Novel imaging biomarkers predict outcomes in stage III unresectable non-small cell lung cancer treated with chemoradiation and durvalumab

Khalid Jazieh,1 Mohammadhadi Khorrami,2 Anas Saad,3 Mohamed Gad,1 Amit Gupta,4 Pradnya Patil,5 Vidya Sankar Viswanathan,2 Prabhakar Rajiah,6 Charles J Nock,7 Michael Gilkey,2 Pingfu Fu,8 Nathan A Pennell,6,9 Anant Madabhushi1,2,9

ABSTRACT

Background The landmark study of durvalumab as consolidation therapy in NSCLC patients (PACIFIC trial) demonstrated significantly longer progression-free survival (PFS) in patients with locally advanced, unresectable non-small cell lung cancer (NSCLC) treated with durvalumab (immunotherapy, IO) therapy after chemoradiotherapy (CRT). In clinical practice in the USA, durvalumab continues to be used in patients across all levels of programmed cell death ligand-1 (PD-L1) expression. While immune therapies have shown promise in several cancers, some patients either do not respond to the therapy or have cancer recurrence after an initial response. It is not clear so far who will benefit of this therapy or what the mechanisms behind treatment failure are.

Methods A total of 133 patients with unresectable stage III NSCLC who underwent durvalumab after CRT or CRT alone were included. Patients treated with durvalumab IO after CRT were randomly split into training (D1=59) and test (D2=59) sets and the remaining 15 patients treated with CRT alone were grouped in D3. Radiomic textural patterns from within and around the target nodules were extracted. A radiomic risk score (RRS) was built and was used to predict PFS and overall survival (OS). Patients were divided into high-risk and low-risk groups based on median RRS.

Results RRS was found to be significantly associated with PFS in D1 (HR=2.67, 95% CI 1.85 to 4.13, p<0.05, C-index=0.78) and D2 (HR=2.56, 95% CI 1.63 to 4, p<0.05, C-index=0.73). Similarly, RRS was associated with OS in D1 (HR=1.89, 95% CI 1.3 to 2.75, p<0.05, C-index=0.67) and D2 (HR=2.14, 95% CI 1.28 to 3.6, p<0.05, C-index=0.69), respectively. RRS was found to be significantly associated with PFS in high PD-L1 (HR=3.01, 95% CI 1.41 to 6.45, p=0.0044) and low PD-L1 (HR=2.74, 95% CI 1.8 to 4.14, p=1.77e-06) groups. Moreover, RRS was not significantly associated with OS in the high PD-L1 group (HR=2.08, 95% CI 0.98 to 4.4, p=0.054) but was significantly associated with OS in the low PD-L1 group (HR=1.61, 95% CI 1.14 to 2.28, p=0.0062). In addition, RRS was significantly associated with PFS (HR=2.77, 95% CI 1.17 to 6.52, p=0.019, C-index=0.77) and OS (HR=2.62, 95% CI 1.25 to 5.51, p=0.01, C-index=0.77) in D3, respectively.

Conclusions Tumor radiomics of pretreatment CT images from patients with stage III unresectable NSCLC were prognostic of PFS and OS to CRT followed by durvalumab IO and CRT alone.

INTRODUCTION

Lung cancer is estimated to be the leading cause of cancer-related deaths in the USA in 2021, and non-small cell lung cancer (NSCLC) accounts for around 85% of lung cancer cases.1 In 2019, an estimated 228,150 patients in the USA were expected to be diagnosed with NSCLC, of which one-third have stage III locally advanced disease at the time of diagnosis.2 Platinum-based, doublet chemotherapy administered with definitive-dose radiotherapy or chemoradiotherapy (CRT) was the standard of care for patients with unresectable stage III NSCLC with good performance status.3

However, despite treatment the overall prognosis remains poor and most patients with unresectable stage III NSCLC develop disease progression after initial CRT, often with additional sites of metastasis, with an overall 5-year survival rates of approximately 15%–30%.4 There was little change in patient management until recently, with CRT being the mainstay of systemic treatment for these patients, rendering a meager median survival improvement from 18 to 23 months.5–7

