

Machine Learning Model Integrating CT Radiomics of the Lung to Predict Checkpoint Inhibitor Pneumonitis in Patients with Advanced Cancer

François Cousin, MD, PhD¹ , Thomas Louis, MS, PhD², Pierre Frères, MD, PhD³, Julien Guiot, MD, PhD⁴, Mariaelena Occhipinti, MD², Fabio Bottari, MS², Wim Vos, MS², and Roland Hustinx, MD, PhD^{1,5}

Abstract

Objective: Checkpoint inhibitor pneumonitis (CIP) is a potentially life-threatening immune-related adverse event. Efficient strategies to select patients at risk are still required. The aim of our study was to assess the utility of a machine learning model, integrating pre-treatment CT lung radiomics features with clinical data, to predict patients at risk of developing CIP.

Methods: In this retrospective study, 116 patients with varied malignancies treated with immune checkpoint inhibitors (ICIs) were included. In this cohort, 35 patients presented with CIP and 81 patients did not. Each lung and its lobes were segmented on pre-treatment CT scans to perform a handcrafted radiomic analysis. Radiomic features were associated with clinical parameters to build generalized linear (GLM) and random forest (RF) models, to predict occurrence of CIP. The models were fine-tuned, validated and tested using a nested 5-fold cross-validation method.

Results: The RF models combining radiomic and clinical features showed the best performances with an area under the ROC curve (AUC) of 0.75 (95%CI:0.62-0.88) on the test set. The most accurate clinical model was a RF model and achieved an AUC of 0.72 (95%CI:0.51-0.92). The best radiomic model was a GLM model and achieved an AUC of 0.71 (95%CI:0.58-0.84).

Conclusions: Our CT-based lung radiomic models showed moderate to good performance at predicting CIP. We demonstrated the potential role of machine learning models associating clinical parameters and lung CT radiomic features to better identify patients treated with ICIs at risk of developing CIP.

Advances in knowledge: Radiomics analysis of the lung parenchyma could be used as a non-invasive tool to select patients at risk of developing immune-checkpoint pneumonitis.

Keywords

immune checkpoint inhibitors, immune-related adverse events, immune checkpoint pneumonitis, radiomics, machine learning

Introduction

In the last decade immune checkpoint inhibitors (ICIs) have demonstrated unprecedented efficacy in the treatment of numerous advanced malignancies, offering more long-term clinical responses.^{1,2} However, as ICIs act by stimulating the immune system in order to limit tumor proliferation, they have been associated with specific autoimmune complications called immune-related adverse events (irAEs).^{3,4} Immune-related adverse events are associated with therapy discontinuation, more frequent hospitalization, emergency room admission, death, and increased health care costs.⁵ Among irAEs, checkpoint inhibitor pneumonitis (CIP), which was first described by Nishino et al in 2015 in three patients with melanoma treated

¹Department of Nuclear Medicine and Oncological Imaging, University Hospital (CHU) of Liège, Liège, Belgium

²Radiomics (Oncoradiomics SA), Liège, Belgium

³Department of Medical Oncology, University Hospital (CHU) of Liège, Liège, Belgium

⁴Department of Respiratory Medicine, University Hospital (CHU) of Liège, Liège, Belgium

⁵GIGA-CRC in vivo imaging, University of Liège, Liège, Belgium

Corresponding Author:

François Cousin, Department of Nuclear Medicine and Oncological Imaging, University Hospital (CHU) of Liège, Liège, Belgium.

Email: fcousin@chuliege.be



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and

distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

with anti-programmed death-1 (PD-1), is particularly worrisome.⁶ A recent meta-analysis revealed that CIP was one of the most fatal irAEs, representing 35% of all anti-PD-1 / programmed death ligand-1 (PD-L1) -related fatalities.⁷ The reported incidence of CIP is comprised within a wide range, varying from 2% up to 19%.^{8–10} Several risk factors have been described in order to identify subgroups of patients that are more likely to present with CIP, among which tumor histology, ICI combination therapies or the presence of preexisting chronic lung disease such as interstitial lung disease (ILD) diagnosed on CT.^{10–14} The latter has been recognized in several studies as one of the most important risk factors, and associated with more high grade CIP, suggesting that lung abnormalities visible on pretreatment CT scans could be used to more accurately predict if a patient is at risk of developing CIP, which is crucial in the therapeutic decision process and the treatment monitoring.^{15–17}

Besides, radiomics is a quantitative, noninvasive imaging technique which extract a large amount of information from standard-of-care imaging. These information can then be used to build machine learning-based models to predict clinical outcomes such as prognosis, response to treatment or treatment toxicity.¹⁸ Radiomic analysis of the lung parenchyma has already shown its ability to predict the onset of radiation pneumonitis.^{19–22} We postulate that radiomic analysis of the lung parenchyma on pre-treatment CT, by depicting unforeseen lung abnormalities, may play a role in the selection of patients who are more likely to develop CIP. Therefore, in this retrospective study, we conducted a CT-based lung radiomic analysis of patients treated with ICIs for an advanced malignancy in order to predict occurrence of CIP.

Materials and Methods

Patients

We identified all patients with advanced malignancy treated with anti PD-1/PD-L1 at our university hospital between December 2015 and December 2021, either as monotherapy or in association with anti-cytotoxic T-lymphocyte antigen-4 (CTLA-4) or chemotherapy. We retrospectively reviewed medical records and all lung CT carried out during the ICI treatment of those patients.

