

# Real-world data prognostic model of overall survival in patients with advanced NSCLC receiving anti-PD-1/PD-L1 immune checkpoint inhibitors as second-line monotherapy

Cristina Julian<sup>1</sup> | Robson J. M. Machado<sup>2</sup> | Sandhya Girish<sup>1</sup> | Pascal Chanu<sup>3</sup> |  
Dominik Heinzmann<sup>4</sup> | Chris Harbron<sup>2</sup> | Anda Gershon<sup>1</sup> | Shannon M. Pfeiffer<sup>1</sup> |  
Wei Zou<sup>1</sup> | Valerie Quarmby<sup>1</sup> | Qing Zhang<sup>1</sup> | Yachi Chen<sup>1</sup>

<sup>1</sup>Genentech, Inc, South San Francisco, California, USA

<sup>2</sup>Roche Products, Welwyn Garden City, UK

<sup>3</sup>Genentech/Roche, Lyon, France

<sup>4</sup>F Hoffmann-La Roche Ltd, Basel, Switzerland

## Correspondence

Yachi Chen, Revolution Medicines, 700 Saginaw Drive, Redwood City, CA 94036, USA.

Email: [ychen@revmed.com](mailto:ychen@revmed.com)

## Present address

Robson J. M. Machado, Drax Group plc, Shelby, UK

Sandhya Girish, Gilead, South San Francisco, California, USA

Shannon M. Pfeiffer, The Parker Institute for Cancer, Immunotherapy, San Francisco, California, USA

## Funding information

F Hoffmann-La Roche Ltd/Genentech, Inc.

## Abstract

**Background and aim:** The objective of this retrospective, observational, non-interventional cohort study was to investigate prognostic factors of overall survival (OS) in patients with advanced non-small cell lung cancer (aNSCLC) and to develop a novel prognostic model.

**Methods:** A total of 4049 patients with aNSCLC diagnosed between January 2011 and February 2020 who received atezolizumab, nivolumab, or pembrolizumab as second-line monotherapy were selected from a real-world deidentified database to build the cohort. Patients could not have received first-line treatment with clinical study drug(s) nor immune checkpoint inhibitors including anti-programmed cell death 1 (PD-1)/programmed death-ligand 1 (PD-L1), and anti-cytotoxic T-lymphocyte-associated protein 4 therapies.

**Results:** Patients had a median age of 69 years; 45% were female, 75% White, 70% had stage IV at initial diagnosis, and 70% had nonsquamous histology. A Cox proportional hazards model with lasso regularization was used to build a prognostic model for OS using 18 baseline demographic and clinical factors based on the real-world data cohort. The risk-increasing prognostic factors were abnormally low albumin and chloride levels, Eastern Cooperative Oncology Group performance status score  $\geq 2$ , and abnormally high levels of alkaline phosphatase and white blood cells. The risk-decreasing prognostic factors were PD-L1 positivity, longer time from advanced diagnosis to start of first-line therapy, and higher systolic blood pressure. The performance of the model was validated using data from the OAK trial, and the c-index for the OAK trial validation cohort was 0.65 and 0.67 for the real-world data cohort.

**Conclusions:** Based on baseline demographic and clinical factors from a real-world setting, this prognostic model was developed to discriminate the risk of death in

Cristina Julian and Robson J. M. Machado shared their first authorship

Qing Zhang and Yachi Chen shared their last authorship

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 F. Hoffmann-La Roche Ltd/Genentech, Inc. *Cancer Reports* published by Wiley Periodicals LLC.



patients with aNSCLC treated with checkpoint inhibitors as second-line monotherapy, and it performed well in the real-world data and clinical trial cohorts.

#### KEYWORDS

advanced NSCLC, immune checkpoint inhibitors, prognostic model, real-world data

## 1 | INTRODUCTION

Lung cancer is the most prevalent type of cancer worldwide. It is estimated that lung cancer contributes to 24% of cancer-related deaths in men and 23% of cancer-related deaths in women in the US.<sup>1</sup> Survival outcomes of patients with lung cancer have improved due to the development of new treatments such as immunotherapy. This improvement is especially pronounced in patients with non-small cell lung cancer (NSCLC).<sup>2</sup> Use of immune checkpoint inhibitors (CPIs) as second-line (2L) treatment for patients with advanced NSCLC (aNSCLC) has improved survival and antitumor response compared with chemotherapy. Specifically, monoclonal antibodies against programmed cell death 1 (PD-1; nivolumab and pembrolizumab) and programmed death-ligand 1 (PD-L1; atezolizumab) as 2L monotherapy have shown a benefit in overall survival (OS) compared with docetaxel.<sup>3-5</sup> However, despite this increase in survival benefit, 40% to 60% of patients do not respond to these therapies.<sup>3,5-7</sup>

Research on prognostic factors in patients with NSCLC receiving CPIs in 2L treatment has mainly focused on one or a few risk factors, such as PD-L1 expression that has been shown to be prognostic for PD-L1 agents in NSCLC.<sup>8</sup> Although PD-L1 expression can help physicians decide on appropriate treatment for a patient, a prognostic model accounting for multiple factors simultaneously could create clinical risk groups for stratifying patients by disease severity. One multivariable prognostic risk model for OS, the Real wOrld PROgnostic (ROPRO) score, was derived across 17 cancer cohorts from 27 demographic, clinical, and routine laboratory parameters and validated using two independent phase I and III trials.<sup>9</sup> In the NSCLC cohort, patients with higher ROPRO scores (upper 10%) had an almost eightfold increased risk of death compared with patients with lower scores (lower 10%).<sup>9</sup> While this model can be applied to different cancer indications, a model with a smaller number of variables tailored to a specific indication, such as NSCLC, may facilitate its application and also minimize overall missing data by using less variables. The objective of this study was to build a prognostic model using clinical and laboratory factors that are prognostic of OS in patients with aNSCLC who received an anti-PD-1/PD-L1 CPI as 2L monotherapy in a real-world setting. An independent clinical trial cohort was used to validate the model.

