>
<td>558 (42%)</td>
<td>53 (49%)</td>
<td>505 (42%)</td>
<td>0.081</td>
</tr>
<tr>
<td>Validation cohort 1</td>
<td>568 (43%)</td>
<td>36 (33%)</td>
<td>532 (44%)</td>
<td></td>
</tr>
<tr>
<td>Validation cohort 2</td>
<td>193 (15%)</td>
<td>20 (18%)</td>
<td>173 (14%)</td>
<td></td>
</tr>
<tr>
<td>RBC</td>
<td>4.13 (3.76, 4.54)</td>
<td>3.86 (3.45, 4.19)</td>
<td>4.16 (3.79, 4.56)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HGB</td>
<td>124 (111, 136)</td>
<td>110 (103, 131)</td>
<td>125 (112, 136)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PLAT</td>
<td>265 (216, 337)</td>
<td>317 (236, 437)</td>
<td>264 (213, 331)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NEUT</td>
<td>5.00 (3.70, 6.70)</td>
<td>7.02 (6.60, 9.60)</td>
<td>4.91 (3.66, 6.47)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LYMP</td>
<td>1.30 (0.93, 1.80)</td>
<td>1.09 (0.80, 1.50)</td>
<td>1.31 (0.96, 1.80)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MONO</td>
<td>0.61 (0.48, 0.85)</td>
<td>0.80 (0.61, 1.07)</td>
<td>0.60 (0.47, 0.81)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HCT</td>
<td>0.38 (0.34, 0.41)</td>
<td>0.34 (0.32, 0.39)</td>
<td>0.38 (0.34, 0.41)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NLR</td>
<td>3.8 (2.5, 6.0)</td>
<td>6.0 (3.5, 9.9)</td>
<td>3.6 (2.4, 5.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PLR</td>
<td>209 (137, 313)</td>
<td>294 (168, 565)</td>
<td>205 (133, 300)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NMR</td>
<td>7.9 (6.2, 10.4)</td>
<td>7.8 (5.9, 11.4)</td>
<td>7.9 (6.2, 10.3)</td>
<td>0.7</td>
</tr>
<tr>
<td>LMR</td>
<td>2.07 (1.34, 3.10)</td>
<td>1.35 (0.83, 1.88)</td>
<td>2.17 (1.43, 3.17)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CRP</td>
<td>14 (4, 42)</td>
<td>45 (15, 111)</td>
<td>12 (4, 38)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ALB</td>
<td>39.0 (35.0, 42.0)</td>
<td>36.9 (31.0, 40.0)</td>
<td>39.0 (36.0, 42.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BILI</td>
<td>6.8 (5.1, 10.0)</td>
<td>6.8 (5.5, 9.4)</td>
<td>6.8 (5.1, 10.0)</td>
<td>0.7</td>
</tr>
<tr>
<td>ALT</td>
<td>18 (13, 26)</td>
<td>22 (13, 34)</td>
<td>18 (13, 25)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AST</td>
<td>21 (17, 27)</td>
<td>25 (17, 36)</td>
<td>21 (17, 27)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LDH</td>
<td>230 (185, 360)</td>
<td>318 (209, 516)</td>
<td>227 (184, 347)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ALP</td>
<td>92 (73, 120)</td>
<td>104 (78, 130)</td>
<td>91 (73, 118)</td>
<td>0.017</td>
</tr>
<tr>
<td>CREAT</td>
<td>75 (64, 90)</td>
<td>72 (60, 80)</td>
<td>76 (65, 91)</td>
<td>0.001</td>
</tr>
<tr>
<td>GLUC</td>
<td>5.61 (5.05, 6.44)</td>
<td>5.71 (5.05, 6.61)</td>
<td>5.60 (5.05, 6.44)</td>
<td>0.6</td>
</tr>
<tr>
<td>TSH</td>
<td>1.44 (0.86, 2.38)</td>
<td>1.34 (0.81, 2.59)</td>
<td>1.45 (0.86, 2.36)</td>
<td>0.9</td>
</tr>
</tbody>
</table>

*n (%) median (IQR), †Pearson’s $\chi^2$ test; Wilcoxon rank-sum test.

ALB, albumin; ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate aminotransferase; BILI, total bilirubin; CREAT, creatinine; CRP, C reactive protein; GLUC, blood glucose; HCT, haematocrit; HGB, haemoglobin; LDH, lactate dehydrogenase; LMR, lymphocyte–monocyte ratio; LYM, lymphocyte count; MONO, monocytes count; NEUT, neutrophil count; NLR, neutrophil–lymphocyte ratio; NMR, neutrophil–monocyte ratio; PLAT, platelet count; PLR, platelet–lymphocyte ratio; RBC, red cell count; TSH, thyroid stimulating hormone.
patients with PD-L1 at least 5% and no brain metastases at baseline.

