Original Article

Radiomic biomarkers from chest computed tomography are assistive in immunotherapy response prediction for non-small cell lung cancer

Kimberly E. Schroeder[1], Luna Acharya[2], Hariharasudan Mani[2], Muhammad Furqan[2,3], Jessica C. Sieren[1,3,4]

[1]Department of Radiology, University of Iowa, Iowa City, IA, USA; [2]Department of Internal Medicine, Hematology, Oncology and Blood and Marrow Transplantation, University of Iowa, Iowa City, IA, USA; [3]Holden Comprehensive Cancer Center, University of Iowa, Iowa City, IA, USA; [4]Roy J. Carver Department of Biomedical Engineering, University of Iowa, Iowa City, IA, USA

Contributions: (I) Concept and design: JC Sieren, M Furqan, KE Schroeder; (II) Administrative support: M Furqan, JC Sieren; (III) Provision of study materials or patients: KE Schroeder, L Acharya, H Mani, M Furqan; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: KE Schroeder, JC Sieren, M Furqan; (VI) Manuscript writing: All authors; (VII) Final approval of Manuscript: All authors.

Correspondence to: Jessica C. Sieren, PhD. Department of Radiology, University of Iowa, 200 Hawkins Dr. CC704GH, Iowa City, IA 52242, USA. Email: jessica-sieren@uiowa.edu.

Background: Immunotherapies, such as programmed death 1/programmed death ligand 1 (PD-1/PD-L1) antibodies have been shown to improve overall and progression-free survival (PFS) in patients with locally advanced or metastatic non-small cell lung cancer (NSCLC). However, not all patients derive a meaningful clinical benefit. Additionally, patients receiving anti-PD-1/PD-L1 therapy can experience immune-related adverse events (irAEs). Clinically significant irAEs may require temporary pause or discontinuation of treatment. Having a tool to identify patients who may not benefit and/or are at risk for developing severe irAEs from immunotherapy will aid in an informed decision-making process for the patients and their physicians.

Methods: Computed tomography (CT) scans and clinical data were retrospectively collected for this study to develop three prediction models using (I) radiomic features, (II) clinical features, and (III) radiomic and clinical features combined. Each subject had 6 clinical features and 849 radiomic features extracted. Selected features were run through an artificial neural network (NN) trained on 70% of the cohort, maintaining the case and control ratio. The NN was assessed by calculating the area-under-the-receiver-operating-characteristic curve (AUC-ROC), area-under-the-precision-recall curve (AUC-PR), sensitivity, and specificity.

Results: A cohort of 132 subjects, of which 43 (33%) had a PFS ≤90 days and 89 (67%) of which had a PFS >90 days was used to develop the prediction models. The radiomic model was able to predict progression-free survival with a training AUC-ROC of 87% and testing AUC-ROC, sensitivity, and specificity of 83%, 75%, and 81%, respectively. In this cohort, the clinical and radiomic combined features did add a slight increase in the specificity (85%) but with a decrease in sensitivity (75%) and AUC-ROC (81%).

Conclusions: Whole lung segmentation and feature extraction can identify those that would see a benefit from anti-PD-1/PD-L1 therapy.

Keywords: Therapy response; computed tomography; machine learning; progression-free survival; pneumonitis

Submitted Oct 24, 2022. Accepted for publication Apr 12, 2023. Published online May 11, 2023.
doi: 10.21037/tlcr-22-763
View this article at: https://dx.doi.org/10.21037/tlcr-22-763

Introduction

Lung cancer, of which non-small cell lung cancer (NSCLC) makes up about 80–85% of all new cases, is the leading cause of cancer-related deaths with a median survival of 13 months (1,2). While there have been recent improvements in early detection of lung cancer, the proportion of disease diagnosed at a localized stage is only 28% (1). The use of programmed death 1/programmed death ligand 1 (PD-1/PD-L1) immune checkpoint inhibitors (ICI) have been shown to improve the outcomes of patients with locally advanced and metastatic NSCLC (3-6).

Pembrolizumab (in KEYNOTE-010, 024, 189, 407 trials) (7-11), nivolumab (Checkmate 017 and Checkmate 057) (6,12), and atezolizumab (IMpower130, OAK trials) (13,14) have been investigated to treat metastatic NSCLC in 1st line and 2nd line settings as monotherapy or in combination with chemotherapy. In 1st line setting, pembrolizumab and atezolizumab when combined with chemotherapy have been shown to improve both progression-free survival (PFS) and overall survival compared to chemotherapy alone. In the KEYNOTE-407 study, the PFS benefit associated with pembrolizumab incrementally improved with increasing PD-L1 tumor proportion score (TPS). A similar trend was noted in the Impower130 trial except for patients with liver metastases and *EGFR/ALK* genomic alterations. In KEYNOTE-189, the hazard ratio for PFS was less than 1.00 across all subgroups of PD-L1 TPS, however, the upper boundary of the 95% confidence interval crossed 1.00 for patients with PD-L1 TPS <1% (9,10,13). In these 1st line clinical trials, where ICIs were combined with chemotherapy, 14–18% of patients developed progression indicating primary refractory disease to both ICIs and chemotherapy. Whereas studies evaluating ICIs as monotherapy have shown an objective response rate of ≤20% in a non-selective patient population (6,12,14) emphasizing the critical need for a biomarker. All the above studies, including KEYNOTE-024, support the use of PD-L1 TPS as a potential biomarker to select patients who are likely to benefit from ICI. However, spatial, and temporal heterogeneity of PD-L1 expression in the tumor along with variability in commercially available assays limits its use as a predictive biomarker (15).

Identifying patients that are unlikely to benefit from ICI before beginning treatment, can be useful in shared decision-making discussions where other downsides of immunotherapy can also be considered. Weighed into this decision is the high cost of treatment ($130,155/QALY) (16), and the risk of severe immune-related adverse events (irAEs) (17). Immunotherapy treatment-related adverse events occur in 64–73% of patients and grade 3 or higher adverse events occur in 7–27% of patients. (6,11,12,18). Pneumonitis is a particularly serious irAE, reported to occur in about 3–5% of patients (19), resulting in a high morbidity and mortality rate for those with a grade 3 or higher (20). Predicting patients who are at risk of developing pneumonitis with immunotherapy could aid in patient monitoring, including increased frequency of imaging.