## 2 | METHODS

### 2.1 | Training cohort

This retrospective, observational, noninterventional cohort study used the nationwide Flatiron Health electronic health record-derived

deidentified database. During the study period, the deidentified data originated from  $\approx 280$  US cancer clinics ( $\approx 800$  sites of care).

The Flatiron Health database is longitudinal and consists of deidentified, patient-level, structured and unstructured data, curated via technology-enabled abstraction.<sup>10,11</sup> Patients with aNSCLC diagnosed between January 1, 2011, and February 1, 2020, who received anti-PD-1/PD-L1 CPIs (atezolizumab, nivolumab, or pembrolizumab) as 2L monotherapy were included. Patients were aged  $\geq 18$  years at the time of advanced diagnosis, not previously diagnosed with other types of cancer, had started any treatment within 90 days after advanced diagnosis, and could not have received first-line (1L) treatment with clinical study drug(s) nor CPIs including anti-PD1/PD-L1 or anti-cytotoxic T-lymphocyte-associated protein 4. Patients were followed until death, last observed patient activity, or the database cutoff date (February 1, 2020).

### 2.2 | Prognostic factors

A literature review on relevant prognostic factors was conducted using PubMed over the past 10 years to identify reviews and trials that evaluated prognostic factors of OS or progression-free survival in patients with lung cancer receiving CPIs.<sup>7,8,12-15</sup> Primary variables of interest in the real-world database included demographic, clinical, and laboratory characteristics that were measured at baseline.

Treatment start was defined as the first date of drug administration in 2L therapy. Sex, age at advanced diagnosis, and age at start of 2L therapy were included in the model. Race was categorized as Asian, Black or African American, Hispanic or Latino, White, or other. Smoking status was defined by previous/current or never smoker. Eastern Cooperative Oncology Group performance status (ECOG PS) scores were used to determine the patient's performance status and were categorized as 0/1 and  $\geq 2$ . NSCLC stage at initial diagnosis was included in the model and categorized as stage  $<IV$  or  $IV$ . Other disease-related variables included histology, body mass index, systolic blood pressure, diastolic blood pressure, tumor PD-L1 expression (positive,  $\geq 1\%$ ; negative,  $<1\%$ ), and mutation status for *KRAS* and *EGFR*.

Laboratory values were determined using standard clinical chemistry methods and then classified as abnormally low, within the normal range, or abnormally high (except for albumin, which was classified as abnormal [low] and normal). The following parameters were assessed: serum albumin, circulating monocyte, lymphocyte, thrombocyte, neutrophil, and platelet counts, white blood cell counts, absolute eosinophil counts, bilirubin, lactate dehydrogenase, alkaline phosphatase (ALP), aspartate aminotransferase, alanine transaminase, chloride,

calcium, creatinine, total serum protein, urea, and hemoglobin. The interactions of neutrophils and lymphocytes, lymphocytes and monocytes, platelets and lymphocytes, neutrophils and white blood cells, and aspartate aminotransferase and alanine transaminase were also investigated. Reference ranges vary across laboratories and were based on instrument sensitivity as well as units of measure for the particular test at the performing laboratory.

### 2.3 | Validation cohort

The developed real-world data (RWD) prognostic model was applied retrospectively to data from the OAK trial (blinded for peer review), a phase III, open-label, randomized controlled trial that assessed the efficacy and safety of atezolizumab monotherapy compared with docetaxel as 2L treatment in patients with aNSCLC.<sup>7</sup> The cohort from the OAK trial was selected based on the following criteria: squamous or nonsquamous NSCLC, age  $\geq 18$  years, measurable disease per Response Evaluation

