Journal for ImmunoTherapy of Cancer

Combined circulating tumor antigen model demonstrates additional prognostic value in first-line treatment of non-small cell lung cancer

Charles W Tran ,[1] Wei Zou,[1] Zoe J Assaf,[1] Anika Mang,[2] Francesco Ranucci ,[3] Jason H Cheng,[1] Ivette Estay,[1] Vy Ma,[1] Minu K Srivastava,[1] David S Shames,[1] Stefan Holdenrieder,[4] Martin Reck ,[5] Vinzent Rolny,[2] Katja Schulze,[1] Birgit Wehnl,[2] Namrata S Patil [1]

To cite: Tran CW, Zou W, Assaf ZJ, *et al*. Combined circulating tumor antigen model demonstrates additional prognostic value in first-line treatment of non-small cell lung cancer. *Journal for ImmunoTherapy of Cancer* 2026;**14**:e014710. doi:10.1136/jitc-2025-014710

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/jitc-2025-014710).

Accepted 01 July 2026

[1]Genentech Inc, South San Francisco, California, USA
[2]Roche Diagnostics GmbH Penzberg, Penzberg, Germany
[3]Roche Diagnostics SL, Sant Cugat del Vallès, Spain
[4]German Heart Centre Munich, Munich, Germany
[5]Thoracic Oncology, LungenClinic Grosshansdorf GmbH, Grosshansdorf, Germany

Correspondence to
Dr Namrata S Patil;
patil.namrata@gene.com

ABSTRACT

Background Circulating tumor antigens (ctA; tumor markers) are blood-based proteins that can offer prognostic value in non-small cell lung cancer (NSCLC) and may serve as potential early surrogates for survival. Given the significantly reduced testing time and cost of ctA compared with circulating tumor DNA (ctDNA), we further explored the utility of ctA using samples collected from over 2,300 patients participating in five clinical trials (IMpower130, 131, 132, 150, and 110).

Methods We analyzed a panel of six ctA (CA125 (cancer antigen 125), CEA (carcinoembryonic antigen), Cyfra21-1 (cytokeratin 19 fragment 21-1; CYFRA), NSE (neuron-specific enolase), SCC (squamous cell carcinoma antigen), and ProGRP (progastrin-releasing peptide)) and CRP (C-reactive protein) from the serum of patients with metastatic NSCLC in these trials, which investigated combinations of atezolizumab (anti-programmed death-ligand 1)±bevacizumab±chemotherapy. Previous work showed that an optimized cut-off using two ctA or a machine learning (ML) model of ctDNA features, both taken at 6 weeks, can stratify patients with stable disease (SD) for survival risk in IMpower150. Building on this approach, we applied an ML model combining ctA features at baseline and at 6 weeks, trained across a much larger aggregate dataset from multiple clinical studies.

Results Previous findings from ctA analysis of IMpower150 were confirmed and found to be applicable to several other trials analyzed in this study. We found that ML model predictions provided similar prognostic performance (c-index of 0.73 and 0.71 in squamous and non-squamous test datasets, respectively), with CYFRA being the top feature for both histologies.

Conclusions While ctA demonstrated limited potential in differentiating treatment effects to inform early drug development, deriving an optimal prediction cut-off for 1-year overall survival (OS) showed that ctA model predictions could effectively stratify patients by radiographic response with 61% sensitivity and 78% specificity, adding significant prognostic value to radiographic imaging. Patients with partial response (PR), progressive disease (PD), or stable disease (SD) at either 6 weeks of treatment or best confirmed overall response could be separated into low-risk or high-risk groups for OS.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The prognostic value of individual tumor antigens such as Cyfra21-1 (cytokeratin 19 fragment 21-1; CYFRA) in non-small cell lung cancer (NSCLC) has been previously reported, although most studies have only examined a baseline time point and have not explored tumor antigens in the specific context of chemoimmunotherapy. Combined circulating tumor antigen (ctA) models in smaller datasets have also been shown to have prognostic utility.

WHAT THIS STUDY ADDS

⇒ In a retrospective analysis of five phase 3 clinical studies of chemoimmunotherapy in NSCLC, we demonstrated that ctA individually (in particular Cyfra21-1), or in combination using a machine learning model, provides significant prognostic value across histologies for overall survival and increased value above radiographic imaging or Eastern Cooperative Oncology Group status. We revealed histology-specific (squamous vs non-squamous) differences in ctA levels and dynamics. In a direct comparison with circulating tumor DNA (ctDNA) for a subset of patients, we showed that ctA has prognostic performance similar to that of ctDNA.

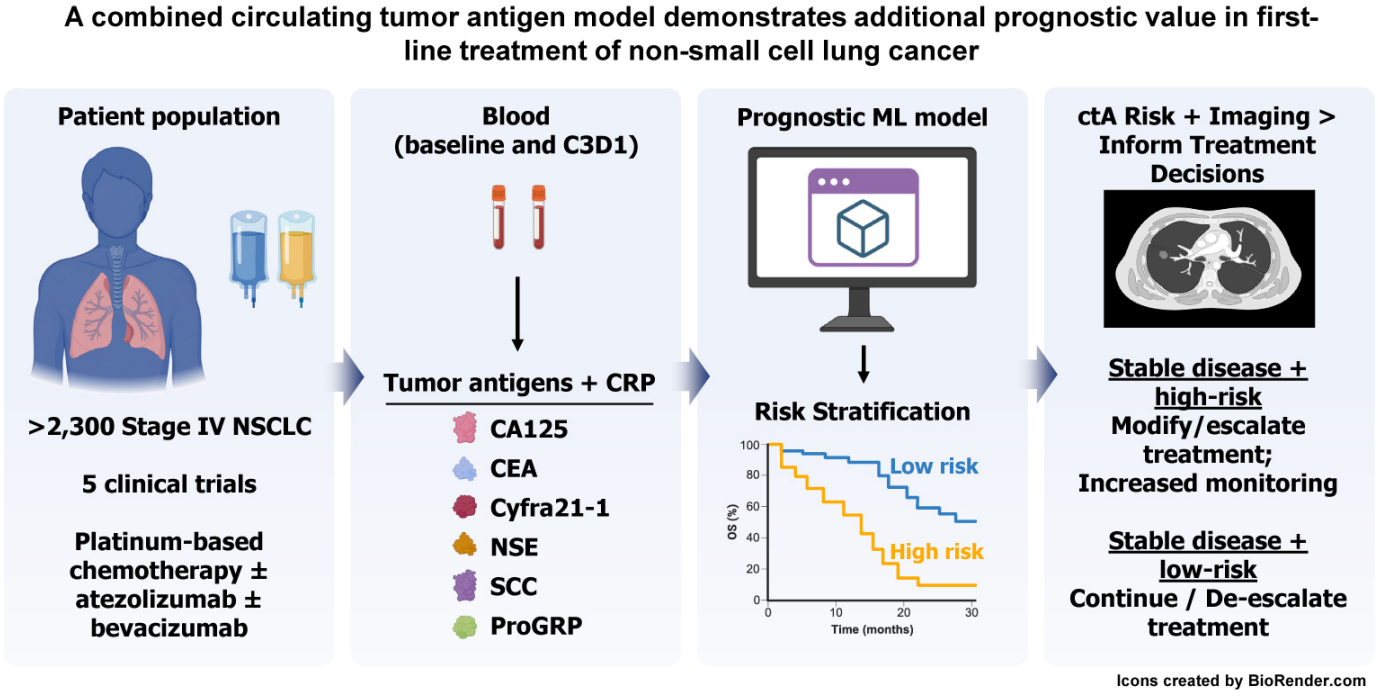
HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our results suggest that Cyfra21-1 alone or a panel of ctA could be used to help inform clinical decision-making, such as for treatment escalation or diversion.