Chest computed tomography (CT) images are the primary imaging modality used to diagnose, stage, and monitor lung cancer patients (21,22); hence, are readily available to use for treatment planning and the ICI selection decision. Prior work has been done using radiomic features extracted from the CT images to develop methods to predict the efficacy of PD-1/PD-L1 and the risk of irAEs in NSCLC (23-35). However, these previous studies have only looked at the lung lesion. This requires an experienced chest radiologist to create a 3D segmentation, or approve of the semi-automated segmentation, leading to a more time-consuming and subjective evaluation that is not as easily incorporated into the clinical setting. Prior work from our group has indicated that radiomic features from the surrounding lung parenchyma in addition to nodule features extracted from CT data are highly effective in distinguishing malignant from benign pulmonary lesions (36-38). Recently, we demonstrated that radiomic features

Highlight box

Key findings

- Radiomic features in conjunction with machine learning have utility in the prediction of progression free survival with immunotherapy in patients with NSCLC.

What is known and what is new?

- Prior research has indicated radiomics from segmented NSCLC lesions can be insightful for immunotherapy response prediction.
- This study demonstrates radiomics from the whole lung parenchyma, instead of the NSCLC lesion, are a useful for immunotherapy response prediction.

What is the implication, and what should change now?

- The use of radiomics derived from automated whole lung segmentations could be used to aid in immunotherapy treatment planning without requiring the expertise and/or time to define and segment tumor boundaries.
- This approach is advantageous for automated incorporation into the clinical workflow.

from automated whole lung segmentation could be used for classification of malignant and benign pulmonary nodules, without requiring specialized segmentation and feature extraction from the lesion (39). In this study, a neural network framework achieved high performance in training and testing (AUC 0.77 and 0.79 respectively), while in comparison, the least absolute shrinkage and selection operator (LASSO) approach had less stable testing performance (training AUC 0.80, *vs.* testing AUC 0.62) (39). Radiomic features from the whole lung quantify markers of lung disease such as chronic obstructive pulmonary disease (COPD). A current review article by Lin *et al* highlighted potential prognostic benefit of ICI treatment in patients with lung cancer and COPD (40), and Noda *et al.* found the Goddard scoring system for severity of emphysema useful in predicting PFS in NSCLC treated with ICI (41). Based on these prior studies, we hypothesize that quantified radiomic features from the whole lung could be useful in predicting response to ICI in NSCLC, without requiring labor intensive image segmentation of every pulmonary lesion.

This study explores the use of selected radiomic features extracted from the whole lung parenchymal tissue from 'baseline' CT images to develop a method to predict progression free survival benefit from immunotherapy. We also explore the utility of these radiomic features to predict the development of pneumonitis irAE. We present the following article in accordance with the TRIPOD reporting checklist (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-763/rc).

Methods

Subjects

For this retrospective study, patients with stage IV NSCLC who received at least one treatment of an immunotherapy agent between January 1, 2013, and March 13, 2021, at the University of Iowa Hospitals and Clinics were reviewed. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013), and approved by the University of Iowa Hospitals and Clinics Institutional Review Board (#202004142). Individual consent for this retrospective analysis was waived with approval from the Institutional Review Board. The study sample included all eligible patients within the retrospective study window with appropriate imaging and outcome data available. Four-fold cross validation was used. In each fold, the cohort was randomly split into 70/30 training and testing sets maintaining the proportion of cases to controls as the entire cohort. *Figure 1* provides an overview of the developed image analysis feature extraction and classification pipeline. Two prediction models were explored by dividing the cohort based on: (I) PFS calculated as the number of days from the start of immunotherapy to progression or death. Those without a recorded date of death or progression were calculated based on the last known date alive. Cases were defined as those with a PFS of less than or equal to 90 days. (II) Pneumonitis diagnosis in the clinical record during or after immunotherapy (cases with pneumonitis *vs.* controls without pneumonitis). REDCap electronic data capture was utilized to control access to patient identifiers and class outcomes.

Clinical features

Demographic, diagnostic, treatment, and clinical features were collected through chart review and managed using REDCap electronic data capture tool (42). These features included age, sex, race, smoking history, use of immunotherapy with/without concurrent chemotherapy, history of radiation treatment to the chest, baseline laboratory results (i.e., lymphocyte count, PD-L1 expression), PFS, and diagnosed irAEs such as pneumonitis (grade ≥2 per Common Terminology Criteria for Adverse Events v4.03).

Radiomic features

Existing standard-of-care CT scans that included the complete chest within ±150 days from the start of immunotherapy treatment were included in the image collection data. Automated lung segmentation was performed using Pulmonary Analysis Software Suite (PASS). PASS extracted 19 histogram features from the lung parenchyma. An additional 11 first-order features, 24 grey level co-occurrence matrix (GLCM) features, 14 gray level dependence matrix (GLDM) features, 16 gray level run length matrix (GLRLM) features, 16 gray level size zone matrix (GLSZM) features, and 5 neighboring gray-tone difference matrix (NGTDM) features were extracted from the segmented lung parenchyma using PyRadiomics (v3.0.1) with default parameters (see Appendix 1 for parameter details) (43). A wavelet filter was applied to the original image to obtain eight wavelet decompositions for which the above-mentioned PyRadiomic features were also pulled with an addition of 7 first-order features.