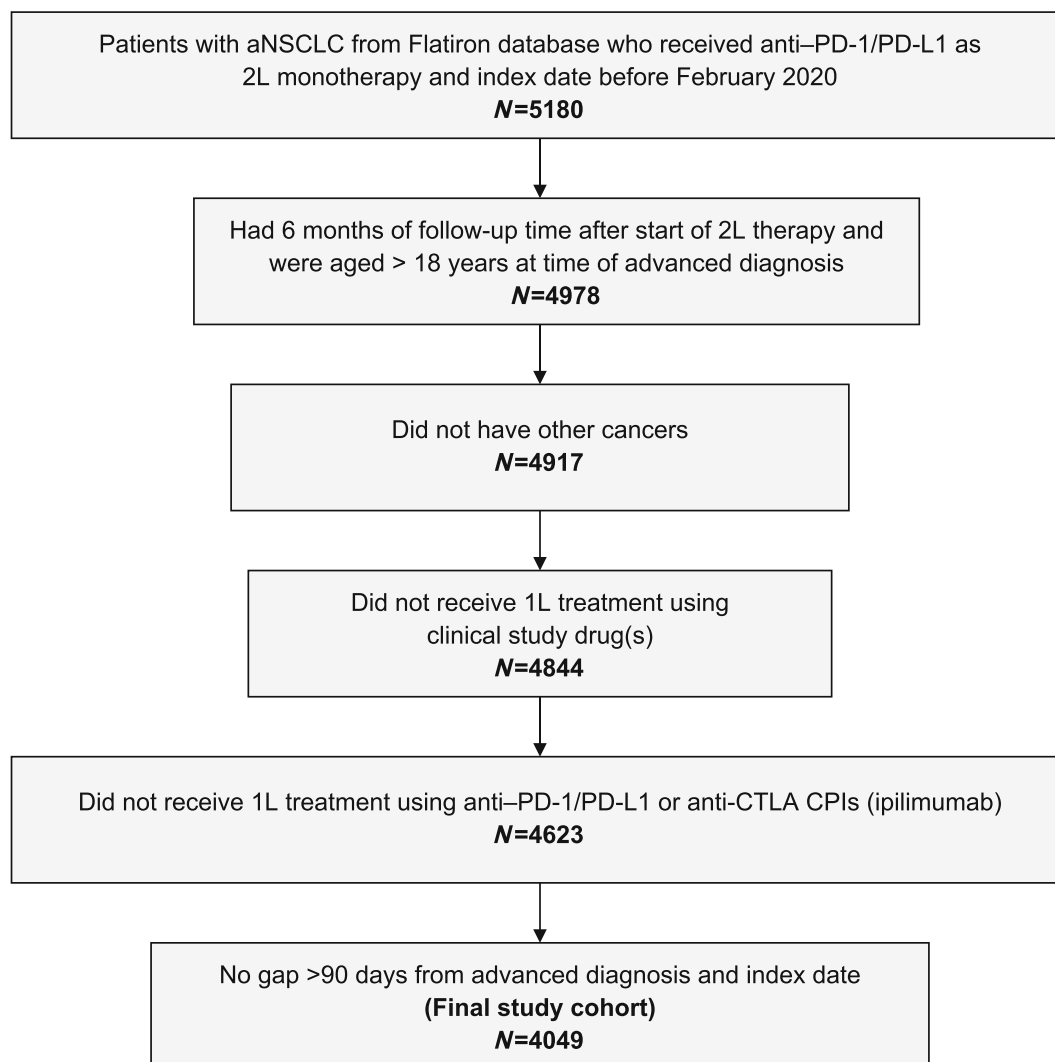
Criteria in Solid Tumors version 1.1, ECOG PS 0/1, previous chemotherapy, and no prior treatment with PD-1/PD-L1 inhibitors.

### 2.4 | Statistical analysis

A Cox proportional hazards model with lasso regularization was used to build the prognostic model based on the aforementioned demographic, clinical, and laboratory factors to select variables that contribute most to OS prediction.<sup>16</sup> The model was estimated using the GLMNET package in R with  $\alpha = 0.9$  and  $\lambda$  tuned using a grid search. The prognostic index (PI) was fitted to predict the risk of death for a patient. The PI was calculated based on the following Cox model:

$$h_i(t) = h_0(t) \exp(\beta^T x_i),$$

where  $h_i(t)$  is the hazard function for a patient,  $h_0(t)$  is an unspecified baseline hazard function,  $x_i$  is a  $p \times 1$  vector of covariates for that



**FIGURE 1** Patient flowchart. 1L, first line; 2L, second line; aNSCLC, advanced non-small cell lung carcinoma; CPI, checkpoint inhibitor; CTLA, cytotoxic T-lymphocyte-associated protein; PD-1, programmed cell death-1; PD-L1, programmed death-ligand 1



**TABLE 1** Demographic and clinical characteristics of the patients included in the prognostic model

Characteristic	Real-world database (n = 4049)	OAK trial (n = 792)
Age at index, median (IQR), years	69.0 (61.0, 76.0)	63.0 (57.0, 70.0)
Age at advanced diagnosis, median (IQR), years	68.0 (60.0, 75.0)	63.0 (56.0, 69.0)
Time from advanced diagnosis to index date, median (IQR), months	7.88 (4.93, 14.36)	13.52 (8.28, 24.44)
Heart rate, median (IQR), bpm	87.0 (76.0, 100.0)	82.0 (72.0, 93.0)
Diastolic blood pressure, median (IQR), mmHg	72.0 (66.0, 80.0)	75.0 (68.0, 80.0)
Systolic blood pressure, median (IQR), mmHg	123.0 (110.0, 137.0)	125.0 (114.0, 135.0)
Smoking status, n (%)		
Previous/current	3685 (91.0)	651 (82.2)
Never	364 (9.0)	141 (17.8)
Stages, n (%)		
Stage <IV	1211 (29.9)	320 (40.4)
Stage IV	2837 (70.1)	472 (59.6)
Sex, n (%)		
Female	1834 (45.3)	309 (39.0)
Male	2215 (54.7)	483 (61.0)
Histology, n (%)		
Nonsquamous	2826 (69.8)	583 (73.6)
Squamous	1223 (30.2)	209 (26.4)
ECOG PS, n (%)		
0–1	2866 (70.8)	792 (100.0)
≥2	1183 (29.2)	0
Race, n (%)		
White	3040 (75.1)	579 (73.1)
Black	445 (11.0)	17 (2.1)
Hispanic or Latino	122 (3.0)	0
Asian	84 (2.1)	174 (22.0)
Other	358 (8.8)	22 (2.8)
BMI, n (%)		
Underweight	348 (8.6)	40 (5.1)
Healthy	1780 (44.0)	364 (46.0)
Overweight	1176 (29.0)	258 (32.6)
Obese	745 (18.4)	130 (16.4)
PD-L1 expression, n (%)		
Negative	1550 (38.3)	359 (45.3)
Positive	2499 (61.7)	433 (54.7)

**TABLE 1** (Continued)

Characteristic	Real-world database (n = 4049)	OAK trial (n = 792)
KRAS mutations, n (%)		
Negative	2958 (73.1)	602 (76.0)
Positive	1091 (26.9)	190 (24.0)
EGFR mutations, n (%)		
Negative	3913 (96.6)	703 (88.8)
Positive	136 (3.4)	89 (11.2)

Abbreviations: BMI, body mass index; bpm, beats per minute; ECOG PS, Eastern Cooperative Oncology Group performance status; EGFR, epidermal growth factor receptor; IQR, interquartile range; PD-L1, programmed death-ligand 1.

patient and  $\beta$  is a  $p \times 1$  vector of coefficients. The number  $\beta^T x_i$  was defined as the PI. Larger values of PI implied shorter survival times, while smaller values of PI implied longer survival times. Survival data were represented by quartiles of PI values. A 10-fold cross-validation with a split of 80% for the training set and 20% for the testing set was performed.

