UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"SCIENZE MEDICHE E CHIRURGICHE TRASLAZIONALI "

CICLO XXXIII

CT-RADIOMICS FOR LUNG CANCER RISK

AND STRATIFICATION OF PROGNOSIS IN ADENOCARCINOMA

Coordinatore:
Chiar.mo Prof. Carlo Ferrari

Tutore:
Chiar.mo Prof. Nicola Sverzellati

Dottorando: Gianluca Milanese

Anni Accademici 2017/2018 – 2019/2020

INDEX

ABSTRACT 2

INTRODUCTION 4

MATERIALS AND METHODS 8

STUDY POPULATION 8

THORACIC CT 10

IMAGE ANALYSIS 10

HISTOLOGICAL STRATIFICATION 12

SURVIVAL ANALYSIS 15

STATISTICAL ANALYSIS 17

RESULTS 18

STUDY POPULATION 18

HISTOLOGICAL STRATIFICATION 23

SURVIVAL ANALYSIS 26

DISCUSSION 29

REFERENCES 39

ABSTRACT

Il tumore del polmone è una delle più importanti cause di morte a livello mondiale, nel cui percorso diagnostico un ruolo centrale è svolto dall'imaging, in particolare dalla tomografia computerizzata (TC). Nella maggior parte dei casi, al primo riscontro di un nodulo polmonare - la più frequente manifestazione di una neoplasia polmonare - non è possibile definire il comportamento biologico di tale lesione, rendendosi necessaria sia una valutazione longitudinale sia ulteriori approfondimenti diagnostici (es. caratterizzazione bioptica, indagini funzionali mediante PET).

Tradizionalmente, la caratterizzazione del nodulo polmonare da parte del Medico Radiologo è basata sull'analisi delle dimensioni e sull'identificazione di caratteristiche morfologiche compatibili con un comportamento biologico maligno, con risultati condizionati dall'esperienza dell'operatore. Una branca sviluppatasi recentemente nell'ambito della ricerca scientifica radiologica – definita Radiomica – prevede l'estrazione di parametri quantitativi, definiti "features radiomiche", che vengono analizzati con l'obiettivo di creare modelli predittivi del comportamento biologico delle lesioni, indipendenti dall'esperienza dell'operatore.

L'obiettivo del presente lavoro di tesi è stato quello di creare un modello predittivo del comportamento biologico di noduli polmonari identificati mediante TC; in particolare, sono stati inclusi e comparati soggetti con riscontro anatomo-patologico di adenocarcinomi polmonari, metastasi da tumori extra-polmonari (rene, colon, endometrio) e lesioni dal comportamento biologico benigno (amartomi).

851 "features radiomiche" sono stati estratte dalle lesioni polmonari mediante software dedicato, e tramite sistemi di machine learning sono stati creati modelli di predizione del comportamento biologico della lesione, per distinguere le lesioni tra adenocarcinomi,

metastasi e amartomi (ROC-AUC pari al 73%). Inoltre, un modello predittivo integrato - comprensivo di parametri clinici e radiomici - ha mostrato una performance più elevata nella stratificazione della prognosi dei soggetti con adenocarcinoma polmonare (C-index pari a 0.74 e 0.81, rispettivamente per soggetti con adenocarcinoma in tutti gli stadi e soggetti con adenocarcinoma in stadio I), rispetto a modelli esclusivamente clinici o radiomici.

INTRODUCTION

Lung cancer (LC) is one of the leading causes of death worldwide, accounting for more than 20% of cancer deaths in Europe [1]. The poor prognosis (5-year survival rate of about 15%) is due to the limited curative options available for the vast majority of cases, as approximately 70% of patients suffer from advanced disease at the time of diagnosis [2,3]. Indeed, LC is an insidious disease with symptoms occurring mostly at advanced stages of disease and being absent or non-specific at early phases.

The detection of a pulmonary nodule (pulmonary lesion with a diameter < 30 mm) is particularly frequent on low-dose computed tomography (LDCT) performed during daily clinical practice [4]. Usually, management of pulmonary nodules is driven by size, as this metric is deemed the most accurate for the evaluation of their biological behavior [5]. Given the clinical relevance of pulmonary nodules, international scientific societies developed guidelines for their management, and defined morphology (i.e. size, either by diameter or volume, and density) as the key parameter [4,6,7]. Size of LC can be measured on CT and represents a metric correlating with decreased overall survival [5,8]. Nonetheless, recent evidences are suggesting that size alone cannot fully explain the mechanism of progression of LC, as even small lesions might metastasize to lymph nodes or distant organs [9]. CT is the imaging modality of choice for the detection and characterization of pulmonary nodules, playing a role of paramount relevance in supporting clinical decision-making related to diagnosis, staging, and treatment-response in LC [4]. In the last decades, qualitative/semantic features were suggested as imaging markers for differentiating benign and malignant lesions. Spiculated margins, perilesional cystic airspaces and nodule density were tested as predictors for identifying lesions with malignant behavior [10]. Furthermore, such qualitative approach is intrinsically limited by the experience of the readers, and is therefore affected by significant inter-observer variability [11].

Tissue sampling represents the most important source of information for tumor classification; however, such approach may be hampered by sampling errors resulting in false negative classifications [12,13]. More accurate and reproducible approaches for the imaging characterization of pulmonary nodules were tested, aiming to overcome the limitation from subjective classification of lesions. Such approach - based on noninvasive analysis of abnormalities suspicious for malignancy – is the basis of the new field of radiology research, the so-called Radiomics, directed toward the analysis of quantitative metrics potentially encompassing information on characterization and prognostication of lesions from different sites [14]. Radiomics emerged as an integrated approach for the interpretation of diagnostic studies: it was defined as the high-throughput mining of quantitative features from medical imaging enabling data to be extracted and applied to improve diagnostic, prognostic, and predictive accuracy [15]. Radiomics features (RF) are image-based descriptors quantitatively capturing shape, size or volume, and texture, potentially associated with heterogeneous tumoral structure. Several tools are available for the extraction of RF, with different approaches and computational systems and in combination with artificial intelligence applications (exploiting workstation with high computational power) they can be included into predictive and prognostic models [16-19]. The workflow of Radiomics is based on a step-wise approach including image acquisition and preprocessing, delineation of either region of interest (ROI) or volume of interest (VOI), RF extraction and selection, followed by model building and validation [18]. RFs act as surrogate endpoints, based on the idea that more information is contained within diagnostic images, potentially correlating imaging features with pathology (i.e. histology, presence of genetic abnormalities) or clinical data (i.e. outcome)[20,21]. Tumors are spatially and temporally heterogeneous, and maybe under sampled during medical procedures where only a portion of the lesion can be completely characterized,

and this could be detrimental for the selection of treatment options [20]. Furthermore, in patients who show resistance to targeted therapies, resampling the neoplasm might be necessary for confirming the development of tumor mutations, in order to optimize their management. However, most cases did not undergo such resampling, as invasive procedures might be contraindicated in potentially fragile patients or because of technical difficulties; furthermore, for the detection of mutations the sensitivity of a single biopsy has been discussed [13,22]. There is an urge for accurate, reproducible and easily/safely obtainable approach for tumor analysis, and RF - non-invasively extracted – might answer this need [12,20,23,24].

Notably, Radiomics has been applied to lesions affecting various organs (e.g. lung, liver, brain, breast, bowel cancer, head and neck, lymphoid system, musculoskeletal system); nonetheless, LC is the most extensively studied and characterized malignancy through radiomics so far [14,25,26]. The up-to-date challenge for Radiomics in thoracic lesions is double-folded: systems must be developed to accurately extract phenotypic characteristics from images and they must determine which RF can assist in the prognosis and clinical management of disease [18,27-29].

The clinical impact of radiomics in the interpretation of a pulmonary nodule incidentally detected on CT can be evaluated from two perspectives: first, the identification of a radiomics signature underlining biological behavior (i.e. differentiating benign nodules from malignant ones) and, for malignant lesions the evaluation of patients' prognosis, expressed as survival time, without the need for longitudinal evaluation and invasive maneuvers.

Hence, the aims of this study were (i) to test the ability of Radiomics to distinguish non-small cell lung cancer (NSCLC) adenocarcinomas from pulmonary metastases and pulmonary

nodules with benign biological behavior at CT scan, and (ii) to test the performance of Radiomics to stratify prognosis of patients with adenocarcinoma.

MATERIALS AND METHODS

STUDY POPULATION

First, we retrospectively selected patients who underwent thoracic CT between December 2006 and December 2015 at the Radiology Department of the University Hospital of Parma showing CT evidence of pulmonary nodules. Patients had to have been surgically treated at the Thoracic Surgery Department, with histological confirmation of adenocarcinoma by the Pathology Department. Subsequently, we retrospectively selected patients with pulmonary lesions, for whom both imaging and pathological information were available, for either secondary extra-thoracic malignant neoplasms (namely, metastases from renal, colon and endometrial cancer) or benign lesions (i.e. hamartomas).