Two groups of patients were selected, a group of patients who presented with CIP (CIP group) and a group of patients who did not develop CIP (control group).

According to the Fleischner Society, the diagnostic criteria used for CIP were : (a) newly identified pulmonary opacities on CT, commonly in a bilateral nonsegmental distribution; (b) temporal association of presentation with the initiation of the treatment; and (c) exclusion of other causes.²³ To validate the diagnosis of CIP all potential cases, previously selected after clinical records and lung CT screening, were reviewed again in a multidisciplinary team composed of two radiologists (4 and 10 years' experience), one oncologist specializing in lung cancer and one pulmonologist specializing in ILDs. The severity of CIP was assessed according to Common Terminology Criteria

for Adverse Events version 5.0. All patients with an alternative diagnosis such as infection, tumor progression or heart failure were excluded using extensive medical file review based on clinical information, laboratory data, bronchoalveolar lavage if available or echocardiography.

Concerning the control group, all patients who received ICIs between 2018 and May 2019 were included. The patient selection period for that group was intentionally limited to one year (situated in the middle of the study period, which we considered to be representative) in order to limit the imbalance of the data. Exclusion criteria were: (a) occurrence of a CIP; (b) ICI treatment duration of less than 3 months.

In the case of both groups, patients without available high-resolution lung CT performed before ICI initiation were also excluded. The end of the follow-up period was December first, 2021.

The following demographic and clinical data were collected: age, sex, smoking history, Eastern Cooperative Oncology Group performance status (PS), type of malignant disease, active auto-immune disease, prior radiotherapy of the lung, ICI molecule, line of treatment and preexisting pulmonary disease visible on CT (including radiation-induced lung injury, emphysema and ILD; assessed by two radiologists of 4 and 10 years' experience). Categorical clinical variables were converted to a numerical format to be implemented in the models.

This study received approval from the institutional review board, and the need for informed consent was waived based on its retrospective design. The reporting of this study conforms to STROBE guidelines.²⁴

CT Image Acquisition and Segmentation

The pre-treatment CT images were acquired with scanners Somatom Emotion, Sensation, Definition, Definition AS or Edge plus (Siemens Healthineers, Erlangen, Germany), Brilliance (Phillips Healthcare, Amsterdam, Netherlands), Brightspeed or Revolution CT (GE Healthcare, Chicago, IL, US), using tube voltage ranging from 80kVp to 130kVp, automatic adjustment of tube current, a 512×512 matrix, with or without intravenous iodine-based contrast injection. All images were reconstructed using a lung convolution kernel with slice thickness comprised between 0.6 and 3 mm. For all patients, each lung and its lobes were semi-automatically segmented using MIM software v7.2.3 (Beachwood, OH, US) by one radiologist with 10 years' experience, taking care to exclude large vessels and gross tumoral tissue.

Radiomic Analysis

Features Selection. Handcrafted radiomic features were extracted from each segmented region of interest (ROI) using the RadiomiX software (Oncoradiomics SA, Liège, Belgium), including first-order and texture features, in compliance with the Image Biomarkers Standardization Initiative.²⁵ Gray value discretization using a fixed bin count of 64 was performed. A minimum redundancy maximum relevance (mRMR) feature-

ranking algorithm was used to select a subset of promising predictors. The feature selection procedure was adapted to the machine learning method considered, the F-test correlation quotient using Analysis of Variance (ANOVA) was used for the generalized linear model (GLM) and the random-Forest correlation quotient based on the Gini Index for the random forest (RF) model.

Model Building. The output to be predicted by the models was the occurrence of CIP. Three types of models were built: (a) clinical models integrating clinical parameters only, (b) radiomic models, integrating radiomic features only, and (c) combined models, associating clinical parameters and radiomic features. Clinical models only integrated a subset of patients for whom all clinical data were available, contrary to radiomics and combined models in which all patients of the cohort were included. The performance of each ROI was first assessed separately regarding the outcome. The patient-level performances of the models were then expressed as the mean area under the ROC curve (AUC) value of all the ROIs belonging to the same patient, using either lung ROIs or lobe ROIs. Due to the lack of an independent validation dataset a stratified 5 × 5-folds nested cross-validation scheme was used. Stratification based on outcome occurrence and cancer origin was used to equally split the data in the 5 outer folds. Each outer train fold was subsequently divided into a subset of 5 inner folds (80/20 split), to optimize hyperparameters and train the models. The best model per outer fold was selected based on the highest mean AUC over the inner folds. The outer folds were used to assess generalizability and robustness of the models. A random down-sampling was carried out on the outer loop test sets to reach a 1:1 outcome ratio. The final performance of the model corresponds to the mean performance of all outer fold signature. Two different types of machine learning model were used for each set of features: a GLM and a RF model. Concerning RF, deeply grown trees were chosen by setting a low minimum node size, leading to deep and unbiased trees with a near-deterministic behavior on training data (AUCs ≈ 1). Overfitting was controlled by applying ensemble averaging across the 5 outer folds of the nested cross-validation. Two hyperparameters were optimized during cross-validation: the number of trees (tested between 100 and 10 000) and the number of features randomly selected at each split (tested between 1 and 50).

Statistical Analysis. Wald test was used for the univariate analysis of the clinical predictors. The performance of each model was evaluated by the area under receiver operating characteristic (ROC) curve (AUC), and its corresponding 95% confidence interval (CI). DeLong test was performed to compare the AUC values of the different models. All the statistical analyses were carried out with R software (version 4.1.1), using the glm function and the caret package (ranger implementation).