For the graphical abstract of this paper see figure 1.

BACKGROUND

Non-small cell lung cancer (NSCLC) represents approximately 80–85% of lung cancer diagnoses and is the leading cause of cancer-related deaths globally.[1] Despite



Authors
Charles W. Tran, Wei Zou, Zoe June Assaf, Anika Mang, Francesco Ranucci, Jason H. Cheng, Ivette Estay, Vy Ma, Minu K. Srivastava, David S. Shames, Stefan Holdenrieder, Martin Reck, Vinzent Rolny, Katja Schulze, Birgit Wehnl, Namrata S. Patil

In Brief
A panel of 6 circulating tumor antigens (ctA) and CRP combined in a machine learning (ML) model provides significant prognostic value for overall survival in metastatic NSCLC. The ctA model (or Cyfra21-1 alone) adds value to radiographic imaging to help inform clinical decision-making, such as for treatment escalation or diversion.

Figure 1 Graphical abstract

multiple advances in treatment,[2] overall patient survival remains low, as nearly 75% of patients are diagnosed in advanced stages of disease. The 5-year overall survival (OS) rate is only 6% in the late-stage metastatic (stage IV) setting, and approximately 26% when averaged across all disease stages.[3]

Given the poor prognosis of advanced NSCLC, novel prognostic biomarkers are needed to rapidly delineate treatment benefit and enhance risk stratification of patients to inform clinical decision-making. Recently, advances in the sensitivity of circulating tumor DNA (ctDNA) assays have increased interest in the use of this technology as a potential tool for informing patient risk, monitoring treatment effect, and for minimal residual disease detection.[4] Both ctDNA positivity and clearance are also being applied prospectively in studies to assess their utility for treatment decisions and as a possible accelerated endpoint.[4] ctDNA assays have been employed in clinical settings for minimal residual disease monitoring, but testing can be cost-prohibitive.[5 6] Moreover, ctDNA assays with the greatest sensitivity and specificity are generally tumor-informed,[7 8] which may increase patient burden by requiring a tumor biopsy for tumor-specific somatic variant identification.

In contrast, circulating tumor antigens (ctA; also known as tumor markers or cancer markers) are circulating proteins found in the blood that are elevated in specific or multiple cancer types. Some ctA that are routinely used in the clinic include prostate serum antigen for prostate cancer and alpha-fetoprotein for hepatocellular carcinoma (HCC). In lung cancer, ctA such as Cyfra21-1 have been reported by multiple previous studies as a sensitive marker for squamous lung cancer,[9] while neuron-specific enolase (NSE) and progastrin-releasing peptide (ProGRP) have been reported as diagnostic markers for small cell lung cancer.[10 11] Cyfra21-1 is the soluble fragment of cytokeratin 19, a key component of the cytoskeleton of epithelial cells, and was identified as a prognostic marker in patients with lung cancer over three decades ago.[12] A meta-analysis revealed that high baseline levels of Cyfra21-1 are negatively associated with both OS and progression-free survival (PFS) across NSCLC disease stages and treatment regimens, including epidermal growth factor receptor (EGFR) inhibitors and chemotherapy combinations.[13] C-reactive protein (CRP) levels have also been investigated for their potential prognostic and predictive potential in NSCLC.[14]

However, despite these and other data demonstrating the potential prognostic value of blood-based proteins for patients with NSCLC, no ctA are currently recommended for use in the routine diagnosis of NSCLC or for informing treatment decisions.[15 16]

Given the low sample volume requirements for ctA and their significantly reduced cost of testing compared with current CE-marked (conformité européenne-marked) ctDNA assays (eg, Natera Signatera, FoundationOne Liquid CDx, Guardant360 CDx), we wished to more closely examine whether ctA can be a valuable tool in drug development or in clinical decision-making. Specifically, we investigated whether ctA could offer additional prognostic value to radiographic imaging, distinguish treatment effects, and/or act as an early surrogate of OS.

To address these questions, we measured and analyzed ctA levels in five separate phase 3 studies of patients with advanced (metastatic) NSCLC. Patients received treatment combinations of atezolizumab±chemotherapy compared with chemotherapy alone (or atezolizumab±bevacizumab+chemotherapy compared with bevacizumab+chemotherapy in IMpower150[17]). Studies included patients across NSCLC histologies, with squamous (IMpower130[18]), non-squamous (IMpower150,[17] IMpower131,[19] and IMpower132[20]), and pan-histology (IMpower110[21]) studies selected for analysis (online supplemental figure 1). To our knowledge, these data represent the largest tumor antigen dataset generated with chemotherapy or chemoimmunotherapy treatment in NSCLC to date, with over 2,300 patients with NSCLC analyzed at both baseline and on-treatment time points.

METHODS

ctA testing

Tumor antigens (CA125, CEA, Cyfra21-1 (or CYFRA), NSE, ProGRP, SCC), and CRP levels were quantified from serum samples from IMpower110, IMpower130, IMpower131, and IMpower132 at baseline and cycle 3 day 1 (C3D1) of treatment using the Roche cobas® e platform. ctA data from IMpower150 were previously reported using the same testing platform.[22] CA125 values are in U/mL; CEA, Cyfra21-1, NSE, and SCC values are in ng/mL, ProGRP is in pg/mL, and CRP is in mg/mL.

Feature selection and engineering

For model training, ctA concentration values less than the 2% quantile (q2), or greater than the 98% quantile (q98) within each of the training or test datasets were censored and replaced with the corresponding q2 or q98 value, respectively, to minimize the effect of outliers. Derived features with $\log_2$(fold change) and percent change for each ctA at C3D1 relative to baseline were then added. Percent change was defined as ((ctA value at C3D1) − (ctA value at baseline)) / (ctA value at baseline) × 100%. All features were normalized using IQR normalization in R. For pan-histology models, histology was one-hot encoded.

Data imputation

Missing ctDNA values were imputed based on the population mean within the training or test datasets as previously described.[23] For ctA data, missing values were not imputed (patients missing either baseline or C3D1 ctA values were excluded from subsequent analysis), with the exception of the 192 patients within the IMpower150 hold-back ctDNA test set. Missing ctA values for these patients were imputed using missForest (V.1.5) in R, with default parameters. These imputed ctA data were used only for the operational characteristic (OC) simulation to enable direct comparison with ctDNA.

Elastic net model

Model training, feature selection, and 10-fold leave-one-out cross-validation were performed as previously described.[23]

ROC plots

ROC (receiver operating characteristic) curves for 3, 6, 9 and 12 month survival with associated AUC (area under the curve) values were plotted using the timeROC package (V.0.4) in R.

Survival plots

Kaplan-Meier survival plots were generated in R (V.4.3.3) using survminer (V.0.4.9) and survival (V.3.8–3). For stratification by ctA median, a global median was calculated for each ctA at each baseline or C3D1 time point within the training or test dataset. ctA values greater than or equal to the corresponding median value were classified as “high” and values below the median were classified as “low”.