A recent landmark development has been the approval of immune checkpoint inhibitors (ICI) for treatment of locally advanced and metastatic (stage III/IV) NSCLC.8 Recently, several ICIs targeting the programmed death-1 (PD-1) receptor and its ligand, PD-L1 (programmed cell death ligand-1), were approved for use in patients with advanced stage NSCLC. These ICIs have an excellent
toxicity profile and the potential to induce durable clinically meaningful response in this patient population.\(^9\) PD-L1 inhibitors are a group of drugs that block PD-L1 binding to PD-1 and allow T cells to recognize and kill tumor cells. Durvalumab is one such novel agent, a human IgG1 anti-PD-L1 monoclonal antibody, which was recently approved by the Food and Drug Administration (FDA) for patients with unresectable stage III NSCLC whose disease has not progressed following concurrent platinum-based chemotherapy and radiation therapy, as per the results from a study of durvalumab as consolidation therapy in NSCLC patients (PACIFIC), a randomized double-blind, placebo-controlled trial.\(^10\) Studies have shown that CRT may increase PD-L1 expression in tumor cells, which is a positive predictive factor for a response to durvalumab.\(^11-13\) Following the results of the PACIFIC trial, the current standard of care for unresectable stage III NSCLC is chemoradiation followed by 12 months of durvalumab consolidation. In the USA, PD-L1 expression has no role in selection of patients and all patients receive durvalumab regardless of PD-L1 expression. However, the response rate to consolidation durvalumab therapy is only 28.4% and significant adverse reactions were seen in approximately 30% of patients, including cough, fatigue, pneumonitis/radiation pneumonitis, upper respiratory tract infections, dyspnea, and rash.\(^10\) At present, there is no reliable parameter or biomarker that can predict response or adverse reaction to durvalumab therapy. An exploratory analysis of the PACIFIC trial outcomes by the extent of PD-L1 expression (divided into <1% expression, 1%–24% expression, and >25% expression) showed that the progression-free survival (PFS) benefit with durvalumab was observed across all PD-L1 subgroups, but the overall survival (OS) benefit was not observed in the group with <1% expression.\(^14\) Consequently, there is an unmet clinical need for accurate and validated predictive biomarkers for identifying and selecting patients with locally advanced NSCLC who will receive clinical benefit from durvalumab consolidation. A robust biomarker would potentially spare these patients from a prolonged course of durvalumab that would not improve their outcomes and just expose them to unnecessary harm, and it may change monitoring strategies if their risk of treatment failure were known to be higher.

CT is an established non-invasive imaging modality routinely used for screening, diagnosis, and staging of lung cancer. In recent years, computational imaging approaches originating from artificial intelligence have achieved success in automatically quantifying radiographic characteristics of tumors.\(^15-17\) Radiomics is an emerging field within medical research that aims to use advanced imaging analysis to analyze tumors and potentially predict treatment outcomes based on radiological features.

In this study, we sought to identify prognostic biomarkers for PFS and OS in patients with NSCLC treated with durvalumab after CRT or CRT alone by interrogating the tumor and peritumoral microenvironment on CT imaging. Our hypothesis is that quantitative subvisual phenotypic differences in NSCLC tumors on CT imaging can be used to develop prognostic biomarkers to improve decision support in stage III NSCLC treatment. Moreover, we aimed to determine a subset of patients who are at increased risk of recurrence compared with those who are at low risk of recurrence and could be spared the toxic side effects of durvalumab immunotherapy (IO).

**MATERIALS AND METHODS**

**Data sets and patient selection**

We did a retrospective multicohort study of patients with unresectable stage III NSCLC admitted to two independent sites treated either with chemoradiation alone or chemoradiation followed by durvalumab consolidation. We included all patients with (1) availability of pathological confirmation of NSCLC, (2) presence of a pretreatment diagnostic chest CT scan, and (3) presence of a solitary pulmonary nodule/mass. The exclusion criteria were applied to remove scans with CT artifacts and poor image quality not suitable for feature extraction.

We reviewed the charts of patients with NSCLC admitted to the Cleveland Clinic Foundation (Cleveland, Ohio, USA). Between July 2017 and July 2019, pharmacological records from the Cleveland Clinic were searched to identify all patients with stage III NSCLC who were treated with durvalumab. After exclusion, 118 patients treated with CRT followed by durvalumab consolidation were identified. The patients were then randomly divided into an equal number for training set (D1) that consisted of 59 patients and a test set (D2) that contained 59 patients.

The second cohort, D3, included 20 eligible patients identified from a chart review of patients with NSCLC admitted at the Cleveland VA Medical Center (Cleveland, Ohio, USA) from August 2018 to August 2020. After exclusion, 15 patients treated with CRT alone were identified and included in this study.

The clinical information of the patients, including their age, race, sex, smoking history, histological subtype of NSCLC, tumor mutational burden and PD-L1 expression, progression status of their cancer (if it recurred or progressed after initiating durvalumab therapy), and survival data, was collected. PD-L1 expression was available for 97 patients from Cleveland Clinic Foundation (CCF), and among them 61 patients had high expression of PD-L1 and 31 patients had low expression of PD-L1. PD-L1 expression data were not available for patients from the VA Medical Center.

**Follow-up**

The primary endpoints of this study were PFS, which is defined as the length of time during and after the treatment that a patient survives without evidence of disease progression or death, whichever occurred earlier, and OS, which was defined as the time from the date of the initiation of treatment until either the date of death or until the date that the patient was last known to be alive.

"Open access"
(censored). The follow-up period for OS was 75 months after the date of diagnosis. The secondary objective was the difference in PFS and OS within high PD-L1 and low PD-L1 expression groups based on different PD-L1 categories, such as 1% and >50% cut-off criteria.

### PD-L1 assessment
PD-L1 IHC 22C3 pharmDx, a qualitative immunohistochemical (IHC) assay that uses Monoclonal Mouse Anti-PD-L1, was used to detect PD-L1 protein in formalin-fixed, paraffin-embedded NSCLC tissue, using EnVision FLEX visualization system for use on Dako Omnis. PD-L1 expression was determined using the tumor proportion score (TPS), which is the percentage of viable tumor cells that show partial or complete membrane staining at any intensity. The specimen is considered to have PD-L1 expression if TPS is ≥1%.