Results

Patient Characteristics

A total of 116 patients were included in the analysis, 35 patients in the CIP group and 81 patients in the control group. A patient flow diagram is shown in Figure 1. A total of 740 ROIs were segmented, composed of 230 lungs (2 patients had a history of left pneumonectomy) and 530 lobes (after lobectomy, 20 ROIs were considered both as lung and lobe). There were significantly more patients with a history of smoking in the CIP group compared to the control group (100% vs 89%, $P = .04$). There were also more patients with a high tumoral PD-L1 expression in the CIP group compared to the control group (51% vs 25%, $P = .03$). There was no other significant difference between the groups. The median time between baseline CT acquisition and ICI initiation was 24 days (range 0-158 days) in the whole dataset, 22 days (range, 1-158 days) in the CIP group and 25 days (range, 0-126 days) in the control group. Contrast-enhanced CTs were performed in 27 patients (77%) of the CIP group, and 68 patients (84%) of the control group. Median time between ICI initiation and CIP occurrence was 113 days (range, 11-1304 days). Grade 1 CIP was observed in 8 patients (23%), grade 2 in 15 (43%), grade 3 in 4 (11%), grade 4 in 4 (11%) and grade 5 in 4 patients (11%). The median treatment duration was 8.3 months (range, 0.9-45 months) in the CIP group and 10.6 months (range, 3-35.7 months) in the control group ($P = .22$). The demographic and clinical characteristics of the patients are presented in Table 1. Tumoral expression of PD-L1 was unknown in 33 patients (28%), including patients with non-lung malignancies and lung malignancies receiving second and further lines treatments, and therefore these patients were not included in the clinical models. Predictive value of the clinical characteristics regarding CIP onset in the 83 patients with known tumoral PD-L1 status are shown in Table 2. In order to use the whole dataset, tumoral PD-L1 status was not integrated in the combined models.

Radiomic Feature Selection

A total of 171 handcrafted radiomic features were extracted from each segmented ROI. A subset of 10 promising predictors was selected for GLM and 50 promising features for RF models using the mRMR feature ranking algorithm. The list of the radiomic and clinical features selected for each model building, and their relevance, can be found in the Supplementary Tables 1-4.

Performance of the Models

Concerning the performances of the clinical models, based on clinical predictors only, the best clinical GLM signature to predict the onset of CIP included 4 parameters (age, smoking history, line of treatment and tumoral PD-L1 status) and achieved an AUC of 0.59 (95%CI:0.38-0.80) on the test set. The best clinical RF model yielded an AUC of 0.72 (95%CI:0.51-0.92).