XGBoost and cross-validation

Models were trained using XGBoost[24] (V.2.1.1) and Python (V.3.10.13) using a Cox proportional hazards (coxph) model for survival data. Model fit was evaluated by minimizing the Cox negative log-likelihood (cox-nloglik) in XGBoost. Hyperparameter optimization was performed using a grid-search approach with five-fold cross-validation using GridSearchCV (scikit-learn V.1.4.2). Parameter optimization and final model training were GPU (graphics processing unit)-accelerated using CUDA (compute unified device architecture) (NVIDIA P6000) on a private HPC (high performance computing). The optimal number of boosting rounds was determined with the number of early stopping rounds set to 5 (ie, the optimal number of boosting rounds equals 1+ the number of boosting rounds corresponding to when the cox-nloglik did not decrease after five rounds). Model training and code are available at https://github.com/cwtran/tran_ctA_2025 Feature importance in XGBoost models was determined using SHAP (Shapley additive explanations) (V.0.45.1).[25]

Model performance

Performance of elastic net or XGBoost models was assessed using Harrell’s Concordance index (c-index).

OC simulations

OC simulations were performed as previously described.[23] In brief, 30 patients from each of the treatment and control arms (ie, active vs control) were chosen at random and compared using the criteria of interest (ctDNA or ctA model prediction, objective response rate (ORR), or PFS) using a one-sided Fisher's test for ORR and a one-sided Wilcoxon test for all other continuous variables. The null hypothesis was that the active arm had a higher average model prediction score or lower ORR frequency or PFS than the control arm, with the alternate hypothesis being the opposite. The comparison was repeated a total of 2,000 times, and the resulting p value from the statistical test for each iteration was recorded.

A "false-go" rate was then established by determining the p value corresponding to the 15% quantile of p values obtained by comparing two groups of 30 patients, each selected from the control arm (ie, control vs control), 2,000 times as above. Finally, the fraction of p values in the active versus control arm comparisons with values lower than the "false-go" p value corresponding to the 15% cut-off was recorded as the frequency of making a "true-go" decision.

Role of the funding source

Study design, data collection, data analysis, and manuscript writing were performed by the authors. The funders did not have any direct role in these aspects, manuscript preparation, or in the decision to submit for publication. Manuscript preparation and the decision to submit for publication were undertaken collectively by all authors in alignment with ICMJE (International Committee of Medical Journal Editors) criteria.

RESULTS

Individual ctA levels and their prognostic value

Individual ctA levels were first analyzed by histology at baseline and at C3D1 of treatment. Consistent with another study, we noted histology-specific differences in baseline levels of ctA.[26] CYFRA and SCC levels were found to be higher at baseline in patients with squamous NSCLC, while CEA, CA125, and CA19-9 (cancer antigen 19-9) levels were higher at baseline in non-squamous NSCLC (online supplemental figure 2a, online supplemental table 1). Some individual ctA were weakly correlated at baseline, including CYFRA and CRP in both squamous and non-squamous (Spearman correlation coefficient, ρ=0.4), while Cyfra21-1 was more strongly correlated with CA125 in non-squamous or with SCC in squamous at both baseline (ρ=0.5 and ρ=0.4, respectively) and C3D1 time points (ρ=0.3 and ρ=0.4, respectively) (online supplemental figure 2b–e).

Next, we investigated whether individual ctA are prognostic for OS. Consistent with previous studies,[22 27–29] patients with high baseline (greater than or equal to median expression) of CYFRA, CA125, and CRP (online supplemental table 1) had a significantly higher risk of death compared with those with low ctA (less than median expression), and with either squamous or non-squamous histology (online supplemental figures 3a, 4a). At C3D1, CYFRA by median split (median of 2.5 and 3.0 ng/mL for non-squamous and squamous, respectively) showed a greater OS separation between risk groups compared with baseline, particularly in squamous NSCLC (online supplemental figures 3b,4b).

Combining ctA in ML models

We then asked whether combining ctA using ML models could improve their prognostic performance for OS. We investigated both elastic net and extreme-gradient boosted (XGBoost) regression ML models because of their overall performance, relative insensitivity to multicollinearity, and resistance to overfitting.[24 30]

Due to differences in both baseline and on-treatment levels and dynamics of ctA between non-squamous and squamous NSCLC (online supplemental figure 2a), we trained and tested histology-specific ML models on the corresponding squamous or non-squamous IMpower dataset(s) (figure 2a,b). XGBoost models were found to have slightly improved c-indices compared with elastic net models (online supplemental table 2, online supplemental figures 5 and 6).

Using a histology-specific XGBoost model, ctA prognostic performance in training was improved in both non-squamous (OS HR: 4.01 (95% CI 3.39 to 4.73)) and squamous models (OS HR: 7.30 (95% CI 5.11 to 10.32)) between high versus low prediction scores by median split compared with CYFRA alone (non-squamous OS HR: 2.35 (95% CI 2.01 to 2.75) and squamous OS HR: 3.11 (95% CI 2.28 to 4.24)) (figures 3a and 4a, online supplemental table 2). Concordance (by Harrell's c-index; c) between model prediction and OS was c=0.75 in training and c=0.71 in test for the non-squamous model and c=0.81 and c=0.73 for training and test, respectively, in the squamous model, with model prediction scores inversely correlated with OS time, as expected (figures 3b,d and 4b,d). The top features contributing to overall model output include C3D1 CYFRA, baseline CRP, C3D1 CA125, and baseline CYFRA (figure 3e,f) for the non-squamous model and C3D1 CYFRA, CA125, CEA, and SCC for the squamous model (figure 4e f).

1-year OS risk stratification

In the metastatic NSCLC setting, over 50% of patients do not survive beyond 1 year. From the ML model predictions, we further investigated if non-squamous patients could be effectively stratified for 1-year OS (ie, predicted as <1 year OS or ≥1 year OS). In the training dataset, we found that ORR is a poor predictor of OS. The majority of responders in this training dataset had a partial response (PR) recorded as their best confirmed overall response (BCOR), but a significant fraction of these PR patients had an OS of less than 1 year (figure 5a). Similarly, a significant fraction of stable disease (SD) patients survive longer than 1 year.

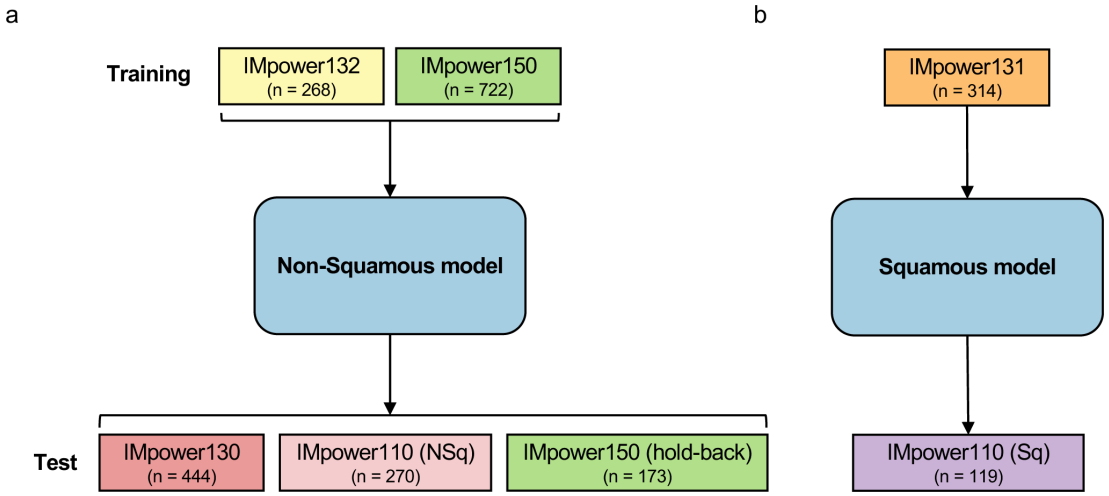


Figure 2 Overview of ctA training and test datasets for ML model development. (a) IMp132 and IMp150 datasets were used to train a non-squamous model with patients from IMp130, IMp110 and IMp150 hold-back serving as the test dataset. (b) IMp131 was selected as the training dataset for a squamous model with squamous patients from IMp110 as the test dataset. ctA, circulating tumor antigen; IMp, IMpower; ML, machine learning; NSq, non-squamous; Sq, squamous.

