

# Radiomics analysis for predicting pembrolizumab response in patients with advanced rare cancers

Rivka R Colen,<sup>1,2</sup> Christian Rolfo,<sup>3</sup> Murat Ak,<sup>1,2</sup> Mira Ayoub,<sup>1,2</sup> Sara Ahmed,<sup>4</sup> Nabil Elshafeey,<sup>5</sup> Priyadarshini Mamindla,<sup>2</sup> Pascal O Zinn,<sup>6</sup> Chaan Ng,<sup>7</sup> Raghu Vikram,<sup>7</sup> Spyridon Bakas,<sup>8</sup> Christine B Peterson,<sup>9</sup> Jordi Rodon Ahnert,<sup>10</sup> Vivek Subbiah,<sup>10</sup> Daniel D Karp,<sup>10</sup> Bettzy Stephen,<sup>10</sup> Joud Hajjar,<sup>11,12</sup> Aung Naing <sup>10</sup>

**To cite:** Colen RR, Rolfo C, Ak M, *et al.* Radiomics analysis for predicting pembrolizumab response in patients with advanced rare cancers. *Journal for ImmunoTherapy of Cancer* 2021;**9**:e001752. doi:10.1136/jitc-2020-001752

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jitc-2020-001752>).

Accepted 12 March 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Aung Naing;  
anaing@mdanderson.org

Dr Rivka R Colen;  
colenrr@upmc.edu

## ABSTRACT

**Background** We present a radiomics-based model for predicting response to pembrolizumab in patients with advanced rare cancers.

**Methods** The study included 57 patients with advanced rare cancers who were enrolled in our phase II clinical trial of pembrolizumab. Tumor response was evaluated using Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 and immune-related RECIST (irRECIST). Patients were categorized as 20 “controlled disease” (stable disease, partial response, or complete response) or 37 progressive disease). We used 3D-slicer to segment target lesions on standard-of-care, pretreatment contrast enhanced CT scans. We extracted 610 features (10 histogram-based features and 600 second-order texture features) from each volume of interest. Least absolute shrinkage and selection operator logistic regression was used to detect the most discriminatory features. Selected features were used to create a classification model, using XGBoost, for the prediction of tumor response to pembrolizumab. Leave-one-out cross-validation was performed to assess model performance.

**Findings** The 10 most relevant radiomics features were selected; XGBoost-based classification successfully differentiated between controlled disease (complete response, partial response, stable disease) and progressive disease with high accuracy, sensitivity, and specificity in patients assessed by RECIST (94.7%, 97.3%, and 90%, respectively;  $p < 0.001$ ) and in patients assessed by irRECIST (94.7%, 93.9%, and 95.8%, respectively;  $p < 0.001$ ). Additionally, the common features of the RECIST and irRECIST groups also highly predicted pembrolizumab response with accuracy, sensitivity, specificity, and  $p$  value of 94.7%, 97%, 90%,  $p < 0.001$  and 96%, 96%, 95%,  $p < 0.001$ , respectively.

**Conclusion** Our radiomics-based signature identified imaging differences that predicted pembrolizumab response in patients with advanced rare cancer.

**Interpretation** Our radiomics-based signature identified imaging differences that predicted pembrolizumab response in patients with advanced rare cancer.

## INTRODUCTION

Rare cancers together account for 22% of all cancer diagnoses<sup>1</sup> and 25% of all adult

cancers,<sup>2</sup> representing a higher percentage than any single common cancer, such as lung (12.9%), breast (15.2%), and colon (8.3%) cancers.<sup>3</sup> The prognosis of rare cancer is poor; the 5-year survival rate for these cancers is 15%–20% lower than for the more common cancers.<sup>4</sup> Immunotherapy, in particular immune checkpoint inhibitors, has revolutionized cancer treatment.<sup>5–8</sup> Recently our group published the first paper on immunotherapy in advanced rare cancer, where we demonstrated a 14% objective response rate in these patients as well varying rates of treatment resistance and toxicity.<sup>9</sup> In general, the reported immunotherapy response rates has been shown to range from 13.3% to 87% depending on the type of cancer.<sup>5 10–13</sup> To date, no single biomarker has been found to accurately predict immunotherapy response.<sup>14</sup> Having the ability to predict treatment failure and therapeutic efficacy would enable clinicians to administer alternate therapeutics and prevent patients who are unlikely to benefit from being exposed to potentially life-threatening toxic effects of the treatment, such as encephalopathy, pneumonitis, hypophysitis, hepatitis, colitis, and pancreatitis.<sup>15</sup> Recently some studies have been published attempting to identify immune biomarkers that can potentially predict response to immunotherapy; these include programmed death protein ligand 1 (*PD-L1*) expression level,<sup>6 16</sup> immunosuppressive factors (Forkhead box P3 (FOXP3), indolamine 2,3-dioxygenase),<sup>17</sup> interferon-gamma,<sup>18</sup> microsatellite instability (mismatch-repair status),<sup>19</sup> and tumor mutation burden.<sup>20</sup> However, the more well-known of these biomarkers, such as *PD-L1* expression in selected tumor types and microsatellite instability, have a low predictive value

in assessing response.<sup>21 22</sup> In addition, these biomarkers require an invasive procedure (surgery or biopsy), and they represent only a small, static “snapshot” of the entire tumor which can be very heterogeneous. Thus, there is an urgent need to develop non-invasive, robust biomarkers that account for tumor heterogeneity to identify patients who will benefit from immunotherapy.

Imaging studies are a non-invasive means of assessing immunotherapy response.<sup>23 24</sup> However, current imaging assessment criteria, including Response Evaluation Criteria in Solid Tumors (RECIST) and immune-related RECIST (irRECIST) are used throughout treatment to assess response only longitudinally during the course of treatment.<sup>23 24</sup> No criteria for robustly predicting immunotherapy response a priori (before receiving therapy) have yet been established.

Radiomics is a high-throughput computational method that unlocks microscale quantitative data hidden within conventional imaging<sup>25</sup> to provide a deeper understanding of cancer biology,<sup>26</sup> genomics,<sup>27</sup> spatial heterogeneity,<sup>27</sup> tumor type,<sup>28</sup> and immune infiltration.<sup>29</sup> Radiomics-based approaches have several advantages; it is non-invasive, and features are extracted from standard medical images, making it ideal for easier clinical translation as a non-invasive method to predict response. Further, radiomics capture the entire three-dimensional tumor landscape inclusive of spatial heterogeneity<sup>27</sup> and allows for better evaluation of the tumor and its microenvironment.<sup>30</sup> Radiomics has been shown to help predict immunotherapy response and outcome in multiple cancers, including melanoma,<sup>31 32</sup> non-small cell lung cancer (NSCLC),<sup>32 33</sup> and advanced solid tumors.<sup>29</sup> However, to the best of our knowledge, no study has been published on radiomics biomarkers predicting immunotherapy response in advanced rare cancers.