### CT acquisition and segmentation
Pretreatment CTs were acquired according to standard scanning protocol at the CCF institution from either Siemens, GE Medical Systems, Philips, or Toshiba machine using a tube voltage of 100–120 kVp and automatic tube current modulation. The data set was composed of images reconstructed with multiple reconstruction kernels, and multiple slice thickness ranged from 0.7 mm to 2 mm. Therefore, precaution was taken to equally distribute cases in D1 and D2 sets and account for variability. Pixel sizes ranged from 0.42±0.42 mm to 0.97±0.97 mm, with an average size of 0.75×0.73 mm.

### Tumor segmentation
Lung nodules on prechemoradiation CT scans were annotated with the 3D SLICER software by a board-certified radiologist with 20 years of experience. The primary tumor site was retrospectively contoured guided by existing treatment planning contours. Using both soft tissues and lung windows, air, vessels, normal tissue, and surrounding organs were subsequently excluded from the contours. Intratumoral texture features were extracted from the segmented nodules. The intratumoral mask was then dilated out to a 20 mm peritumoral radius. The choice of peritumoral compartment size was determined based on the National Comprehensive Cancer Network guidelines, which suggest a resection margin of 2 cm or greater in patients with NSCLC.

### Computer-extracted texture and shape features
In total, 13 Haralick texture features, 25 Laws features, 25 Laws Laplacian, and 48 Gabor features were extracted from the nodule and investigated in this study. All features were extracted from two-dimensional contours in a slice-by-slice basis to pick all representative slices that had the tumor. These features capture textural pattern and variation in tumor microarchitecture, heterogeneity, and local appearance of nodules. Haralick features were extracted from gray-level co-occurrence matrix and show entropy of intensity values among local pixel neighborhoods and capture variation in tumor microarchitecture.

Laws is a filter-based descriptor that can capture combinations of five uncommon textural patterns, such as levels (L), edges (E), spots (S), waves (W), or ripples (R). The Gabor filter captures six different spatial frequencies (f=0, 2, 4, 8, 16, or 32) within the image at eight directional orientations (θ=0, π/8, π/4, 3π/8, π/2, 5π/8, 3π/4, 7π/8). The first-order statistics (mean, median, SD, skewness, kurtosis) for each family features were computed within the tumor and each peritumoral region, resulting in 495 statistical features per region. Feature extraction was performed using an inhouse MATLAB V.2019 (MathWorks, Natick, Massachusetts, USA) toolbox. Additionally, a total of 24 shape features were also automatically extracted. All extracted radiomics feature intensity values were then normalized to lie down between −1 and 1.

### Statistical analysis
The least absolute shrinkage and selection operator (LASSO) Cox regression model, which is suitable for regression of high-dimensional data, was used to select the most prognostic features to PFS in the training data set. After selecting the top prognostic features, the corresponding LASSO coefficients were used for radiomic signature risk score (RRS) construction. RRS was built based on a linear combination of non-zero coefficients-selected features weighted by their corresponding coefficients. The LASSO Cox regression was performed using the ‘glmnet’ package in R. The association of RRS with PFS was first assessed in D1 and then validated in D2 using Kaplan-Meier survival analysis, log-rank test, HR (95% CI), and Harrell’s concordance index (C-index). The association of constructed RRS with OS was also evaluated in D1 and D2, respectively. According to the median RRS threshold, patients were classified into high risk or low risk. To evaluate if our model has similar prognostic impact on patients who received CRT alone, the association of RRS (calculated from D1) was assessed in D3 using survival analysis.

The multivariable Cox regression on PFS and OS with RRS and clinicopathological biomarker was also employed in D1 (n=41) and D2 (n=51) sets (PD-L1 expression was only available for 41 patients in D1 and 51 patients in D2 sets). In addition, prognosticating outcome of RRS in predicting PFS and OS for low and high PD-L1 groups was also evaluated.

The clinicopathological factors and RRS were incorporated into a multivariable Cox regression analysis to develop the nomogram. The nomogram model was generated by use of R with the ‘rms’ package (regression modeling strategies). To determine the clinical usefulness of the radiomics nomogram, a decision curve analysis (DCA) which can quantify the net benefits at different threshold probabilities was performed. The net benefit was defined as summing the benefits (true-positive results) and subtracting the harms (false-positive results) weighted by a factor related to the relative harm of not identifying a high-risk patient who might have low PFS compared with the harm of subjecting a lower-risk patient
to more aggressive therapy, when the more intense therapy is not required. A model is said to be better compared with another at the chosen threshold probability if its net benefit surpasses the net benefit of the other model for that value of probability. In addition, differences between clinical categories were assessed using Fisher’s exact test, while a two-sided t-test was used for continuous variables.

### RESULTS

#### Patient analysis

Of the 118 patients treated with CRT followed by durvalumab IO, 50% were male, 79% were white, 97% were former or current smokers, and about 48% of the patients had adenocarcinoma. The median number of patients in each cohort are shown in Table 1.