We next examined the overall sensitivity and specificity of the XGBoost model for predicting 1-year OS in non-squamous patients. In the training dataset, the model had an AUC of 0.79 (95% CI 0.75 to 0.82) but reduced AUC of 0.67 (95% CI 0.61 to 0.72) in the test (figure 5b,c), potentially reflective of differences in treatment regimens (ie, specific chemotherapy combination) and/or patient (sub)populations between training and test. AUC values were comparable for OS at shorter intervals (6 or 9 months), demonstrating the robustness of the model for identifying high-risk patients.

A threshold cut-off was established from the XGBoost model prediction to best classify patients for 1-year OS. The cut-off value was varied, and the resulting positive predictive value (PPV) and negative predictive value (NPV) were determined at each corresponding prediction cut-off value. The final prediction cut-off was selected by maximizing both the PPV and NPV (figure 5d). Using this optimized cut-off value, the PPV and NPV were 69.6% and 70.2%, respectively, with a sensitivity of 60.9% and a specificity of 77.7% (online supplemental table 3). Applying this cut-off value, patients could be effectively stratified for OS risk by their week 6 radiographic response in both the training (figure 5e) and test (figure 5f) datasets or by their best confirmed overall response (online supplemental figure 7a,b). Importantly, these results show that high-risk and low-risk subgroups exist for both week 6 and BCOR SD and PR patients, which is consistent with previous findings,[22 31] demonstrating that ctA can offer significant prognostic value above radiographic imaging or ECOG status (online supplemental figure 8).

Comparing ctA to ctDNA as an early OS surrogate in an OC simulation

To explore whether ctA could also provide value as an early surrogate for OS and as an accelerated endpoint for drug development, we compared ctA with previous results from ctDNA and other traditional criteria such as ORR and PFS from IMpower150 in an OC simulation (figure 6a; see also Methods[23 32]). IMpower150 showed a statistically significant OS benefit in patients with metastatic NSCLC receiving atezolizumab+bevacizumab+chemotherapy (ABCP) or atezolizumab+chemotherapy (ACP) compared with patients receiving bevacizumab+chemotherapy (BCP; control) in the ctA or ctDNA biomarker-evaluable populations (figure 6b), similar to the intent-to-treat population.[17]

Previous work showed that an ML model developed on ctDNA data at 6 weeks could improve the frequency of making a correct "true-go" decision to proceed from a phase 2 to phase 3 study using a subset of patients from IMpower150 to simulate a phase 2 trial.[23]

To improve the performance of ctDNA in the OC simulation, we applied ML model predictions as a continuous variable rather than the previous approach of binary categorization of patients by high or intermediate/low molecular risk. Using this modified approach, ctDNA achieved a 52.6% true-go rate in the ABCP arm compared with the BCP control arm, and a 62.4% true-go rate in the ACP arm compared with the BCP arm (figure 6d) in the OC simulation. Consistent with previous results, ctDNA significantly outperformed week 6 radiographic imaging (ORR_6wk: 17.1% and 3.2% true-go for ABCP vs BCP and ACP vs BCP, respectively) and PFS at 6 weeks (PFS_6wk: 15.5% and 14.3% true-go, ABCP vs BCP, and ACP vs BCP, respectively).

Next, we tested the ctA ML model predictions for patients with matched ctDNA data from IMpower150 in the OC simulation. Using the larger available ctA non-squamous biomarker evaluable population, we applied the ML model to the OC simulation and compared it with ctDNA (figure 6c). We found that ctA significantly underperformed ctDNA in the OC simulation (5.7% and 3.0% true-go rate for ABCP

