

## Cardiothoracic Imaging

## Radiogenomics in personalized management of lung cancer patients: Where are we?



Jose Arimateia Batista Araujo-Filho<sup>a,b,1</sup>, Maria Mayoral<sup>a,c,\*1</sup>, Natally Horvat<sup>a</sup>, Fernando C. Santini<sup>d</sup>, Peter Gibbs<sup>a,e</sup>, Michelle S. Ginsberg<sup>a</sup>

<sup>a</sup> Department of Radiology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

<sup>b</sup> Department of Radiology, Hospital Sirio-Libanes, Rua Adma Jafet 91, São Paulo, SP 01308-050, Brazil

<sup>c</sup> Medical Imaging Department, Hospital Clinic of Barcelona, 170 Villarroel street, Barcelona 08036, Spain

<sup>d</sup> Thoracic Oncology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

<sup>e</sup> Department of Medical Physics, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

## ARTICLE INFO

## Keywords:

Lung cancer

Radiomics

Radiogenomics

## ABSTRACT

With the rise of artificial intelligence, radiomics has emerged as a field of translational research based on the extraction of mineable high-dimensional data from radiological images to create “big data” datasets for the purpose of identifying distinct sub-visual imaging patterns. The integrated analysis of radiomic data and genomic data is termed radiogenomics, a promising strategy to identify potential imaging biomarkers for predicting driver mutations and other genomic parameters. In lung cancer, recent advances in whole-genome sequencing and the identification of actionable molecular alterations have led to an increased interest in understanding the complex relationships between imaging and genomic data, with the potential of guiding therapeutic strategies and predicting clinical outcomes. Although the integration of the radiogenomics data into lung cancer management may represent a new paradigm in the field, the use of this technique as a clinical biomarker remains investigational and still necessitates standardization and robustness to be effectively translated into the clinical practice. This review summarizes the basic concepts, potential contributions, challenges, and opportunities of radiogenomics in the management of patients with lung cancer.

## 1. Introduction

In the current era of precision medicine, clinical genomics is considered a crucial source of information to select personalized management strategies according to individual genetic alterations.<sup>1</sup> While this applies to many clinical settings, it is particularly vital in cancer care. Nevertheless, to date, the implementation of clinical genomics in the management of patients with cancer is challenged by intra- and inter-tumor heterogeneity, which is known to affect disease progression and lead to treatment failure in some individuals. In this context, multiple biopsies and sequential invasive procedures are sometimes necessary to depict molecular and genomic biomarkers throughout the course of the disease.<sup>2</sup>

With the rapid development of artificial intelligence (AI) and genome sequencing technologies, the integration of imaging and

genomic data has been considered a promising strategy to non-invasively stratify patients across different types of cancer, with the potential of guiding therapeutic strategies and predicting clinical outcomes.<sup>3</sup> Radiomics has emerged as a field of translational research based on the extraction of mineable high-dimensional data from radiological images—in particular, computed tomography (CT), magnetic resonance (MR), and positron emission tomography (PET) images—to create “big data” datasets for the purpose of identifying distinct sub-visual imaging patterns.<sup>4</sup> The mining of radiomic data to detect correlations between imaging and genomic patterns is called radiogenomics, an approach which has provided new insights into the complex relationships between imaging features and underlying tumor biology along with clinical and therapeutic outcomes, but that still cannot replace genomics analysis based on invasive tissue sampling in oncologic practice.<sup>5</sup>

In lung cancer research, radiomics has been investigated for its use in

\* Corresponding author at: Department of Radiology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA; Medical Imaging Department, Hospital Clinic of Barcelona, 170 Villarroel street, Barcelona 08036, Spain.

E-mail address: [mmayoral@clinic.cat](mailto:mmayoral@clinic.cat) (M. Mayoral).

<sup>1</sup> Both authors contributed equally to this work.

early detection, treatment efficacy evaluation, and prediction of treatment-related outcomes in non-small cell lung cancer (NSCLC) patients, who are traditionally treated with chemotherapy and/or radiotherapy.<sup>6–8</sup> With the emergence of targeted therapy and immunotherapy in the era of precision medicine, radiogenomics has been considered a potential tool to identify patients who can benefit from such therapies.<sup>9,10</sup> Although at present, the relationships between genomics and radiomic phenotyping in lung cancer are not completely understood and larger studies and further clinical validation are required, the integration of radiogenomics data into lung cancer management may indeed represent a new paradigm and revolutionize the management of patients with lung cancer in the near future. This review summarizes the basic concepts, potential contributions, challenges, and opportunities of radiogenomics in the personalized management of patients with lung cancer.

## 2. Radiomics and radiogenomics workflow

### 2.1. Radiomics workflow

Radiomics analysis is a multistep process where medical images are processed to generate a large number of quantitative imaging features, i.e., radiomic features, from the voxels of segmented images (Fig. 1).

#### 2.1.1. Image segmentation and pre-processing

Segmentation is a critical component of radiomics workflow that consists of delineating the borders of the tumor volume with computer-assisted contouring after the region or volume of interest is identified. Image segmentation can be done manually, semi-automatically, or automatically, using different open-source software packages.