In this study, we sought to identify a CT based radiomics texture signature that can robustly predict treatment response to immunotherapy, specifically pembrolizumab, in patients with advanced rare solid cancers. We used CT-based radiomics features to create a non-invasive predictive model for classifying 57 patients with advanced rare cancers as likely “controlled disease” or “progressive disease” to immunotherapy. To the best of our knowledge, this is the first study to predict pembrolizumab treatment response in advanced rare cancers.

## MATERIALS AND METHODS

### Study design and patient population

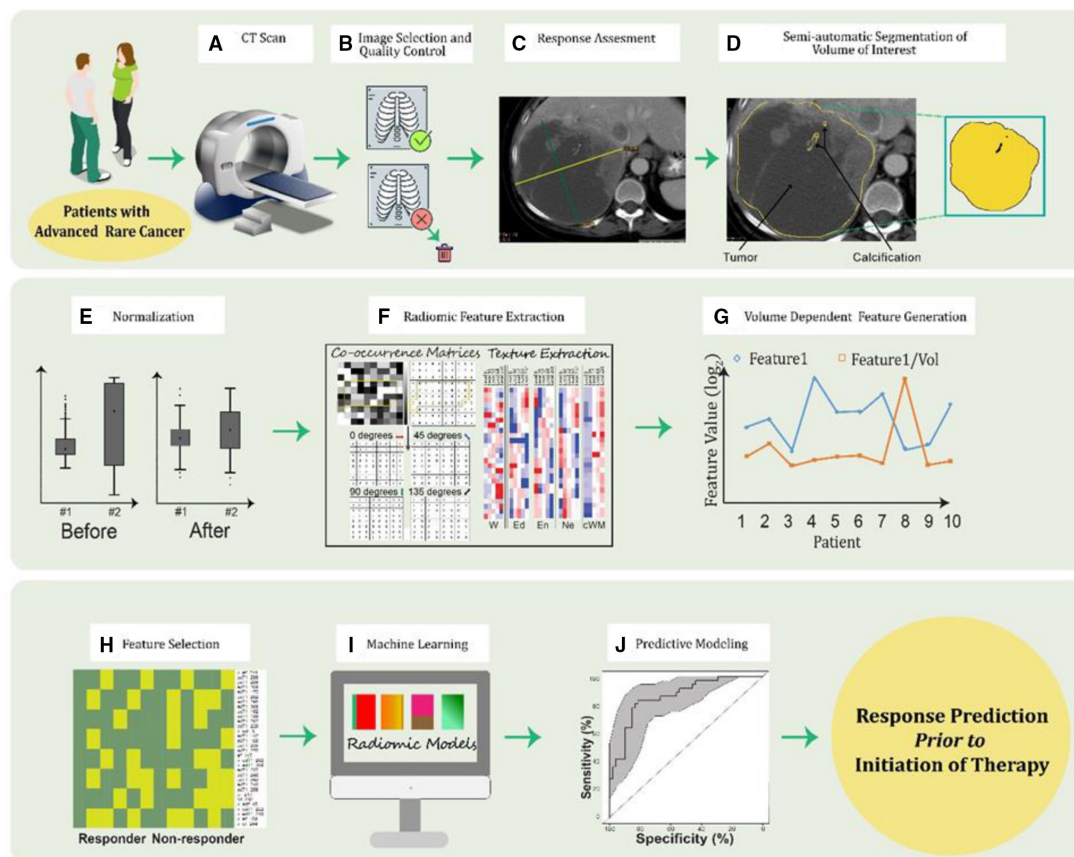
This retrospective study was conducted in compliance with US Health Insurance Portability and Accountability Act regulations. All patients included in the study were enrolled in our phase II clinical trial of pembrolizumab (ClinicalTrials.gov number NCT02721732) at UTMDACC between August 2016 and July 2018; the safety and efficacy outcomes for that trial were reported previously.<sup>9</sup> Inclusion criteria for the present study were age  $\geq 18$  years; histopathologically proven rare cancer; availability

of contrast-enhanced CT scans obtained prior to anti-programmed cell death protein 1 (PD-1) treatment; and absence of CT image artifacts (such as motion or streak artifacts) that would hinder imaging analysis. Written informed consent was obtained before the initiation of the clinical trial from all patients, and the patients' clinical, radiological, and histopathological data were retrieved from the medical records.

### Imaging analysis

During the clinical trial, tumor response to pembrolizumab was evaluated using both RECIST 1.1 and irRECIST. At that time, patients were categorized as either controlled disease (defined as patients demonstrating stable disease, partial response, or complete response) or progressive disease.

For the radiomics analysis, target lesions on the contrast-enhanced CT scans obtained before anti-PD-1 treatment were segmented using the 3D Slicer 4.8.1 module (<http://slicer.kitware.com/midas3/item/330416>). Lesion segmentation was performed and subsequently reviewed by a board-certified radiologist (R.R.C.; >10-year experience). Each target lesions were segmented slice by slice and given a color label using 3D Slicer v4.8.1. In patients who had more than one target lesion, all lesions were segmented in the same manner and given the same color label in order to predict response of target lesion tumor burden. Additional volumes of interest were segmented for within-phase normalization; the normal pectoralis major muscle was segmented for chest lesion normalization, and the normal paraspinal muscle was segmented for abdominal/pelvic lesion normalization. The contour outlines of the segmented tumors and the normal muscle were saved as an output label volume for radiomics analysis. Radiomics analysis was performed as we described previously.<sup>27</sup> Intensity-level histogram and gray-level co-occurrence matrices (GLCMs) were used for radiomics feature extraction. In brief, a total of 10 first-order features were extracted based on Papoulis *et al*<sup>34</sup> and include: minimum, maximum, mean, SD, skewness, kurtosis, and four percentiles (1%, 5%, 95%, 99%). A total of 300 GLCM second-order features were extracted based on Haralick *et al*,<sup>35</sup> Clausi,<sup>36</sup> and Soh *et al*,<sup>37</sup> and include: autocorrelation, contrast, correlation, cluster shade, cluster prominence, dissimilarity, energy, entropy, homogeneity, maximum probability, variance, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation 1 and 2, inverse difference moment, and normalized inverse difference moment. Each feature is calculated in four angles (angle between the voxel of interest and the in-plane neighbor) of 0°, 45°, 90°, and 135°; rotation-invariant features were calculated using average, median and range resulting in 60 rotation-invariant features. To increase the signal-to-noise ratio, original images discretized into five grey-levels (8, 16, 32, 64, and 256 grey-levels) at which we obtain the 60 rotation-invariant features. Finally, because the GLCM



**Figure 1** Radiomic pipeline for advanced rare tumor.

is not dependent on the volume of interest (VOI), we computed 300 volume-dependent second-order features by dividing each second-order feature by the volume of the segmented VOI; thus, a total of 610 radiomics features were obtained. Radiomics analysis was performed using our in-house software in matlab (version 2017b; MathWorks; Natick, Massachusetts, USA) (figure 1).