Clinical information (age, gender, smoking history, results of pulmonary function tests – PFT, previous history of oncologic disease) was retrieved from the database of the Pathology Department, integrated with data from the Thoracic Surgery Department.

The patients included were stratified into three groups (**Figure 1**), as follows:

a. Subjects with lung cancer (pathologically proven adenocarcinoma) surgically resected
    - Predominant subtype of LC was assessed using IASLC/ATS/ERS classification of lung adenocarcinoma and tumor staging was performed on the basis of the 8th edition of TNM classification for lung cancer [30]
b. Subjects with pulmonary metastases from extra-thoracic neoplasms
    - Patients with renal cancer
    - Patients with endometrial cancer
    - Patients with colon cancer
c. Subjects with benign disease

- Patients with surgically resected hamartomas

The study was approved by the Institutional Review Board (Ethical Committee) of the University Hospital of Parma (973/2018/OSS/UNIPR). Informed consent was waived.

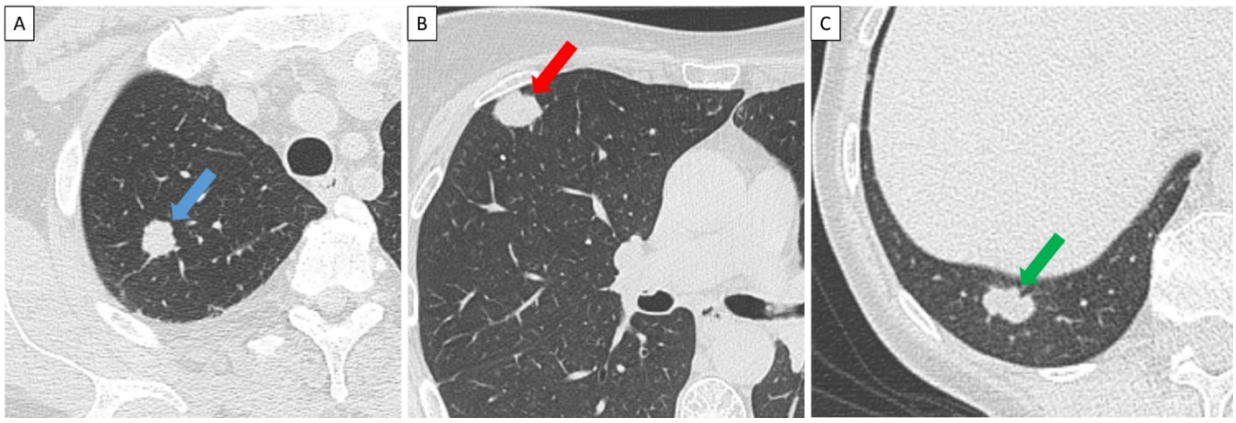

Figure 1. Examples of pulmonary nodules included in the study. (A) Computed Tomography of a 72-year-old male patient with a pulmonary nodule in the right upper lobe (blue arrow). Pathological analysis after surgery showed evidence of adenocarcinoma. (B) Computed Tomography of a 48-year-old female patient with previous history of breast cancer showing a pulmonary nodule in the right upper lobe (red arrow). Pathological analysis after surgery reported hamartoma. (C) Computed Tomography of a 63-year-old female patient with a pulmonary nodule in the right lower lobe (green arrow). Patient underwent thoracic CT during staging of colon cancer, and pathological analysis after surgery reported metastases from colon cancer.

THORACIC CT

Eligible baseline CTs (i.e. first CT scans available at the Picture Archive and Communicating System – PACS - of the University Hospital of Parma depicting the presence of the pulmonary lesion selected for the study) were performed with four different scanners, on a single-source 6-slice (Emotion 6; Siemens Healthcare, Forchheim, Germany), on a 16-slice (Emotion 16; Siemens Healthcare), 64-slice (Sensation Cardiac 64; Siemens Healthcare) and on a second-generation dual-source 128-slice CT scanner (Somatom Definition Flash; Siemens Healthcare). CT scans were performed at end-inspiration, capturing the entire lung volume from the apices to the costo-diaphragmatic recesses. CT scans included in the present study were performed both with and without intravenous administration of contrast media.

CT datasets were reconstructed with filtered back projection (FBP) algorithm, using different medium-sharp tissue convolution kernel with lung window settings (window width, 1200 Hounsfield Unit, HU; window level, −600 HU); slice thickness of reconstructed images ranged 1 mm to 2.5 mm.

IMAGE ANALYSIS

Images were transferred from PACS to an external workstation and analyzed using dedicated software for lesions contouring and advanced image analysis (3dSlicer 4.9.0, www.slicer.org). Pulmonary lesions were semi-automatically delineated every three slices by means of manually drawn ROIs and a dedicated interpolation algorithm tool was launched to calculate a VOI encompassing the whole lesions. In case of extra-thoracic metastatic disease with multiple pulmonary nodules, only the lesion with the greatest size was segmented. Two readers - blinded to clinical outcomes - independently performed the analyses on different subset of the study population: they were instructed to modify VOI boundaries in case of inadequate segmentation. RFs extraction was performed by a dedicated function embedded

into the segmentation software (SlicerRadiomics), and included RFs from different domains, as follows [14,31]:

- Shape-based
- Gray Level Dependence Matrix (GLDM)
- Gray Level Co-occurrence Matrix (GLCM)
- First Order Statistic
- Gray Level Run Length Matrix (GLRLM)
- Gray Level Size Zone Matrix (GLSZM)
- Neighboring Gray Tone Difference Matrix (NGTDM)
- Wavelet Transformed Features

HISTOLOGICAL STRATIFICATION

The outcome was defined as a three-levels categorical variable including (1) patients with adenocarcinoma, (2) patients with metastasis (independently from the site of origin) and (3) patients with hamartoma.

For being included in the histological stratification analysis, pulmonary lesions had to fulfill the definition of the Fleischner Society for nodule (diameter up to 30 mm; diameter was derived from the semi-automated measurements obtained from the VOI) [32]. To take into account the potential impact of calcifications on textural analysis and RFs, hamartomas displaying calcifications were excluded [33].

Classification analyses were performed, and results compared after cross-validation, which allows for all observations to serve for both training and test [34]. Monte Carlo cross-validation was used to create multiple random splits of the dataset into training and test data. For each split, the model was fit to the training data and predictive accuracy was assessed using the test data, with final results deriving from the average of the splits. The advantage of this method (over k-fold cross validation) is that the proportion of the training/test split is not dependent on the number of iterations (i.e., the number of partitions). The disadvantage is that some observations may never be selected in the test subsample, whereas others may be selected more than once.

To eliminate redundant RFs, from the overall study population highly correlated RFs were removed by calculating their Spearman Rho correlation coefficient: RFs with a statistically significant coefficient greater than 0.99 were excluded from the subsequent analyses.

Overall, for the purpose of differentiating pulmonary lesions according to their biological behavior, a total of 2,000 iterations was performed. Then, for each iteration, the

study population was randomly split – as per Monte Carlo cross-validation - in the recommended ratio of 2/3 training and 1/3 testing.

Feature selection and model development

The least absolute shrinkage and selection operator (LASSO) logistic regression model allows interpretation and analysis of large database and was in order to identify the most useful prognostic RF [35]. LASSO is a method of regression analysis that performs feature selection and regularization to improve the prediction accuracy via penalized estimation functions. Penalized logistic regression imposes a penalty to the logistic model to reduce RF number. Such approach – named regularization process - shrinks the coefficients of the less contributive RFs which are forced to be exactly zero within the logistic regression. LASSO was performed to choose the optimal hyper parameter (λ), which is used as a regularization term acting as parameter for tuning penalization. Based on the LASSO logistic regression analysis, we selected the most predictive RFs.

Our predictive model was developed in the training set; subsequently, prediction analyses were made based on the data of the test set.

Feature stability

Given that our approach was based on 2,000 iterations, each iteration could select a different subset of RFs predictive the outcome during each single iteration. To account for such intrinsic variability between iterations, we described the stability of our model by computing the following stability indexes [36 37]:

- Average Normalize Hamming Distance (ANHD)
- DICE
- Adjusted stability measure (ASM)

- Weighted consistency measures (CW)

Model Performances

Model performances were computed for both training and test set, to evaluate the amount of the differences between the performances from the two models. The main metric was the Receiver Operating Characteristic – Area Under the Curve (ROC-AUC) of each model and performance results were averaged over the 2,000 iterations.

Further metrics testing models' performance were precision (i.e. positive predictive value) and recall (i.e. sensitivity) of each individual category in both training and test set: hence, average values of 2,000 iterations for both precision and recall were derived for each categorical variable on training and test set.

SURVIVAL ANALYSIS

The outcome variable was overall survival (OS), calculated from the date of baseline CT until either death from any cause (event), or until the date (30/01/2020) that Patients were last known to be alive (censored). For the purposes of survival analysis, the analyses were repeated twice:

- patients with adenocarcinoma, independently from stage (n = 285);
- patients with stage I adenocarcinoma (n = 174).