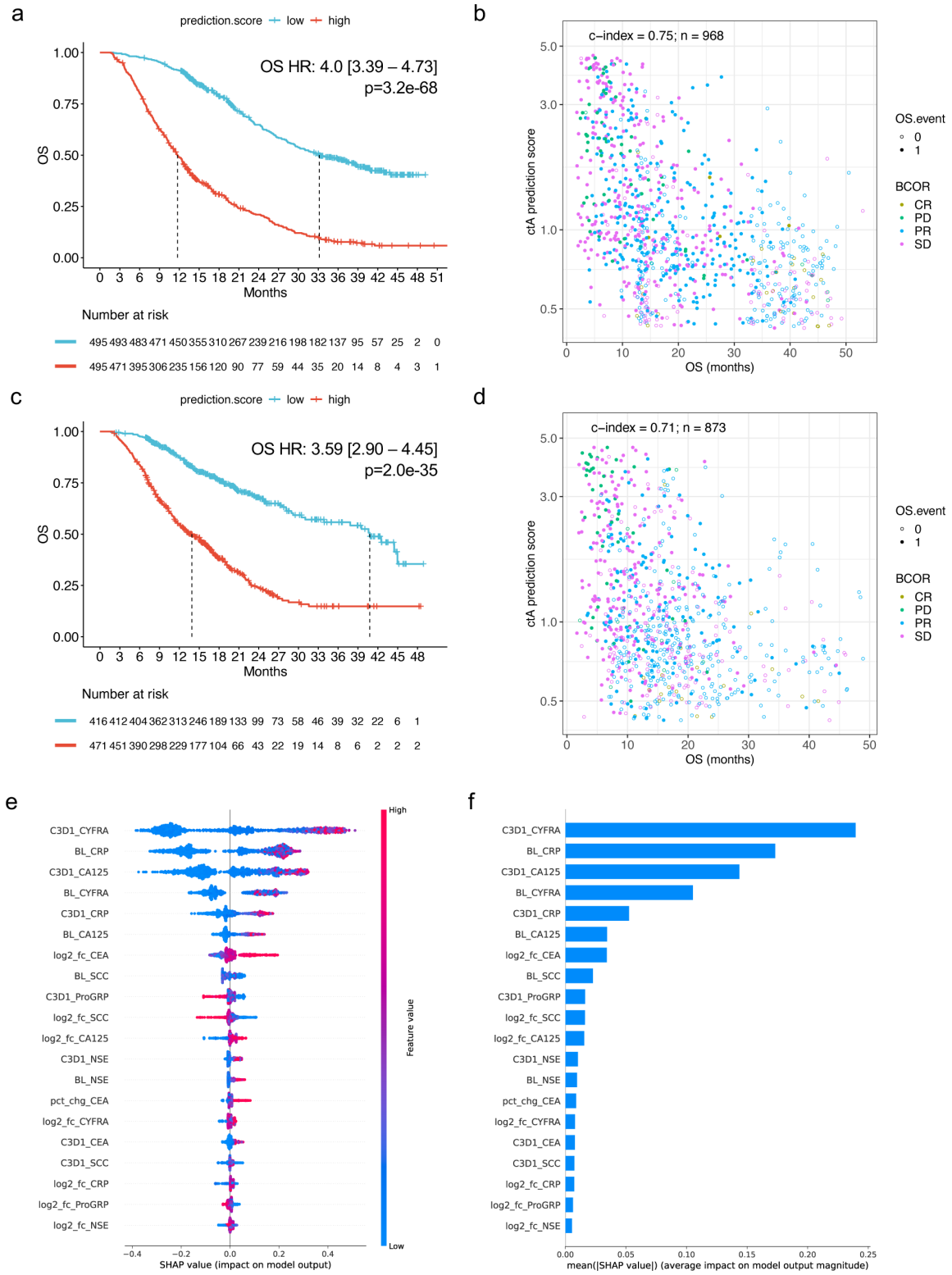


Figure 3 A XGBoost model trained on IMpower132–150 (non-squamous) data selects CYFRA, CRP and CA125 as top features. Kaplan-Meier plot of OS in the (a) training dataset by median split for low/high risk or (c) test (pooled IMpower130, IMpower110 non-squamous and IMp150 hold-back) dataset using the training median value for low/high risk. OS HR and 95% CI are indicated. XGBoost model prediction scores and OS time in (b) training or (d) test. Each point represents one patient, colored by BCOR. Not all patients with OS data have recorded BCOR data. Filled points represent patients with an OS event (death). (e) Distribution of SHAP values for each feature and relative impact to model output. (f) Global feature importance with mean absolute SHAP values plotted. BCOR, best confirmed overall response; CA125, cancer antigen 125; CR, complete response; CRP, C-reactive protein; ctA, circulating tumor antigen; CYFRA, cytokeratin 19 fragment 21-1 (Cyfra21-1); IMp, IMpower; OS, overall survival; PD, progressive disease; PR, partial response; SD, stable disease; SHAP, Shapley additive explanations.

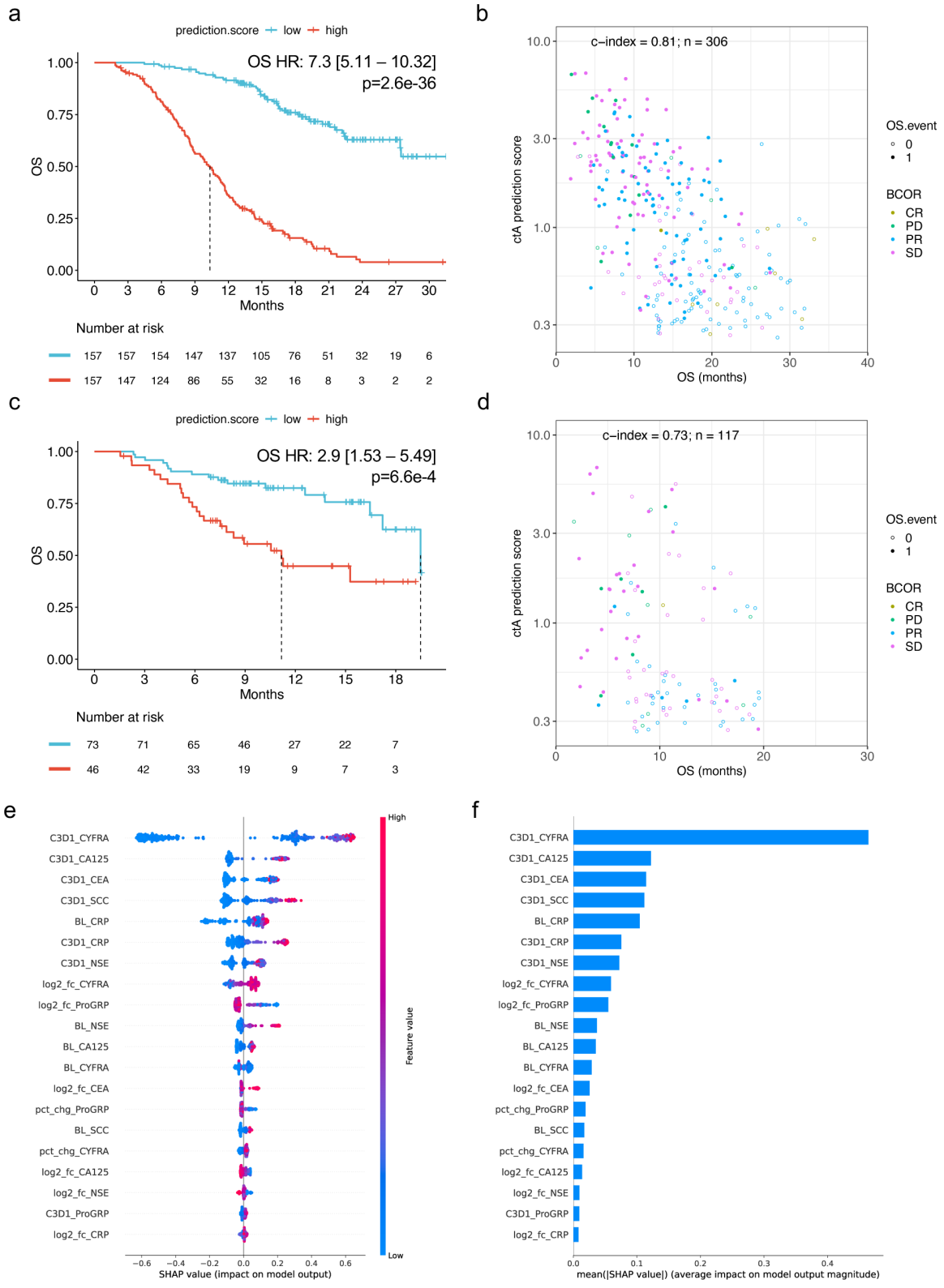


Figure 4 A XGBoost model trained on IMpower131 squamous data selects CYFRA, CA125 and CEA as top features. Kaplan-Meier plot of OS in the (a) training dataset by median split for low/high risk or (c) test (IMpower110 squamous) dataset using the training median value for low/high risk. OS HR and 95% CI are indicated. XGBoost model prediction scores and OS time in (b) training or (d) test. Each point represents one patient, colored by BCOR. Not all patients with OS data have recorded BCOR data. Filled points represent patients with an OS event (death). (e) Distribution of SHAP values for each feature and relative impact to model output. (f) Global feature importance with mean absolute SHAP values plotted. BCOR, best confirmed overall response; CA125, cancer antigen 125; CEA, carcinoembryonic antigen; CR, complete response; ctA, circulating tumor antigen; CYFRA, cytokeratin 19 fragment 21-1 (Cyfra21-1); OS, overall survival; PD, progressive disease; PR, partial response; SD, stable disease; SHAP, Shapley additive explanations.