**Risk factors associated with FP**

In the combined cohort (n=1319), laboratory parameters and clinical characteristics were compared according to the presence of FP. The density plot for those laboratory parameters is shown in online supplemental figures 1–3, except total bilirubin, blood glucose and thyroid stimulating hormone, all other 18 blood tests were significantly different between FP and non-FP patients (table 1). For clinical characteristics, ECOG-PS (p=0.004, χ² test), bone and liver metastases at baseline (p=0.017 and p=0.049, respectively, χ² test), number of metastatic sites at baseline (p=0.030, χ² test) were significantly associated with FP (online supplemental table 1). No statistically significant associations were observed for age, sex, smoking history, brain metastases, tumour histology, PD-L1 expression and the rest of the investigated clinical variables.

**Development and optimisation ML models for FP prediction**

Initially, all the 21 laboratory variables were put into the seven ML frameworks to develop the primary 21-marker prediction models. Detailed information and set of seven ML algorithms were listed in online supplemental table 2. All seven ML methods achieved good predictive performance with AUC values ranging from 0.629 to 0.923 in the training cohort (OAK); among them, the XGBoost and SVM methods obtained the highest AUC values in the validation cohorts 1 (AUC=0.708) and 2 (AUC=0.818), respectively (online supplemental figure 4). Other performances, for example, accuracy, precision, recall, F1 and specificity, support those findings (online supplemental figure 5).

Next, we added five clinical characteristics (ECOG-PS, baseline body mass index, number of metastatic sites at baseline, bone and liver metastases at baseline), which were previously found to have significant FP predictive power, to construct the combined model consisting of laboratory biomarkers and clinical parameters. Results showed that compared with the primary 21-marker models, the combined model did not significantly improve the performances, especially the AUC values. (p=0.892, two-way analysis of variance test, online supplemental figures 5–8).

To facilitate clinical application, the FP prediction models are optimised by reducing the number of laboratory variables without decreasing diagnosis performances. Based on the relative importance scores of the 21 markers in each ML approach at training cohort (OAK study) (online supplemental figure 9), the most frequent, high-ranking variables in terms of importance scores (online supplemental table 3) were extracted, tested in different combinations, including 9-biomarker panel (online supplemental figures 10 and 11) and 6-biomarker panel (online supplemental figures 12 and 13) reduced one by one to obtain the optimal model. Eventually, CRP, neutrophil count (NEUT), lactate dehydrogenase (LDH) and alanine transaminase (ALT), were selected for FP prediction in the final optimal models (figure 2).
Comparing the AUC value, accuracy, precision, recall, F1-score and specificity of each model including 21-biomarker panel, 9-biomarker panel, 6-biomarker panel, 4-biomarker panel applied to the training and validation cohorts, respectively, when using different laboratory variables (online supplemental table 4). Finally, we found the 4-biomarker panel with a minimum of laboratory parameters without a significant decrease of the AUC value in each cohort (online supplemental figure 14). The 4-biomarker panel was identified as optimal panel to predict FP in the training and validation cohorts. The SVM method with the 4-biomarker panel demonstrated relatively good performance: the AUC obtained from the training cohort and validation cohorts 1 and 2 were 0.908, 0.666 and 0.776, respectively. The accuracy, precision, recall, F1-score and specificity of the SVM reached a good level, suggesting that this 4-biomarker panel was robust to predict FP in two validation cohorts (figure 3). In the combined cohort which pooled all patients, it was demonstrated that the SVM obtained the highest AUC value (AUC=0.805, online supplemental figure 15).

Figure 2  Receiver operating characteristic (ROC) analysis predicting fast progression (FP) with the area under curve (AUC) for seven machine learning (ML) methods with 4-biomarker panel (optimal panel) of three cohorts. (A) ROC cures and related AUC of seven FP predicting ML methods in training cohort (OAK); (B) ROC cures and related AUC of seven FP predicting ML methods in validation cohorts 1 (BIRCH); (C) ROC cures and related AUC of seven FP predicting ML methods in validation cohorts 2 (merged POPLAR and FIR); (D) boxplot showing AUC values in three cohorts of seven ML methods. Data are means±SE. DT, Decision Tree; GBM, Gradient-Boosted Machine; GLM, Generalised Linear Models; LASSO, Least Absolute Shrinkage and Selection Operator; RF, Random Forest; SVM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting.