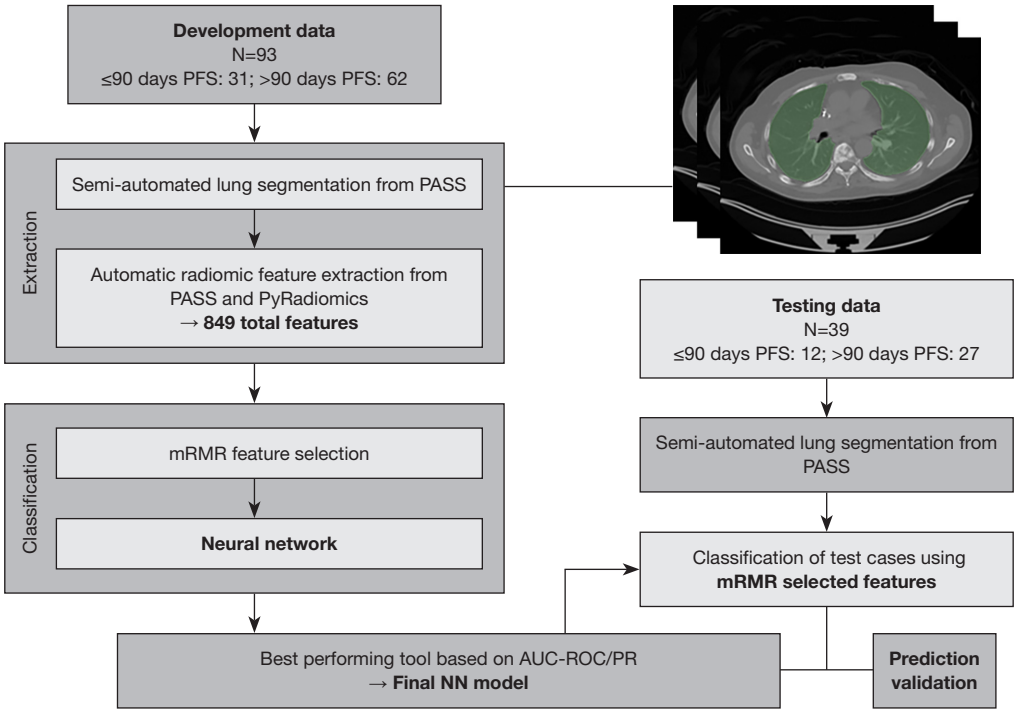

Figure 1 Pipeline outline for the development of the PFS prediction model. PFS, progression-free survival; PASS, Pulmonary Analysis Software Suite; mRMR, minimum Redundancy Maximum Relevance; AUC, area under the curve; ROC, receiver operating characteristic; PR, precision recall; NN, neural network.

A total of 849 radiomic features were extracted for each segmentation.

Neural network (NN)

Four cross validation sets were created from the data, each with a 70/30 training and testing split that maintained the case *vs.* control proportions. Using the 70% training set, minimum Redundancy Maximum Relevance (mRMR) (44), feature reduction and selection was performed with the R package for parallelized mRMR ensemble from De Jay *et al.* (45). Three prediction models were developed using (I) radiomic features, (II) clinical features, (III) radiomic and clinical (radiomic + clinical) features combined. The selected features were then run through an artificial NN trained on the 70% training cohort. Each epoch left out 10% of the subjects to validate the NN before updating and refining the weights. The receiver operating characteristic (ROC) or precision recall (PR) in conjunction with Youden's index was used to determine the optimum dichotomous cut-off point for the NN output (which ranges from 0–1). The remaining 30% of subjects were then used to test the trained NN, in which the Youden's index from training was used to categorize test subjects based on the NN output value. The classification was explored for two use cases: (I) progression-free survival prediction (cases with ≤90 days PFS, *vs.* controls with >90 days PFS) and (II) pneumonitis development during immunotherapy (cases with pneumonitis *vs.* controls without pneumonitis).

Statistical analysis

Demographic data were analyzed using GraphPad Prism (version 8.3.1). Normality was tested using the Shapiro-Wilk test. Unpaired *t*-tests with Welch's correction or Mann-Whitney test were used for parametric or non-parametric data, respectively. Chi-squared test was used for discrete variables. As inclusion in the study required appropriate image data availability, model construction (feature reduction, selection and NN) utilized complete case analysis. The NN performance was assessed using the area-under-the-receiver-operating-characteristic curve (AUC-ROC), area-under-the-precision-recall curve (AUC-PR), sensitivity, and specificity. Confidence intervals were

Table 1 Clinical variables

Parameter	All	≤90 days PFS (cases)	>90 days PFS (controls)	P value
N	132	43	89	–
Age, years (mean ± SD)	64.1±10.4	63.4±9.7	63.8±12.6	0.5743
Sex (Male:Female)	67:65	24:19	43:46	0.4193
Race (White: African American/Black:Asian:Declined)	126:3:2:1	41:1:1:0	85:2:1:1	0.8589
Pack years (>20 yrs:<20 yrs:never smoker)	100:16:16	34:4:5	66:12:11	0.7698
Immunotherapy type (Pemb:Nivo:Avel)	97:30:5	27:14:2	70:16:3	0.1473
Programmed death ligand 1 (PD-L1) expression (<1%:1–49%:≥50%:Unknown)	53:29:49:11	22:6:9:6	31:23:30:5	0.0501
Immunotherapy concurrent with chemotherapy (Yes:No)	28:104	6:37	18:71	0.3813
Lymphocyte count (per μL) (mean ± SD)	1,921±8,168	940±647	2,395±9,921	0.0147
Pneumonitis (Yes:No)	15:117	1:42	14:75	0.0230
Line of treatment (Naïve:2nd Line:3rd Line:4th Line)	77:44:7:4	20:17:4:2	57:27:3:2	0.1918
Prior chest radiation (Yes:No)	54:78	18:25	36:53	0.8772
Days PFS (mean ± SD)	288.6±317.5	47.7±23.0	404.5±326.7	<0.0001

PFS, progression-free survival; SD, standard deviation; Pemb, pembrolizumab; Nivo, nivolumab; Avel, avelumab; yrs, years.

calculated using Student's *t* distribution.

Results

A total of 132 subjects were included in the study; 43 (33%) of which had a PFS less than or equal to 90 days and 89 (67%) of which had a PFS greater than 90 days. Between the two groups, statistical difference was found for lymphocyte count (P=0.0147) and pneumonitis diagnosis (P=0.0230) as seen in *Table 1*. PD-L1 expression and acquisition results are further described in Table S1. *Table 2* describes the CT variables for the cohort. No statistical differences were found between the two PFS groups. The clinical demographics for the pneumonitis classification are presented in Table S2.