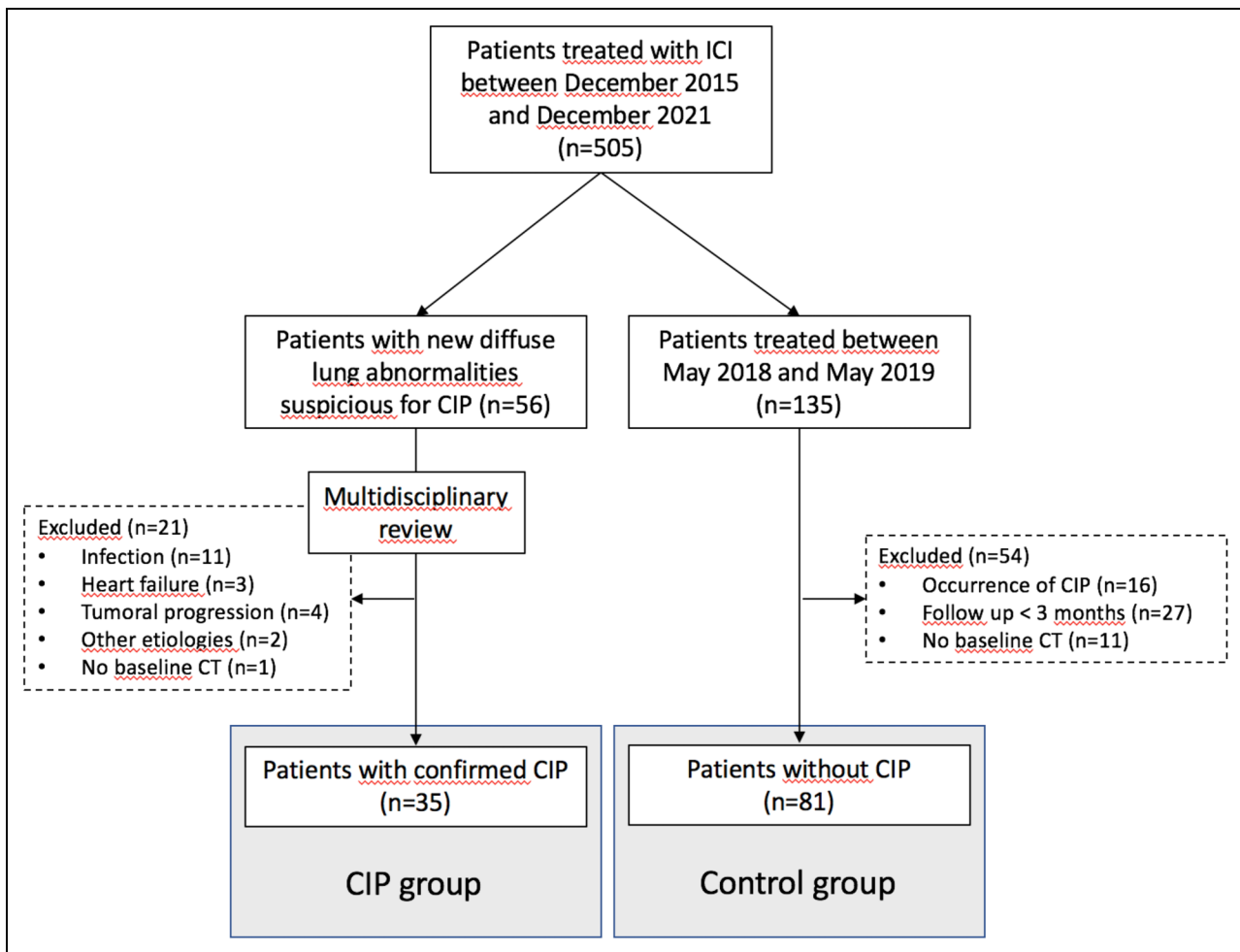


Figure 1. Patients Flow Diagram.

Concerning radiomic models, including radiomic features only, the GLM lobe-based approach achieved an AUC of 0.78 (95%CI:0.77-0.79) and 0.71 (95%CI:0.58-0.84) on the training and validation dataset; and 0.75 (95%CI: 0.74-0.76) and 0.61 (95%CI: 0.46-0.76) for the lung-based analysis, respectively. The RF approach achieved an AUC of 0.61 (95%CI: 0.43-0.79) for lobes and 0.68 (95%CI: 0.57-0.79) for lungs on the testing dataset.

Regarding the performances of the combined models, associating clinical parameters and radiomic features to predict the onset of CIP, the GLM lobe-based approach achieved an AUC of 0.83 (95%CI: 0.82-0.84) and 0.75 (95%CI: 0.64-0.86) on the training and validation dataset; and an AUC of 0.78 (95%CI: 0.75-0.81) and 0.62 (95%CI: 0.49-0.75) for the lung-based analysis, respectively. The RF lobe-based approach achieved an AUC of 0.75 (95%CI: 0.68-0.82) and the lung-based approach an AUC of 0.75 (95%CI: 0.62-0.88) on the testing dataset. Regarding RF model's performances on training sets, the deeply grown trees of the RF models applied led to near-deterministic behavior on training data with AUCs ≈ 1 .

A summary of the models' performance is presented in Table 3 and the corresponding ROC curves are shown in Figure 2.

No statistically significant difference was noted between AUC values of the different models following the DeLong test. All radiomic and clinical features used to build the signatures are available in the Supplementary Tables 1-4.

Discussion

Immune checkpoint inhibitors offer unprecedented tumoral response in a subset of patients with advanced cancer, but are also associated with specific and potentially life-threatening adverse events, among which is CIP. Thus, in the era of personalized medicine, a more accurate understanding of these irAEs, and the reliable identification of patients at risk, are crucial to aid the best treatment selection for an individual patient.²⁶

In this study, we used radiomics to extract texture features from pre-treatment lung CT to predict the onset of CIP in patients with solid malignancies. These features were combined to clinical parameters and analyzed using machine learning algorithms to develop predictive models. Our best combined patient-based model was the lobe-based RF model, achieving

Table 1. Demographic and Clinical Characteristics of the Patients.

Characteristics	All patients (N = 116)	Control group (N = 81)	CIP group (N = 35)	P-value
Age, median (range)	66 (38-87)	66 (38-83)	69 (42-87)	.56
Sex				.09
Male	66 (57%)	42 (52%)	24 (68%)	
Female	50 (43%)	39 (48%)	11 (32%)	
Current or former smoker	107 (92%)	72 (89%)	35 (100%)	.04
Cancer type				.70
Lung	94 (81%)	65 (80%)	29 (82%)	
Melanoma	14 (12%)	11 (14%)	3 (9%)	
Other ^a	8 (7%)	5 (6%)	3 (9%)	
ECOG PS				.69
0-1	104 (90%)	72 (89%)	32 (91%)	
≥2	12 (10%)	9 (11%)	3 (9%)	
Tumoral PD-L1 expression				.03
Unknown	33 (28%)	24 (30%)	9 (26%)	
<1%	21 (18%)	18 (22%)	3 (9%)	
1%-49%	24 (21%)	19 (23%)	5 (14%)	
≥50%	38 (33%)	20 (25%)	18 (51%)	
ICI molecule				.29
Pembrolizumab	49 (43%)	34 (42%)	15 (43%)	
Nivolumab	28 (24%)	19 (23%)	9 (25%)	
Durvalumab	21 (18%)	13 (16%)	8 (23%)	
Atezolizumab	14 (12%)	13 (16%)	1 (3%)	
Nivolumab / Ipilimumab	4 (3%)	2 (3%)	2 (6%)	
Treatment line				.67
First line	43 (37%)	29 (36%)	14 (40%)	
Further lines	73 (63%)	52 (64%)	21 (60%)	
Treatment type				.67
ICI monotherapy	102 (88%)	72 (89%)	30 (86%)	
ICI association	4 (3%)	2 (2%)	2 (6%)	
ICI/CT association	10 (9%)	7 (9%)	3 (8%)	
Autoimmune condition	7 (6%)	4 (5%)	3 (8%)	.44
Previous lung irradiation	60 (52%)	39 (48%)	21 (60%)	.24
Preexisting radiation-induced lung injury	31 (27%)	19 (23%)	12 (34%)	.22
Preexisting ILD on CT	5 (4%)	2 (3%)	3 (8%)	.13
Treatment duration ^b , median (range)	9.8 (0.9-45)	10.6 (3-35.7)	8.3 (0.9-45)	.22

^aOther includes renal, head and neck, urothelial, esophageal and penile cancers.