#### 2.1.2. Feature extraction

The extraction of radiomic features is performed by dedicated software and extracted radiomic features can be categorized into several classes.

- 1) **Morphologic features.** These features reflect the volume, shape, and 3D geometric properties of a lesion, including variables such as surface-to-volume ratio (an indirect measurement to assess spiculation), compactness, and sphericity.
- 2) **First-order statistics features.** First-order statistics features are based on statistical modelling and reflect the distribution of individual grayscale voxel values in a histogram curve, without concern for spatial relationships but reporting on the distribution of the voxel intensities in the image (maximum, minimum, mean, and median values), as well as their uniformity, kurtosis (flatness/magnitude of pixel distribution), skewness (asymmetry of a histogram), and entropy (irregularity/randomness of the intensities).<sup>4</sup>
- 3) **Second-order statistics features.** Also based on statistical modelling, second-order statistics features consider the statistical interrelationships between neighboring voxels and provide an evaluation of intra-lesion heterogeneity through the spatial arrangement of voxel intensities<sup>4</sup> (Fig. 2). Using gray level dependence matrices,

these features are classified into three further classes: gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), and gray level size zone matrix (GLSZM). While GLCM takes into account the incidence of two voxels with particular gray values at a predetermined distance along different directions, GLRLM considers runs of any length for pixels with the same intensities along different directions and GLSZM considers pixel areas of any size with the same intensity for analysis.<sup>11</sup>

- 4) **Superior or higher-order statistics features.** These features have the advantage of considering the relationship between neighboring voxels and are obtained using neighborhood gray tone difference matrices that examine the location and relationships between three or more pixels.<sup>12</sup> Filters or mathematical transforms are used to extract repetitive or non-repetitive patterns, including Wavelet and Laplacian transforms (to identify coarse texture patterns), fractal analysis (to assess the irregularity of a surface), and Minkowski functionals (to evaluate voxels whose intensity is above a determined threshold).<sup>5</sup>

#### 2.1.3. Feature selection

Once features have been extracted, a wide range of statistical models are commonly used to select a subset of optimal features that are associated with the outcome specified by the predefined hypothesis, such that all selected features are reproducible, non-redundant, and informative.<sup>4</sup> Many extracted features are in fact redundant, and to reduce redundancy, either supervised feature selection (dimensionality reduction) or unsupervised feature selection (association analysis) is applied. In supervised feature selection, univariate or multivariate statistical models are often used, while cluster analysis and principal component analysis are useful strategies in unsupervised feature selection. Machine learning and deep learning techniques are also emerging as useful tools that may lead to a more accurate selection of features.<sup>13,14</sup>

#### 2.1.4. Prediction modelling

Once optimal features are identified, models including these features are built to predict the outcome specified by the predefined hypothesis. Models can be built using different classifiers e.g., generalized linear model, random forests, support vector machines, or neural networks.<sup>8,15</sup> Of note, when the number of extracted features is larger than the study sample, this can contribute to model overfitting; in this scenario, regularization methods can be applied to control the complexity of the model.<sup>15</sup> Currently, an effective means to properly interpret the myriad data derived from radiomics analysis is still lacking, which therefore necessitates a balanced interpretation of radiomics results.<sup>16</sup>

#### 2.1.5. Validation

Radiomics models should be validated to confirm their value for clinical application. Models that are validated on independent, external validation sets are considered stronger and more reliable compared with prediction models that are only internally validated.