### Radiomic features from pretreatment CT scans were associated with PFS and OS in patients with NSCLC treated with CRT followed by durvalumab IO

The median PFS was 14.5 months (95% CI 12.7 to 15.5). A univariable Cox regression analysis identified that PFS was not significantly different for gender (male vs female; HR: 1.4, 95% CI 0.8 to 2.46, p=0.24, C-index=0.55), race (HR: 1.4, 95% CI 0.63 to 3.12, p=0.41, C-index=0.51), smoking status (HR: 1.11, 95% CI 0.52 to 1.82, p=0.86, C-index=0.48), tumor stage (IIIA, IIIB, IIIC; HR: 1.01, 95% CI 0.01 to 5.25, p=0.11, C-index=0.56), tumor type (adenocarcinoma, large cell, squamous; HR: 0.71, 95% CI 0.16 to 3, p=0.76, C-index=0.51), and lymph node status (N0, N1, N2, N3; HR: 1.21, 95% CI 0.35 to 4.12, p=0.41, C-index=0.55), but was significantly different with PD-L1 status (high vs low; HR: 0.18, 95% CI 0.076 to 0.44, p=0.00013, C-index=0.65). Figure 1A–G illustrates the Kaplan-Meier curves of the different clinical factors.

The optimum cut-off value (the median) for RRS was found to be 0.118, and patients were divided into high-risk and low-risk groups based on this value. A univariable Cox regression analysis developed using textural features indicated that RRS was significantly associated with PFS in D1 (HR: 2.76, 95% CI 1.85 to 4.13, p=7.3e-7, C-index=0.78) and D2 (HR: 2.56, 95% CI 1.63 to 4, p=4.6e-5, C-index=0.73), respectively. In a multivariable analysis using a combination of clinicopathological and radiomic signatures, the radiomic risk score and PD-L1 expression were found to be significantly associated with PFS in D1 (risk score: HR: 2.3, 95% CI 1.46 to 3.63, p=0.0003; PD-L1: HR: 0.31, 95% CI 0.081 to 0.96, p=0.038, C-index=0.81) and D2 (risk score: HR: 2.56, 95% CI 1.75 to 4, p=8.7e-5; PD-L1: HR: 0.27, 95% CI 0.048 to 0.58, p=0.005, C-index=0.77). The corresponding Kaplan-Meier survival curves showed a significant difference in PFS between patients with low and high RRS both in D1 and D2 sets (p<0.0001). The Kaplan-Meier survival curves for D1 and D2 are shown in FIGURE 1H and FIGURE 1I, respectively.

**Figure 2** illustrates the discriminability of the intratumoral Haralick entropy texture feature for one low-risk patient and one high-risk patient before therapy. As it

---

**Table 1** Demographics and clinical characteristics of 133 patients

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>D1 cohort (n=59)</th>
<th>D2 cohort (n=59)</th>
<th>D3 cohort (n=15)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, n (%)</td>
<td>Male</td>
<td>28 (47)</td>
<td>31 (53)</td>
<td>14 (93)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>31 (53)</td>
<td>28 (47)</td>
<td>1 (7)</td>
</tr>
<tr>
<td>Age, median (range)</td>
<td></td>
<td>68 (48–89)</td>
<td>68 (50–84)</td>
<td>72 (60–82)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>White</td>
<td>41 (69)</td>
<td>52 (88)</td>
<td>14 (93)</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>18 (31)</td>
<td>7 (12)</td>
<td>1 (7)</td>
</tr>
<tr>
<td>Smoking, n (%)</td>
<td>Never</td>
<td>2 (3)</td>
<td>2 (3)</td>
<td>13 (87)</td>
</tr>
<tr>
<td></td>
<td>Former/current</td>
<td>41 (69)</td>
<td>57 (97)</td>
<td>2 (13)</td>
</tr>
<tr>
<td>Stage, n (%)</td>
<td>IIIA</td>
<td>34 (58)</td>
<td>23 (39)</td>
<td>10 (67)</td>
</tr>
<tr>
<td></td>
<td>IIIB</td>
<td>20 (34)</td>
<td>32 (54)</td>
<td>4 (27)</td>
</tr>
<tr>
<td></td>
<td>IIIC</td>
<td>5 (8)</td>
<td>4 (7)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Histology, n (%)</td>
<td>Adenocarcinoma</td>
<td>28 (47)</td>
<td>29 (49)</td>
<td>10 (67)</td>
</tr>
<tr>
<td></td>
<td>Squamous cell</td>
<td>28 (47)</td>
<td>27 (46)</td>
<td>5 (33)</td>
</tr>
<tr>
<td></td>
<td>Large cell</td>
<td>3 (6)</td>
<td>3 (5)</td>
<td>0</td>
</tr>
<tr>
<td>Median follow-up (PFS) Months</td>
<td>14.5 (0.23–30.2)</td>
<td>14.9 (1.6–30.5)</td>
<td>10.5 (0.2–34.8)</td>
<td></td>
</tr>
<tr>
<td>Median follow-up (OS) Months</td>
<td>16.5 (3.4–30.2)</td>
<td>16.5 (2.5–30.5)</td>
<td>10.5 (0.2–35.7)</td>
<td></td>
</tr>
</tbody>
</table>

OS, overall survival; PFS, progression-free survival.
may be observed, the texture heat maps appear to suggest higher textural entropy values for the high-risk compared with the low-risk patient.