We performed the analyses on patients displaying the whole set of variables: those with missing data were excluded. The results of the analyses derived from a total of 50 iterations and compared after Monte-Carlo cross-validation, with each single iteration split into the recommended ratio of 2/3 training and 1/3 testing.

Clinical Prognostic Model

The Clinical Prognostic Model (CPM) acted as standard of reference. Univariate Cox regression was used to identify clinical variables which were more predictive for OS, including the type of surgical approach. In the final clinical multivariate Cox regression model, only features that showed p-value of their hazard ratio (HR) < 0.3 in more than 60% of the iterations were included as predictors.

Radiomics Prognostic Model

Highly correlated RFs (Spearman Rho correlation coefficient greater than 0.99) were excluded. Z-scores were computed for each RF to normalize data and subsequently, univariate Cox regression was used to identify those RFs predictive for OS. Only variables showing p-values of HR lesser than 0.05 were considered significant and included for the development of the Radiomics Prognostic Model (RMP).

We adopted supervised principal component analysis (s-PCA) for further reducing the number of covariates for the final multivariate Cox proportional hazard model: s-PCA is similar to conventional PCA, and it can be applied to regression problem and survival analysis, by considering a subset of predictors that are better associated with the outcome [38]. The number of principal components was set to three (s-PCA_1, s-PCA_2, s-PCA_3), which were used for the final multivariate Cox hazard model based on RFs. Median value of OS was used to discretize the survival endpoint in short-term and long-term survivors. The ROC analysis was performed to determine the optimal threshold for the three components of the s-PCA, in particular the value that yielded the highest Youden Index was selected as optimal cut-off for discretization.

Clinical-Radiomics Prognostic Model

Predictors from CPM and RPM were included in a comprehensive model based on both clinical and radiomics analysis (Clinical-Radiomics Prognostic Model, CRPM).

Finally, the metric that was used to evaluate the performance of incremental value of CPM, RPM and CRPM was the concordance index (*C*-index), computing its average value (and SD) over 50 iterations.

STATISTICAL ANALYSIS

All statistical analysis was performed on R software (version 4.0.2, R Foundation for Statistical Computing, Vienna, Austria) in this study. The following R packages were used:

- The "*glmnet*" package was used for LASSO
- The "*pROC*" packages were used to draw and analyze ROC-AUC curves, in particular by using the "*multiclass.roc*" function, and to compute the Youden Index
- The "*survival*" package was used for Cox regression analyses
- The "*superpc*" package was used for s-PCA

RESULTS

STUDY POPULATION

Overall, 285 patients with surgically resected lung cancer, 63 with pulmonary metastases and 41 with hamartomas were included in this study, resulting in the overall study population of 389 patients, whose characteristics are summarized in **Table 1** and **Table 2**.

Out of 389 patients, 185 nodules showed a maximum diameter - automatically computed by the software - lesser than 30 mm (105 LC; 50 metastases – 34 from colon cancer, 5 from endometrial cancer, 11 from renal cancer; and 30 hamartomas). There were 9 hamartomas displaying calcifications that were excluded from the analyses. Mean values for the major axis for the 185 nodules included in the histological analysis, according to their group, were 16.3 mm for ADK, 11.7 mm for metastases and 12.7 mm for hamartomas.

	Adenocarcinoma	*Metastasis*	*Hamartoma*	*p-value*
Total	285	63	41	
Age (years)	67.3 (±9.7)	67.6 (±9.6)	61 (±11.6)	***<0.05***
Sex	F=100 (35) M=185 (65)	F=36 (57) M=27 (43)	F=16 (39) M=25 (61)	***<0.05***
Lesion characteristics				
Diameter	29.8 mm (±20.3)	22.2 mm (±18.8)	21.7 mm (±15.2)	***<0.05***

Table 1. Study population characteristics. Continuous data are expressed as mean ± SD. Categorical data are presented as number with percentage in parentheses.

Patients with ADK

Stage (8 TNM)		Frequency	Deceased
	- IA1/IA2/ IA3	112 (39%)	29 (26%)
	- IB	62 (22%)	32 (52%)
	- IIA	7 (2%)	4 (57%
	- IIB	46 (16%)	29 (63%)
	- IIIA	34 (12%)	31 (91%)
	- IIIB	11 (4%)	11 (100%)
	- IV	7 (2%)	5 (71%)
	- Missing	7 (2%)	-
Histologic subtype	- Lepidic: 39 (14%)		
	- Papillary: 31 (11%)		
	- Micropapillary: 27 (9%)		
	- Acinar: 80 (28%)		
	- Glandular: 7 (2%)		
	- Solid: 97 (34%)		
	- Missing: 4 (1%)		
Oncologic history	71		
Smoking Status	- Never: 46		
	- Former: 74		
	- Current: 159		
	- Missing: 6		
Pulmonary function test	- DLCO: 3,3 (2,2 – 4,2)		

(median, IQR)	- DLCO%: 87 (73.5 – 109) - FEV1: 2,2 (1,9 – 2,7) - FEV1%: 92 (75 – 104,6) - FVC: 3,2 (2,7 – 3,8) - FVC%: 100 (87,7 – 114) - Tiffenau: 72 (65 – 77)
Surgery	- Lobectomy: 250 (number of deaths: 124) - Segmentectomy: 3 (number of deaths: 1) - Bi-lobectomy/Pneumonectomy: 16 (number of deaths: 12) - Wedge resection: 12 (number of deaths: 4) - Missing: 2
Follow-up *(months; median, IQR)*	- Overall: 64 (32 – 97) o Deceased: 63 (32 – 95) - Alive: 69 (36 – 99)
Relapse (Y:N)	111:34 - Regional/Local: 74 - Distant: 37
Mortality	- Overall mortality: 141 deaths - Lung cancer specific mortality: 105/141 (74.5%) - Other causes o Cardiovascular disease: 5 o Extra-pulmonary neoplasm: 5 o Other causes: 5

	○ N/A: 21

Table 2. Clinical characteristics of patients with adenocarcinoma.

HISTOLOGICAL STRATIFICATION

The number of RFs extracted for each pulmonary lesion was 851. After the redundancy analysis which allowed to exclude highly correlated RFs, the number was reduced to 530 (62%) (**Figure 2**). After LASSO regularization the average number of RF selected for the model was 10 (±3) (**Figure 3**).

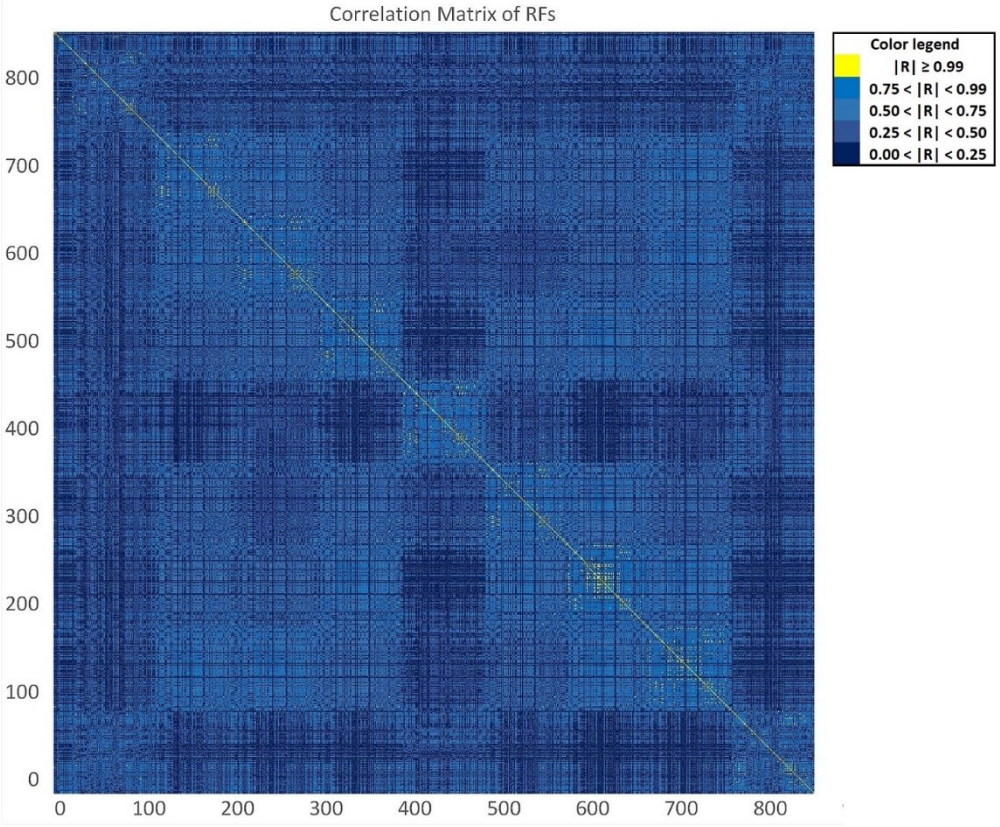

Figure 2. Correlation matrix displaying color-coded Spearman correlation coefficients between RFs (Correlation matrix). For couples of RFs with correlation coefficient equal or greater than 0.99 (yellow cells), one of the two RFs was discarded.