The validity of the model developed from the real-world database was tested in the OAK data by assessing the model's discrimination. Kaplan–Meier OS curves, stratified into quartiles by the PI, were generated for both the RWD and trial cohorts. A concordance index (c-index), defined as the proportion of concordant pairs divided by the total number of possible evaluation pairs, was then calculated for both the real-world database and OAK.

## 2.5 | Missing data

Missing values were imputed using the chained equation method and predictive mean matching as previously described in the literature.<sup>17</sup> The imputation model included all variables used in the analysis model, including the outcome variables. In this case, the cumulative hazards and censor indicator were used for each patient.<sup>18</sup> Given that the real-world database and OAK data sets have different missing data mechanisms, separate imputation models for each data set were developed, with exactly the same model specification.

## 3 | RESULTS

### 3.1 | Patient characteristics

A total of 4049 patients from the real-world database and 792 patients from the OAK trial were included. Initially, 5180 patients with aNSCLC were selected from the RWD who had received anti-PD-1/PD-L1 CPIs as 2L treatment and diagnosed between January 1, 2011, and February 1, 2020 (Figure 1).

**TABLE 2** Laboratory characteristics of patients included in the prognostic model

Characteristic	Real-world database (n = 4049)	OAK trial (n = 792)
White blood cells, n (%)		
High	824 (20.4)	124 (15.7)
Low	323 (8.0)	26 (3.3)
Normal	2902 (73.3)	642 (81.1)
Hemoglobin, n (%)		
High	10 (0.2)	0
Low	2812 (69.4)	792 (100)
Normal	1227 (30.3)	0
Protein total, n (%)		
High	86 (2.1)	38 (4.8)
Low	700 (17.3)	55 (6.9)
Normal	3263 (80.6)	699 (88.3)
Creatinine, n (%)		
High	550 (13.6)	101 (12.8)
Low	707 (17.5)	57 (7.2)
Normal	2792 (69.0)	634 (80.1)
Bilirubin, n (%)		
High	70 (1.7)	7 (0.9)
Low	361 (8.9)	28 (3.5)
Normal	3618 (89.4)	757 (95.6)
Aspartate aminotransferase, n (%)		
High	332 (8.2)	64 (8.1)
Low	197 (4.9)	15 (1.9)
Normal	3520 (83.4)	713 (90.0)
Alanine transaminase, n (%)		
High	230 (5.7)	40 (5.1)
Low	279 (6.9)	21 (2.7)
Normal	3540 (87.4)	731 (92.3)
Alkaline phosphatase, n (%)		
High	788 (19.5)	153 (19.3)
Low	39 (1.0)	11 (1.4)
Normal	3222 (79.6)	628 (79.3)
Lactate dehydrogenase, n (%)		
High	1599 (39.5)	336 (42.4)
Low	151 (3.7)	16 (2.0)
Normal	2299 (56.8)	440 (55.6)
Calcium, n (%)		
High	153 (3.8)	791 (99.9)
Low	518 (12.8)	1 (0.1)
Normal	3378 (83.4)	0
Chloride, n (%)		
High	97 (2.4)	22 (2.8)
Low	823 (20.3)	117 (14.8)
Normal	3129 (77.3)	653 (82.4)

(Continues)

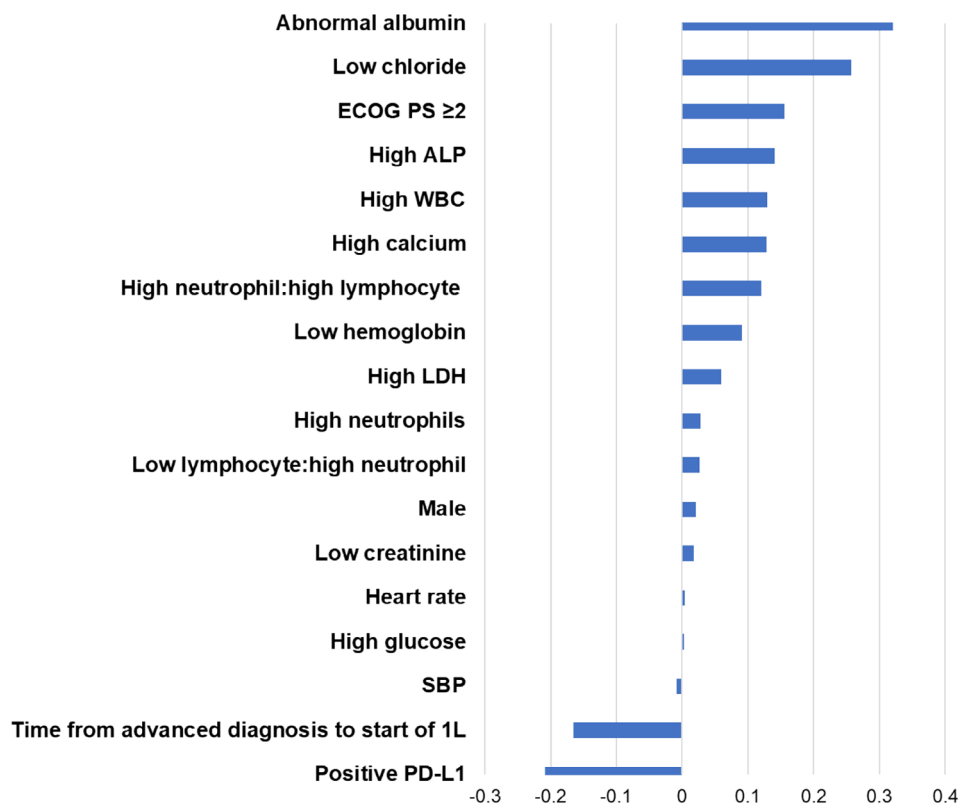
**TABLE 2** (Continued)