Automated segmentation of the chest CT data to identify the lung parenchyma with PASS was successful for all cases. It was found that the automated algorithm excluded lesions and consolidation attached to the chest wall. All other pulmonary lesions without attachment to the chest wall were included in the segmentation. To explore the performance of the algorithm without requiring manual interaction, no editing of the automated lung segmentation was performed before feature extraction.

To prevent overfitting of the PFS prediction model, based on the number of cases in our study, the selected features were restricted to a maximum of 9 features (1 feature per 5 cases). The clinical features model incorporated age, sex, pack-years, prior chest radiation, concurrent chemotherapy, lymphocyte count, and PD-L1 expression. *Table 3* highlights the individual features that were selected in the radiomic and radiomic + clinical prediction models. From the features included in the prediction models, 4 features were statistically significant between the groups: wavelet.LLL_glszm_HighGrayLevelZoneEmphasis (P=0.0005), wavelet.HHL_glcm_ClusterShade (P=0.0134), wavelet.HHH_firstorder_Skewness (P=0.0368), and lymphocyteCount (P=0.0147). The three significant radiomic features were selected in both the radiomic and radiomic + clinical models as well as the other 4 non-statistically significant radiomic features. Table S3 details the equivalent feature table for the pneumonitis prediction model. Cross validation cohorts also chose majority wavelet features. Seven wavelet features were selected by at least two of the folds: wavelet.LLL_glszm_HighGrayLevelZoneEmphasis, wavelet.HHH_firstorder_Skewness, wavelet.LLL_firstorder_Minimum, wavelet.HHH_glrlm_GrayLevelVariance, wavelet.HHH_firstorder_Mean, wavelet.HHH_glszm_LargeAreaEmphasis, wavelet.HLH_glszm_LargeAreaHighGrayLevelEmphasis.

The PFS prediction models' performances are listed in

Table 2 CT characteristics

Parameter	All	≤90 days PFS (cases)	>90 days PFS (controls)	P value
N	132	43	89	–
Days ± from start of immunotherapy and CT (mean ± SD)	25.11±37.54	16.74±14.36	29.15±44.14	0.1520
Reconstruction Kernel (B31f:I31f:Bf40:B30f:I30f:Qr40d)	23:50:26:28:2:3	5:14:9:11:1:3	18:36:17:17:1:0	0.1167
Contrast enhanced (Yes:No)	121:11	40:3	81:8	0.6951
Slice thickness, mm (mean ± SD)	2.967±0.24	2.90±0.42	3±0	0.1044
Voltage, kV (mean ± SD)	109.84±12.72	111.62±14.62	108.98±11.68	0.5056
Tube current, mA (mean ± SD)	454.007±269.84	414.06±291.11	473.30±258.43	0.1629
Siemens Scanner Model (Definition AS+:Definition AS:Force:Biograph 40:Biograph 64)	17:56:29:13:17	5:14:12:4:8	12:42:17:9:9	0.3787

PFS, progression-free survival; CT, computed tomography; SD, standard deviation; kV, kilovoltage; mA, milliampere.

Table 3 Feature reduction and selection results used for PFS prediction model

Feature	Radiomic + Clinical	Radiomic	≤90 days PFS	>90 days PFS	P value
wavelet.LLL_glszm_HighGrayLevelZoneEmphasis	X	X	2,446±497.1	2,062±695.3	0.0005
wavelet.HHH_firstorder_TotalEnergy	X	X	$1.48\times10^{8}\pm9.64\times10^{8}$	$1.94\times10^{8}\pm1.62\times10^{8}$	0.1805
wavelet.HHL_glcm_ClusterShade	X	X	$-9.55\times10^{-2}\pm1.52\times10^{-1}$	$-4.01\times10^{-2}\pm1.05\times10^{-1}$	0.0134
wavelet.LLH_ngtdm_Strength	X	X	$2.24\times10^{-2}\pm1.69\times10^{-2}$	$3.91\times10^{-2}\pm5.62\times10^{-2}$	0.1019
wavelet.LLH_glszm_ZoneVariance	X	X	$3.48\times10^{6}\pm8.77\times10^{6}$	$1.02\times10^{6}\pm9.74\times10^{5}$	0.1205
wavelet.HHH_firstorder_Skewness	X	X	$5.83\times10^{-2}\pm1.14\times10^{-1}$	$1.69\times10^{-2}\pm8.94\times10^{-2}$	0.0368
concurrentChemo (Yes:No)	X		6:37	18:71	0.3813
lymphocyteCount	X		940±646.5	2,395±9,921	0.0147
wavelet.LLL_firstorder_Minimum	X	X	–2,894±77.32	–2,869±107.4	0.1927
wavelet.HHL_glszm_LargeAreaEmphasis		X	$1.208\times10^{8}\pm1.103\times10^{8}$	$2.012\times10^{8}\pm4.879\times10^{8}$	0.2386
wavelet.HHH_firstorder_Median		X	$-8.39\times10^{-3}\pm1.09\times10^{-2}$	$-7.78\times10^{-3}\pm1.264\times10^{-2}$	0.7725

Data are shown as mean ± SD. PFS, progression-free survival; SD, standard deviation.

Table 4. The best performance was seen in the radiomic model with a testing AUC-ROC, sensitivity, and specificity of 83% (95% CI: 65–101%), 83%, and 81%, respectively (Figure S1). Two clinical models were run, one with PD-L1 expression and one without. The model with PD-L1 expression resulted in a testing AUC-ROC of 44% (95% CI: 38–50%) *vs.* 47% (95% CI: 34–60%) in the model without and sensitivity and specificity of 17% *vs.* 50% and 81% *vs.* 56%, respectively. The pneumonitis prediction model results are featured in Table S4, for which the best performing model was radiomic + clinical with a sensitivity of 75% and specificity of 71% however, there was a large drop in AUC-PR from training (60%; 95% CI: 42–78%) to testing (24%; 95% CI: 11–37%), indicating instability.