Abbreviations: ECOG PS, Eastern Cooperative Oncology Group performance status; PD-L1, programmed death-ligand 1; ICI, immune checkpoint inhibitor; CT, chemotherapy; ILD, interstitial lung disease; CIP, Checkpoint inhibitor pneumonitis.

^bin months.

moderate to good performances at stratifying the risk of developing CIP, yielding an AUC of 0.75 (95%CI: 0.68-0.82).

The presence of underlying chronic lung disease such as ILD has been recognized as an important risk factor for CIP.^{13,15,16} This has been confirmed in two recent meta-analyses looking at the occurrence of CIP in non-small cell lung cancer (NSCLC), showing an increased incidence of CIP in patients with ILD, with a pooled OR of 3.23 (95%CI: 2.06-5.06) and 3.68 (95% CI: 2.49-5.44).^{27,28} In the KEYNOTE 001 study, an association was shown between occurrence of pneumonitis and a history of asthma or chronic obstructive pulmonary disease in patients with NSCLC treated with pembrolizumab.²⁹ In a study by Atchley et al, the development of CIP in a cohort of patients with lung cancer treated with ICIs was independently associated with several chronic pulmonary diseases such as fibrosis, COPD and emphysema.⁹ Reuss et al have also demonstrated that NSCLC patients treated with ICIs who subsequently

developed CIP had lower pre-treatment forced expiratory volume in 1 s, suggesting that preexisting lung disease resulting in decreased lung function may be a CIP risk factor.³⁰ All these data suggest that beside risk factors linked to patient characteristics and treatment modalities, preexisting lung disease is also to be considered as a risk factor for CIP.¹⁴ The underlying mechanism is not well known but some authors suggest that the presence of inflammatory pulmonary comorbidities leads to predisposition to acute inflammatory lung response.^{4,31,32}

Although CT radiomics has already proven its ability to assess interstitial lungs abnormalities, few studies have evaluated its role in the identification of patients at risk of developing CIP.³³⁻³⁵ A proof of concept study by Colen et al based on two patients with CIP was the first attempt to show the potentiality of CT radiomics of the lung parenchyma at predicting CIP onset.³⁶ More recently, Tan et al have developed a deep learning model based on clinical data and pre-treatment CT images of the lungs to predict patients

Table 2. Univariate Analysis of the Clinical Predictors.

Clinical predictors	CIP occurrence	
	AUC (95%CI)	P-value
Age	0.60 (0.48-0.72)	.086
Sex	0.58 (0.49-0.67)	.098
Smoking history	0.55 (0.52-0.58)	.990
ECOG PS ≥ 2	0.51 (0.45-0.57)	.681
Tumoral PD-L1 expression > 50%	0.60 (0.51-0.69)	.062
ICI molecule		
Anti-PD-1	0.51 (0.42-0.60)	.743
Anti-PD-L1	0.53 (0.44-0.62)	.492
ICI association	0.52 (0.47-0.57)	.393
Line of treatment	0.53 (0.43-0.63)	.559
Autoimmune condition	0.52 (0.46-0.58)	.456
Lung cancer	0.51 (0.44-0.58)	.742
Previous lung irradiation	0.57 (0.47-0.67)	.148
Preexisting radiation-induced lung injury	0.55 (0.46-0.64)	.229
Treatment duration	0.57 (0.44-0.70)	.774

Abbreviations: ICI, immune checkpoint inhibitor; AUC, area under the ROC curve; OS, overall survival; CIP, Checkpoint inhibitor pneumonitis; ROC, receiver operating characteristic; ECOG PS, Eastern Cooperative Oncology Group performance status.

Table 3. Test set Performances of the Different Machine Learning Models to Predict CIP Occurrence.

Types of models	AUC (95%CI)
Clinical models	
GLM	0.59 (0.38-0.80)
RF	0.72 (0.51-0.92)
Radiomic Lobe-based models	
GLM	0.71 (0.58-0.84)
RF	0.61 (0.43-0.79)
Radiomic Lung-based models	
GLM	0.61 (0.46-0.76)
RF	0.68 (0.57-0.79)
Combined Lobe-based models	
GLM	0.75 (0.64-0.86)
RF	0.75 (0.68-0.82)
Combined Lung-based models	
GLM	0.62 (0.49-0.75)
RF	0.75 (0.62-0.88)

Abbreviations: CIP, checkpoint inhibitor pneumonitis; AUC, area under the ROC curve; ROI, region of interest; GLM, generalized linear model; RF, random forest; ROC, receiver operating characteristic.