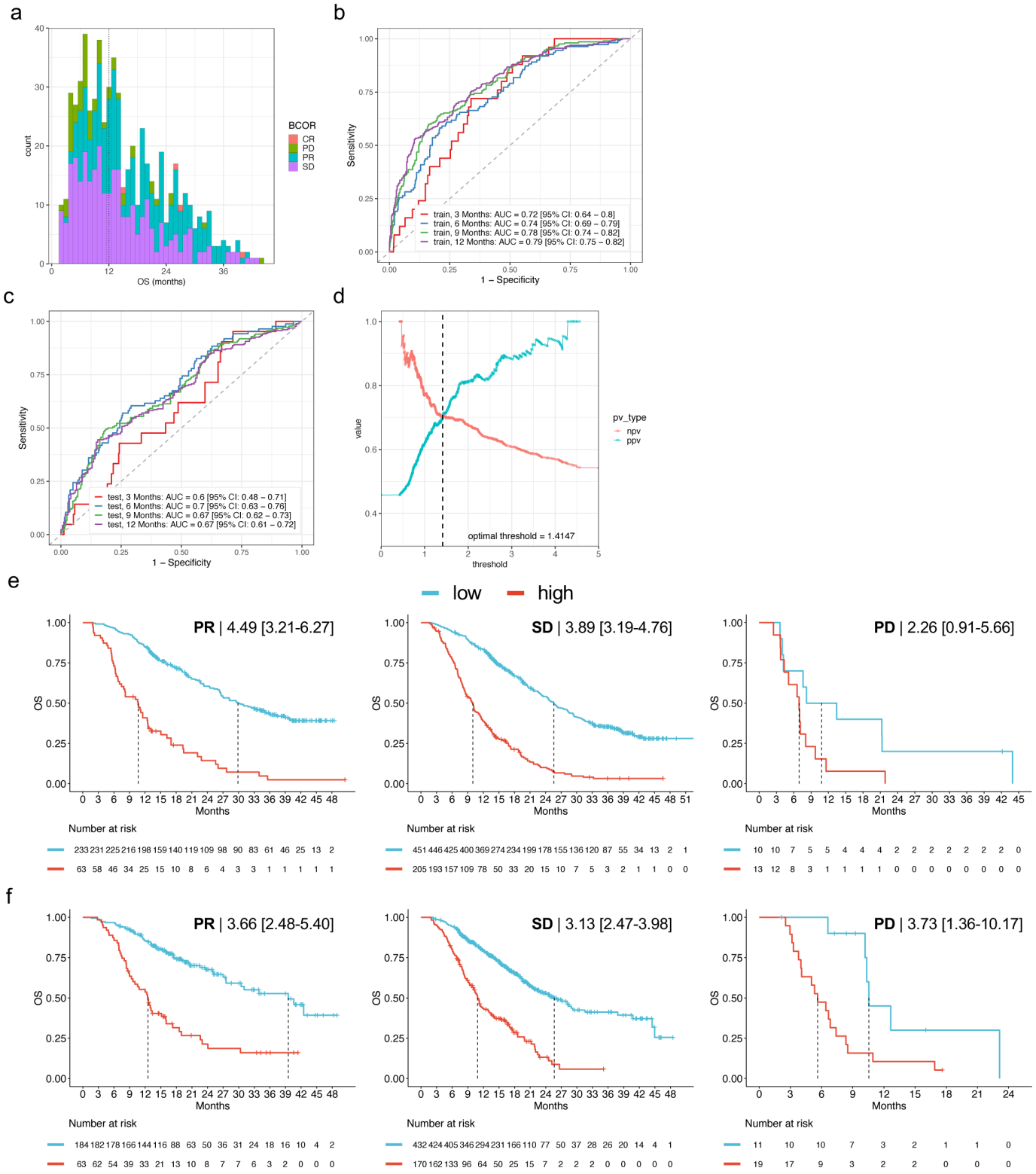


Figure 5 1-year OS predictions in the non-squamous dataset using a XGBoost model and risk stratification by radiographic response. (a) Histogram of BCOR and OS time in the training dataset for patients with an OS event (n=624). (b, c) ROC plots of 3, 6, 9, and 12 month OS as a function of XGBoost model predictions in the (b) training or (c) test dataset. AUC values are indicated for each dataset. (d) PPV and NPV as the threshold cut-off for predicting 1-year OS was adjusted. The vertical dashed line indicates the optimized cut-off value that equally maximizes both PPV and NPV. Using the optimized cut-off for 1-year OS, patients with PR, SD or PD at week 6 in the (e) training or (f) test dataset were risk-stratified and OS plotted. Radiographic response assessed by RECIST and OS HR (95% CI) are indicated on each plot. AUC, area under the curve; BCOR, best confirmed overall response; CR, complete response; NPV, negative predictive value; OS, overall survival; PD, progressive disease; PPV, positive predictive value; PR, partial response; RECIST, response evaluation criteria in solid tumors; ROC, receiver operating characteristic; SD, stable disease.