A radiomics model incorporating the developed radiomic signature with clinicopathological biomarkers was chosen as the best model predicting PFS with C-index in D1 of 0.66 (95% CI 0.63 to 0.7) (figure 3A). In D2, the C-index was 0.65 (95% CI 0.62 to 0.69). The calibration plot of D2 (figure 3B) demonstrated an optimal consistency between nomogram-predicted and the actual observed PFS. The DCA was used to demonstrate clinical decision utility of the combined nomogram. Figure 3C shows the DCA for three models (clinical model, radiomic model, and integrated model). As can be seen, the integrated model had the highest net benefit in predicting which high-risk patients should receive more aggressive treatment compared with low-risk patients.

Radiomic features from pretreatment CT scans were also associated with OS. The median OS was 16.5 months (95% CI 16.0 to 18.35). A univariable Cox regression

Figure 1  (A–G) Kaplan-Meier PFS curves for race, gender, smoking status, clinical stage, tumor type, lymph node status, and PD-L1 expression. Kaplan-Meier PFS curves based on the (H) training set (D1) and (I) the test set (D2). A significant association of the radiomic risk score with PFS was shown in the D1 and D2 sets. PD-L1, programmed cell death ligand-1; PFS, progression-free survival.
analysis identified that OS was not significantly different for gender (male vs female; HR: 0.9, 95% CI 0.44 to 1.84, p=0.77, C-index=0.50), race (HR: 1.4, 95% CI 0.48 to 4, p=0.53, C-index=0.52), smoking status (HR: 1.72, 95% CI 0.7 to 4.22, p=0.23, C-index=0.57), tumor stage (HR: 1, 95% CI 0.01 to 5, p=1, C-index=0.55), tumor type (HR: 1.7, 95% CI 0.38 to 7.7, p=0.47, C-index=0.54), and lymph node status (HR: 0.97, 95% CI 0.22 to 4.37, p=0.97, C-index=0.55), but was significantly different with PD-L1 status (HR: 0.32, 95% CI 0.12 to 0.87, p=0.025, C-index=0.63).

A univariable analysis showed that RRS generated with PFS was also associated with OS in D1 (HR: 1.89, 95% CI 1.3 to 2.75, p=0.00077, C-index=0.67) and D2 (HR: 2.14, 95% CI 1.28 to 3.6, p=0.0038, C-index=0.69), respectively. The corresponding Kaplan-Meier survival curves showed a significant difference in OS between patients with low and high RRS both in D1 and D2 sets (D1: p=0.00059, D2: p=0.00018). The Kaplan-Meier survival curves for D1 and D2 are shown in Figure 4A,B, respectively.

Table 2 shows the distribution of each biomarker, such as RRS, age, sex, tumor histology, tumor stage, PD-L1 expression, and smoking status, based on low-risk and high-risk groups.

Radiomic features from pretreatment CT scans were associated with PFS and OS in patients with NSCLC treated with CRT alone

The median PFS for patients treated with CRT alone was 10.53 months (95% CI 6.11 to 16.98) vs 14.6 months in patients treated with CRT plus durvalumab (p=0.01), and the median OS was 10.5 months (95% CI 6.12 to 16.84) vs 16.5 months in CRT plus durvalumab group (p=0.004).

The patients were divided into high-risk and low-risk group based on the median RRS constructed from baseline CT images of patients who received CRT plus durvalumab. A univariable Cox regression analysis indicated that RRS was significantly associated with PFS (HR: 2.77, 95% CI 1.17 to 6.52, p=0.019, C-index=0.77) and OS (HR: 2.62, 95% CI 1.25 to 5.1, p=0.01 C-index=0.77) in D3, respectively.

In addition, when D2 and D3 were combined, RRS was accurately able to discriminate 87% (13 of 15) of the patients treated with CRT alone, assigning them to the high-risk group.

Predicting PFS and OS in low and high PD-L1 groups

A significant difference was observed in PFS and OS for patients with low and high PD-L1 based in the 50% cut-off criteria (PD-L1 <50% as low PD-L1 and ≥50% as high PD-L1), while no significant difference was observed in the 1% cut-off criteria.

RRS was found to be significantly associated with PFS (between low and high risk) in the high PD-L1 (HR: 3.01, 95% CI 1.41 to 6.45, p=0.0044) and low PD-L1 (HR: 2.74, 95% CI 1.8 to 4.14, p=1.77e-06) groups based on the 50% cut-off criteria. The corresponding Kaplan-Meier survival curves (figure 5A,B) show these differences for PFS and OS within the low and high PD-L1 groups.