at risk of developing CIP in a lung cancer population, achieving an excellent performance with an AUC of 0.92.³⁷ However, the dataset they used to train their deep learning model was still limited, including only 24 patients with CIP among 48 patients with lung cancer, leading to a risk of overfitting. In comparison, we used more understandable machine learning processes to build models based on a dataset composed of 116 patients presenting with different types of cancers. We also performed our radiomic analysis on two anatomical levels, the lobes and the lungs, postulating that some textural parameters could be different when considering the whole lung rather than a part of it. Even if no statistically significant difference was noted between the different models with the DeLong test,

which could be due to the limited size of our dataset, our lobe-based linear models showed better performances compared to the lung-based analyses. The lobe-based approach could be more efficient to depict non-uniformly distributed lung abnormalities. However, looking at the RF analysis, no difference in performance was noted between the two approaches. In this study the combined models slightly outperformed both radiomic and clinical models. However, tumoral PD-L1 expression, which was the clinical factor most strongly associated with CIP with an AUC of 0.60 (95%CI, 0.51-0.69), was not available in 28% of patients and was therefore not included in the combined models, contrary to the clinical models. It means that the clinical parameters used to train these models were slightly different and that these models could not be directly compared. Nonetheless, performances of the combined models without PD-L1 expression were still better than the clinical models including PD-L1 expression, underlining the importance of developing integrative models. In this specific indication, models integrating lung textural analysis is a way of incorporating quantitative lung parenchymal assessment in the CIP risk stratification process, along with patient and treatment-based risk factors.

Our study has some limitations. First, it is a monocentric study and an external validation dataset was lacking, which could be a problem to assess the generalizability of the models. However, to compensate for this limitation we used a nested cross-validation approach to fine-tune parameters and test the models on different data. This approach was taken in order to limit the effect of an optimistically biased evaluation of the models and to obtain a more accurate assessment of their true performance. Also, our patient population is heterogeneous, including different tumor histologies, treatment types and lines, and patient profiles; a more accurate selection of a sub-population of patients might have resulted in better performance. The median treatment duration was slightly shorter in the CIP group (8.3 months) compared to the control group

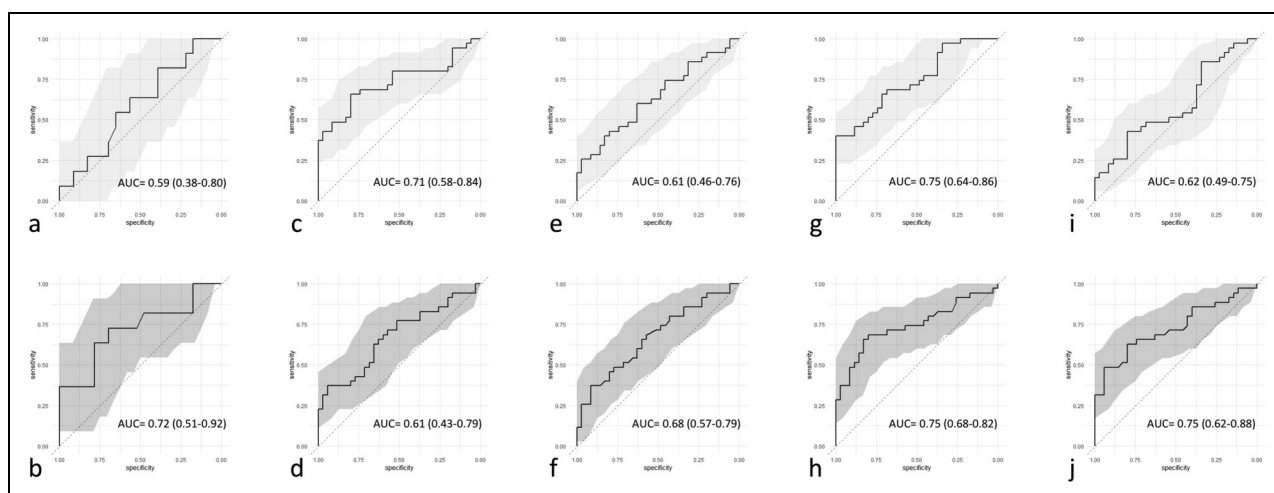


Figure 2. ROC Curves of the Clinical Models. (a, b), radiomic lobe-based models (c, d), radiomic lung-based models (e, f), combined lobe-based models (g, h) and lung-based models (i, j). For each model, GLM (a, c, e, g, i) and RF (b, d, f, h, j) analyses with confidence intervals are represented in light and dark grey respectively. Abbreviations: GLM, generalized linear model; ROC, receiver operating characteristic.

(10.6 months), but that difference was not statistically significant ($P=.22$). This difference is probably due to the fact that treatment was withdrawn in case of CIP of grade 2 and higher, as part of side effect management. The distribution of classes within our dataset was imbalanced with a 30/70 proportion of patients with and without CIP. Even if this could be considered as a light imbalance, this could have led to bias toward the majority class.

Additionally, tumoral PD-L1 expression was not included in the combined models, consequently, its performances could not be directly compared with the performances of the clinical model. Finally, differences in acquisition parameters and contrast injection protocols between CTs could potentially have affected radiomic features extraction and analysis.

Conclusion

This study demonstrated the potential role of a machine learning model associating clinical parameters and lung CT radiomics features to more accurately identify patients with advanced malignancies treated with ICIs at risk of developing CIP.