### Bioinformatic analysis

Given the large multi-dimensional dataset, we used regression analysis with the least absolute shrinkage and selection operator (LASSO) feature selection method to find the most meaningful predictors of response.<sup>38</sup> The selected features were used to build a classification model for the prediction of tumor response to pembrolizumab; with XGBoost (eXtreme Gradient Boosting) machine-learning algorithm as it offers various potential benefits such as implementing tree-boosting technique, enabling regularization (L1–L2), cross-validation, parallel computation, missing value imputation, and tree pruning.<sup>39</sup> It is recognized as the most popular and widely used tool in many competitions held worldwide primarily due to its nature in handling the overfitting of the model through L2 or also called as ridge regularization. The tree-boosting techniques in XGBoost are implemented using gradient boosting machines through which the boosted trees are said to continuously learn from the new data that are added

during the training of the model. All these attributes within the XGBoost algorithm eventually accounts for the improvement in the model's performance and computational speed. To look for the effectiveness of the XGBoost algorithm, we built classification models with other classifiers like Support Vector Machines, Decision Trees and Random Forest, and compared the XGBoost's model performance with the model performance from other classifiers.

Since, we had a small dataset with 57 patients, we used leave-one-out cross validation (LOOCV) method to evaluate the robustness of the estimates as the cross-validation technique for building our radiomics models. In the LOOCV process, we are holding out a single patient sample as the test sample and then fitting a model on the remaining patient samples that serves as the training test using the radiomics texture features from the LASSO feature selection. After the training, we are obtaining the prediction on the held-out patient sample or the test sample. In order that we develop our model for the future clinical implementation, we thus selected our final model to be the one that was built with least number of texture features and the one that required less computation time. Each model's accuracy was predicted using the area under the receiver operating characteristic [ROC] curve (AUC), sensitivity, and specificity; p value was also obtained.



Additionally, we performed LASSO feature selection and LOOCV simultaneously for prediction of the response and calculated AUC, sensitivity, specificity, and p value. R software (V.3.4.0, R Foundation for Statistical Computing, Vienna, Austria) was used for all of the analysis: XGBoost package (V.0.6.4.1) and the Machine Learning package mlr (V.2.11) were used for feature selection and model building and finally, ROC analysis was performed using pROC package (V.1.9.1).

### Univariate and multivariate analysis

Shapiro-Wilk's test was applied to analyze the distribution of the continuous variables (top 10 LASSO features from RECIST radiomics model, top 10 LASSO features from irRECIST radiomics model and age) before deciding the appropriate test for univariate and multivariate analysis.

For multivariate analysis, multivariate ANOVA test was performed on 57 patients by considering all the top 10 LASSO features collectively from the RECIST and irRECIST models to see if there are any statistically significant differences between the patient's groups in terms of clinical variables (gender, race, and number of prior therapies).

Univariate analysis was performed to assess the differences between the clinical covariates for the features used in radiomics models. One-way ANOVA test was used for clinical variables (ie, gender, race, number of prior therapies), and Pearson correlation coefficient was performed for the continuous variable (age).

Additionally, Welch two sample t-test was applied to analyze if the radiomics features used for model generation were independently associated with the treatment response.

### Univariate and multivariate analysis with overall survival and radiomics texture features

A univariate analysis was performed using univariate Cox regression model to assess the significant difference and independent association of each of the 10 radiomics texture features that were used in our predictive RECIST and irRECIST models with overall survival (OS). The OS for the 57 patients were calculated in months and the OS status (0=alive and 1=deceased) was assigned for each patient. The radiomics texture features that were significant from the univariate Cox regression model were later used for constructing a multivariate Cox regression model to predict the OS outcome. Further, we categorized the 10 radiomics texture features from the RECIST and irRECIST models into high versus low groups using their median as the cut-off. Using the univariate Kaplan-Meier curves and the log-rank test, we analyzed for the association between high versus low groups of each radiomics texture feature with OS.

## RESULTS

### Patient population

Of the 127 patients with advanced rare cancers enrolled in the pembrolizumab clinical trial, 70 patients were excluded; contrast-enhanced CT scans obtained prior to anti-PD-1 treatment was not available for 48 patients

**Table 1** Patients' demographic characteristics (N=57)

Characteristic	No. of Patients (%)
<b>Age at diagnosis, years</b>	
Mean±SD	47.16±16.70
Median (range)	53.00 (22–75)
<b>Sex</b>	
Female	26 (46)
Male	31 (54)
<b>Race/ethnicity</b>	
African American	4 (7)
Asian	2 (4)
White	42 (74)
Other	8 (14)
Unknown	1 (2)
<b>ECOG performance status</b>	
1	49 (86)
0	8 (14)
<b>Number of prior therapies</b>	
≤2	37 (65)
>2	20 (35)

All data are no. of patients (%) unless otherwise noted.

ECOG, Eastern Cooperative Oncology Group.

and 22 patients had CT image artifacts that would hinder imaging analysis. Fifty-seven patients met the inclusion criteria for the present study. Their demographic data are outlined in [table 1](#). A total of 37 patients were classified as controlled disease and 20 patients as progressive disease per evaluation by RECIST. Similarly, 33 patients were classified as progressive disease and 24 patients as controlled disease per evaluation by irRECIST. Patients' responses based on RECIST and irRECIST by cancer type are given in [table 2](#).

### Radiomics predicts response

A total of 610 features (10 histogram-based features and 600 second-order texture features) were calculated from each extracted VOI. Among these 610 radiomics features, LASSO feature selection identified 44 features that were most discriminatory between the response outcomes in patients assessed by RECIST ([table 2](#)). Of these 44 selected features, the most relevant, most significant 10 features ([table 3](#)) robustly discriminated controlled disease and progressive disease using a classification model based on the XGBoost machine-learning method. LOOCV demonstrated that our radiomics model had high accuracy, sensitivity, and specificity (94.7%, 97.3%, and 90%, respectively;  $p < 0.001$ ) ([figure 2A](#)). When patients were grouped according to irRECIST ([table 2](#)), LASSO feature selection identified 56 features. XGBoost classification model using the 10 top discriminatory features ([table 4](#)) predicted controlled disease and progressive disease with

**Table 2** Pembrolizumab response status based on RECIST and irRECIST for 57 patients with rare cancers

Primary cancer	Based on RECIST		Based on irRECIST	
	Controlled disease (N=20)	Non-responders (N=37)	Controlled disease (N=24)	Non-responders (N=33)
Adrenocortical carcinoma	5 (25)	4 (11)	6 (25)	3 (9)
Carcinoma of unknown primary	2 (10)	6 (16)	4 (12)	4 (12)
Medullary renal carcinoma	1 (5)	3 (8)	1 (4)	3 (9)
Other rare histologies*	5 (25)	5 (15)	5 (20)	5 (15)
Paraganglioma-pheochromocytoma	3 (15)	2 (5)	3 (13)	2 (6)
Penile carcinoma	0 (0)	1 (3)	0 (0)	1 (3)
Small cell malignancies of non-pulmonary origin	0 (0)	3 (8)	1 (4)	2 (6)
Spindle cell sarcoma of retroperitoneum†	0 (0)	1 (3)	0 (0)	1 (3)
Squamous cell carcinoma of the skin	1 (5)	3 (8)	1 (4)	3 (9)
Testicular carcinoma/germ cell tumor	2 (10)	7 (19)	2 (8)	7 (21)
Vascular sarcoma	1 (5)	2 (5)	1 (4)	2 (6)

All data are no. of patients (%).