Characteristic	Real-world database (n = 4049)	OAK trial (n = 792)
Albumin, n (%)		
Abnormal (low)	1352 (33.4)	138 (17.4)
Normal/high	2697 (66.6)	654 (82.6)
Urea, n (%)		
High	627 (15.5)	789 (99.6)
Low	208 (5.1)	0
Normal	3214 (79.4)	3 (0.4)
Absolute eosinophil count, n (%)		
High	160 (4.0)	45 (5.7)
Low	67 (1.7)	45 (5.7)
Normal	3822 (94.4)	702 (88.6)
Monocyte, n (%)		
High	951 (23.5)	163 (20.6)
Low	89 (2.2)	7 (0.9)
Normal	3009 (74.3)	622 (78.5)
Lymphocyte, n (%)		
High	41 (1.0)	3 (0.4)
Low	1263 (31.2)	218 (27.5)
Normal	2745 (67.8)	571 (72.1)
Neutrophil, n (%)		
High	996 (24.6)	125 (15.8)
Low	84 (2.1)	10 (1.3)
Normal	2969 (73.3)	657 (83.0)
Platelet, n (%)		
High	406 (10.0)	83 (10.5)
Low	459 (11.3)	29 (3.7)
Normal	3184 (78.6)	680 (85.9)
Thrombocytosis, n (%)		
No	3643 (90.0)	709 (89.5)
Yes	406 (10.0)	83 (10.5)

Patients selected from the RWD had a median age of 69 years, 70% had stage IV NSCLC at diagnosis, and 70% had nonsquamous histology (Table 1). In the validation cohort from the OAK trial, median age was 63 years, ECOG PS was 0/1, for all patients per the trial inclusion parameters, and 74% had nonsquamous histology. The majority of patients in the RWD and OAK cohorts had laboratory values in the normal range, except the majority of patients in both cohorts having low hemoglobin levels and most patients in OAK had high calcium and urea levels (Table 2).

### 3.2 | Prognostic model

The final model selected 18 variables (Figure 2). The top risk-increasing prognostic factors were abnormally low albumin and



**FIGURE 2** Summary of selected model coefficients. 1L, first line; ALP, alkaline phosphatase; ECOG PS, Eastern Cooperative Oncology Group performance status; LDH, lactate dehydrogenase; PD-L1, programmed death-ligand 1; SBP, systolic blood pressure; WBC, white blood cell

chloride levels, ECOG PS  $\geq 2$ , and abnormally high levels of ALP and white blood cells. The top risk-decreasing prognostic factors were PD-L1 positivity, longer time from advanced diagnosis to start of 1L, and higher systolic blood pressure.

### 3.3 | Model performance

When looking at the association of OS with the PI quartiles in both cohorts, OS was clearly differentiated among the PI quartiles (Figure 3). The c-index calculated for the RWD cohort OS curves was 0.67, and the c-index for the OAK trial validation cohort was 0.65. When using only ECOG PS to calculate the c-index for the RWD, it was 0.56. Performance was similar across the CPI molecules (atezolizumab, nivolumab, and pembrolizumab; Figure S1).