Discussion

In this study, it was observed that baseline chest CT imaging features from the whole lung parenchyma can predict PFS in patients receiving ICI, perhaps better than clinical features alone. The PFS prediction model using only radiomic features was able to predict patients that

Table 4 Performance results from the developed PFS prediction models using the image and/or clinical features.

PFS prediction model	Training AUC (95% CI)	Testing AUC (95% CI)	Sensitivity	Specificity
Radiomic + clinical without PD-L1	0.86 (0.82–0.90)	0.81 (0.61–1.01)	0.75	0.85
Radiomic	0.87 (0.85–0.89)	0.83 (0.65–1.01)	0.83	0.81
Clinical without PD-L1	0.65 (0.58–0.72)	0.47 (0.34–0.60)	0.50	0.56
Clinical with PD-L1	0.63 (0.61–0.65)	0.44 (0.38–0.50)	0.17	0.81

PFS, progression-free survival; AUC, area under the curve; PD-L1, programmed death ligand 1 expression.

did see a meaningful benefit from immunotherapy in the form of >90 days PFS (81% specificity) and was slightly more effective in identifying those with ≤90 days PFS (83% sensitivity). The addition of clinical features did add a slight increase in the ability to predict those with a beneficial outcome (85% specificity) but with a decrease in the ability to predict unbeneficial outcomes (75% sensitivity).

Due to the small number of patients with pneumonitis (N=15), it is difficult to train a model that can effectively predict this important irAE. The developed pneumonitis prediction models were not stable in their prediction performance as demonstrated by the drop from a training AUC-PR of 63% to a testing AUC-PR of only 24%. Unlike the PFS prediction models, the addition of clinical features to the radiomic features resulted in an increased ability to predict those who developed pneumonitis (50% to 75% sensitivity) and decreased ability to predict those who did not (86% to 71% specificity).

As highlighted in the introduction, some prior results have been published with a focus on radiomics features for predicting immunotherapy outcomes. A unique aspect of the work in this paper is that the feature extraction was performed from the whole lung parenchyma as opposed to the segmented pulmonary lesions. Whole lung segmentation and feature extraction does not require an expert user identification and segmentation of pulmonary lesions. Hence, our approach is highly suited to background processing of the data without expert input and may be more suitable for clinical workflow adoption. Despite the difference in the area from which features were extracted, the incorporation of wavelet decomposition features has been found useful in our study and others for survival prediction of lung cancer patients (46). The inclusion of wavelet features has also been used to predict outcomes in immunotherapy by other studies (25,27,28,35). Wavelets were the most selected radiomic features for both PFS and pneumonitis prediction, reinforcing them as a non-invasive, informative radiomic marker for immunotherapy efficacy.

Development of immune-related pneumonitis in NSCLC has been associated with an increase in the efficacy of immunotherapy (23), however, some studies (24,25) found that the development of pneumonitis in NSCLC patients is linked to significantly lower survival. Concurrent chemotherapy, chest radiation, or history of prior chest radiation, pre-existing interstitial lung disease, use of tyrosine kinase inhibitor before or in combination with ICI, and peripheral blood eosinophilia are associated with increased risk of ICI-related pneumonitis (47).

Lymphocyte count is one clinical feature that has been associated with OS (48); in this cohort, it was found to be statistically different between the two PFS groups and was selected by the mRMR feature selection method for the PFS prediction model (Radiomic + Clinical) but not for the pneumonitis prediction model (Radiomic + Clinical). Another clinical feature, PD-L1 TPS, was found to have no statistically significant difference and adding PD-L1 TPS to the Clinical model did not increase its performance. Though PD-L1 TPS is the only Food and Drug Administration-approved biomarker for ICI therapy in advanced NSCLC, it is not an ideal biomarker. Besides the inherent issues in using PD-L1 TPS as a biomarker i.e., spatial, and temporal tumor heterogeneity in PD-L1 expression, our study may have additional limitations explaining why PD-L1 TPS did not help predict patients' outcomes. Determination of PD-L1 TPS in the majority of our patients (n=92, 70%) was made by utilizing E1L3N antibody. In addition, 31% of the samples that were used to assess PD-L1 TPS were obtained via fine-needle aspirates (E1L3N =29, 22C3 =12).

Additional limitations of this study should be acknowledged. First, this was a retrospective study with data collected from only a single institution leading to selection biases. Second, this study had a small sample size. The proportion of positive pneumonitis cases (1.1%) in this cohort is less than what has been seen in others.

The promising results from this small, single institution study indicate the potential of this approach and support future investigation including multi-center data and an external validation cohort. Lastly, to obtain the maximum study cohort, restrictions on the inclusion of CT images had to be minimal. This led to diverse CT parameters and acquisitions. About 10% of the >90 days PFS subjects' scans and about 13% of <90 days PFS subjects' scans were non-contrast. The PFS prediction models which included radiomics features (Radiomic and Radiomic + Clinical) both correctly predicted all non-contrast scans. This suggests that the radiomics features selected by the prediction models are not heavily influenced by the presence or absence of contrast enhancement, such that either CT scan data can be applied to the prediction model. Of the diverse acquisitions, only 6 of the 30 PET/CT scans were misclassified by the PFS prediction models (Radiomic and/or Radiomic + Clinical). The number of days between the start of immunotherapy and the CT scan used for analysis was not statistically different for the PFS groups, despite differences in the ranges (>90 days PFS subjects had an SD of 44 days, ≤90 days PFS subjects had SD of 14 days). Again, this suggests a robust prediction model that can be used effectively in most cases.

Conclusions

To the authors' knowledge, this is the first immunotherapy efficacy prediction model to have used whole-lung parenchymal features, without requiring segmentation of pulmonary lesions. Future work could see an increase in performance with the addition of lesion features. As the use of immunotherapy increases, the selection of patients that would see a clinically relevant benefit over traditional treatments needs to be defined. For patients and physicians, the cost-effectiveness of this type of treatment, risk of severe irAEs, and survival benefit are important factors to consider in treatment planning. While PD-L1 TPS is clinically utilized, diversity in the method and sample source for this biomarker in the real-world setting may limit its predictive value, a gap that can be filled by more readily available and validated radiomics biomarkers.