\*Vaginal squamous cell carcinoma and others.

†The one case of spindle cell sarcoma of retroperitoneum was initially diagnosed as adrenocortical carcinoma.

irRECIST, immune-related RECIST; RECIST, Response Evaluation Criteria in Solid Tumors.

robust accuracy, sensitivity, and specificity (94.7%, 93%, and 95%, respectively;  $p < 0.001$ ) (figure 2B).

Fifteen features were common to both the group of 44 discriminatory features for patients assessed by RECIST and the group of 56 discriminatory features for patients assessed by irRECIST (figure 3B, table 5). The predictive model generated with these 15 features, effectively discriminated between controlled disease and progressive disease and achieved high accuracy, sensitivity, and specificity in patients assessed by RECIST (94.7%, 97%, and 90%, respectively;  $p < 0.001$ ) and those assessed by irRECIST (96%, 96%, and 95%, respectively;  $p < 0.001$ ) (figure 3A,C).

In the radiomics analysis with LASSO and LOOCV performed concurrently, we achieved an accuracy, sensitivity, and specificity in patients assessed by RECIST (89.7%, 70%, and 100%, respectively;  $p < 0.001$ ) and those assessed by irRECIST (87.7%, 87.5%, and 87.8%,

respectively;  $p < 0.001$ ) (online supplemental figure S1A,B, respectively).

Online supplemental table 1 shows the accuracy, sensitivity, specificity, AUC, and  $p$  value for the models generated by using other classifiers (support vector machines, decision tree, and random forest) to predict treatment response based on RECIST and irRECIST.

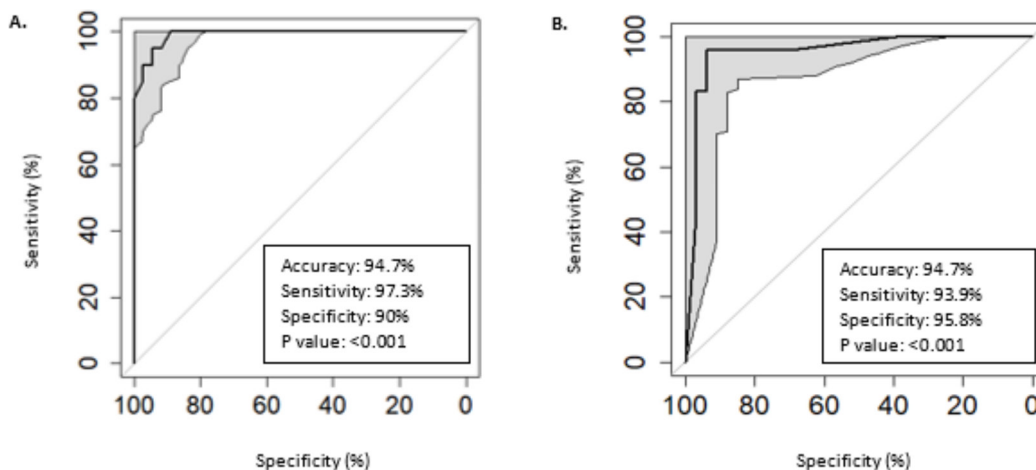
### Univariate and multivariate analysis

Shapiro-Wilk's test demonstrated that 10 lasso features from RECIST and irRECIST model are normally distributed whereas the age is not normally distributed (online supplemental table 2).

Multivariate analysis determined that all the top 10 LASSO features (collectively) from the RECIST model does not account for any statistically significant ( $p$  value  $< 0.05$ ) differences between the patient groups with regard to number of prior therapies, race, and gender (online

**Table 3** Top 10 features (from 44 features extracted in patients assessed by Response Evaluation Criteria in Solid Tumors)

Order No	Feature	Level	Feature name
1	PV_F109	16	Angular variance of homogeneity
2	PV_F35	8	Range of difference variance
3	Ar_F152	32	Range of sum average
4	P_F269	32	Angular variance of homogeneity
5	Ar_F154	32	Range of sum entropy
6	P_F84	16	Range of cluster prominence
7	PV_F89	16	Range of homogeneity
8	P_F289	256	Angular variance of homogeneity
9	Ar_F238	64	Angular variance of information measure of correlation 2
10	P_F7	8	Average of energy



**Figure 2** (A) Receiver operating characteristic (ROC) curve representing the performance of the predictive model when using the top 10 least absolute shrinkage and selection operator (LASSO) features in the Response Evaluation Criteria in Solid Tumors (RECIST) group. (B) ROC curve representing the performance of the predictive model when using the top 10 LASSO features in the immune-related RECIST group.

supplemental table 3A). Multivariate analysis showed that all the top 10 LASSO features (collectively) from the irRECIST model does not account for any statistically significant ( $p$  value  $<0.05$ ) differences between the patient groups in terms of number of prior therapies and race however; these features account for statistically significant difference between men and women as demonstrated in online supplemental table 3B.

Univariate analysis over clinical variables versus discriminatory 10 radiomics features (from 44 features extracted in patients assessed by RECIST) that used for radiomics model showed that there are no statistically significant differences with respect to the clinical variables ( $p$  value  $<0.05$ ). However, for the features; range of cluster prominence (P\_F 84) and range of sum average (Ar\_F152), one-way ANOVA test showed that these features are statistically significant in terms of the gender (online supplemental table 4A). Similarly, there are no significant differences between the meaningful 10 radiomics features (from 56 features extracted in patients assessed by irRECIST) that used for radiomics model with respect to the clinical

variables ( $p$  value  $<0.05$ ). Except, for the range of sum average (Ar\_F152 feature), ANOVA test demonstrated that this feature is statistically significant in the sense of the gender (online supplemental table 4B).