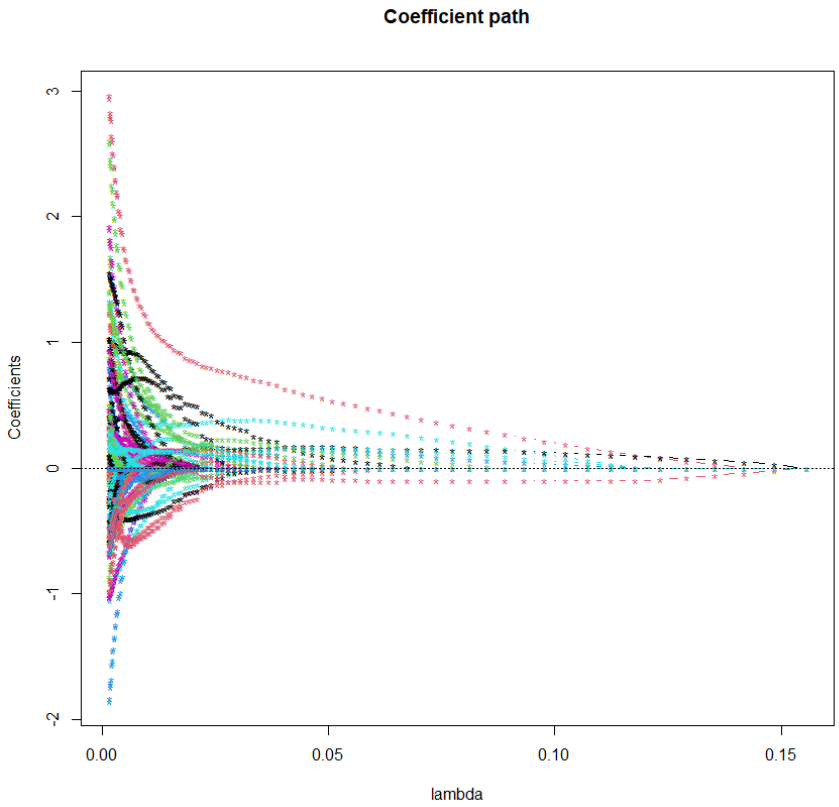

Figure 3. Least Absolute Shrinkage and Selection Operator (LASSO) logistic regression model coefficient profiles of the 530 RFs. As the tuning hyper parameter (λ) increased, more coefficient profiles tended to 0: the plot shows RF with coefficients values going to zero, as the penalty in the objective function of the LASSO is increased.

Model performance

Stability of the feature selection method

Stability of the model according to the stability indexes was substantially robust for all the four metrics computed (ANHD <0.001; DICE 0.6; ASM 0.7; CW 0.6).

Diagnostic Performance Rate

ROC-AUC values were 78% (± 0.06) for training set and 73% (± 0.05) for the test set. Precision in the training set was 0.70 (± 0.07), 0.58 (± 0.08) and 0.63 (± 0.24) for adenocarcinoma, metastasis, and hamartoma, respectively. In the test set, precision was 0.66 (± 0.08), 0.58 (± 0.21) and 0.34 (± 0.29) for adenocarcinoma, metastasis, and hamartoma, respectively. Recall in the training set was 0.98 (± 0.02), 0.32 (± 0.18) and 0.17 (± 0.12) for adenocarcinoma, metastasis and hamartoma, respectively. In the test set, precision was 0.96 (± 0.04), 0.28 (± 0.16) and 0.14 (± 0.13) for adenocarcinoma, metastasis and hamartoma, respectively (**Figure 4**).

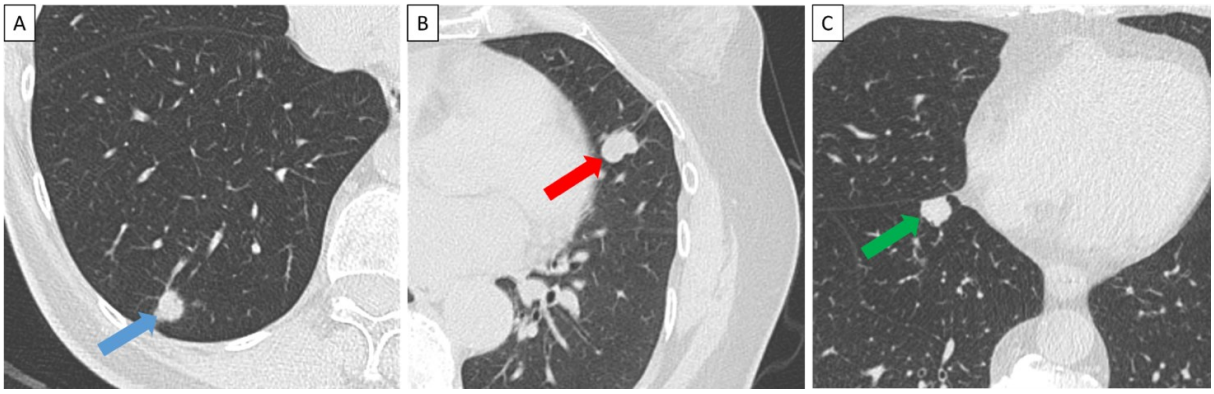

Figure 4. Examples of pulmonary nodules with correct histological classification obtained by the Radiomics approach. (A) Computed Tomography of a 64-year-old female patient with a pulmonary nodule in the right lower lobe (blue arrow). Pathological analysis after surgery showed evidence of adenocarcinoma. (B) Computed Tomography of a 61-year-old female patient showing a pulmonary nodule in the left upper lobe (red arrow). Pathological analysis after surgery reported hamartoma. (C) Computed Tomography performed in a 76-year-old female patient with colon cancer showing a pulmonary nodule in the right lower lobe (green arrow). Pathological analysis after surgery reported metastases.

Most frequent Features

The ten most frequent RF selected throughout the 2,000 iterations are reported in **Table 3**.

	Radiomics Feature	*Frequency*	*Percentage*
Original	First Order - Kurtosis	1575	79
Wavelet $_{LHL}$	GLCM - MCC	1498	75
Original	Shape - Sphericity	1484	74
Wavelet $_{HLL}$	First Order - Mean	1198	60
Original	Shape - Maximum2D-DiameterSlice	1147	57
Original	First Order – 10 Percentile	1078	54
Wavelet $_{LLL}$	GLCM - Correlation	1045	52
Original	First Order - Skewness	839	42
Wavelet $_{LLL}$	GLCM - MCC	566	28
Wavelet $_{HLL}$	First Order - Median	502	25

Table 3. Ten most frequent Radiomic Features selected throughout the 2,000 iterations.

SURVIVAL ANALYSIS

The overall study population was composed by 285 individuals who underwent LC surgery. The analyses were repeated twice, first on the whole database of 285 patients (all stages ADK), subsequently on 174 patients with stage I disease (stage I ADK).

Among clinical variables, the only two parameters that were included in the CPM for both study populations were FVC% and gender, as they displayed a frequency of inclusion into

the model (based on p-value of the corresponding HR) of 84% (FVC%) and 94% (gender) for all stages ADK patients, and 70% (FVC%) and 96% (gender) for stage I ADK patients. The other clinical variables did not reach the threshold of 60% (**Table 4**).

	All stages ADK	**Stage I ADK**
Age	26	28
Gender	**94**	**96**
Oncologic history	42	44
Smoking status	34	38
DLCO	46	28
DLCO%	24	20
FEV1	18	18
FEV1%	58	58
FVC	22	26
FVC%	**84**	**70**
Tiffenau	24	8
Tiffenau%	6	2
Surgery	8	14

Table 4. Frequency of occurrence of clinical variables; data are expressed as percentages.

The s-PCA showed that the frequency of HR with p-values lesser than 0.05 in multivariate Cox regression analysis was higher for all stages ADK patients, with a 74% of runs with at least one component out of the three s-PCA (s-PCA1, s-PCA2, or S-PCA3) under the threshold of 0.05. On the other hand, for stage I ADK patients, there was at least one component under the threshold of 0.05 for p-value of HR in 22% of times.

C-indexes showed a minor increase from CPM to RPM, while the integrated model, including both clinical variables and RFs yielded the highest predictive performance. Such observation was true for both groups of patients, and the C-index reached its highest value for the CRPM of the study population of patients with stage I ADK (0.81, ± 0.04) (**Table 5**) (**Figure 5**).