Acknowledgments

Funding: This study was supported in part by the National Institute of Health (UL1TR002537, R01CA267820 and P30CA086962) along with internal research funds from the University of Iowa.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-763/rc

Data Sharing Statement: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-763/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-763/coif). All authors report that this work was supported in part by the National Institute of Health (UL1TR002537, R01CA267820 and P30CA086962) along with internal research funds from the University of Iowa. JCS reports that she serves as a study section chair in the National Institute of Health and receives payment from the National Institute of Health. Her husband has stocks and paid consultant for VIDA Diagnostics (this company was not involved in any way in this study). The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). It was approved by the University of Iowa Hospitals and Clinics Institutional Review Board (#202004142) and individual consent for this retrospective analysis was waived with approval from the Institutional Review Board.

References

1. Siegel RL, Miller KD, Fuchs HE, et al. Cancer statistics, 2022. CA Cancer J Clin 2022;72:7-33.
2. Key Statistics for Lung Cancer. American Cancer Society. 2023. Available online: https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html. 2023.
3. 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.
4. Hellmann MD, Ciuleanu TE, Pluzanski A, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. N Engl J Med 2018;378:2093-104.
5. Gandhi L, Rodríguez-Abreu D, Gadgeel S, et al. Pembrolizumab plus Chemotherapy in Metastatic Non-Small-Cell Lung Cancer. N Engl J Med 2018;378:2078-92.
6. 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;373:123-35.
7. Brahmer JR, Rodríguez-Abreu D, Robinson AG, et al. Health-related quality-of-life results for pembrolizumab versus chemotherapy in advanced, PD-L1-positive NSCLC (KEYNOTE-024): a multicentre, international, randomised, open-label phase 3 trial. Lancet Oncol 2017;18:1600-9.
8. 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;387:1540-50.
9. Gadgeel S, Rodríguez-Abreu D, Speranza G, et al. Updated Analysis From KEYNOTE-189: Pembrolizumab or Placebo Plus Pemetrexed and Platinum for Previously Untreated Metastatic Nonsquamous Non-Small-Cell Lung Cancer. J Clin Oncol 2020;38:1505-17.
10. Paz-Ares L, Vicente D, Tafreshi A, et al. A Randomized, Placebo-Controlled Trial of Pembrolizumab Plus Chemotherapy in Patients With Metastatic Squamous NSCLC: Protocol-Specified Final Analysis of KEYNOTE-407. J Thorac Oncol 2020;15:1657-69.
11. Reck M, Rodríguez-Abreu D, Robinson AG, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. N Engl J Med 2016;375:1823-33.
12. 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;373:1627-39.
13. 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.
14. Mazieres J, Rittmeyer A, Gadgeel S, et al. Atezolizumab Versus Docetaxel in Pretreated Patients With NSCLC: Final Results From the Randomized Phase 2 POPLAR and Phase 3 OAK Clinical Trials. J Thorac Oncol 2021;16:140-50.
15. Hong L, Negrao MV, Dibaj SS, et al. Programmed Death-Ligand 1 Heterogeneity and Its Impact on Benefit From Immune Checkpoint Inhibitors in NSCLC. J Thorac Oncol 2020;15:1449-59.
16. Huang M, Lopes GL, Insinga RP, et al. Cost-effectiveness of pembrolizumab versus chemotherapy as first-line treatment in PD-L1-positive advanced non-small-cell lung cancer in the USA. Immunotherapy 2019;11:1463-78.
17. Lynch TJ, Bondarenko I, Luft A, et al. Ipilimumab in combination with paclitaxel and carboplatin as first-line treatment in stage IIIB/IV non-small-cell lung cancer: results from a randomized, double-blind, multicenter phase II study. J Clin Oncol 2012;30:2046-54.
18. Park K, Özgüroğlu M, Vansteenkiste J, et al. Avelumab Versus Docetaxel in Patients With Platinum-Treated Advanced NSCLC: 2-Year Follow-Up From the JAVELIN Lung 200 Phase 3 Trial. J Thorac Oncol 2021;16:1369-78.
19. Khunger M, Rakshit S, Pasupuleti V, et al. Incidence of Pneumonitis With Use of Programmed Death 1 and Programmed Death-Ligand 1 Inhibitors in Non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis of Trials. Chest 2017;152:271-81.
20. Suresh K, Voong KR, Shankar B, et al. Pneumonitis in Non-Small Cell Lung Cancer Patients Receiving Immune Checkpoint Immunotherapy: Incidence and Risk Factors. J Thorac Oncol 2018;13:1930-9.
21. Tsim S, O'Dowd CA, Milroy R, et al. Staging of non-small cell lung cancer (NSCLC): a review. Respir Med 2010;104:1767-74.
22. Tanoue LT. Staging of non-small cell lung cancer. Semin Respir Crit Care Med 2008;29:248-60.
23. He B, Dong D, She Y, et al. Predicting response to immunotherapy in advanced non-small-cell lung cancer using tumor mutational burden radiomic biomarker. J Immunother Cancer 2020;8:e000550.
24. Liu Y, Wu M, Zhang Y, et al. Imaging Biomarkers to