RRS was not significantly associated with OS in the high PD-L1 group (HR: 2.08, 95% CI 0.98 to 4.4, p=0.054) but was significantly associated with OS in the low PD-L1 group (HR: 1.61, 95% CI 1.14 to 2.28, p=0.0062) based on the 50% cut-off criteria.
DISCUSSION

NSCLC is the leading cause of cancer-related mortality in the USA and worldwide. NSCLC accounts for approximately 85% of the lung cancer cases in the USA, and about one-third of them present with unresectable, locally advanced (stage III) disease.23

Until 2017, the standard treatment for unresectable stage III NSCLC was mainly platinum-based, doublet chemotherapy concurrent with radiotherapy (CRT). However, recent data from the PACIFIC trial demonstrated an OS benefit with durvalumab consolidation, a monoclonal antibody against PD-L1, after CRT in these patients. In the PACIFIC trial, patient enrollment was not restricted based on PD-L1 expression. Nevertheless, irrespective of PD-L1 expression, PFS and OS with durvalumab improved versus placebo. An unplanned exploratory post-hoc analysis of the PACIFIC trial based on PD-L1 expression24 demonstrated a PFS benefit regardless of PD-L1 expression, but OS benefit was seen only in patients with PD-L1 expression ≥1%. However, an imbalanced small sample size with too few events prevented robust conclusions regarding OS in patients with PD-L1 expression <1%.

In light of these findings, in February 2018, the US FDA approved durvalumab as a standard of care for patients with locally advanced, unresectable NSCLC whose disease...
has not progressed after CRT, irrespective of PD-L1 expression. Later in September durvalumab was approved by the European Medicines Agency (EMA) only for patients with PD-L1 expression more than 1% of tumor cells.\textsuperscript{10,25,26}

Despite the promising clinical activity of durvalumab, this IO drug is often accompanied by autoimmune toxicities. Unfortunately, the current gold standard of tissue-based biomarkers using PD-L1 expression is imperfect in selecting patients for treatment with IO. Moreover, its role is uncertain when such agents are given in sequence or in combination with other therapies like CRT. Consequently, there is an unmet clinical need for accurate, validated, and non-invasive predictive biomarkers for identifying and selecting patients with unresectable stage III NSCLC who will receive maximum clinical benefit from IO after CRT, since CRT alone is curative for some and consolidation with IO for those who are unlikely to benefit from it puts them at risk of side effects without impacting their clinical course.

In this study, we demonstrated that a radiographic image-based biomarker (known as radiomics) on baseline CT scans is significantly associated with PFS and OS (specifically prognostic within individual PD-L1 categories) in patients with NSCLC treated with durvalumab therapy after CRT or CRT alone by using radiomic texture patterns within and outside the NSCLC tumor cells.\textsuperscript{27}

Table 2 Distribution of each biomarker, such as RRS, age, sex, race, tumor histology, tumor stage, PD-L1 expression, and smoking status, based on low-risk and high-risk groups

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>High-risk</th>
<th>Low-risk</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRS</td>
<td>Median=0.47 95% CI 0.53 to 0.92</td>
<td>Median=0.6 95% CI 0.87 to 0.59</td>
<td>1.14e-21</td>
</tr>
<tr>
<td>PD-L1</td>
<td>8 high PD-L1 38 low PD-L1 14 N/A</td>
<td>25 high PD-L1 23 low PD-L1 11 N/A</td>
<td>0.0012</td>
</tr>
<tr>
<td>Smoking</td>
<td>16 current smoker 45 never or former</td>
<td>18 current smoker 41 never or former</td>
<td>0.68</td>
</tr>
<tr>
<td>Pack per year</td>
<td>Median=45.5 95% CI 42.6 to 56.9</td>
<td>Median=42 95% CI 38.55 to 51.27</td>
<td>0.32</td>
</tr>
<tr>
<td>Age</td>
<td>Median=66 95% CI 64.8 to 68.6</td>
<td>Median=69 95% CI 67.23 to 71.95</td>
<td>0.63</td>
</tr>
<tr>
<td>Sex</td>
<td>23 male 28 female</td>
<td>28 male 31 female</td>
<td>0.58</td>
</tr>
<tr>
<td>Race</td>
<td>49 white 11 black</td>
<td>46 white 13 black</td>
<td>0.65</td>
</tr>
<tr>
<td>Tumor histology</td>
<td>28 squamous 27 adenocarcinoma 5 large cell carcinoma</td>
<td>27 squamous 32 adenocarcinoma 1 large cell carcinoma</td>
<td>0.24</td>
</tr>
<tr>
<td>Stage</td>
<td>25 stage IIIA 32 stage IIIB 3 stage IIIC</td>
<td>29 stage IIIA 27 stage IIIB 3 stage IIIC</td>
<td>0.72</td>
</tr>
</tbody>
</table>

N/A, not available; PD-L1, programmed cell death ligand-1; RRS, radiomic signature risk score.

Structural heterogeneity of tumor as captured by intratumoral radiomic features has been shown to correlate with heterogeneous blood supply in the tumor, which results in a hypoxic tumor environment that is known to be a barrier and associated with suboptimal response to CRT in NSCLC.\textsuperscript{28,29} Similarly, in this study we found that the entropy of intratumoral Haralick feature has a higher expression in short-term PFS compared with