	All stages ADK	**Stage I ADK**
Clinical Predictive Model	0.68 (0.06)	0.66 (0.06)
Radiomics Predictive Model	0.70 (0.04)	0.70 (0.06)
Clinical-Radiomics Predictive Model	0.74 (0.05)	0.81 (0.04)

Table 5. Average values of C-index (and standard deviation) after 50 runs, for clinical predictive model, radiomics predictive model and the comprehensive clinical-radiomics predictive model. Results are reported for the overall study population of patients with surgically resected adenocarcinomas and for patients with stage I adenocarcinoma.

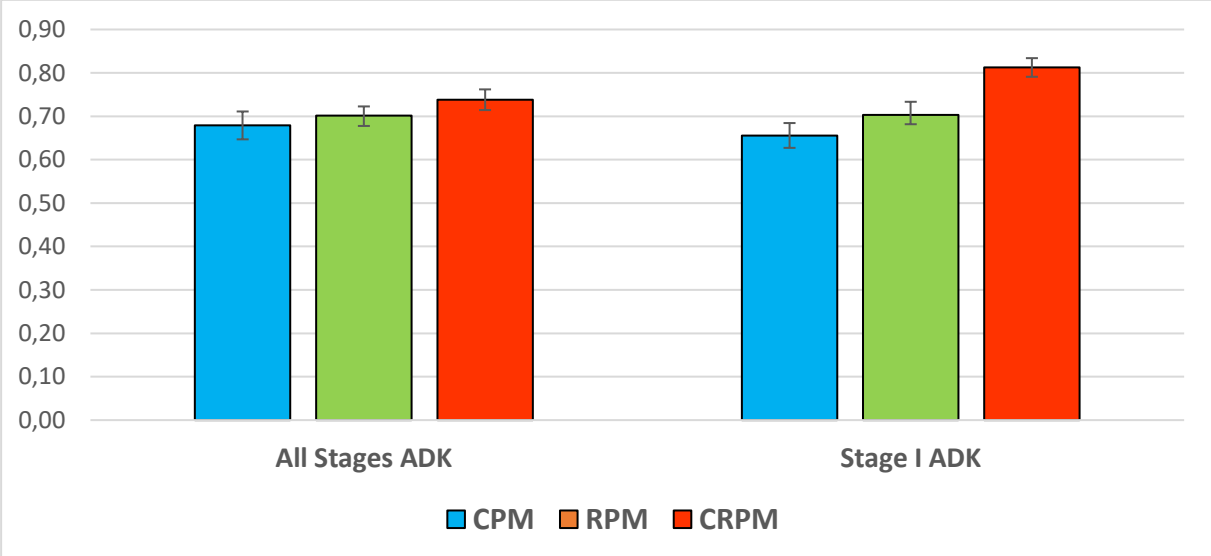

Figure 5. Histograms of the C-indexes of clinical predictive model (CPM), radiomics predictive model (RPM) and the comprehensive clinical-radiomics predictive model (CRPM). Results are reported for the overall study population of patients with all stages adenocarcinoma and for patients with stage I adenocarcinoma

DISCUSSION

This study tested a CT-Radiomics one-stop shop approach for characterization of lung cancer risk for pulmonary nodules, and for the stratification of prognosis for subjects with surgically resected adenocarcinoma.

The results of our study are two folded: first, we showed that CT-Radiomics may identify pulmonary lesions with different histology. Indeed, in a group of homogeneously selected pulmonary nodules (i.e. with maximum diameter < 30 mm, without pathognomonic patterns of calcifications), the CT-Radiomics approach yielded a ROC-AUC of 73% for the discrimination of the biological behavior of pulmonary nodules. Subsequently, we compared prognostic models for patients with adenocarcinoma based on clinical variables (CPM), RFs (RPM), or on an integrated clinical-radiomics model (CRPM). Such analysis was performed on two sets of patients: first, the overall study population; subsequently, only patients with stage I adenocarcinoma. Our study demonstrated that integration of clinical and RF (CRPM) yielded the highest predictive power.

This research takes place in the field of advanced analysis of CT images, testing whether quantitative parameters (RFs) might play a role as non-invasive biomarkers predicting biological behavior and prognosis. Notably, the first implementation of artificial intelligence within thoracic imaging was computer-aided detection (CAD) of pulmonary nodules; however, CAD tools were directed toward delivering a single judgment as a second opinion and do not provide prognostic data or therapy decision support [20,39]. Radiomics is a further development of such approach, and associates quantitative data (advanced analysis of shape, texture, intensity) with clinical features to define decision support models for precision medicine [20,40]. Radiomics is the high-throughput extraction of features imperceptible to the human eyes, such as pixel values (e.g. HU in CT images), variation of values within ROIs

(or VOIs), and edge strengths [14,21]. The workflow of Radiomics is designed in a multi-steps fashion, starting from image acquisition, delineation/contouring of ROIs/VOIs, and extraction and analysis of RF creating imaging-derived biomarkers defining biological behavior of pulmonary lesions [24,41]. The impact of scanning parameters has been explored within the field of quantitative analysis as different manufacturers, acquisition protocols and reconstruction parameters were shown to influence RFs [42-44]. Milanese et al tested texture analysis and a machine learning approach on patients suffering from systemic sclerosis (SSc), comparing RFs extracted from CT scans with different acquisition parameters and reconstruction algorithms, showing that both texture-based diagnosis of SSc and classification between cases and controls was possible [34]. Noteworthy, for the implementation of Radiomics within clinical practice classification and stratification of prognosis of neoplasms should be unaffected by technical parameters, to allow for reproducible analyses between centers [45]. However, in daily routine, standardization of image acquisition and reconstruction parameters - although desirable - will be difficult to be achieved [42,43,46]. In this project, we tested the impact of machine learning on the classification of malignant and benign pulmonary lesions (i.e. hamartomas), with CT acquired from different scanners and reconstruction parameters, thus reflecting a real-world scenario.

Our first analysis aimed at evaluating LC risk of subjects displaying pulmonary nodules, by extracting RFs from nodules at the time of first detection. Early (or immediate) assessment of biological behavior of a pulmonary lesion could reduce subjects' anxiety as well as costs for the health systems, by limiting the number of both invasive (e.g. biopsies) or non-invasive (18F-FDG-PET scans) procedures [47,48]. In this study, we included pulmonary lesions with a maximum diameter lesser than 30 mm, such as surgically resected primary LC (adenocarcinoma) and pulmonary metastases (from colon carcinoma, renal cell carcinoma,

and endometrial carcinoma) as malignant lesions, and surgically resected hamartomas as benign lesions. To exclude the impact of hamartomas' calcifications, as the increased HU values because of calcifications would differentiate their texture also on a visual basis, a preliminary analysis allowed us to select only non-calcified hamartomas. Our model yielded a ROC-AUC in the test population of 73%. Previously, Beig et al reported that RFs extracted from both intranodular and perinodular regions may discriminate NSCLC and benign granulomas (ROC-AUC of 0.80 by combining RF from the two regions), with higher performance for RF analysis as compared to human readers [49]. Notably, we included in our quantitative analyses not only primary LC and pathologically proven benign lesions (hamartomas), but also lesions representing secondary pulmonary involvement (i.e. metastases from renal, endometrial and colon cancer). We do recognize that clinical suspicion of metastatic disease could be potentially easy in case of nodules scattered throughout the pulmonary parenchyma of patients with known extra-thoracic neoplasms; however, our approach was directed toward the differentiation of pulmonary nodules based on their RFs, and we foster future studies testing Radiomics as a tool to discriminate between metastases of different neoplasms. Such approach may be useful to optimize subsequent diagnostic workflow for patients (i.e. endoscopic evaluation for subjects for whom RFs analysis would suggest that a – potentially single – pulmonary nodule might represent a metastasis from colon cancer). Because the strict inclusion criteria selected for this study, the number of eligible subjects was limited, therefore the number of observations was not enough to discriminate metastases based on their origin. Nonetheless, our results are in keeping with Kirienko et al., who trained a model (ROC-AUC of 79%) to discriminate primary and secondary lung lesions [50]. Yang et al. trained a model (ROC-AUC of 77%) differentiating between invasive and non-invasive adenocarcinoma [51]. Based on data from the National Lung Screening Trial, Huang et al realized a matched case-control study

testing features extracted from intranodular, perinodular and extranodular areas to increase the positive predictive value of pulmonary nodules detected in a lung cancer screening scenario [52,53].

Our results showed that among the ten most frequent RFs selected throughout the 2,000 iterations for the histological prediction model, 50% of them was composed by first order features, both from original and wavelet transformed groups. The latter are based on histograms describing distribution of voxel intensities [14,54] and include metrics (e.g. mean, median, standard deviation, kurtosis, skewness, energy, entropy, uniformity, and variance) that – along with features describing the texture of the contoured lesions - appear to be reproducible and useful for predictive models [55-58]. Furthermore, metrics describing size and shape of nodules were included among the most frequently selected RFs, thus underscoring the role of morphological descriptors as predictors for biological behavior, not only from qualitative (i.e. evaluation performed by radiologists, taking into account size of the nodules) but also from quantitative (Radiomics-based) approaches [59]. In our study, recall values (representing the number of positive cases selected from the algorithm out of the overall positive cases) were above 0.95 for adenocarcinomas in both training and test sets. On the other hand, precision and recall values were lower for metastases (0.58 and 0.28, respectively) and hamartomas (0.34 and 0.14, respectively). This might be due the low number of patients with metastases and hamartomas. Notably, we tested a Radiomics approach evaluating more than a binary variable (malignant/benign), as we included two different types of malignant pulmonary lesions (i.e. primary adenocarcinoma and secondary diseases).