Figure 3  Performance of four biomarkers (optimal panel) for predicting fast progression (FP) in the training and validation datasets. Training cohort: OAK, validation cohorts 1: BIRCH, validation cohorts 2: merged POPLAR and FIR. DT, Decision Tree; GBM, Gradient-Boosted Machine; GLM, Generalised Linear Models; LASSO, Least Absolute Shrinkage and Selection Operator; RF, Random Forest; SVM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting.
Four-biomarker panel predicted FP patients associated with poor survival
To examine the prognostic predictability of 4-biomarker panel using SVM, we performed survival analysis within each cohort and the combined cohort. Based on the cut-off scores or probabilities of each ML method, patients have been unrandomised divided into predicted FP patients and predicted non-FP patients. Survival analyses were performed between these two groups of patients. Results confirmed that FP patients predicted by SVM with 4-biomarker panel were associated with poorer OS and progression-free survival (PFS). Compared with the predicted non-FP patients, predicted FP with poorer OS and PFS in each cohort (training cohort: HR$_{OS}$=HR 5.51 (95% CI 4.08 to 7.45, p<0.0001, log-rank test); HR$_{PFS}$=3.22 (95% CI 2.43 to 4.28, p<0.0001, log-rank test); validation cohort 1: HR$_{OS}$=2.23 (95% CI 1.67 to 2.99, p<0.0001, log-rank test); HR$_{PFS}$=1.58 (95% CI 1.28 to 1.96, p<0.0001, log-rank test); validation cohort 2: HR$_{OS}$=2.50 (95% CI 1.66 to 3.75, p<0.0001, log-rank test); HR$_{PFS}$=2.41 (95% CI 1.72 to 3.37, p<0.0001, log-rank test); combined cohort HR$_{OS}$=2.66 (95% CI 2.23 to 3.18, p<0.0001, log-rank test); HR$_{PFS}$=1.76 (95% CI 1.52 to 2.04, p<0.0001, log-rank test)) (figure 4). Hence, the SVM with 4-biomarker panel not only predicted the occurrence of FP, but also predicted the prognosis of patients treated with atezolizumab. Similarly, the other ML methods predicted non-FP patients with shorter OS and PFS (the details are shown in online supplemental figures 16–21).

Four-biomarker panel for different PD-L1 expression subgroup of atezolizumab-treated patients
PD-L1 expression affects the effect of immunotherapy, but whether PD-L1 expression could affect the predictive performance of our model was unknown. Therefore, we tested the FP prediction models in the training and validation cohorts after separating the patients into PD-L1 negative and positive subgroups. In the PD-L1 positive subgroup, the AUC values of SVM were 0.897, 0.666 and 0.788 in the training (n=311) and validation cohorts 1 (n=567) and 2 (n=150), respectively (online supplemental figure 22), whereas the AUC value of SVM

Figure 4 Kaplan-Meier curve comparing overall survival (OS) and progression-free survival (PFS) between Support Vector Machine (SVM) predicted (fast progression) FP and predicted non-FP patients in each cohort. (A) Kaplan-Meier curve comparing OS between predicted FP and predicted non-FP patients in training cohort (OAK) (HR$_{OS}$=HR 5.51 (95% CI 4.08 to 7.45, p<0.0001, log-rank test)). (B) Kaplan-Meier curve comparing PFS between predicted FP and predicted non-FP patients in training cohort (OAK) (HR$_{PFS}$=3.22 (95% CI 2.43 to 4.28, p<0.0001, log-rank test)). (C) Kaplan-Meier curve comparing OS between predicted FP and predicted non-FP patients in validation cohort 1 (BIRC) (HR$_{OS}$=1.58 (95% CI 1.28 to 1.96, p<0.0001, log-rank test)). (D) Kaplan-Meier curve comparing PFS between predicted FP and predicted non-FP patients in validation cohort 1 (BIRC) (HR$_{PFS}$=2.23 (95% CI 1.67 to 2.99, p<0.0001, log-rank test)). (E) Kaplan-Meier curve comparing OS between predicted FP and predicted non-FP patients in validation cohort 2 (merged POPLAR and FIR) (HR$_{OS}$=5.00 (95% CI 1.66 to 3.75, p<0.0001, log-rank test)). (F) Kaplan-Meier curve comparing PFS between predicted FP and predicted non-FP patients in validation cohort 2 (merged POPLAR and FIR) (HR$_{PFS}$=2.41 (95% CI 1.72 to 3.37, p<0.0001, log-rank test)). (G) Kaplan-Meier curve comparing OS between predicted FP and predicted non-FP patients in combined cohort (HR$_{OS}$=2.66 (95% CI 2.23 to 3.18, p<0.0001, log-rank test)); combined HR and CIs based on stratified Cox models are shown along with log-rank p values, and statistical tests were two-sided.
ICI treatment was reported to have lower response rate, neutrophil at pretreatment and after the first cycle of CRP was reported to be a prognostic biomarker for ICI. Neutrophil and CRP were correlated with ICI resistance and conferring resistance to ICI.29 On the other hand, immune exclusion and adaptive immune cell suppression on ICI therapeutic efficacy is associated with neutrophils on ICI therapeutic efficacy is associated with neutrophils on ICI, which may be a potential biomarker associated with HPD in patients with NSCLC treated with programmed cell death protein 1/PD-L1 blockade. Most existing studies supported immune biomarkers that may successfully predict HPD, but there was no successful predictive model with peripheral blood biomarkers for clinical decisions.