Figure 4 Kaplan-Meier survival curves based on the (A) training set (D1) and (B) test set (D2). A significant association of the radiomic risk score with overall survival was shown in the D1 set (p=0.00059) and D2 set (p=0.00018).
long-term PFS. Angiogenesis is another major factor that plays a crucial role in the invasiveness/metastatic potential of a tumor and is a significant predictor of worse PFS and OS, which may be captured by peritumoral Laws features. This may be explained by the fact that angiogenesis reduces the tumor’s accessibility to therapeutic drugs, which in turn leads to tumor growth, and blockage of tumor angiogenesis stops tumor proliferation. The rationale behind peritumoral radiomics is that these textural patterns could be capturing the degree of immune response, which in turn is known to be correlated with the likelihood that the tumor will have a favorable response. The tumor immune microenvironment and associated vascularity play a crucial role in immune homeostasis. Recent studies showed that there is a positive feedback between antitumor T helper type 1 (Th-1) immune response and tumor vessel normalization. Disruption of tumor vasculature can lead to impaired T cell infiltration. Interferon gamma (INF-γ-mediated) gene expression signatures promote chemotactic response and could improve T cell infiltration-promoting antitumor immune responses. Patients with endogenous spontaneous antitumor immune responses with the ability to generate an INF-γ Th-1 immune response and tumor vessel normalization benefit from ICI. The presence of a permissive tumor vasculature and immune contexture predicts response to ICI therapy. The peritumoral texture features can capture vessels and vessel irregularity around the tumor.

PD-L1 is the current gold standard biomarker used for patient selection for therapy with ICI. However, across a range of clinical trials, the performance of PD-L1 expression as a predictive biomarker is quite poor. Using PD-L1 expression in clinical practice is fraught with many challenges, such as the use of different thresholds of expression in various trials, different scoring systems, and the dynamic and heterogeneous nature of PD-L1 expression. PD-L1 testing is currently the only routinely used biomarker for ICI but is not accurate in identifying responders (27% in PD-L1 positive in the first-line setting and 45% in PD-L1 high subgroup, and 19% in the second-line setting).

There are some data to suggest that patients with NSCLC with a high tumor expression of PD-L1 and CD8+ TIL (tumor infiltrating lymphocytes) level have higher response rate and better OS than those with lower PD-L1 expression level treated with CRT plus durvalumab. Furthermore, evidence from a previous study revealed that specimens of NSCLC histology had higher numbers of tumor-associated inflammatory cells in the peritumoral compartment as compared with the intratumoral region. Also, ICIs mechanistically act by allowing the immune system to be activated in the tumor microenvironment by blocking the PD-1/PD-L1 axis. Given the higher Gabor texture feature expression preferentially in the peritumoral region in this study, we postulate that this feature captures the increment of PD-L1 and lymphocyte level selectively in peritumoral regions.

Finally, in the present study, a radiomic signature was significantly associated with PFS and OS in patients with stage III NSCLC treated with CRT followed by durvalumab IO or CRT alone. We showed that radiomic features that are associated with PFS in patients treated with CRT followed by durvalumab IO are also associated with PFS and OS in patients treated with CRT alone. These preliminary results suggest that there could be a role for our radiomic signature in identifying which patients would receive maximum benefit from durvalumab.

We developed a nomogram that integrated radiomic risk score with clinical biomarkers to further improve its prognostic accuracy. The integrated radiomic and clinical nomogram exhibited a good C-index value in both D1 and D2 cohorts. This CT image-derived radiomic model may provide a less invasive and more
accurate biomarker for prognostication of PFS and OS in patients with NSCLC. We also evaluated the radiomic nomogram of patients by DCA and calculated the net benefit of our model. The decision curve indicated that radiomic signature had a higher overall net benefit in predicting patients at higher risk of receiving inappropriately aggressive treatment than the clinical-pathological measurements across a number of threshold probability values.

By having capability to assess the risk of PFS non-invasively, the oncologist can be empowered to estimate therapeutic outcome for a given patient and has the potential to avoid ineffective treatment and toxicity associated with durvalumab consolidation in patients who are unlikely to respond to therapy and could pursue a more intensive monitoring for patients at higher risk of recurrence. This seems to hold true regardless of PD-L1 expression, in that even patients with high PD-L1 expression would be at risk of treatment failure if they had a high radiomic risk score and that patients with low PD-L1 expression but a low radiomic risk score had good outcomes with treatment and should continue to be offered durvalumab consolidation as it was proposed in the second experiment.

However, our study has its limitations. The cohort sample sizes were relatively small from two institutions, which limited our ability to perform extensive analyses. Additionally, a few recent studies have investigated the influence of convolution kernels, reconstruction algorithms, and slice thickness on radiomic features for characterization of lung nodules on CT. We did not explicitly consider the influence of these parameters on the extracted texture features, but randomly distributed the cases with different image acquisition parameters between training and validation sets to account for variability. The other limitation is the retrospective nature of our study, not a prospective study. In addition, patients meet specific inclusion criteria and belong to a particular stage of disease. We analyzed the association of radiomic features only from pretreatment CRT with patient outcomes. We have not analyzed either of the change of radiomic features between pretreatment and post-treatment (delta radiomics) or the association of post-CRT images with patient outcomes. In addition, due to lack of routine testing for actionable driver oncogenes in patients with stage III NSCLC, we were not able to examine the associations between genetic biomarkers and high-risk/low-risk groups. Although there is some evidence suggesting limited activity of durvalumab consolidation in patients with non-resectable stage III NSCLC with driver genomic alterations, the current FDA-approved indication does not exclude these patients and more studies are needed to define the role of consolidation immunotherapy in this heterogeneous population. We hope to address the limitations shown in this study in future work.