Reducing the number of manual inputs might be mandatory for Radiomics to be implemented in the clinical workflow [60]. We performed our analyses with a semi-automated approach for the delineation of a three-dimensional VOI. As completely manual contouring of

lesions can be time consuming, the process should be fast and reproducible for radiomics of pulmonary nodules to be implemented in daily practice. Pavic et al reported that the reproducibility of 1,404 RF extracted from patients affected by different tumors (i.e. head and neck squamocellular carcinoma, NSCLC, and malignant pleural mesothelioma) was influenced by inter-observer variability [61]. A semi-automated segmentation may reduce inter-observer variability and reproducibility of RF from various categories (shape, histogram-derived, GLCM and GLRLM) extracted from CT of patients with NSCLC, were more robust and reproducible when semi-automatically extracted, as compared to manual extraction [62,63]. Notably, RFs from GLCM and GLRML depend on the accuracy of the delineation of boundaries and on tumors' irregularity, which is better depicted through a semi-automated approach [31,64]. Among the RFs included in our model for histological stratification, three out of the ten most frequent RFs selected in the 2,000 iterations were from GLCM matrix, and two were extracted from the same wavelet group (LLL). The GLCM matrix measures heterogeneity, and we postulate that based on such metric our CT-Radiomics approach stratifies lesions with different biological behavior, in keeping with previous observation on heterogeneity-descriptors RFs associated to malignant lesions with aggressive behavior correlating with outcome [65,66].

RFs were traditionally differentiated between agnostic and semantic features; the latter are derived from the analysis of the reading radiologist (e.g. morphology, metrics for size), whilst agnostic RFs derives from mathematical analyses [20,67]. Integrating semantic features into the predictive models may increase their performance, although with such approach the experience of the readers might affect the classification processes [68,69]. In our predictive model we included only agnostic RFs derived from one of the most used tools for Radiomics (based on PyRadiomics, implemented in a user-friendly software for contouring of lesions). PyRadiomics is recognized as a tool for the extraction of RF that satisfies the criteria

from the Imaging Biomarker Standardization Initiative (IBSI) [16]. Therefore, one of the main aim of studies based on radiomics is the generalizability of the results and using a software that has a large distribution is of paramount relevance for reproducing and testing Radiomics signatures [16]. The integration of RFs with metrics assessing smoke-related pulmonary damage might increase the performance of discrimination of pulmonary lesions with different biological behavior [41,70]. We foster further studies testing such integrated multi-parametric approach, albeit it could request large populations to account for overfitting of data and high-end computational units to develop such complex models.

Subsequently, we evaluated models predicting survival of patients with adenocarcinoma: both clinical and radiomics were predictors of OS in two study populations (patients with all stage disease and patients with stage I disease). However, the best performance was from the integrated CRPM (0.74 in all stage disease, 0.81 in stage I disease). Nonetheless, beyond this absolute value lesser iterations with at least one of the s-PCA components having significant HR were found in stage I patients, because of the lower number of cases included in this subset. For confirmation of the predictive role of CRPM, large multicentric studies should be performed to test such integrated approaches for stratification of prognosis. Xie et al developed a Radiomics signature to categorize risk of disease recurrence (high-risk and low-risk) for 237 subjects with stage I adenocarcinoma included in their validation cohort. Notably, patients with stage I adenocarcinoma suffer from a substantial risk of recurrence, which might develop even after complete resection [71]. Incidence of N2 disease in clinical stage I NSCLC ranges 4.0 to 16.8%, with CT showing very poor sensitivity and specificity in the detection of mediastinal lymph node involvement [72-74]. Yang et al tested a CT-based radiomics approach to predict N2 disease in 1.212 individuals with stage I lung adenocarcinoma, and their radiomics signature (including RFs describing size, density and

heterogeneity) showed better predictive performance as compared to clinical parameters [75]. The efficacy of adjuvant chemotherapy reducing recurrence rate in stage I disease is being tested, and a radiomics approach may integrate prognostic prediction, along with prediction of stage of the disease or progression-free survival [9].

For patients with advanced LC, prognosis is influenced by the presence of molecular and genetic alterations targeted by specific drugs [76]. Previous studies tested the impact of integrated clinical and radiological semantic features to discriminate between anaplastic lymphoma kinase (ALK) rearrangement from epidermal growth factor receptor (EGFR) mutations in advanced lung adenocarcinoma [77]. Indeed, the detection of molecular alterations gained utmost relevance, and to overcome possible limitations of molecular tests based on analysis of tumor samples derived from biopsies and/or cytological specimens, non-invasive evaluation of lesions could be of particular clinical relevance [78]. Sun et al tested a Radiomics signature in correlation with lymphocyte infiltrates in an heterogeneous study population encompassing different neoplasms treated by immunotherapy, reporting that homogeneous tumors correlated with increased overall survival [79]. Similarly, Mazzaschi et al explored the potential role of RFs to non-invasively define the tumor immune microenvironment (TIME): on a study population composed by 100 surgically resected NSCLC, a Radiomics signature encompassing different RFs outperformed the prognostic value of the standard pTNM staging system in NSCLC [68,69]. Ravanelli et al reported that an approach focused on prediction of survival on patients treated with Nivolumab after failure of platinum-based chemotherapy based on a low-throughput approach producing a low number of parameters, namely the filtered-histogram approach, can be used to predict outcome of NSCLC patients [80].

Efforts are being made to define descriptive and unstandardized terms to establish a uniform lexicon for semantic annotations [81,82]. To overcome the limitation represented by

variability in detecting and quantifying semantic features, Radiomics - extracting and analyzing quantitative and reproducible imaging biomarkers - may non-invasively and rapidly predict tumor phenotypes [83]. Nonetheless, a recent systematic review evaluating the role of imaging-based models reported that further validation is required before non-invasive approaches could replace traditional molecular pathology testing, as studies performed under strict adherence to quality guidelines are requested [83].

In a study performed on 68 malignant LC, Ferreira Junior et al reported that Radiomics yielded ROC-AUC of 0.75, 0.71, and 0.81 for detecting lymph nodal metastasis, distant metastasis, and histopathology pattern recognition [84]. Notably, they reported that the integration of clinical data with RFs showed high potential on predicting nodal metastasis, distant metastasis and histopathological recognition. Similarly, Zhang et al developed models including CT-RFs extracted from brain lesions to discriminate between different histology of LC: the combination of clinical parameters (such as age and gender) increased the performance of the radiomics models [85]. This is in keeping with our results, as CRPM showed highest performance in predicting OS in both study populations. Choe et al recently developed a CT-Radiomics approach for resectable lung adenocarcinoma patients, showing incremental value over clinical-pathological models in predicting disease free survival, except in prediction of OS in the validation cohort [86]. Parmar et al reported that RFs are cancer-specific and that they can play a role as prognostic or predicitive imaging biomarkers [87]. Furthermore, RFs might predict pathologic response in patients with LC undergoing chemo-radiation therapy [88].

Notably, the impact of machine learning expands beyond prognostication of pulmonary nodules; indeed, it may increase detectability of lesions simultaneously allowing the extraction of metrics that are currently used to guide management of lesions (i.e.

measurement of volume): a post-processing tool removing vascular, bronchial and fissural structures can facilitate detection of pulmonary nodules [89].

Any identification of reliable and meaningful quantitative imaging biomarkers for tumors must be reproducible and efforts should be made to standardize dose administration, image acquisition, image reconstruction, and value normalization [90,91]. Furthermore, the methodology used to extract RFs is subject to variability, too [92-94]. Because it is essentially high-throughput data mining, Radiomics must incorporate appropriate feature selection strategies that enhance the performance of its predictive models while minimizing the overfitting of those models to increase generalizability [95,96]. In addition, predictive and prognostic models with high accuracy, reliability, and efficiency are vital factors driving the success of Radiomics. In this study, we included analyses of the stability of our models by the random repetition of training and test splitting and stability metrics were included for the description of the histology predictive model. Such approach was not possible for the survival analysis because of the sPCA design. Data must be replicated in other institutions, and biomarkers must be validated independently to reliably generalized the results [91,97,98].