Abbreviations

ICIs	immune checkpoint inhibitors
irAEs	immune-related adverse events
CIP	checkpoint inhibitor pneumonitis
PD-I	programmed death-I
PD-L1	programmed death ligand-I
ILD	interstitial lung disease
CTLA-4	cytotoxic T-lymphocyte antigen-4
CTCAE	Common Terminology Criteria for Adverse Events
PS	performance status
ROI	region of interest
IBSI	Imaging Biomarkers Standardization Initiative

mRMR	minimum redundancy maximum relevance
GLM	generalized linear model
RF	random forest
ROC	receiver operating characteristic
AUC	area under the ROC curve
CI	confidence interval
NSCLC	non-small cell lung cancer

ORCID iD

François Cousin  <https://orcid.org/0000-0003-4014-0297>

Ethical Considerations

The approval from the Institutional Review Board (*Comité d'Ethique Hospitalo-Facultaire Universitaire de Liège*) was obtained, and the need for informed consent was waived based on its retrospective design (Reference number: 2019/191; date of acceptance : 08/08/2019).

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author.

Declaration of Conflicting Interests

Julien Guiot reports personal fees for advisory board, work and lectures from Boehringer Ingelheim, Janssen, SMB, GSK, Roche, Astra Zeneca, and Chiesi, non-financial support for meeting attendance from Chiesi, MSD, Roche, Boehringer Ingelheim and Janssen. He is in the permanent SAB of Radiomics (Oncoradiomics SA) for the SALMON trial without any specific consultancy fee for this work. He is co-inventor of one issued patent on radiomics licensed to

Radiomics (Oncoradiomics SA).

Thomas Louis and Fabio Bottari are salaried employees of Radiomics (Oncoradiomics SA). Mariaelena Occhipinti reports personal fees from Radiomics, outside the submitted work. Wim Vos has shares in the company Radiomics.

The rest of the authors does not have conflicts to declare.

Supplemental Material

Supplemental material for this article is available online.

References

- Topalian SL, Hodi FS, Brahmer JR, et al. Five-Year survival and correlates among patients with advanced melanoma, renal cell carcinoma, or non-small cell lung cancer treated with nivolumab. *JAMA Oncol*. 2019 Oct 1;5(10):1411.
- Reck M, Rodríguez-Abreu D, Robinson AG, et al. Five-Year outcomes with pembrolizumab versus chemotherapy for metastatic non-small-cell lung cancer with PD-L1 tumor proportion score $\geq 50\%$. *J Clin Oncol*. 2021 Jul 20;39(21):2339–2349.
- Michot JM, Bigenwald C, Champiat S, et al. Immune-related adverse events with immune checkpoint blockade: A comprehensive review. *Eur J Cancer*. 2016 Feb;54:139–148.
- Burke KP, Grebinoski S, Sharpe AH, Vignali DAA. Understanding adverse events of immunotherapy: A mechanistic perspective. *J Exp Med*. 2021 Jan 4;218(1):e20192179.
- George S, Bell EJ, Zheng Y, et al. The impact of adverse events on health care resource utilization, costs, and mortality among patients treated with immune checkpoint inhibitors. *Oncologist*. 2021 Jul 1;26(7):e1205–e1215.
- Nishino M, Sholl LM, Hatabu H, Ramaiya NH, Hodi FS. Anti-PD-1-related pneumonitis during cancer immunotherapy. *N Engl J Med*. 2015 Jul 16;373(3):288–290.
- Wang DY, Salem JE, Cohen JV, et al. Fatal toxic effects associated with immune checkpoint inhibitors: A systematic review and meta-analysis. *JAMA Oncol*. 2018 Dec 1;4(12):1721.
- Tiu BC, Zubiri L, Iheke J, et al. Real-world incidence and impact of pneumonitis in patients with lung cancer treated with immune checkpoint inhibitors: A multi-institutional cohort study. *J Immunother Cancer*. 2022 Jun;10(6):e004670.
- Atchley WT, Alvarez C, Saxena-Beem S, et al. Immune checkpoint inhibitor-related pneumonitis in lung cancer. *Chest*. 2021 Aug;160(2):731–742.
- Suresh K, Voong KR, Shankar B, et al. Pneumonitis in non-small cell lung cancer patients receiving immune checkpoint immunotherapy: Incidence and risk factors. *J Thorac Oncol*. 2018 Dec;13(12):1930–1939.
- Naidoo J, Wang X, Woo KM, et al. Pneumonitis in patients treated with anti-programmed death-1/programmed death ligand 1 therapy. *J Clin Oncol*. 2017 Mar 1;35(7):709–717.
- Nishino M, Giobbie-Hurder A, Hatabu H, Ramaiya NH, Hodi FS. Incidence of programmed cell death 1 inhibitor-related pneumonitis in patients with advanced cancer: A systematic review and meta-analysis. *JAMA Oncol*. 2016 Dec 1;2(12):1607.
- Cho JY, Kim J, Lee JS, et al. Characteristics, incidence, and risk factors of immune checkpoint inhibitor-related pneumonitis in patients with non-small cell lung cancer. *Lung Cancer*. 2018 Nov;125:150–156.
- Kim S, Lim JU. Immune checkpoint inhibitor-related interstitial lung disease in patients with advanced non-small cell lung cancer: Systematic review of characteristics, incidence, risk factors, and management. *J Thorac Dis*. 2022 May;14(5):1684–1695.
- Kanai O, Kim YH, Demura Y, et al. Efficacy and safety of nivolumab in non-small cell lung cancer with preexisting interstitial lung disease: Nivolumab in patients with ILD. *Thorac Cancer*. 2018 Jul;9(7):847–855.
- Yamaguchi T, Shimizu J, Hasegawa T, et al. Pre-existing pulmonary fibrosis is a risk factor for anti-PD-1-related pneumonitis in patients with non-small cell lung cancer: A retrospective analysis. *Lung Cancer*. 2018 Nov;125:212–217.
- Shimoji K, Masuda T, Yamaguchi K, et al. Association of preexisting interstitial lung abnormalities with immune checkpoint inhibitor-induced interstitial lung disease among patients with nonlung cancers. *JAMA Netw Open*. 2020 Nov 2;3(11):e2022906.
- Frix AN, Cousin F, Refaie T, et al. Radiomics in lung diseases imaging: State-of-the-art for clinicians. *J Pers Med*. 2021 Jun 25;11(7):602.
- Jiang W, Song Y, Sun Z, Qiu J, Shi L. Dosimetric factors and radiomics features within different regions of interest in planning CT images for improving the prediction of radiation pneumonitis. *Int J Radiat Oncol*. 2021 Jul;110(4):1161–1170.
- Krafft SP, Rao A, Stingo F, et al. The utility of quantitative CT radiomics features for improved prediction of radiation pneumonitis. *Med Phys*. 2018 Nov;45(11):5317–5324.
- Hirose Ta, Arimura H, Ninomiya K, Yoshitake T, Fukunaga Ji, Shioyama Y. Radiomic prediction of radiation pneumonitis on pretreatment planning computed tomography images prior to lung cancer stereotactic body radiation therapy. *Sci Rep*. 2020 Dec;10(1):20424.
- Zhang Z, Wang Z, Yan M, et al. Radiomics and dosiomics signature from whole lung predicts radiation pneumonitis: A model development study with prospective external validation and decision-curve analysis. *Int J Radiat Oncol*. 2023 Mar;115(3):746–758.
- Johkoh T, Lee KS, Nishino M, et al. Chest CT diagnosis and clinical management of drug-related pneumonitis in patients receiving molecular targeting agents and immune checkpoint inhibitors: A position paper from the fleischner society. *Radiology*. 2021 Mar;298(3):550–566.
- Elm V, Altman E, Egger DG, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Ann Intern Med*. 2007 Oct 16;147(8):573–577.
- Zwanenburg A, Leger S, Vallière M, Löck S. Image Biomarker Standardisation Initiative [Internet]. 2019. <https://ibsi.readthedocs.io/en/latest/>
- Nakagawa N, Kawakami M. Choosing the optimal immunotherapeutic strategies for non-small cell lung cancer based on clinical factors. *Front Oncol*. 2022 Aug 12;12:952393.
- Zhang M, Fan Y, Nie L, Wang G, Sun K, Cheng Y. Clinical outcomes of immune checkpoint inhibitor therapy in patients with advanced non-small cell lung cancer and preexisting interstitial lung diseases. *Chest*. 2022 Jun;161(6):1675–1686.