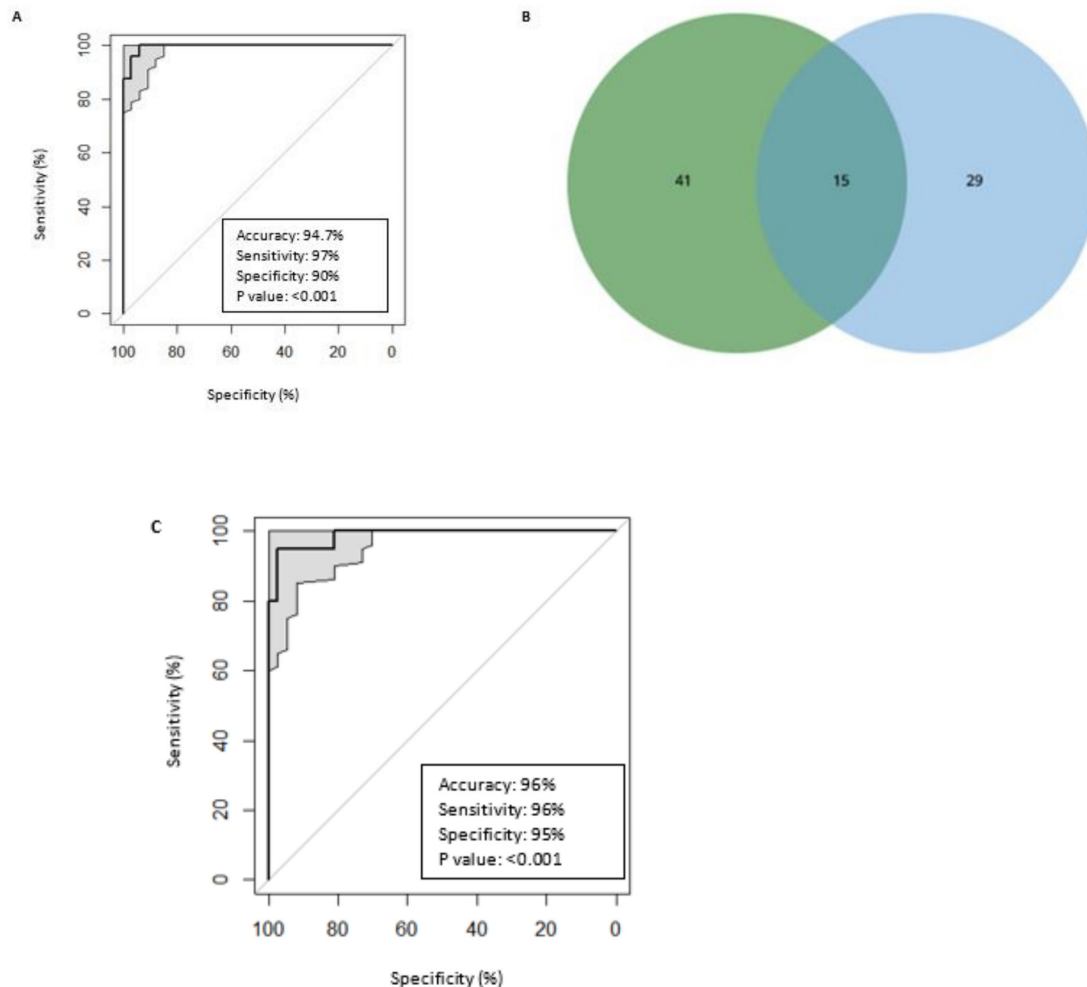
Welch two sample t-test showed that 5 out of the 10 LASSO features from RECIST model (P\_F269, P\_F289, PV\_F35, PV\_F89, and PV\_F109) and 4 out of the 10 LASSO features from irRECIST model (P\_F89, P\_F269, P\_F289, and PV\_F81) are statistically significant ( $p$  value  $<0.05$ ) and are independently associated with treatment response ( $p$  value  $<0.05$ ) as shown in online supplemental table 5.

#### Univariate and multivariate analysis with OS and radiomics texture features

Of the 57 patients, 13 patients were identified to be alive and 44 patients were deceased. The univariate Cox regression models built by 10 radiomics texture features from RECIST model with OS had identified 6 out of 10 texture features to have significant association with OS as shown in online supplemental table 6. The  $p$  value for all

**Table 4** Top 10 features (from 56 features extracted in patients assessed by immune-related Response Evaluation Criteria in Solid Tumors)

Order no.	Feature	Level	Feature name
1	PV_F24	8	Range of cluster prominence
2	Ar_F152	32	Range of sum average
3	PV_F81	16	Range of autocorrelation
4	P_F89	16	Range of homogeneity
5	PV_F270	256	Range of maximum probability
6	ArV_F247	256	Average of energy
7	P_F187	64	Average of energy
8	P_F289	256	Angular variance of homogeneity
9	P_F269	256	Range of homogeneity
10	ArV_F36	8	Range of difference entropy



**Figure 3** (A) Receiver operating characteristic (ROC) curve representing the performance of the predictive model when using the common 15 least absolute shrinkage and selection operator (LASSO) features in the Response Evaluation Criteria in Solid Tumors (RECIST) group. (B) Fifteen common LASSO features between 44 features and 56 features obtained from 57 patients assessed by RECIST and immune-related RECIST (irRECIST), respectively. (C) ROC curve representing the performance of the predictive model when using the common 15 LASSO features in the irRECIST group.

three overall tests (likelihood (p value=0.001), Wald (p value=0.03), and score (p value=0.02)) from the multivariate Cox regression model are significant, indicating that the multivariate Cox regression model created with the six significant radiomics texture features from RECIST model with OS was significant, as shown in online supplemental table 7. From the multivariate Cox analysis as shown in online supplemental table 7, the texture feature PV\_F35 remain significant (p value=0.0493). The univariate log-rank test for the high versus low groups of the 10 radiomics texture features from RECIST model with OS showed that, 4 out of 10 texture features; P\_F289 (p value=0.0014), PV\_F35 (p value=0.0045), PV\_F89 (p value=0.016), and Ar\_F238 (p value=0.011) were found to be significantly associated with high versus low groups texture feature with OS, as shown in online supplemental table 8. The Kaplan-Meier survival curves and p values from the log-rank test for the four significant radiomics texture features from RECIST can be seen in figure 4A.

The univariate Cox regression models generated by 10 radiomics texture features from irRECIST model with OS had determined 5 out of 10 texture features to have significant association with OS; as demonstrated in online supplemental table 9. The p value for all three overall tests (likelihood (p value=8e-05), Wald (p value=0.04), and score (p value=6e-04)) from the multivariate Cox regression model are significant, indicating that the multivariate Cox regression model built using the five significant radiomics texture features from irRECIST model with OS was significant, as detailed in online supplemental table 10. From the multivariate Cox analysis as shown in online supplemental table 10, two of the texture features; PV\_F81 (p value 0.00696) and ArV\_F36 (p value=0.02417) remain significant. The univariate log-rank test for the high versus low groups of the 10 radiomics texture features from irRECIST model with OS showed that, 3 out of 10 texture features; P\_F289 (p value=0.0014), PV\_F81 (p value=0.0071), and PV\_F270 (p value=0.048) were found to be significantly associated with high versus low groups



**Table 5** Shared 15 features (between features obtained from patients assessed by Response Evaluation Criteria in Solid Tumors (RECIST) and immune-related RECIST)

Order no.	Feature	Level	Feature name
1	Ar_F152	32	Range of sum average
2	ArV_F247	256	Average of energy
3	P_F289	256	Angular variance of homogeneity
4	P_F269	256	Range of homogeneity
5	Ar_F154	32	Range of sum entropy
6	P_F153	32	Range of sum variance
7	Ar_F188	64	Average of entropy
8	ArV_F51	8	Angular variance of sum of squares: variance
9	P_F114	16	Angular variance of sum entropy
10	ArV_F248	256	Average of entropy
11	ArV_F130	32	Average of maximum probability
12	ArV_F14	8	Average of sum entropy
13	Ar_F238	64	Angular variance of information measure of correlation 2
14	Ar_F275	256	Range of difference variance
15	Ar_F258	256	Average of information measure of correlation 2

texture feature with OS, as shown in online supplemental table 11. The Kaplan-Meier survival curves and p values from the log-rank test for the three significant radiomics texture features from irRECIST can be seen in [figure 4B](#).