Predict and Evaluate the Effectiveness of Immunotherapy in Advanced Non-Small-Cell Lung Cancer. Front Oncol 2021;11:657615.
25. Mu W, Tunali I, Qi J, et al. Radiomics of (18)F Fluorodeoxyglucose PET/CT Images Predicts Severe Immune-related Adverse Events in Patients with NSCLC. Radiol Artif Intell 2020;2:e190063.
26. Nardone V, Tini P, Pastina P, et al. Radiomics predicts survival of patients with advanced non-small cell lung cancer undergoing PD-1 blockade using Nivolumab. Oncol Lett 2020;19:1559-66.
27. Shen L, Fu H, Tao G, et al. Pre-Immunotherapy Contrast-Enhanced CT Texture-Based Classification: A Useful Approach to Non-Small Cell Lung Cancer Immunotherapy Efficacy Prediction. Front Oncol 2021;11:591106.
28. Trebeschi S, Drago SG, Birkbak NJ, et al. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. Ann Oncol 2019;30:998-1004.
29. Tunali I, Gray JE, Qi J, et al. Novel clinical and radiomic predictors of rapid disease progression phenotypes among lung cancer patients treated with immunotherapy: An early report. Lung Cancer 2019;129:75-9.
30. Yang B, Zhou L, Zhong J, et al. Combination of computed tomography imaging-based radiomics and clinicopathological characteristics for predicting the clinical benefits of immune checkpoint inhibitors in lung cancer. Respir Res 2021;22:189.
31. Yang Y, Yang J, Shen L, et al. A multi-omics-based serial deep learning approach to predict clinical outcomes of single-agent anti-PD-1/PD-L1 immunotherapy in advanced stage non-small-cell lung cancer. Am J Transl Res 2021;13:743-56.
32. Khorrami M, Prasanna P, Gupta A, et al. Changes in CT Radiomic Features Associated with Lymphocyte Distribution Predict Overall Survival and Response to Immunotherapy in Non-Small Cell Lung Cancer. Cancer Immunol Res 2020;8:108-19.
33. Xie D, Xu F, Zhu W, et al. Delta radiomics model for the prediction of progression-free survival time in advanced non-small-cell lung cancer patients after immunotherapy. Front Oncol 2022;12:990608.
34. Lin Q, Wu HJ, Song QS, et al. CT-based radiomics in predicting pathological response in non-small cell lung cancer patients receiving neoadjuvant immunotherapy. Front Oncol 2022;12:937277.
35. Tankyevych O, Trousset F, Latappy C, et al. Development of Radiomic-Based Model to Predict Clinical Outcomes in Non-Small Cell Lung Cancer Patients Treated with Immunotherapy. Cancers (Basel) 2022;14:5931.
36. Uthoff J, Stephens MJ, Newell JD Jr, et al. Machine learning approach for distinguishing malignant and benign lung nodules utilizing standardized perinodular parenchymal features from CT. Med Phys 2019;46:3207-16.
37. Uthoff J, Nagpal P, Sanchez R, et al. Differentiation of non-small cell lung cancer and histoplasmosis pulmonary nodules: insights from radiomics model performance compared with clinician observers. Transl Lung Cancer Res 2019;8:979-88.
38. Dilger SK, Uthoff J, Judisch A, et al. Improved pulmonary nodule classification utilizing quantitative lung parenchyma features. J Med Imaging (Bellingham) 2015;2:041004.
39. Uthoff JM, Mott SL, Larson J, et al. Computed Tomography Features of Lung Structure Have Utility for Differentiating Malignant and Benign Pulmonary Nodules. Chronic Obstr Pulm Dis 2022;9:154-64.
40. Lin M, Huang Z, Chen Y, et al. Lung cancer patients with chronic obstructive pulmonary disease benefit from anti-PD-1/PD-L1 therapy. Front Immunol 2022;13:1038715.
41. Noda Y, Shiroyama T, Masuhiro K, et al. Quantitative evaluation of emphysema for predicting immunotherapy response in patients with advanced non-small-cell lung cancer. Sci Rep 2022;12:8881.
42. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377-81.
43. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res 2017;77:e104-7.
44. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. J Bioinform Comput Biol 2005;3:185-205.
45. De Jay N, Papillon-Cavanagh S, Olsen C, et al. mRMRe: an R package for parallelized mRMR ensemble feature selection. Bioinformatics 2013;29:2365-8.
46. Soufi M, Arimura H, Nagami N. Identification of optimal mother wavelets in survival prediction of lung cancer patients using wavelet decomposition-based radiomic features. Med Phys 2018;45:5116-28.
47. Yin J, Wu Y, Yang X, et al. Checkpoint Inhibitor Pneumonitis Induced by Anti-PD-1/PD-L1 Therapy in Non-Small-Cell Lung Cancer: Occurrence and

Mechanism. Front Immunol 2022;13:830631.
48. Li F, Li C, Cai X, et al. The association between CD8+ tumor-infiltrating lymphocytes and the clinical outcome of cancer immunotherapy: A systematic review and meta-analysis. EClinicalMedicine 2021;41:101134.

Cite this article as: Schroeder KE, Acharya L, Mani H, Furqan M, Sieren JC. Radiomic biomarkers from chest computed tomography are assistive in immunotherapy response prediction for non-small cell lung cancer. Transl Lung Cancer Res 2023;12(5):1023-1033. doi: 10.21037/tlcr-22-763

Supplementary

Appendix 1

PyRadiomics settings were specified in the parameter file which used the default binWidths = 25 for the original image and the wavelet decomposition extraction. No kernel, image normalization or resampling options were used. Images were filtered through all the ["coif1"] wavelet decomposition levels (LLL, LLH, LHL, HLL, HLH, HHL, & HHH) with the default settings for 'start_level' and 'level'. For both the original and wavelet images the following feature classes were included: first order, shape, GLCM, GLRLM, GLSZM, GLDM, and NGTDM. Shape features were excluded in feature selection/reduction process since the whole lung was segmented.