Our study tested seven ML methods to predict FP in advanced NSCLC patients treated with immunotherapy. First, we covered 21 laboratory variables to train a suitable model for FP prediction. All ML methods reached a respectable level of performance, except DT and RF. We constructed the combined model with laboratory and clinical parameters but did not increase performance. In other words, those liquid biomarkers without clinicopathological variables could generate a robust model for predicting FP.

To improve clinical utility and minimise liquid parameters, the most frequent, high-ranking variables in importance scores were extracted to reduce dimensionality of the liquid parameters. We reduced dimensionality from 21, 9, 6, to 4 parameters. Finally, the 4-biomarker panel using SVM with substantial performances, including neutrophil, CRP, LDH, and ALT was defined as optimal FP predictive model. This model not only predicted the occurrence of FP regardless of PD-L1 expression, but also served as a prognostic predictor for patients treated with immunotherapy.

Among the four laboratory variables, two inflammatory biomarkers, CRP and neutrophil have been reported in previous research of ICI therapy on NSCLC.25–26 High neutrophil at pretreatment and after the first cycle of ICI treatment was reported to have lower response rate, shorter OS, and PFS in NSCLC patients,27–28 in support of our findings. The potential biological mechanism of neutrophils on ICI therapeutic efficacy is associated with immune exclusion and adaptive immune cell suppression, which confer resistance to ICI.29 On the other hand, CRP was reported to be a prognostic biomarker for ICI therapy in NSCLC. Our study identified CRP as a robust biomarker for FP prediction, in concordance with our previous findings.30 CRP is a marker of systemic inflammation and immune activation. It fosters cancer progression by promoting cell proliferation, angiogenesis, and cancer cell migration.31

In the present study, two liver enzyme parameters were also associated with FP. Previous studies reported that pretreatment LDH served as a predictive biomarker for advanced NSCLC patients treated with ICI,32 not only because it is a key enzyme involved in cancer metabolism, but also because it allows neoplastic cells to suppress and evade the immune system by altering the tumour microenvironment.33 Additionally, our results demonstrated that high ALT correlated with FP, suggesting that proinflammatory status or other organ immunotoxicity may help identify patients at a higher likelihood of ICI resistance.34 Generally, patients with proinflammatory status or poor ECOG-PS have elevated risk of FP. Hence, physicians are expected to be cautious about the use of atezolizumab in these patients predicted with FP.

The 4-biomarker panel is not only practical for application in general clinical settings, but also the prediction of atezolizumab treatment outcome could guide the regimen choice before treatment. This is the first study to develop and validate FP risk prediction models in advanced NSCLC patients treated with immunotherapy. The convenience and feasibility of sample collection from routine blood tests ensures the prospective application of this 4-biomarker panel to assess the benefits and risks of immunotherapy.

Many failed external validations could have been foreseen by rigorous internal validation, saving time and resources.35 To decrease the volatility of biomarker panel, by increasing the transferability to real-world clinical application, we selected OAK as training dataset, BIRCH and merged FIR+POPLAR as external validation datasets. We constructed a pipeline that tested seven different ML approaches to predict FP at training and validation datasets, each of which devises different algorithms and hence avoids algorithmic bias.