## DISCUSSION

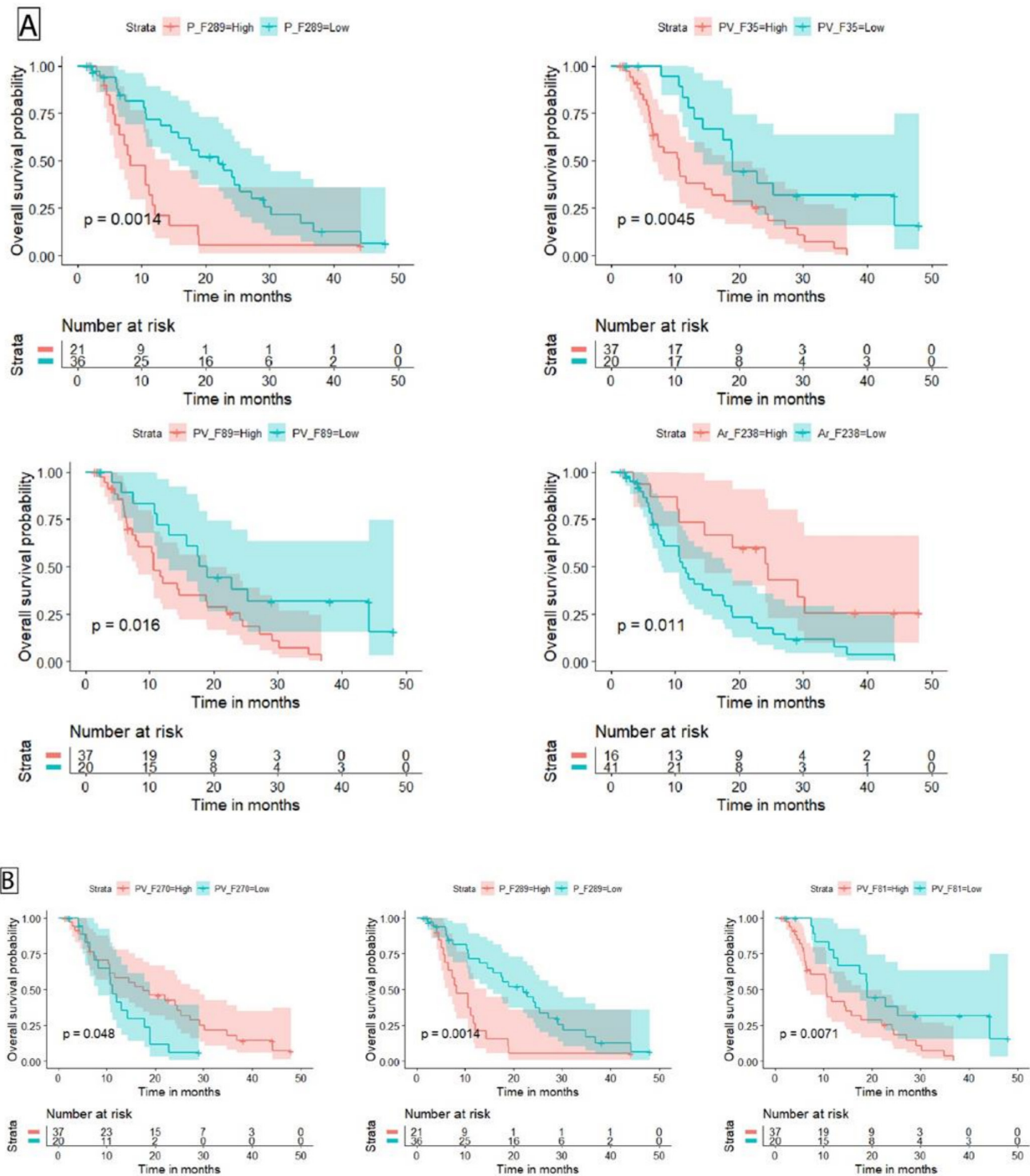
The results of our study demonstrate that quantitative radiomics texture analysis on standard-of-care CT scans obtained prior to the initiation of therapy can successfully predict pembrolizumab response. To our knowledge, this is the first study to use a radiomics-based approach to predict immunotherapy response in patients with advanced rare solid cancers. Radiomic signatures based on top 10 LASSO-selected features predicted immunotherapy response with high accuracy, sensitivity, and specificity in patients assessed by RECIST (94.7%, 97.3%, and 90%, respectively) and in patients assessed by irRECIST (94.7%, 93.9%, and 95.8%, respectively). Furthermore, the shared features of the RECIST and irRECIST groups also highly predicted pembrolizumab response with the accuracy, sensitivity, specificity, and p value of 94.7%, 97%, 90%,  $p < 0.001$  and 96%, 96%, 95%,  $p < 0.001$ , respectively.

The future clinical use of our radiomics-based approach to predict immunotherapy response is supported by previous studies in patients with other diseases.<sup>29 32 40</sup> Sun *et al*<sup>29</sup> generated a radiomics-based treatment response prediction model using machine learning. They found that their baseline radiomics score was able to predict CD8 gene expression. Patients with higher radiomics score had higher CD8 gene expression and better response to immunotherapy compared with patients with progressive disease or stable disease ( $p = 0.049$ ); the baseline radiomic score however, was not significantly higher in the patients with controlled disease (stable disease, partial response,

complete response) compared with patients with progressive disease ( $p = 0.05$ ). However, that latter study was for the most part correlative and the sensitivity, specificity, and accuracy of their radiomic model was not reported. In another retrospective study, Mu *et al*<sup>40</sup> described a multi-parametric radiomic signature based on positron emission tomography (PET)/CT images to predict durable clinical benefit of immunotherapy response in patients with advanced NSCLC (AUC ranging from 0.81 to 0.86). They found that compared with that of the multi-parametric radiomics signature, the AUCs of the CT-only-based features used to predict response were lower (AUCs ranging from 0.64 to 0.69). These AUCs are also lower compared with our study (0.95 and 0.99 in the irRECIST and RECIST groups, respectively), likely because the CT images used from PET/CT studies have lower resolution than the diagnostic CT studies as used in the present study and were non-contrast enhanced CTs.