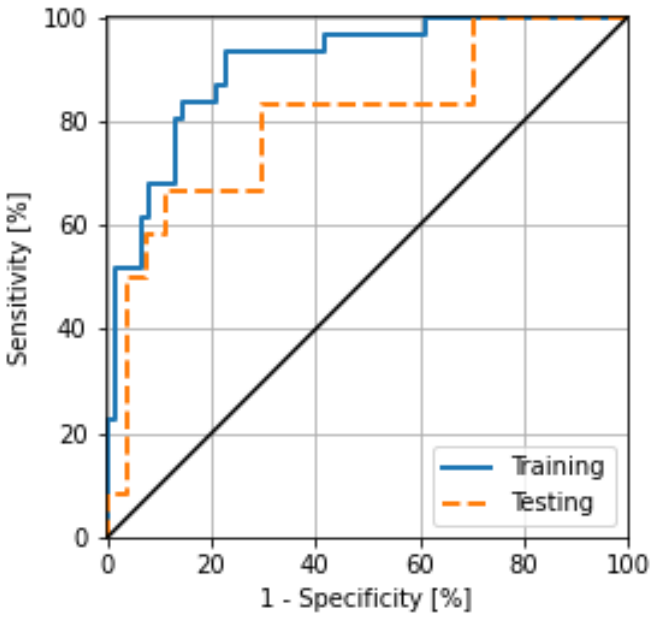

Figure S1 Area under the curve (AUC) performance results from the developed PFS prediction model using CT based radiomic features. PFS, progression-free survival.

Table S1 Programmed death ligand 1 (PD-L1) expression and acquisition results

	PFS		Pneumonitis	
	≤90 days PFS (cases)	>90 days PFS (controls)	Pneumonitis (cases)	No Pneumonitis (controls)
PD-L1 TPS (Biopsy) (<1%: 1-49%: ≥50%)	16:4:7	19:12:22	5:2:3	30:14:26
PD-L1 TPS (Cytology) (<1%: 1-49%: ≥50%)	6:2:2	12:11:8	0:4:0	18:9:10
PD-L1 TPS (All samples) (<1%: 1-49%: ≥50%: Unknown)	22:6:9:6	31:23:30:5	5:6:3:1	48:23:36:10
PD-L1 Assay (E1LN3:22C3:Unknown)	29:8:6	63:21:4	12:2:1	81:27:9

PFS, progression-free survival; SD, standard deviation.

https://dx.doi.org/10.21037/tlcr-22-763

Table S2 Pneumonitis clinical demographics

Parameter	All	Pneumonitis (cases)	No Pneumonitis (controls)	P value
N	132	15	117	-
Age, years (mean ± SD)	64.1±10.4	64.3±8.06	64.1±10.7	>0.9999
Sex (Male: Female)	67:65	5:10	62:55	0.1516
Race (White:African American/Black: Asian: Declined)	126:3:2:1	13:2:0:0	113:1:2:1	0.0220
Pack years (>20 yrs:<20 yrs:never smoker)	100:16:16	10:2:3	90:14:13	0.5851
Immunotherapy type (Pemb: Nivo: Avel)	97:30:5	8:6:1	89:24:4	0.1714
Programmed death ligand 1 (PD-L1) TPS (<1%:1-49%:≥50%:Unknown)	54:28:40:10	5:6:3:1	48:23:36:10	0.3507
Immunotherapy concurrent with chemotherapy (Yes:No)	28:104	3:12	25:92	0.9029
Lymphocyte count (per uL) (mean ± SD)	1,921±8,168	994.3±773.2	2,040±8,669	0.1390
Pneumonitis (Yes:No)	15:117	15:0	0:117	<0.0001
Line of treatment (Naïve:2nd Line:3rd Line:4th Line)	77:44:7:4	6:6:1:2	71:38:6:2	0.0661
Prior chest radiation (Yes:No)	54:78	9:6	45:72	0.1102
Days PFS (mean ± SD)	288.6±317.5	427.2±418.9	270.7±299.6	0.0405

PFS, progression-free survival; SD, standard deviation; Pemb, pembrolizumab; Nivo, nivolumab; Avel, avelumab.

Table S3 Feature reduction of selection results for pneumonitis prediction

Feature (mean ± SD)	PNEUMO_Radiomic + Clinical	PNEUMO_ Radiomic	Pneumonitis	No Pneumonitis	P value
wavelet.HHH_glcm_ClusterProminence	X	X	0.9599±0.5444	0.9369±3.595	0.0001
original_glcm_ClusterShade	X	X	5106±3050	3899±1482	0.0409
wavelet.HLH_glcm_ClusterShade	X	X	0.8151±0.5282	0.2964±0.4593	0.0003
wavelet.LHH_firstorder_Mean	X	X	12.15±2.315	12.29±2.988	0.4995
HistoryofSystemicRadiation	X		9:6	45:72	0.1102
wavelet.HLH_gldm_SmallDependence-HighGrayLevelEmphasis	X	X	12.29±15.26	5.638±13.87	<0.0001
wavelet.LLH_glcm_Idmn	X	X	0.9959±0.0022	0.9966±0.0015	0.6182
wavelet.LHH_glrlm_GrayLevelVariance	X		6.551±1.636	4.492±1.692	<0.0001
Smoker	X		10:2:3	90:14:13	0.5851
wavelet.LHL_gldm_SmallDependenceEmphasis		X	0.0598±0.0187	0.0475±0.0189	0.0522
wavelet.HLH_glcm_InverseVariance		X	0.4729±0.0062	0.4756±0.0065	0.2118
wavelet.LHH_glcm_ClusterShade		X	0.7727±0.7312	0.1391±0.3272	<0.0001

https://dx.doi.org/10.21037/tlcr-22-763

Table S4 Performance results of pneumonitis prediction from developed models using image and/or clinical features

Pneumonitis prediction model	Training AUC (95% CI)	Testing AUC (95% CI)	Sensitivity	Specificity
Pneumo_Radiomic + Clinical (AUC-PR)	0.60 (0.42–0.78)	0.24 (0.11–0.37)	0.75	0.71
Pneumo_Radiomic (AUC-PR)	0.63 (0.5–0.76)	0.24 (0.03–0.45)	0.5	0.86
Pneumo_Clinical without PD-L1 (AUC-PR)	0.19 (0.14–0.24	0.13 (0.03–0.23)	0.75	0.54
Pneumo_Clinical with PD-L1 (AUC-PR)	0.13 (–0.49–0.75)	0.10 (0.05–0.15)	0.25	0.74

AUC, area under the curve; PR, precision recall; PD-L1, programmed death ligand 1 expression.

https://dx.doi.org/10.21037/tlcr-22-763