Limitations

This study suffers from various limitations. As a retrospective study, not all variables were available for the study population. The number of subjects was limited, particularly for patients included in the metastases and hamartomas groups. We plan to test our Radiomics approach in external populations, as well as in more advanced stage disease. Clinical predictors were included in the CPM when p-values of their HR was below 0.3 in more than 60% of the iterations: such higher-than standard threshold was set to include parameters selected as potential predictors in more than a single iteration. Indeed, by selecting a lower threshold for p-values, we could have introduced into the model variables for which the

significance could derive from a single iteration, thus reducing robustness of the analysis. We evaluated all-cause mortality, without limiting the study population to patients with lung cancer-mortality, as this could result in poorer performance because of the limited number of events.

In conclusion, we developed a CT-Radiomics approach defining lung cancer risk for subjects with pulmonary nodules, to differentiate lesions with malignant and benign biological behavior. Furthermore, stratification of prognosis for patients with adenocarcinoma can be increased by the combination of clinical and radiomics features.

REFERENCES

1. Silva M, Pastorino U, Sverzellati N. Lung cancer screening with low-dose CT in Europe: strength and weakness of diverse independent screening trials. Clin Radiol 2017;72:389-400.

2. Wong MCS, Lao XQ, Ho KF, Goggins WB, Tse SLA. Incidence and mortality of lung cancer: global trends and association with socioeconomic status. Sci Rep 2017;7:14300.

3. Ferreira-Junior JR, Koenigkam-Santos M, Magalhaes Tenorio AP, et al. CT-based radiomics for prediction of histologic subtype and metastatic disease in primary malignant lung neoplasms. Int J Comput Assist Radiol Surg 2020;15:163-72.

4. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017. Radiology 2017;284:228-43.

5. Silva M, Milanese G, Sestini S, et al. Lung cancer screening by nodule volume in Lung-RADS v1.1: negative baseline CT yields potential for increased screening interval. Eur Radiol 2020.

6. Callister ME, Baldwin DR, Akram AR, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. Thorax 2015;70 Suppl 2:ii1-ii54.

7. Henschke CI, Yankelevitz DF, Mirtcheva R, et al. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 2002;178:1053-7.

8. Devaraj A, van Ginneken B, Nair A, Baldwin D. Use of Volumetry for Lung Nodule Management: Theory and Practice. Radiology 2017;284:630-44.

9. Padole A, Singh R, Zhang EW, et al. Radiomic features of primary tumor by lung cancer stage: analysis in BRAF mutated non-small cell lung cancer. Transl Lung Cancer Res 2020;9:1441-51.

10. Truong MT, Sabloff BS, Ko JP. Multidetector CT of solitary pulmonary nodules. Radiol Clin North Am 2010;48:141-55.

11. Gierada DS, Pilgram TK, Ford M, et al. Lung cancer: interobserver agreement on interpretation of pulmonary findings at low-dose CT screening. Radiology 2008;246:265-72.

12. Papanikolaou N, Matos C, Koh DM. How to develop a meaningful radiomic signature for clinical use in oncologic patients. Cancer Imaging 2020;20:33.

13. Marchiano AV, Cosentino M, Di Tolla G, et al. FNA and CNB in the diagnosis of pulmonary lesions: a single-center experience on 665 patients, comparison between two periods. Tumori 2017;103:360-6.

14. Parekh V, Jacobs MA. Radiomics: a new application from established techniques. Expert Rev Precis Med Drug Dev 2016;1:207-26.

15. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017.

16. Fornacon-Wood I, Mistry H, Ackermann CJ, et al. Reliability and prognostic value of radiomic features are highly dependent on choice of feature extraction platform. Eur Radiol 2020;30:6241-50.

17. Castiglioni I, Gallivanone F, Soda P, et al. AI-based applications in hybrid imaging: how to build smart and truly multi-parametric decision models for radiomics. Eur J Nucl Med Mol Imaging 2019;46:2673-99.

18. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: The promise of radiomics. Phys Med 2017;38:122-39.

19. Shannon CE. The mathematical theory of communication. 1963. MD Comput 1997;14:306-17.

20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.

21. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine Learning for Medical Imaging. Radiographics 2017;37:505-15.

22. Ichihara E, Hotta K, Kubo T, et al. Clinical significance of repeat rebiopsy in detecting the EGFR T790M secondary mutation in patients with non-small cell lung cancer. Oncotarget 2018;9:29525-31.

23. Digumarthy SR, Padole AM, Gullo RL, Sequist LV, Kalra MK. Can CT radiomic analysis in NSCLC predict histology and EGFR mutation status? Medicine (Baltimore) 2019;98:e13963.

24. 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.

25. Lian C, Ruan S, Denoeux T, Jardin F, Vera P. Selecting radiomic features from FDG-PET images for cancer treatment outcome prediction. Med Image Anal 2016;32:257-68.

26. Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. Transl Lung Cancer Res 2017;6:86-91.

27. Hassani C, Varghese BA, Nieva J, Duddalwar V. Radiomics in Pulmonary Lesion Imaging. AJR Am J Roentgenol 2019;212:497-504.

28. Tailor TD, Schmidt RA, Eaton KD, Wood DE, Pipavath SN. The Pseudocavitation Sign of Lung Adenocarcinoma: A Distinguishing Feature and Imaging Biomarker of Lepidic Growth. J Thorac Imaging 2015;30:308-13.

29. Koenigkam Santos M, Muley T, Warth A, et al. Morphological computed tomography features of surgically resectable pulmonary squamous cell carcinomas: impact on prognosis and comparison with adenocarcinomas. Eur J Radiol 2014;83:1275-81.

30. Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011;6:244-85.

31. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.

32. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Muller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. Radiology 2008;246:697-722.

33. Landini N, Milanese G, Zambrini E, et al. Computed tomography - histology correlations of unusual lung tumors. Pathologica 2016;108:110-9.

34. Milanese G, Mannil M, Martini K, Maurer B, Alkadhi H, Frauenfelder T. Quantitative CT texture analysis for diagnosing systemic sclerosis: Effect of iterative reconstructions and radiation doses. Medicine (Baltimore) 2019;98:e16423.

35. Wu L, Wang C, Tan X, et al. Radiomics approach for preoperative identification of stages I-II and III-IV of esophageal cancer. Chin J Cancer Res 2018;30:396-405.

36. Khaire UM, Dhanalakshmi R. Stability of feature selection algorithm: A review. Journal of King Saud University - Computer and Information Sciences 2019.

37. Somol P, Novovicova J. Evaluating stability and comparing output of feature selectors that optimize feature subset cardinality. IEEE Trans Pattern Anal Mach Intell 2010;32:1921-39.

38. Barshan E, Ghodsi A, Azimifar Z, Zolghadri Jahromi M. Supervised principal component analysis: Visualization, classification and regression on subspaces and submanifolds. Pattern Recognition 2011;44:1357-71.

39. Giger ML. Machine Learning in Medical Imaging. J Am Coll Radiol 2018;15:512-20.

40. Yan M, Wang W. Radiomic Analysis of CT Predicts Tumor Response in Human Lung Cancer with Radiotherapy. J Digit Imaging 2020.

41. Silva M, Milanese G, Seletti V, Ariani A, Sverzellati N. Pulmonary quantitative CT imaging in focal and diffuse disease: current research and clinical applications. Br J Radiol 2018:20170644.

42. Larue R, van Timmeren JE, de Jong EEC, et al. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study. Acta Oncol 2017:1-10.

43. Mackin D, Fave X, Zhang L, et al. Measuring Computed Tomography Scanner Variability of Radiomics Features. Invest Radiol 2015;50:757-65.

44. Ahn SJ, Kim JH, Lee SM, Park SJ, Han JK. CT reconstruction algorithms affect histogram and texture analysis: evidence for liver parenchyma, focal solid liver lesions, and renal cysts. Eur Radiol 2018.

45. Berenguer R, Pastor-Juan MDR, Canales-Vazquez J, et al. Radiomics of CT Features May Be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. Radiology 2018:172361.

46. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of A Method to Compensate Multicenter Effects Affecting CT Radiomics. Radiology 2019:182023.

47. Abreu C, Grilo A, Lucena F, Carolino E. Oncological Patient Anxiety in Imaging Studies: the PET/CT Example. J Cancer Educ 2016.

48. Yang Y, Czernin J. Contribution of imaging to cancer care costs. J Nucl Med 2011;52 Suppl 2:86S-92S.

49. Beig N, Khorrami M, Alilou M, et al. Perinodular and Intranodular Radiomic Features on Lung CT Images Distinguish Adenocarcinomas from Granulomas. Radiology 2019;290:783-92.

50. Kirienko M, Cozzi L, Rossi A, et al. Ability of FDG PET and CT radiomics features to differentiate between primary and metastatic lung lesions. Eur J Nucl Med Mol Imaging 2018;45:1649-60.

51. Yang B, Guo L, Lu G, Shan W, Duan L, Duan S. Radiomic signature: a non-invasive biomarker for discriminating invasive and non-invasive cases of lung adenocarcinoma. Cancer Manag Res 2019;11:7825-34.