Finally, Trebeschi *et al*,<sup>32</sup> looked at radiomics to assess immunotherapy response (rule out progression) in patients with NSCLC and melanoma. The model was able to predict progressive disease in NSCLC primary tumor with an AUC of 0.79 and  $p = 0.05$ . But, their radiomics based prediction model achieved poor performance on both pulmonary and hepatic melanoma lesions (AUC of 0.55). The authors did not report the sensitivity and specificity of the models. These results are consistent with our study in terms of highlighting the potential of CT-based radiomics features in predicting response to immunotherapy. In our cohort, we assessed immunotherapy response by analyzing each patient's entire (whole body) tumor burden. That was achieved by segmentation of all target lesions (defined by RECIST 1.1 and irRECIST) and further extraction of distinct radiomics features





**Figure 4** (A) The Kaplan-Meier curves for overall survival (OS) with radiomic texture features from Response Evaluation Criteria in Solid Tumors (RECIST) model. (B) The Kaplan-Meier curves for OS with radiomic texture features from immune-related RECIST model.

as illustrated in online supplemental figures S2-4. The high accuracy of our model for classifying controlled disease and progressive disease in the RECIST and irRECIST based groups suggests that radiomics analysis has a potential role in making clinical decisions and in selecting patients who would likely benefit from therapy prior to the initiation of such treatment. Additionally, the radiomics approach is cost and time-effective, as it uses

standard-of-care images that are already being obtained for diagnosis and follow-up.

The present study had some inherent limitations. The rare cancer study had a heterogeneous patient population with biologically different diseases. However, despite the latter, our radiomics model was highly predictive of patient response to anti-PD-1 therapy. In addition, given the nature of the clinical trial study, the sample size was

relatively small. Finally, although promising, our single-institution cohort was retrospectively evaluated; therefore, validation is required, preferably prospectively with a larger data set acquired from multiple institutions. These validation studies using a larger prospective data set in a multi-institutional patient population are currently underway at our institution to ensure reliability and generalizability of our model.

In summary, this study presents radiomics-based signature that can successfully predict immunotherapy response in patients with advanced rare cancers. The proposed predictive radiomics-based signature is cost-effective, non-invasive and can potentially play a significant role in the clinical setting, allowing oncologists to predict treatment response a priori and enable better patient stratification, management, and clinical decision-making. Although our predictive models had high accuracy in retrospectively discriminating controlled disease and progressive disease to pembrolizumab, prospective studies on a larger data set in a more homogeneous patient population are needed to ensure its reliability and generalizability.

#### Author affiliations

<sup>1</sup>Department of Radiology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

<sup>2</sup>Hillman Cancer Center, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA

<sup>3</sup>Department of Thoracic Medical Oncology, University of Maryland, Baltimore, Maryland, USA

<sup>4</sup>Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>5</sup>Department of Breast Imaging, Diagnostic Radiology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>6</sup>Department of Neurosurgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA

<sup>7</sup>Abdominal Imaging Department, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>8</sup>Radiology, Pathology, and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

<sup>9</sup>Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>10</sup>Department of Investigational Cancer Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>11</sup>Section of Immunology, Allergy and Retrovirology, Baylor College of Medicine, Houston, Texas, USA

<sup>12</sup>William T Shearer Center for Human Immunobiology, Texas Children's Hospital, Houston, Texas, USA

**Twitter** Nabil Elshafeey @naelshafeey@mdanderson.org and Aung Naing @AningMD

**Acknowledgements** Sunita Patterson and Joseph Munch, Department of Scientific Publications at MD Anderson, provided editorial assistance. We thank the patients and their families and caregivers for participating in the study.

**Contributors** RRC contributed to conception and design, provision of study materials or patients, collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. CR contributed to conception and design, provision of study materials or patients, collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. MAk contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. MAYoub contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. SA contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. NE contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. PM

contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. POZ contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. CN contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. RV contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. SB contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. CP contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. JRA contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. VS contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. DDK contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. BS contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. JH contributed to manuscript writing, final approval of manuscript, and is accountable for all aspects of the work. AN contributed to conception and design, provision of study materials or patients, collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript, and is accountable for all aspects of the work.

**Funding** Merck Sharp & Dohme Corp., a subsidiary of Merck & Co. provided the study drug and funded the study. This research was funded in part by startup funding from The University of Pittsburgh Hillman Cancer Center (RRC) and the Hillman Cancer Center's NCI Cancer Center Support Grant (CCSG) P30CA047904 (RRC). Support was also provided by the National Institutes of Health/National Cancer Institute under award number P30CA016672 to The University of Texas MD Anderson Cancer Center (used the Biostatistics Resource Group) and The University of Texas MD Anderson Cancer Center Institutional Research Grant (IRG) program.

**Competing interests** CN reports grant support and personal fees from General Electric Healthcare, outside the submitted work. SB reports grant support from National Institutes of Health, outside the submitted work. JRA reports personal fees from Novartis, Eli Lilly, Orion Pharmaceuticals, Servier Pharma, Peptomyc, and Merck Sharpe, on the advisory board for Novartis, Eli Lilly, Orion Pharmaceuticals, Servier Pharma, Peptomyc, Merck Sharpe & Dome, Kelun Pharma/Klus Pharma, Pfizer, Roche Pharma, and Elipses Pharma, research funding from Bayer, Novartis, Spectrum Pharmaceuticals, Tocagen, Symphogen, BioAtla, Pfizer, GenMab, CytomX, KELUN-BIOTECH, Takeda-Millennium, GLAXOSMITHKLINE, Ipsen, from null, outside the submitted work. VS reports clinical trial research funding from Novartis, Bayer, GlaxoSmithKline, Nanocarrier, Veogenics, Celgene, Northwest Biotherapeutics, Berghealth, Incyte, Fujifilm, Pharmamar, D3, Pfizer, Multivir, Amgen, Abbvie, Alfa-sigma, Agensys, Boston Biomedical, Idera Pharma, Inhibrx, Exelixis, Blueprint medicines, Loxo oncology, Takeda and Roche/ Genentech, National Comprehensive Cancer Network, NCI-CTEP and UT MD Anderson Cancer Center, outside the submitted work. JH reports grants from Immune Deficiency Foundation, Jeffery Modell Foundatoin and Chao Physician-Scientist, and Baxalta, and has served as an advisory board member for Takeda, CSL Behring, and Horizin Pharma outside the submitted work. AN reports research support and non-financial support from Merck Sharp & Dohme Corp., grants from NCI/NIH, research support from The University of Texas MD Anderson Cancer Center, during the conduct of the study; grants from NCI, research support from EMD Serono, MedImmune, Healios Onc. Nutrition, Attercor, Amplimmune, ARMO BioSciences, Karyopharm Therapeutics, Incyte, Novartis, Regeneron, Merck, Bristol Myers Squibb, Pfizer, CytomX Therapeutics, Neon Therapeutics, Calithera BioSciences, TopAlliance BioSciences, Eli Lilly, Kymab, PsiOxus, Arcus Biosciences, NeolmmuneTech, ImmuneOncia, and Surface Oncology, non-financial support for travel and accommodation from ARMO BioSciences, and has served as an advisory board member for Novartis, CytomX Therapeutics, Genome and Company, STCube Pharmaceuticals, OncoSec KEYNOTE-695, and Kymab outside the submitted work. RRC, CR, MAK, MAYoub, SA, NE, PM, POZ, RV, CP, BS, DDK declare no competing interests.