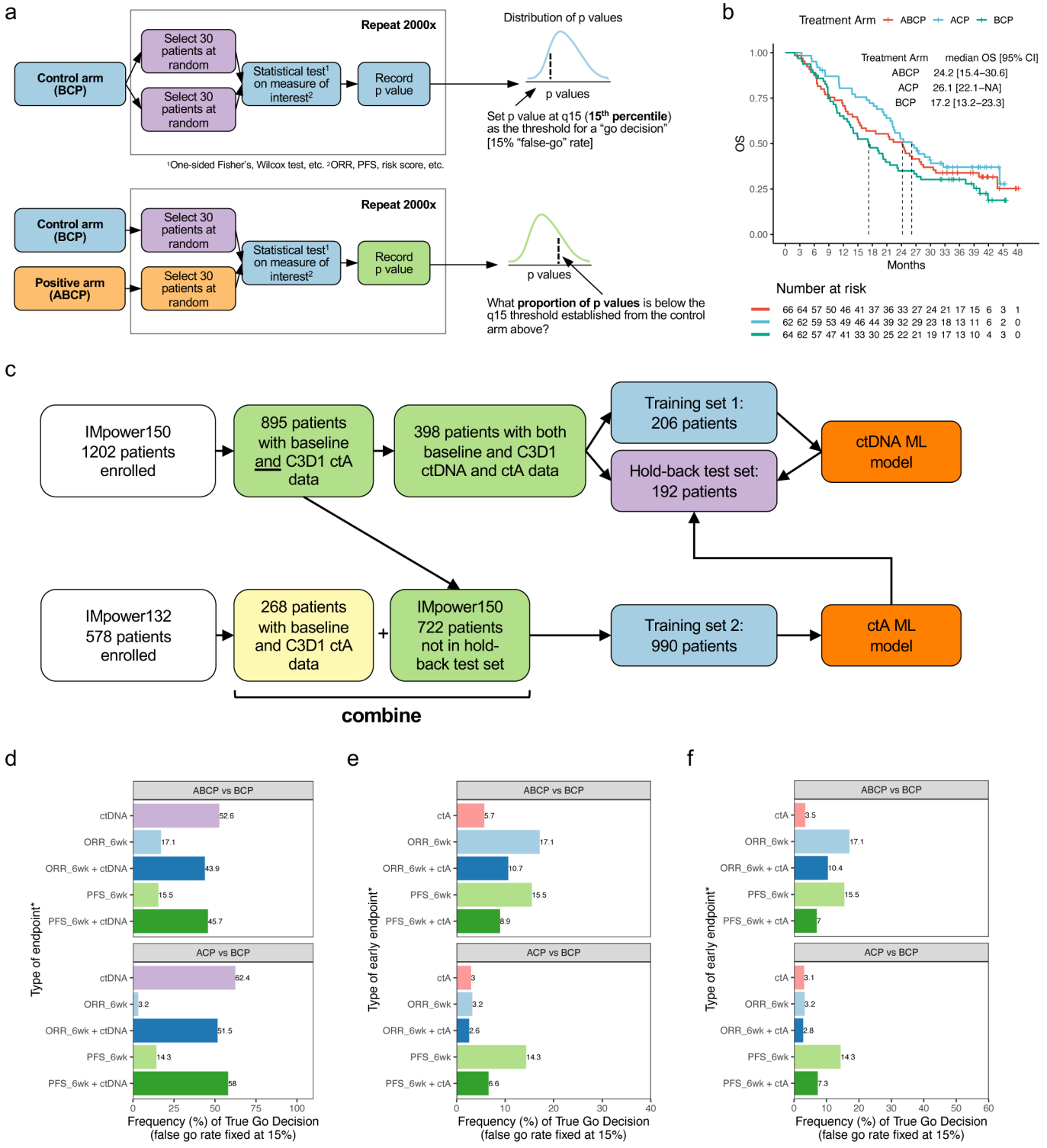


Figure 6 OC simulation with IMp150 ctDNA and ctA. (a) Overview of OC simulation. (b) OS time by treatment arm in IMpower150 hold-back test patients with ctDNA and ctA data. (c) IMpower150 ctA and ctDNA BEP for ML model development. Missing ctA data in the IMpower150 ctDNA hold-back test set was imputed specifically for the OC simulation. (d) Frequency of true-go decision in an OC simulation using various assessment criteria and separated by treatment arm. (e) OC simulation results using the ctA XGBoost model trained on IMpower132–150 and applied to the IMpower150 hold-back test set. (f) OC simulation results using a modified XGBoost model trained only on C3D1 features and relative change features at C3D1 relative to baseline (log_2(fold change) or percent change). ABCP, atezolizumab+bevacizumab+chemotherapy; ACP, atezolizumab+chemotherapy; BCP, bevacizumab+chemotherapy; BEP, biomarker evaluable population; C3D1, cycle 3 day 1; ctA, circulating tumor antigen; ctDNA, circulating tumor DNA; IMp, IMpower; ML, machine learning; OC, operational characteristic; ORR, objective response rate (responders defined as CR/PR and non-responders as SD/PD); OS, overall survival; PD, progressive disease; PFS, progression-free survival; PR, partial response; SD, stable disease; wk, week.

vs BCP and ACP vs BCP, respectively), and that combining ctA with traditional response criteria, such as ORR or PFS, reduced, rather than improved, their performance in the simulation (figure 6e). Because the ctA ML model was trained on both baseline and C3D1 ctA features, baseline prognostic features that are prognostic for OS but do not reflect any treatment effect may contribute to underperformance in the OC simulation. Therefore, we re-trained a ctA ML model with only features that capture changes in ctA at C3D1 relative to baseline. The revised ctA ML model had similar performance for ABCP versus BCP and ACP versus BCP (3.5% and 3.1% true-go rate, respectively) (figure 6f). Consistent with these results, analysis of individual ctA at baseline and C3D1 in the IMpower150 training or hold-back test data revealed no significant differences between treatment arms in mean levels of CRP or any of the ctA tested (online supplemental figure 9). Overall, these data show that while ctA are strongly prognostic for OS, they are unable to capture differences in treatment effects compared with ctDNA in the OC simulation.

DISCUSSION

Our data support previous findings that individual ctA, particularly Cyfra21-1, offers strong prognostic value in addition to radiographic imaging. These results are consistent with multiple meta-analyses which revealed that high baseline levels of Cyfra21-1 are negatively associated with both OS and PFS across NSCLC disease stages and treatment regimens.[33–35] Our findings also confirm previous work showing histology-specific differences in baseline levels of ctA including Cyfra21-1 and SCC.[26] While ML models combining the ctA tested improved prognostic performance in training, they offered similar performance to Cyfra21-1 alone in the test datasets. These results suggest that Cyfra21-1 levels at 6 weeks post therapy are a sufficient, pan-histology prognostic marker for patients with NSCLC and can add value to radiographic imaging and help inform clinical decision making. For example, SD patients by radiographic imaging who are prognostically high-risk by the ctA model or Cyfra21-1 could potentially benefit from treatment escalation and increased monitoring, while SD patients who are prognostically low-risk may have treatment de-escalation. While multiple prospective studies would be needed for implementation into clinical guidelines, our results suggest that ctA like Cyfra21-1 can be valuable tools to support treatment decisions.

In the OC simulation, we subsampled patients from a positive phase 3 study with a known, statistically significant benefit in OS in experimental treatment arms compared with the control to simulate a phase 2 clinical trial. At 6 weeks, a correct "go decision" to proceed from phase 2 to phase 3 was observed in over 50% of the simulations using either traditional response criteria, such as ORR, or molecular criteria, such as ctDNA. However, ctA significantly underperformed ctDNA, with the frequency of true-go decisions made in less than 5% of the total simulations. The OC simulation results suggest that while ctA levels at both baseline and 6 weeks into treatment have a strong association with OS, ctA are less able to distinguish treatment arms (at least with ABCP against BCP) and less able to capture treatment modulatory effects at week 6 compared with ctDNA or traditional criteria such as PFS used for clinical development of drug candidates.

Interestingly, a previous report suggested that baseline ctDNA positivity was associated with elevated Cyfra21-1 levels.[36] We observed some modest associations between baseline Cyfra21-1 and ctDNA levels in our limited IMpower150 dataset ($\rho = 0.4$; online supplemental figure 10), and further validation in additional studies is needed to support this finding.

For the comparisons between ctDNA and ctA, these data are also limited by the number of patients from IMpower150 with both ctDNA and ctA data. Since available, clinical grade (IVD-CE (in vitro diagnostic conformité européenne)) assays were prioritized in the selection of ctA assays to maximize immediate translatability of these findings, the number of ctA tested is also limited in this study, and ctA that were not tested here may potentially better capture on-treatment effects and/or offer superior prognostic value.

Further work will focus on the continued investigation of blood-based biomarkers, including the use of unbiased discovery approaches to identify pharmacodynamic biomarkers that best reflect treatment effect to support gating decisions in drug development and to further improve prognostic risk stratification to help inform clinical decision-making (treatment escalation/diversion or de-escalation). Baseline ctA levels may also be an important patient stratification factor in future clinical trials; however, further investigation is required. Overall, this work will help support efforts to shorten development timelines and reduce costs for new therapies, bringing potentially transformative medicines to patients faster.

Acknowledgements This study was sponsored by Roche Diagnostics GmbH Mannheim and Genentech. The authors would like to thank the patients for their contributions to this study. We thank Ezgi Kalay for final proofreading of this manuscript and coordination of review at Roche Diagnostics. The assays used in this analysis are currently not intended to be used to aid in the management of patients with NSCLC in the USA or in CE markets. All other product names and trademarks are the property of their respective owners.

Contributors BW, DSS, and NSP conceived of and designed the study. JHC, IE, and VM coordinated sample testing. CWT, WZ, ZJA, AM, and FR performed data analysis. CWT drafted the manuscript and AM, BW, VR, KS, and NSP provided critical feedback. All authors have read and approved the final version of the manuscript. CWT, WZ, and AM had access to and verified the data in the study. NSP is the guarantor of this study.

Funding This work was funded by Roche Diagnostics GmbH and Genentech, a member of the Roche group.

Competing interests CWT, WZ, ZJF, JHC, IE, VM, MKS, KS, and NSP are employees of Genentech, a member of the Roche group, and hold stocks/stock options in F. Hoffmann-La Roche. DSS is a previous employee of Genentech and holds stocks/stock options in F. Hoffmann-La Roche. AM, BW, FR, and VR are employees of Roche Diagnostics GmbH, and BW and VR hold stocks/stock options in F. Hoffmann-La Roche. SH received funding or honoraria from Roche Diagnostics, Bristol Myers Squibb, Merck KgaA, Sysmex Inostics, and Volition SPRL. MR received honoraria for lectures and consultancy from Roche. The cobas testing platform and associated reagents are manufactured and sold by Roche Diagnostics.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data were generated by the authors and deposited in a repository. Anonymized patient-level tumor antigen data can be found in the European Genome Phenome Archive (EGA). Users must agree to the Data Access Agreement detailed in the EGA entry before they can access these data. Data access requests are reviewed and adjudicated by the Genentech DevSci Data Access Committee (devsci-dac-d@gene.com).