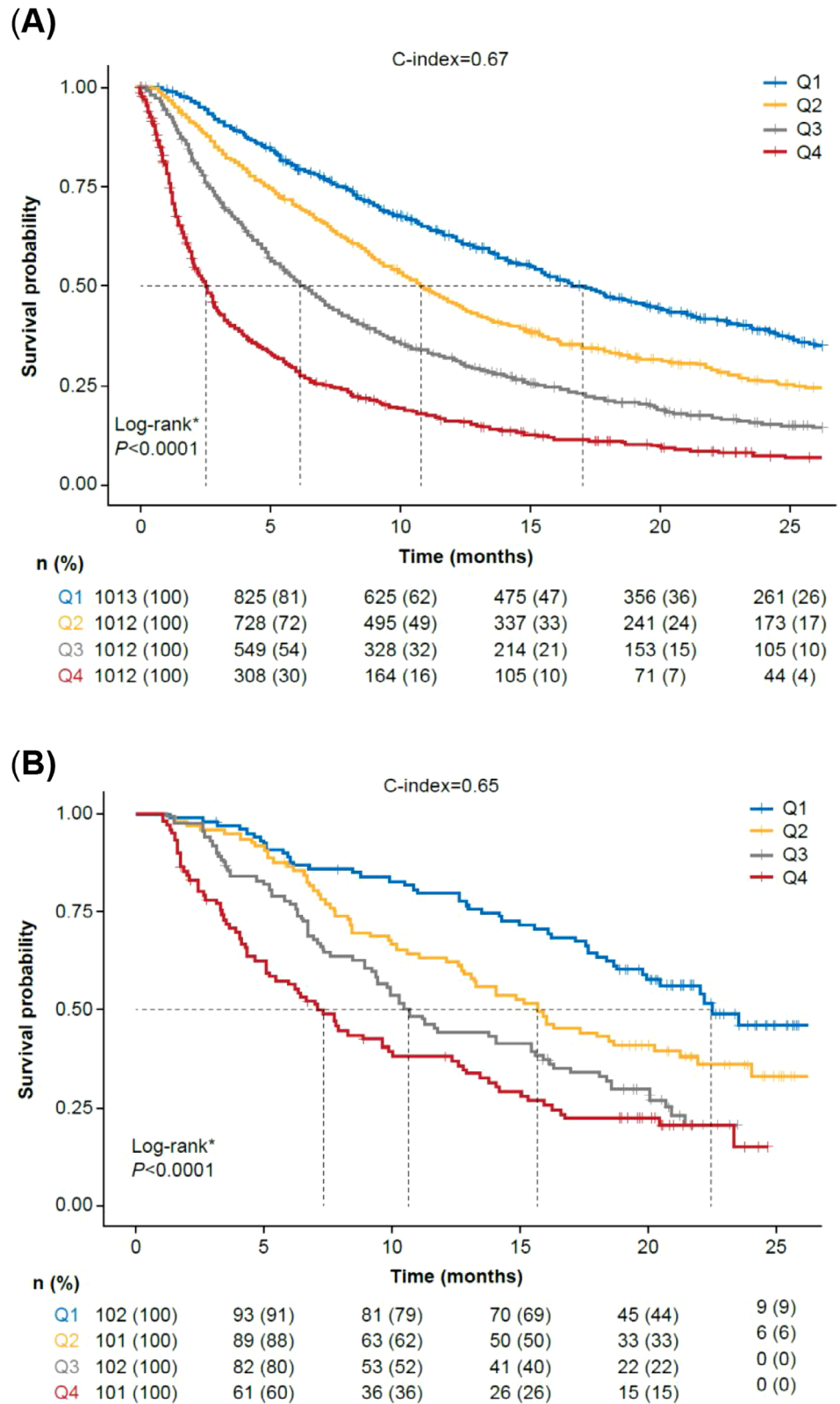
## 4 | DISCUSSION

To our knowledge, this is the first study developing a prognostic model in patients with aNSCLC receiving anti-PD-1/PD-L1 CPIs as 2L monotherapy. This study evaluated factors that may influence patient survival and found several baseline demographic and clinical characteristics of prognostic value. A Cox proportional hazards model with lasso regularization was used to estimate the prognostic variables. This method limits overfitting due to either collinearity of the covariates or high dimensionality. The model was then validated using data from the OAK clinical trial.

This approach of evaluating multiple prognostic factors was improved compared with only using single prognostic factors, such as ECOG PS. The c-index for ECOG PS was 0.56, whereas the c-index of the prognostic model that incorporated multiple variables was 0.67. Compared with ROPRO, this model specifically focused on patients with aNSCLC receiving anti-PD-1/PD-L1 CPIs as 2L monotherapy vs a more general, pooled cohort of 17 different tumor types who had received 1L treatment.<sup>9</sup> This model incorporated 18 variables of which most were not defined by units of measurement (Figure 2) versus 27 variables defined by specific units in ROPRO.<sup>9</sup> For laboratory values, levels were defined qualitatively (e.g., high vs. low) in this prognostic model compared with deriving a score based on the weighted sum of the patients' differences from the respective reference mean for each variable as performed in ROPRO.<sup>9</sup> Simplifying how variables were defined (units or not) and the total number required could make this prognostic model easier to implement in clinical practice.

Risk-decreasing factors included PD-L1 positivity, longer time to advanced diagnosis, and higher systolic blood pressure. Being PD-L1 positive was shown to decrease the risk, and patients who are PD-L1 positive are known to respond better to PD-1/PD-L1 therapies.<sup>19</sup> Recent research suggested that this could be a key factor for reducing cancer progression with immune checkpoint inhibition.<sup>20</sup> Longer time from advanced diagnosis to start of 1L treatment was recognized as a risk-decreasing factor. This finding should be interpreted with caution, as it is possible that the model assigned lower risk to patients having proved that they can survive for some time. This should not be interpreted as an argument for starting treatment later, because this is not a causal relationship. A higher systolic blood pressure was also observed as a risk-

**FIGURE 3** Association of prognostic index quartiles and overall survival in patients from (A) the real-world database (derivation data) and (B) the OAK trial (validation data). Q1 represents the highest PI and Q4 represents the lowest PI. \*The log-rank test evaluated the null hypothesis of no difference in survival between the different PI groups. PI, prognostic index



decreasing factor. Previous studies have evaluated the use of hypertension treatments, such as angiotensin-converting enzyme inhibitors, and the risk of cancer.<sup>21,22</sup> It could be hypothesized that systolic blood pressure and/or the use of hypertension treatments could be associated with cancer survival; although this has not been evaluated in the literature.

When evaluating prognostic factors that were risk increasing, ECOG PS  $\geq 2$  was associated with a worse PI. This finding was not surprising, as ECOG PS is a well-known prognostic factor of cancer and indicative of a patient's frailty, comorbidity, disease progression, and effects on daily life.<sup>23</sup> Laboratory characteristics that were strong risk-increasing



prognostic factors included high ALP and white blood cell counts, and abnormally low albumin and low chloride levels. Both ROPRO and this model identified high ALP as a risk-increasing prognostic factor, and ROPRO also noted that higher levels of albumin and chloride were protective, which aligns with this model's findings.<sup>9</sup> Other index models have also recognized high white blood cell counts and ALP levels as well as low albumin levels as prognostic risk factors.<sup>23-26</sup> A retrospective analysis with a multivariate logistic regression analysis indicated that ALP was an independent risk factor for bone metastases in patients with bladder cancer.<sup>27</sup> A retrospective study in Turkish patients with aNSCLC (stage IIIB) found that low pretreatment serum albumin level was an independent poor prognostic factor in patients with aNSCLC, which was associated with a reduced response to 1L therapy and decreased survival rates.<sup>28</sup> A systematic review also reported that low pretreatment serum albumin was associated with poor survival and could be used to define baseline patient risk.<sup>24</sup> Only limited studies have explained the low chloride findings in aNSCLC. In colorectal cancer, a retrospective, single-center study found that hypochloremia was associated with shorter OS and disease-free survival in patients with early- to later-stage disease who had resection.<sup>29</sup> Further research is warranted to understand the role of chloride levels in aNSCLC outcomes.