28. Chen X, Li Z, Wang X, Zhou J, Wei Q, Jiang R. Association of pre-existing lung interstitial changes with immune-related pneumonitis in patients with non-small lung cancer receiving immunotherapy. *Support Care Cancer*. 2022 Aug;30(8):6515–6524.
29. Sul J, Blumenthal GM, Jiang X, He K, Keegan P, Pazdur R. FDA Approval summary: Pembrolizumab for the treatment of patients with metastatic non-small cell lung cancer whose tumors express programmed death-ligand 1. *Oncologist*. 2016 May 1;21(5):643–650.
30. Reuss JE, Brigham E, Psoter KJ, et al. Pretreatment lung function and checkpoint inhibitor pneumonitis in NSCLC. *JTO Clin Res Rep*. 2021 Oct;2(10):100220.
31. Suazo-Zepeda E, Bokern M, Vinke PC, Hiltermann TJN, de Bock GH, Sidorenkov G. Risk factors for adverse events induced by immune checkpoint inhibitors in patients with non-small-cell lung cancer: A systematic review and meta-analysis. *Cancer Immunol Immunother CII*. 2021 Nov;70(11):3069–3080.
32. Nakanishi Y, Masuda T, Yamaguchi K, et al. Pre-existing interstitial lung abnormalities are risk factors for immune checkpoint inhibitor-induced interstitial lung disease in non-small cell lung cancer. *Respir Investig*. 2019 Sep;57(5):451–459.
33. De Giacomo F, Raghunath S, Karwoski R, Bartholmai BJ, Moua T. Short-term automated quantification of radiologic changes in the characterization of idiopathic pulmonary fibrosis versus nonspecific interstitial pneumonia and prediction of long-term survival. *J Thorac Imaging*. 2018 Mar; 33(2):124–131.
34. Martini K, Baessler B, Bogowicz M, et al. Applicability of radiomics in interstitial lung disease associated with systemic sclerosis: Proof of concept. *Eur Radiol*. 2021 Apr;31(4): 1987–1998.
35. Refaee T, Salahuddin Z, Frix AN, et al. Diagnosis of idiopathic pulmonary fibrosis in high-resolution computed tomography scans using a combination of handcrafted radiomics and deep learning. *Front Med*. 2022;9:915243.
36. Colen RR, Fujii T, Bilen MA, et al. Radiomics to predict immunotherapy-induced pneumonitis: Proof of concept. *Invest New Drugs*. 2018 Aug;36(4):601–607.
37. Tan P, Huang W, Wang L, et al. Deep learning predicts immune checkpoint inhibitor-related pneumonitis from pretreatment computed tomography images. *Front Physiol*. 2022 Jul 25;13:978222.