ORCID iDs
Charles W Tran https://orcid.org/0000-0001-5818-818X
Francesco Ranucci https://orcid.org/0009-0009-8370-1403
Martin Reck https://orcid.org/0000-0002-5348-4462
Namrata S Patil https://orcid.org/0000-0001-6156-684X

REFERENCES

1 Society, A.C. Cancer facts & figures, 1930. 2025. Available: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2025/2025-cancer-facts-and-figures-acs.pdf
2 Meyer M-L, Fitzgerald BG, Paz-Ares L, *et al*. New promises and challenges in the treatment of advanced non-small-cell lung cancer. *Lancet* 2024;404:803–22.
3 Ganti AK, Klein AB, Cotarla I, *et al*. Update of Incidence, Prevalence, Survival, and Initial Treatment in Patients With Non-Small Cell Lung Cancer in the US. *JAMA Oncol* 2021;7:1824–32.
4 Normanno N, Morabito A, Rachiglio AM, *et al*. Circulating tumour DNA in early stage and locally advanced NSCLC: ready for clinical implementation? *Nat Rev Clin Oncol* 2025;22:215–31.
5 de Abreu AR, Op de Beeck K, Laurent-Puig P, *et al*. The Position of Circulating Tumor DNA in the Clinical Management of Colorectal Cancer. *Cancers (Basel)* 2023;15:1284.
6 Dang DK, Park BH. Circulating tumor DNA: current challenges for clinical utility. *J Clin Invest* 2022;132:e154941.
7 Chen K, Shen H, Wu S, *et al*. Abstract 5916: Tumor-informed patient-specific panel outperforms tumor-naïve and tumor-informed fixed panel for circulating tumor DNA (ctDNA)-based postoperative monitoring of non-small cell lung cancer (NSCLC). *Cancer Res* 2022;82:5916.
8 Santonja A, Cooper WN, Eldridge MD, *et al*. Comparison of tumor-informed and tumor-naïve sequencing assays for ctDNA detection in breast cancer. *EMBO Mol Med* 2023;15:e16505.
9 van der Gaast A, Schoenmakers CH, Kok TC, *et al*. Evaluation of a new tumour marker in patients with non-small-cell lung cancer: Cyfra 21.1. *Br J Cancer* 1994;69:525–8.
10 Oh H-J, Park H-Y, Kim K-H, *et al*. Progastrin-releasing peptide as a diagnostic and therapeutic biomarker of small cell lung cancer. *J Thorac Dis* 2016;8:2530–7.
11 Zha Z, Li D, Zhang P, *et al*. Neuron specific enolase promotes tumor metastasis by activating the Wnt/β-catenin pathway in small cell lung cancer. *Transl Oncol* 2021;14:101039.
12 Pujol JL, Grenier J, Daurès JP, *et al*. Serum fragment of cytokeratin subunit 19 measured by CYFRA 21-1 immunoradiometric assay as a marker of lung cancer. *Cancer Res* 1993;53:61–6.
13 Xu Y, Xu L, Qiu M, *et al*. Prognostic value of serum cytokeratin 19 fragments (Cyfra 21-1) in patients with non-small cell lung cancer. *Sci Rep* 2015;5:9444.
14 Patil NS, Zou W, Mocci S, *et al*. C-reactive protein reduction post treatment is associated with improved survival in atezolizumab (anti-PD-L1) treated non-small cell lung cancer patients. *PLoS One* 2021;16:e0246486.
15 Riely GJ, Wood DE, Ettinger DS, *et al*. Non-Small Cell Lung Cancer, Version 4.2024, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2024;22:249–74.
16 Planchard D, Popat S, Kerr K, *et al*. Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2018;29:iv192–237.
17 Socinski MA, Jotte RM, Cappuzzo F, *et al*. Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC. *N Engl J Med* 2018;378:2288–301.
18 West H, McCleod M, Hussein M, *et al*. Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol* 2019;20:924–37.
19 Jotte R, Cappuzzo F, Vynnychenko I, *et al*. Atezolizumab in Combination With Carboplatin and Nab-Paclitaxel in Advanced Squamous NSCLC (IMpower131): Results From a Randomized Phase III Trial. *J Thorac Oncol* 2020;15:1351–60.
20 Nishio M, Barlesi F, West H, *et al*. Atezolizumab Plus Chemotherapy for First-Line Treatment of Nonsquamous NSCLC: Results From the Randomized Phase 3 IMpower132 Trial. *J Thorac Oncol* 2021;16:653–64.
21 Herbst RS, Giaccone G, de Marinis F, *et al*. Atezolizumab for First-Line Treatment of PD-L1-Selected Patients with NSCLC. *N Engl J Med* 2020;383:1328–39.
22 Mang A, Zou W, Rolny V, *et al*. Combined use of CYFRA 21-1 and CA 125 predicts survival of patients with metastatic NSCLC and stable disease in IMpower150. *Tumour Biol* 2024;46:S177–90.
23 Assaf ZJF, Zou W, Fine AD, *et al*. A longitudinal circulating tumor DNA-based model associated with survival in metastatic non-small-cell lung cancer. *Nat Med* 2023;29:859–68.
24 Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, New York, NY, USA: Association for Computing Machinery, (KDD '16); 2016:785–94. Available: https://dl.acm.org/doi/10.1145/2939672.2939785
25 Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Proceedings of the 31st international conference on neural information processing systems, Red Hook, NY, USA: Curran Associates Inc, (NIPS'17); 2017:4768–77.
26 Trulson I, Klawonn F, von Pawel J, *et al*. Improvement of differential diagnosis of lung cancer by use of multiple protein tumor marker combinations. *Tumour Biol* 2024;46:S81–98.
27 Nisman B, Biran H, Heching N, *et al*. Prognostic role of serum cytokeratin 19 fragments in advanced non-small-cell lung cancer: association of marker changes after two chemotherapy cycles with different measures of clinical response and survival. *Br J Cancer* 2008;98:77–9.
28 Pujol J-L, Molinier O, Ebert W, *et al*. CYFRA 21-1 is a prognostic determinant in non-small-cell lung cancer: results of a meta-analysis in 2063 patients. *Br J Cancer* 2004;90:2097–105.
29 Huang J, Xiao Y, Zhou Y, *et al*. Baseline serum tumor markers predict the survival of patients with advanced non-small cell lung cancer receiving first-line immunotherapy: a multicenter retrospective study. *BMC Cancer* 2023;23:812.
30 Zou H, Hastie T. Regularization and Variable Selection Via the Elastic Net. *J R Stat Soc Ser B Stat Methodol* 2005;67:301–20.
31 Yang Y, Jiang X, Liu Y, *et al*. Elevated tumor markers for monitoring tumor response to immunotherapy. *eClinicalMedicine* 2022;46:101381.
32 Fridlyand J, Kaiser LD, Fyfe G. Analysis of tumor burden versus progression-free survival for Phase II decision making. *Contemp Clin Trials* 2011;32:446–52.
33 Xu Y, Xu L, Qiu M, *et al*. Prognostic value of serum cytokeratin 19 fragments (Cyfra 21-1) in patients with non-small cell lung cancer. *Sci Rep* 2015;5.
34 Zhang ZH, Han YW, Liang H, *et al*. Prognostic value of serum CYFRA21-1 and CEA for non-small-cell lung cancer. *Cancer Med* 2015;4:1633–8.
35 Yu Z, Zhang G, Yang M, *et al*. Systematic review of CYFRA 21-1 as a prognostic indicator and its predictive correlation with clinicopathological features in Non-small Cell Lung Cancer: A meta-analysis. *Oncotarget* 2017;8:4043–50.
36 Buresova M, Benesova L, Minarik M, *et al*. Circulating Tumor DNA correlates with Lactate Dehydrogenase, CYFRA 21-1, and CRP levels in patients with advanced NSCLC. *J Cancer* 2023;14:1–8.