52. Huang P, Park S, Yan R, et al. Added Value of Computer-aided CT Image Features for Early Lung Cancer Diagnosis with Small Pulmonary Nodules: A Matched Case-Control Study. Radiology 2018;286:286-95.

53. Papapietro VR, Milanese G, Borghesi A, Sverzellati N, Silva M. Look around your target: a new approach to early diagnosis of lung cancer. Ann Transl Med 2018;6:S77.

54. Bae KT, Fuangtharnthip P, Prasad SR, Joe BN, Heiken JP. Adrenal masses: CT characterization with histogram analysis method. Radiology 2003;228:735-42.

55. Yang J, Zhang L, Fave XJ, et al. Uncertainty analysis of quantitative imaging features extracted from contrast-enhanced CT in lung tumors. Comput Med Imaging Graph 2016;48:1-8.

56. Kamiya A, Murayama S, Kamiya H, Yamashiro T, Oshiro Y, Tanaka N. Kurtosis and skewness assessments of solid lung nodule density histograms: differentiating malignant from benign nodules on CT. Jpn J Radiol 2014;32:14-21.

57. E L, Lu L, Li L, Yang H, Schwartz LH, Zhao B. Radiomics for Classifying Histological Subtypes of Lung Cancer Based on Multiphasic Contrast-Enhanced Computed Tomography. J Comput Assist Tomogr 2019;43:300-6.

58. Baek HJ, Kim HS, Kim N, Choi YJ, Kim YJ. Percent change of perfusion skewness and kurtosis: a potential imaging biomarker for early treatment response in patients with newly diagnosed glioblastomas. Radiology 2012;264:834-43.

59. Ko JP, Berman EJ, Kaur M, et al. Pulmonary Nodules: growth rate assessment in patients by using serial CT and three-dimensional volumetry. Radiology 2012;262:662-71.

60. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and Prognosis of Quantitative Features Extracted from CT Images. Transl Oncol 2014;7:72-87.

61. Pavic M, Bogowicz M, Wurms X, et al. Influence of inter-observer delineation variability on radiomics stability in different tumor sites. Acta Oncol 2018;57:1070-4.

62. Mitchell DM, Perry L, Smith S, et al. Assessing the effect of a contouring protocol on postprostatectomy radiotherapy clinical target volumes and interphysician variation. Int J Radiat Oncol Biol Phys 2009;75:990-3.

63. Fuller CD, Nijkamp J, Duppen JC, et al. Prospective randomized double-blind pilot study of site-specific consensus atlas implementation for rectal cancer target volume delineation in the cooperative group setting. Int J Radiat Oncol Biol Phys 2011;79:481-9.

64. Orlhac F, Boughdad S, Philippe C, et al. A Postreconstruction Harmonization Method for Multicenter Radiomic Studies in PET. J Nucl Med 2018;59:1321-8.

65. Bowen SR, Yuh WTC, Hippe DS, et al. Tumor radiomic heterogeneity: Multiparametric functional imaging to characterize variability and predict response following cervical cancer radiation therapy. J Magn Reson Imaging 2018;47:1388-96.

66. Thawani R, McLane M, Beig N, et al. Radiomics and radiogenomics in lung cancer: A review for the clinician. Lung Cancer 2018;115:34-41.

67. Park JE, Kim HS. Radiomics as a Quantitative Imaging Biomarker: Practical Considerations and the Current Standpoint in Neuro-oncologic Studies. Nucl Med Molec Imag 2018;52:99-108.

68. Mazzaschi G, Milanese G, Pagano P, et al. Integrated CT imaging and tissue immune features disclose a radio-immune signature with high prognostic impact on surgically resected NSCLC. Lung Cancer 2020;144:30-9.

69. Mazzaschi G, Milanese G, Pagano P, et al. Dataset on the identification of a prognostic radio-immune signature in surgically resected Non Small Cell Lung Cancer. Data Brief 2020;31:105781.

70. Sverzellati N, Randi G, Spagnolo P, et al. Increased mean lung density: another independent predictor of lung cancer? Eur J Radiol 2013;82:1325-31.

71. Xie D, Wang TT, Huang SJ, et al. Radiomics nomogram for prediction disease-free survival and adjuvant chemotherapy benefits in patients with resected stage I lung adenocarcinoma. Transl Lung Cancer Res 2020;9:1112-23.

72. Defranchi SA, Cassivi SD, Nichols FC, et al. N2 disease in T1 non-small cell lung cancer. Ann Thorac Surg 2009;88:924-8.

73. Zhang Y, Sun Y, Xiang J, Zhang Y, Hu H, Chen H. A prediction model for N2 disease in T1 non-small cell lung cancer. J Thorac Cardiovasc Surg 2012;144:1360-4.

74. Carnevale A, Milanese G, Sverzellati N, Silva M. A novel prediction model for the probability of mediastinal lymph node metastases detected by endobronchial ultrasound-transbronchial needle aspiration in non-small cell lung cancer: possible applications in clinical decision-making. Mediastinum 2018;2.

75. Yang M, She Y, Deng J, et al. CT-based radiomics signature for the stratification of N2 disease risk in clinical stage I lung adenocarcinoma. Transl Lung Cancer Res 2019;8:876-85.

76. Park H, Sholl LM, Hatabu H, Awad MM, Nishino M. Imaging of Precision Therapy for Lung Cancer: Current State of the Art. Radiology 2019;293:15-29.

77. Miao Y, Zhu S, Li H, et al. Comparison of clinical and radiological characteristics between anaplastic lymphoma kinase rearrangement and epidermal growth factor receptor mutation in treatment naive advanced lung adenocarcinoma. J Thorac Dis 2017;9:3927-37.

78. Mazzaschi G, Pagano P, Milanese G, et al. Advanced CT imaging features reflect distinct tissue immune profiles and exhibit high prognostic impact on NSCLC. ESMO IMMUNO-ONCOLOGY CONGRESS 2018. Geneva, Switzerland2018.

79. 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.

80. Ravanelli M, Agazzi GM, Milanese G, et al. Prognostic and predictive value of histogram analysis in patients with non-small cell lung cancer refractory to platinum treated by nivolumab: A multicentre retrospective study. Eur J Radiol 2019;118:251-6.

81. Opulencia P, Channin DS, Raicu DS, Furst JD. Mapping LIDC, RadLex, and lung nodule image features. J Digit Imaging 2011;24:256-70.

82. Aerts HJ. The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review. JAMA Oncol 2016;2:1636-42.

83. Ninatti G, Kirienko M, Neri E, Sollini M, Chiti A. Imaging-Based Prediction of Molecular Therapy Targets in NSCLC by Radiogenomics and AI Approaches: A Systematic Review. Diagnostics (Basel) 2020;10.

84. Ferreira Junior JR, Koenigkam-Santos M, Cipriano FEG, Fabro AT, Azevedo-Marques PM. Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. Comput Methods Programs Biomed 2018;159:23-30.

85. Zhang J, Jin J, Ai Y, et al. Differentiating the pathological subtypes of primary lung cancer for patients with brain metastases based on radiomics features from brain CT images. Eur Radiol 2020.

86. Choe J, Lee SM, Do KH, et al. Outcome prediction in resectable lung adenocarcinoma patients: value of CT radiomics. Eur Radiol 2020;30:4952-63.

87. Parmar C, Leijenaar RT, Grossmann P, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. Sci Rep 2015;5:11044.

88. Chong Y, Kim JH, Lee HY, et al. Quantitative CT variables enabling response prediction in neoadjuvant therapy with EGFR-TKIs: are they different from those in neoadjuvant concurrent chemoradiotherapy? PLoS One 2014;9:e88598.

89. Milanese G, Eberhard M, Martini K, Vittoria De Martini I, Frauenfelder T. Vessel suppressed chest Computed Tomography for semi-automated volumetric measurements of solid pulmonary nodules. Eur J Radiol 2018;101:97-102.

90. Zhao B, Tan Y, Tsai WY, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. Sci Rep 2016;6:23428.

91. Lee G, Lee HY, Park H, et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. Eur J Radiol 2017;86:297-307.

92. Leijenaar RT, Nalbantov G, Carvalho S, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. Sci Rep 2015;5:11075.

93. Fave X, Mackin D, Yang J, et al. Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? Med Phys 2015;42:6784-97.

94. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. J Med Imaging (Bellingham) 2015;2:041002.

95. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts H. Machine Learning methods for Quantitative Radiomic Biomarkers. Sci Rep 2015;5:13087.

96. Reyes-Aldasoro CC, Bhalerao A. Volumetric texture description and discriminant feature selection for MRI. Inf Process Med Imaging 2003;18:282-93.

97. Alic L, Niessen WJ, Veenland JF. Quantification of heterogeneity as a biomarker in tumor imaging: a systematic review. PLoS One 2014;9:e110300.

98. O'Connor JP, Rose CJ, Waterton JC, Carano RA, Parker GJ, Jackson A. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. Clin Cancer Res 2015;21:249-57.