It is important to highlight that extracting clinically relevant data using electronic health records is a complex process and is a limitation of the retrospective nature of the present study. This database lacked complete information on other possible prognostic factors such as C-reactive protein, steroid use, tumor-infiltrating lymphocytes, pro-inflammatory cytokine levels, and T-cell (CD3+), cytotoxic T-cell, and memory T-cell counts. For laboratory-based data, values were categorized based on the reference range supplied by the laboratory in the real-world setting. These ranges can vary across laboratories, and accordingly, "cutoff" minimum and maximum values may differ across methods/patients. While laboratory-to-laboratory variability exists, this heterogeneity can allow more flexibility when implementing the prognostic model in populations in which different instruments, techniques, or reference ranges have been used across multiple clinical sites.

In summary, these analyses used a large cohort of patients with aNSCLC from a real-world setting to create a prognostic model, which was validated using clinical trial data. By using baseline demographic, clinical, and laboratory factors, this prognostic model could discriminate the risk of death among patients with aNSCLC receiving anti-PD1/PDL1 CPIs in 2L therapy. Further research into prognostic models to aid treatment selection is a critical element of ensuring optimal support for decision-making in real practice.

## ACKNOWLEDGMENTS

This work was sponsored by F Hoffmann-La Roche Ltd/Genentech, Inc. Support for third-party writing assistance, furnished by Health Interactions, Inc., was provided by F Hoffmann-La Roche Ltd.

## CONFLICT OF INTEREST

Cristina Julian, Pascal Chanu, Chris Harbron, Anda Gershon, Dominik Heinzmann, Wei Zou, and Valerie Quarmby are employees and shareholders of Genentech/Roche. Cristina Julian, Robson J. M. Machado,

Sandhya Girish, Pascal Chanu, Dominik Heinzmann, Chris Harbron, Valerie Quarmby, Qing Zhang, and Yachi Chen have a planned/issued/pending patent with Roche. Chris Harbron also owns stock in AstraZeneca. Shannon M. Pfeiffer received payment from Genentech for a 10-month internship. Yachi Chen, Robson J. M. Machado, and Sandhya Girish were employees of Genentech/Roche at the time of the study; Yachi Chen is a shareholder of Genentech/Roche. Qing Zhang is an employee of Genentech/Roche and owns stock in AbbVie, AC Immune, Bristol Myers Squibb, Regeneron, and Roche.

## AUTHOR CONTRIBUTIONS

All authors had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing-original draft, Writing-review & editing, C.J.;* *Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing-original draft, Writing-review & editing, R.J.M.M.;* *Conceptualization, Investigation, Project administration, Supervision, Visualization, Writing-original draft, Writing-review & editing, S.G.;* *Conceptualization, Investigation, Methodology, Project administration, Supervision, Visualization, Writing-original draft, Writing-review & editing, P.C.;* *Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Writing-original draft, Writing-review & editing, D.H.;* *Formal analysis, Methodology, Writing-original draft, Writing-review & editing, C.H.;* *Conceptualization, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing-original draft, Writing-review & editing, A.G.;* *Formal analysis, Methodology, Software, Writing-original draft, Writing-review & editing, S.M.P.;* *Conceptualization, Writing-original draft, Writing-review & editing, W.Z.;* *Writing-original draft, Writing-review & editing, V.Q.;* *Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing-original draft, Writing-review & editing, Q.Z.;* *Conceptualization, Methodology, Supervision, Writing-original draft, Writing-review & editing, Y.C.*

## ETHICS STATEMENT

Institutional review board approval of the study protocol was obtained prior to study conduct and included a waiver of informed consent. The data are deidentified and subject to obligations to prevent re-identification and protect patient confidentiality.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study have been originated by Flatiron Health, Inc. These deidentified data may be made available upon request and are subject to a license agreement with Flatiron Health; interested researchers should contact <dataaccess@flatiron.com> to determine licensing terms. Qualified researchers may request access to individual patient-level data through the clinical study data request platform (<https://vivli.org/>). Further details on Roche's criteria for eligible studies are available here (<https://vivli.org/members/>



ourmembers/). For further details on 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/research\\_and\\_development/who\\_we\\_are\\_how\\_we\\_work/clinical\\_trials/our\\_commitment\\_to\\_data\\_sharing.htm](https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm)).