**Patient consent for publication** Not required.

**Ethics approval** The protocol was approved by the FDA and the Institutional Review Board at The University of Texas MD Anderson Cancer Center. The study was conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonization Good Clinical Practice guidelines. All the study participants provided written informed consent before enrollment.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request and approval from study sponsor according to available guidelines at time of request.

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

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

#### ORCID iD

Aung Naing <http://orcid.org/0000-0002-4803-8513>

#### REFERENCES

- Bogaerts J, Sydes MR, Keat N, *et al*. Clinical trial designs for rare diseases: studies developed and discussed by the International rare cancers initiative. *Eur J Cancer* 2015;51:271–81.
- Greenlee RT, Goodman MT, Lynch CF, *et al*. The occurrence of rare cancers in U.S. adults, 1995–2004. *Public Health Rep* 2010;125:28–43.
- Institute NC. Cancer STAT facts. Surveillance, epidemiology, and end results program, 2020. Available: <https://seercancer.gov/statfacts/>
- Komatsubara KM, Carvajal RD. The promise and challenges of rare cancer research. *Lancet Oncol* 2016;17:136–8.
- Ansell SM, Lesokhin AM, Borrello I, *et al*. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med* 2015;372:311–9.
- Reck M, Rodriguez-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.
- Ribas A, Hamid O, Daud A, *et al*. Association of pembrolizumab with tumor response and survival among patients with advanced melanoma. *JAMA* 2016;315:1600–9.
- Goldberg SB, Gettinger SN, Mahajan A, *et al*. Pembrolizumab for patients with melanoma or non-small-cell lung cancer and untreated brain metastases: early analysis of a non-randomised, open-label, phase 2 trial. *Lancet Oncol* 2016;17:976–83.
- Naing A, Meric-Bernstam F, Stephen B, *et al*. Phase 2 study of pembrolizumab in patients with advanced rare cancers. *J Immunother Cancer* 2020;8:e000347.
- Rosenberg JE, Hoffman-Censits J, Powles T, *et al*. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* 2016;387:1909–20.
- Nghiem PT, Bhatia S, Lipson EJ, *et al*. PD-1 blockade with pembrolizumab in advanced Merkel-cell carcinoma. *N Engl J Med* 2016;374:2542–52.
- Ferris RL, Blumenschein G, Fayette J, *et al*. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl J Med* 2016;375:1856–67.
- Eroglu Z, Zaretsky JM, Hu-Lieskovan S, *et al*. High response rate to PD-1 blockade in desmoplastic melanomas. *Nature* 2018;553:347–50.
- Fujii T, Naing A, Rolfo C, *et al*. Biomarkers of response to immune checkpoint blockade in cancer treatment. *Crit Rev Oncol Hematol* 2018;130:108–20.
- Puzanov I, Diab A, Abdallah K, *et al*. Managing toxicities associated with immune checkpoint inhibitors: consensus recommendations from the Society for immunotherapy of cancer (SITC) toxicity management Working group. *J Immunother Cancer* 2017;5:95.
- Taube JM, Klein A, Brahmer JR, *et al*. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 2014;20:5064–74.
- Hamid O, Schmidt H, Nissan A, *et al*. A prospective phase II trial exploring the association between tumor microenvironment biomarkers and clinical activity of ipilimumab in advanced melanoma. *J Transl Med* 2011;9:204.
- Higgs BW, Morehouse C, Streicher K, *et al*. Relationship of baseline tumoral IFN $\gamma$  mRNA and PD-L1 protein expression to overall survival in durvalumab-treated NSCLC patients. *Journal of Clinical Oncology* 2016;34:3036.
- Le DT, Uram JN, Wang H, *et al*. Pd-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015;372:2509–20.
- Hellmann MD, Callahan MK, Awad MM, *et al*. Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. *Cancer Cell* 2018;33:853–61.
- Marcus L, Lemery SJ, Keegan P, *et al*. FDA approval summary: pembrolizumab for the treatment of microsatellite Instability-High solid tumors. *Clin Cancer Res* 2019;25:3753–8.
- Davis AA, Patel VG. The role of PD-L1 expression as a predictive biomarker: an analysis of all US food and drug administration (FDA) approvals of immune checkpoint inhibitors. *J Immunother Cancer* 2019;7:278.
- Bohnsack O, Hoos A, Ludajic K. Adaptation and modification of the immune related response criteria (IRRC): IrRECIST. *Journal of Clinical Oncology* 2014;32:e22121-e:e22121.
- Wolchok JD, Hoos A, O'Day S, *et al*. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin Cancer Res* 2009;15:7412–20.
- Lambin P, Rios-Velazquez E, Leijenaar R, *et al*. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441–6.
- Braman N, Prasanna P, Whitney J, *et al*. Association of Peritumoral Radiomics With Tumor Biology and Pathologic Response to Preoperative Targeted Therapy for HER2 (ERBB2)-Positive Breast Cancer. *JAMA Netw Open* 2019;2:e192561.
- Zinn PO, Singh SK, Kotrotsou A, *et al*. A Coclinical Radiogenomic validation study: conserved magnetic resonance radiomic appearance of Periostin-Expressing glioblastoma in patients and xenograft models. *Clin Cancer Res* 2018;24:6288–99.
- Knief HC, Madesta F, Schneider T, *et al*. Radiomics of brain MRI: utility in prediction of metastatic tumor type. *Radiology* 2019;290:479–87.
- Sun R, Limkin EJ, Vakalopoulou M, *et al*. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* 2018;19:1180–91.
- Braman N, Prasanna P, Whitney J, *et al*. Association of Peritumoral Radiomics With Tumor Biology and Pathologic Response to Preoperative Targeted Therapy for HER2 (ERBB2)-Positive Breast Cancer. *JAMA Netw Open* 2019;2:e192561-e:e192561.
- Colen RR, Ologun GO, Zinn P, *et al*. Radiomic signatures to predict response to targeted therapy and immune checkpoint blockade in melanoma patients (PTS) on neoadjuvant therapy. *Journal of Clinical Oncology* 2020;38:10067.
- Trebeschi S, Drago SG, Birkbak NJ, *et al*. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol* 2019;30:998–1004.
- 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:canimm.0476.2019.
- Papoulis A. *Probability, random variables, and stochastic processes*, 1965.
- Haralick RM, Shanmugam K, Dinstein Its'Hak, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;SMC-3:610–21.
- Clausi DA. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing* 2002;28:45–62.
- Soh L-K, Tsatsoulis C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing* 1999;37:780–95.
- Tibshirani R. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B* 1996;58:267–88.
- Chen T, Guestrin C. Xgboost: a scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016:785–94.
- Mu W, Tunali I, Gray JE, *et al*. Radiomics of  $^{18}\text{F}$ -FDG PET/CT images predicts clinical benefit of advanced NSCLC patients to checkpoint blockade immunotherapy. *Eur J Nucl Med Mol Imaging* 2020;47:1168–82.