The predictive model for FP risk in advanced NSCLC patients treated with ICIs holds practical promise. It aids clinicians in pretreatment patient risk assessment, guiding personalised treatment decisions for improved outcomes. By optimising resource allocation and influencing clinical trial design, the model contributes to more efficient healthcare delivery and research. Prospective validation and seamless integration into clinical systems enhance accessibility. However, its applicability should be restricted to patients with specified treatment histories for accurate predictions and responsible clinical decisions.

Nevertheless, this study had limitations. First, the present study only included anti-PD-L1 (atezolizumab) monotherapy, in the absence of validation of immunotherapy combined with chemotherapy or radiotherapy, as well as double ICI therapy. Second, most of the patients had been previously treated with chemotherapy, so this model needs to be validated in treatment-naïve patients with NSCLC. Third, given that FP and poor prognosis were 0.893 in the combined cohort with PD-L1 negative (n=291) (online supplemental figure 23). The subgroup analysis has shown this 4-biomarker panel predicting the FP with robust performances regardless of PD-L1 expression.
come hand-in-hand, the current analysis could not distinguish between these biomarkers’ predictive and prognostic values. Last but not least, even though our analysis included PD-L1 positive and negative patient subgroups, subgroup analyses based on TMB, and other tumour microenvironment factors have not been conducted. While our data were derived from three independent clinical studies, improvements can be performed if data from large multi-centre real-world datasets are used as external validation.

CONCLUSIONS
To summarise, SVM trained using 4-biomarker panel performs well in predicting the occurrence of FP regardless of PD-L1 expression. By identifying FP, peripheral blood biomarkers based on ML approaches improve prognosis prediction and personalised therapeutic decision-making in immunotherapy for NSCLC.

Author affiliations
1Department of Oncology, The second affiliated hospital of Zunyi Medical University, Zunyi, People’s Republic of China
2Translational Radiobiology, Department of Radiation Oncology, Universitätsklinikum Erlangen & Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
3Department of Radiation Oncology, Universitätsklinikum Erlangen & Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
4Comprehensive Cancer Center Erlangen—EMN, Erlangen, Germany
5FAU Profile Center Immunomedicine (FAU I-MED), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
6Department of Pulmonary Surgery, Guangdong Lung Cancer Institute, Guangdong Provincial People’s Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, People’s Republic of China
7Thoracic Surgery Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA
8AW Medical Co Ltd, Macau, People’s Republic of China
9Department of Analytics, Harrisburg University of Science & Technology, Harrisburg, Pennsylvania, USA
10Oncology Biometrics, AstraZeneca, Gaithersburg, Maryland, USA
11Department of Radiation Oncology, Searle University Medical Center, Homburg, Germany

Acknowledgements
We would like to thank all of the patients, investigators and staff involved in the FIR, BIRCH, POPLAR and OAK studies who released and shared their data. This publication is based on research using data from data contributors, Roche, that has been made available through Vivli, Inc (Data Request ID: 5935; lead investigator: J-GZ). Vivli has not contributed to or approved, and is not in any way liable for the edition (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

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

ORCID iDs
Jian-Guo Zhou http://orcid.org/0000-0002-5021-3739
Benjamin Frey http://orcid.org/0000-0001-6743-3351
Markus Hecht http://orcid.org/0000-0003-2082-216X
Udo Gaip http://orcid.org/0000-0001-6375-5470

Competing interests The authors declare no relevant conflict of interest regarding this manuscript. MH reports collaborations with Merck Serono (advisory role, speakers’ bureau, honoraria, travel expenses, research funding); MSD (advisory role, speakers’ bureau, honoraria, travel expenses, research funding); AstraZeneca (research funding); Novartis (research funding); BMS (advisory role, honoraria, speakers’ bureau); Teva (travel expenses). UG and RF received support for presentation activities for Dr. Sennewald Medizintechnik GmbH, have received support for investigator initiated clinical studies (IIts) from MSD and AstraZeneca and contributed at Advisory Boards Meetings of AstraZeneca and Bristol-Myers Squibb.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and this study provides secondary analysis for trial datasets and was deemed to be of negligible risk and was approved by the Institutional Review Board of the Second Affiliated Hospital, Zunyi Medical University (No. YXLL(KY-R)-2021-010). Written informed consent for participation was not required for this study in accordance with national legislation and institutional requirements. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Qualified researchers may request access to individual patient level data through the clinical study data request platform (Vivli, https://vivli.org/). For further details please refer to Roche’s Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents; see https://www.roche.com/innovation/process/clinical-trials/data-sharing/ and https://vivli.org/ourmember/roche/

Our source codes for the prediction of FP are available at https://github.com/JianGuoZhou3/ML_ICI_fast.progression.

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

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

REFERENCES