## REFERENCES

- American Cancer Society. Cancer facts & figures 2020. Accessed April 5, 2021. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf>
- Hirsch FR, Scagliotti GV, Mulshine JL, et al. Lung cancer: current therapies and new targeted treatments. *Lancet*. 2017;389(10066):299–311. doi:10.1016/S0140-6736(16)30958-8
- Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N Engl J Med*. 2015;373(2):123–135. doi:10.1056/NEJMoa1504627
- Garon EB, Rizvi NA, Hui R, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. 2015;372(21):2018–2028. doi:10.1056/NEJMoa1501824
- Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med*. 2015;373(17):1627–1639. doi:10.1056/NEJMoa1507643
- Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2016;387(10027):1540–1550. doi:10.1016/S0140-6736(15)01281-7
- Rittmeyer A, Barlesi F, Waterkamp D, 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(10066):255–265. doi:10.1016/S0140-6736(16)32517-X
- Prelaj A, Tay R, Ferrara R, Chaput N, Besse B, Califano R. Predictive biomarkers of response for immune checkpoint inhibitors in non-small-cell lung cancer. *Eur J Cancer*. 2019;106:144–159. doi:10.1016/j.ejca.2018.11.002
- Becker T, Weberpals J, Jegg AM, et al. An enhanced prognostic score for overall survival of patients with cancer derived from a large real-world cohort. *Ann Oncol*. 2020;31(11):1561–1568. doi:10.1016/j.annonc.2020.07.013
- Ma X, Long L, Moon S, Adamson BJS, Baxi SS. Comparison of population characteristics in real-world clinical oncology databases in the US: flatiron health, SEER, and NPCR. *medRxiv*. Accessed April 5, 2020. <https://www.medrxiv.org/content/10.1101/2020.03.16.20037143v2.full-text>.
- Birnbaum B, Nussbaum N, Seidl-Rathkopf K, et al. 2001 Model-assisted cohort selection with bias analysis for generating large-scale cohorts from the EHR for oncology research. *ArXiv:2001.09765*. Accessed April 5, 2020. <https://arxiv.org/abs/2001.09765>
- Paesmans M. Prognostic and predictive factors for lung cancer. *Breathe*. 2012;9:112–121.
- Garde-Noguera J, Martin-Martorell P, De Julián M, et al. Predictive and prognostic clinical and pathological factors of nivolumab efficacy in non-small-cell lung cancer patients. *Clin Transl Oncol*. 2018;20(8):1072–1079. doi:10.1007/s12094-017-1829-5
- Lara MS, Brunson A, Wun T, et al. Predictors of survival for younger patients less than 50 years of age with non-small cell lung cancer (NSCLC): a California cancer registry analysis. *Lung Cancer*. 2014; 85(2):264–269. doi:10.1016/j.lungcan.2014.04.007
- Taniguchi Y, Tamiya A, Isa SI, et al. Predictive factors for poor progression-free survival in patients with non-small cell lung cancer treated with nivolumab. *Anticancer Res*. 2017;37(10):5857–5862. doi:10.21873/anticancer.12030
- Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw*. 2011;39(5):1–13. doi:10.18637/jss.v039.i05
- Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modeling strategies for improved prognostic prediction. *Stat Med*. 1984; 3(2):143–152. doi:10.1002/sim.4780030207
- White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med*. 2011;30(4): 377–399. doi:10.1002/sim.4067
- Xu Y, Wan B, Chen X, et al. The association of PD-L1 expression with the efficacy of anti-PD-1/PD-L1 immunotherapy and survival of non-small cell lung cancer patients: a meta-analysis of randomized controlled trials. *Transl Lung Cancer Res*. 2019;8(4):413–428. doi:10.21037/tlcr.2019.08.09
- Tseng JS, Yang TY, Wu CY, et al. Characteristics and predictive value of PD-L1 status in real-world non-small cell lung cancer patients. *J Immunother*. 2018;41(6):292–299. doi:10.1097/CJI.0000000000000226
- Yoon C, Yang HS, Jeon I, Chang Y, Park SM. Use of angiotensin-converting-enzyme inhibitors or angiotensin-receptor blockers and cancer risk: a meta-analysis of observational studies. *CMAJ*. 2011; 183(14):E1073–E1084. doi:10.1503/cmaj.101497
- Hicks BM, Filion KB, Yin H, Sakr L, Udell JA, Azoulay L. Angiotensin converting enzyme inhibitors and risk of lung cancer: population based cohort study. *BMJ*. 2018;363:k4209. doi:10.1136/bmj.k4209
- Ulas A, Turkoz FP, Silay K, et al. A laboratory prognostic index model for patients with advanced non-small cell lung cancer. *PLoS One*. 2014;9(12):e114471. doi:10.1371/journal.pone.0114471
- Gupta D, Lis CG. Pretreatment serum albumin as a predictor of cancer survival: a systematic review of the epidemiological literature. *Nutr J*. 2010;9:69. doi:10.1186/1475-2891-9-69
- Zhang K, Lai Y, Axelrod R, et al. Modeling the overall survival of patients with advanced-stage non-small cell lung cancer using data of routine laboratory tests. *Int J Cancer*. 2015;136(2):382–391. doi:10.1002/ijc.28995
- Sculier JP, Chansky K, Crowley JJ, Van Meerbeeck J, Goldstraw P, International Staging Committee and Participating Institutions. The impact of additional prognostic factors on survival and their relationship with the anatomical extent of disease expressed by the 6th edition of the TNM classification of malignant tumors and the proposals for the 7th edition. *J Thorac Oncol*. 2008;3(5):457–466. doi:10.1097/JTO.0b013e31816de2b8
- Huang P, Lan M, Peng AF, et al. Serum calcium, alkaline phosphatase and hemoglobin as risk factors for bone metastases in bladder cancer. *PLoS One*. 2017;12(9):e0183835. doi:10.1371/journal.pone.0183835
- Tanriverdi O, Avci N, Oktay E, et al. Pretreatment serum albumin level is an independent prognostic factor in patients with stage IIIB non-small cell lung cancer: a study of the Turkish descriptive oncological researches group. *Asian Pac J Cancer Prev*. 2015;16(14):5971–5976. doi:10.7314/apjcp.2015.16.14.5971
- Li Q, Dai W, Jia H, et al. Prognostic impact of hypochloremia in patients with stage I to III colorectal cancer after radical resection. *Dis Colon Rectum*. 2018;61(11):1273–1280. doi:10.1097/DCR.0000000000001196

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Julian C, Machado RJM, Girish S, et al. Real-world data prognostic model of overall survival in patients with advanced NSCLC receiving anti-PD-1/PD-L1 immune checkpoint inhibitors as second-line monotherapy. *Cancer Reports*. 2022;5(10):e1578. doi:10.1002/cnr2.1578