CONCLUSIONS
Our results suggest that radiomic texture features from pretreatment CT images of patients with stage III NSCLC treated with chemoradiation followed by durvalumab or chemoradiation alone were prognostic of PFS and OS. In addition, we showed that there is a subset of patients with high PD-L1 expression who are at high risk of recurrence and simultaneously a subset of patients with low PD-L1 expression who are at low risk of recurrence. While the FDA has approved durvalumab for all patients with locally advanced NSCLC regardless of PD-L1 expression and EMA has approved it for patients with PD-L1 1% of tumor cells, our finding showed that radiomic feature may be a better predictor for selecting patients for durvalumab therapy after CRT. Additional prospective validation of these novel quantitative imaging-based approaches is warranted to accurately define their clinical utility in deciding who should get durvalumab consolidation.

Author affiliations
1Department of Internal Medicine, Cleveland Clinic, Cleveland, Ohio, USA
2Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio, USA
3Heart and Vascular Institute, Cleveland Clinic, Cleveland, Ohio, USA
4Department of Radiology, University Hospitals Cleveland Medical Center, Cleveland, Ohio, USA
5Department of Hematology and Medical Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio, USA
6Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA
7Louis Stokes Cleveland VA Medical Center Mental Health Services, Cleveland, Ohio, USA
8Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio, USA

Twitter Amit Gupta @AmitGupta_83 and Nathan A Pennell @n8pennell

Contributors Conceptualization: KJ, MK, NP, and AM. Methodology: KJ, MK, and AM. Software: MK. Validation: KJ, MK, and AM. Formal analysis: MK, KJ, and PF. Investigation: MK and KJ. Resources: MK. Data curation: KJ, VSV, PP, PR, AS, MGa, CJN, MGb, and NP. Writing - original draft preparation: KJ, MK, and AG. Writing - review and editing: KJ, MK, AG, PR, PF, NP, and AM. Visualization: MK. Supervision: AM. All authors have read and agreed to the published version of the manuscript.

All authors confirm that they have full access to all the data in the study and accept responsibility to submit for publication. AM act as the guarantor of the study.

Funding Research reported in this publication was supported by the National Cancer Institute under award numbers 1U24CA199374-01, R01CA249992-01A1, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, R01CA257612-01A1, U01CA239055-01, 1U01CA248226-01, 1U54CA254566-01, National Heart, Lung and Blood Institute 1R01HL15127701A1, RO1HL15807101A1, National Institute of Biomedical Imaging and Bioengineering 1R43EB028736-01, National Center for Research Resources under award number 1 C06 RR12463-01, VA Merit Review Award BX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service the Office of the Assistant Secretary of Defense for Health Affairs, through the Breast Cancer Research Program (W81XWH-19-1-0668), the Prostate Cancer Research Program (W81XWH-15-1-0558, W81XWH-20-1-0851), the Lung Cancer Research Program (W81XWH-18-1-0440, W81XWH-20-1-0955), the Peer Reviewed Cancer Research Program (W81XWH-18-1-0404), the Kidney Precision Medicine Project (KPM) Glue Grant, the Ohio Third Frontier Technology Validation Fund, the Clinical and Translational Science Collaborative of Cleveland (UL1TR0002548) from the National Center for Advancing Translational Sciences (NCATS) component of the National Institutes of Health and NIH roadmap for Medical Research. The Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering at Case Western Reserve University. Sponsored research agreements from Bristol Myers-Squibb, Boehringer-Ingelheim, and AstraZeneca. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.


Open access
Institutes of Health, the U.S. Department of Veterans Affairs, the Department of Defense, or the United States Government.

Competing interests AM is an equity holder in Eleucid Bioimaging and in Inspirata.

In addition, he has served as a scientific advisory board member for Inspirata, AstraZeneca, Bristol Meyers Squibb and Merck. Currently he serves on the advisory board of Afioria and currently consults for Caris, Roche and Afioria. He also has sponsored research agreements with Philips, AstraZeneca, Boehringer Ingelheim and Bristol Meyers Squibb. His technology has been licensed to Eleucid Bioimaging. He is also involved in a NIH U24 grant with PathCore, and three different R01 grants with Inspirata. Other authors declare no potential conflicts of interest.

Patient consent for publication Not required.

Ethics approval This study was conducted in full accordance with the Health Insurance Portability and Accountability Act (HIPAA) regulations after approval from the Institutional Review Board at Case Western Reserve University (Cleveland, Ohio) and the IRB waived the requirements for informed consent of all patients. Informed consent was not required due to the retrospective nature of the study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Access to data sets from the Cleveland Clinic Foundation and Cleveland VA Medical Center (used with permission for this study) should be requested directly from these institutions via their data access request forms. Subject to the institutional review boards’ ethical approval, unidentified data would be made available as a test subset. However, we are including all the codes used for analyses following feature extraction on GitHub. All experiments and implementation details are described thoroughly in the Materials and methods section so they can be independently replicated with non-proprietary libraries.

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

ORCID iDs Nathan A Pennell http://orcid.org/0000-0002-1458-0064 Anant Madabhushi http://orcid.org/0000-0002-5741-0399

REFERENCES