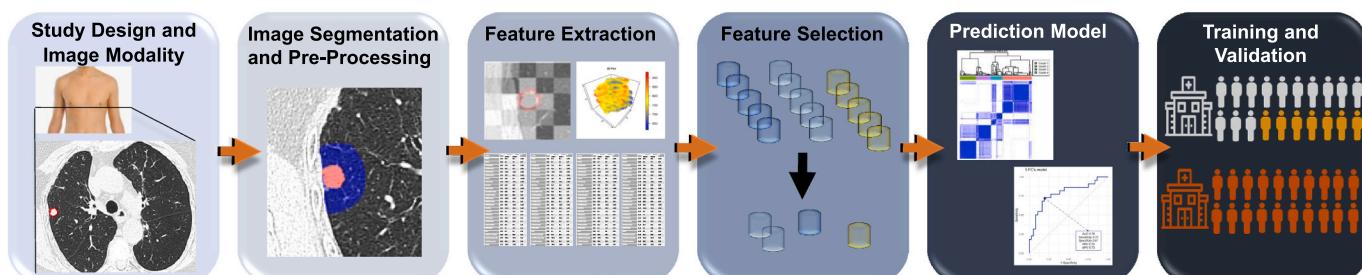
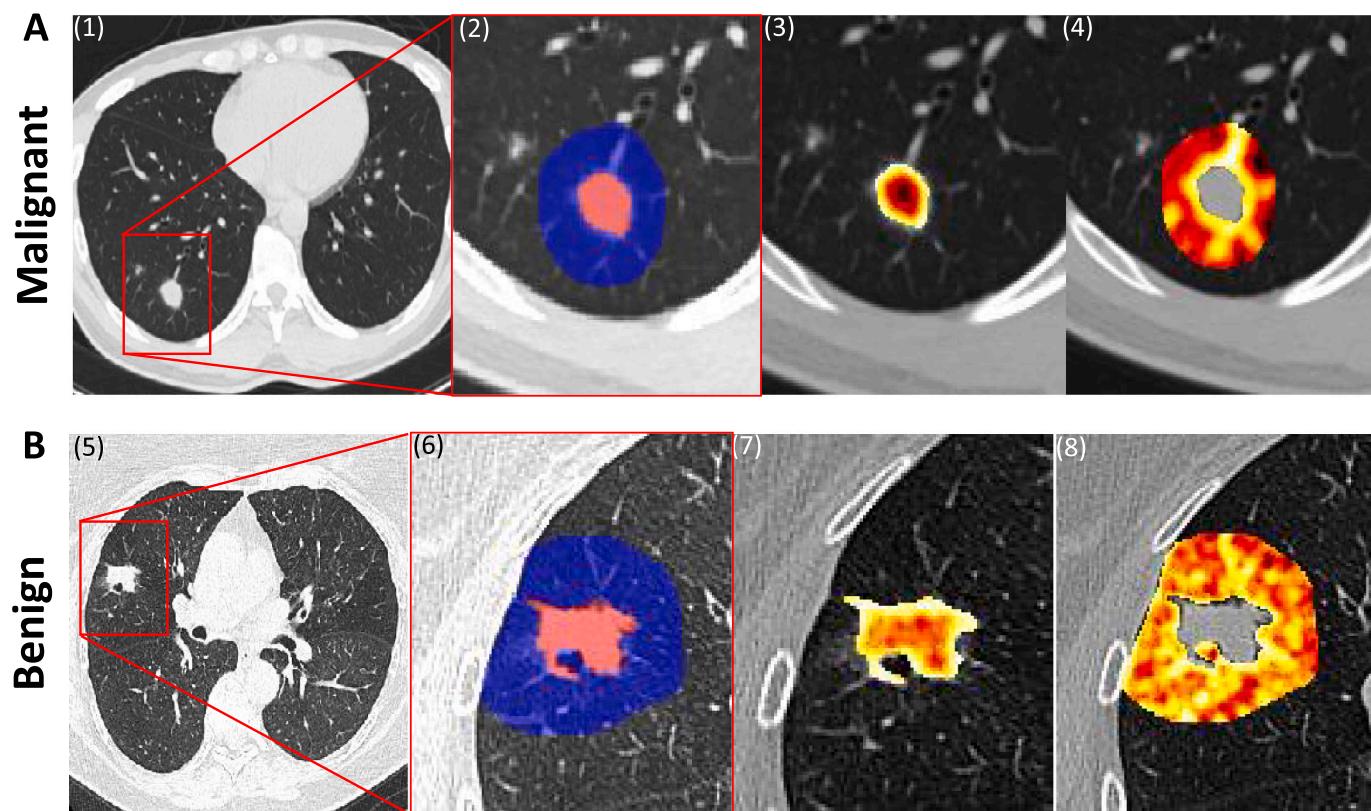


Fig. 1. Illustration demonstrating the step-by-step radiomics process.



**Fig. 2.** Entropy maps illustrating lesion heterogeneity for malignant versus benign lung nodules. A) Non-contrast CT scan in the axial view (1) of an adenocarcinoma tumor in a 62-year-old man. Segmentation images (2) and entropy map extracted from intra (3) and perinodular (4) regions illustrating the high heterogeneity of this malignant nodule. B) Non-contrast CT scan in the axial view (5) of a benign lung nodule in a 50-year-old man with pulmonary cryptococciosis. Segmentation images (6) and entropy map extracted from intra (7) and perinodular (8) regions illustrating different entropy patterns from the referred adenocarcinoma.

## 2.2. Radiogenomics workflow

A radiogenomics dataset contains both genomic and radiomics data. It is important to highlight that radiomics focuses on the use of computational methods to identify imaging traits not necessarily seen by the naked eye, whereas radiogenomics focuses on analyzing the relationships between radiomics features and genomic data. To enable accurate diagnosis and allow personalized cancer treatment, other data may also be integrated into the analysis, e.g., clinical data, histopathologic data, as well as other ‘omics’ data including proteomics and transcriptomics data (Fig. 3). Analysis involving high dimensional and heterogeneous multi-omics data remains one of the main challenges in the era of precision medicine, usually requiring powerful analytical methods and tools.

## 3. Targeting actionable alterations in the management of lung cancer patients: basic concepts

In the past decade, two breakthrough therapies have emerged for the treatment of lung cancer: checkpoint blockade immunotherapy and targeted therapy. Advances in the understanding of cancer evolution along with the emergence of clinical genomics were critical for the development of these new therapies. While checkpoint blockade immunotherapy acts on a common mechanism of immune evasion, targeted therapy targets specific molecular alterations in lung cancer, bringing precision medicine to this disease.<sup>17</sup>

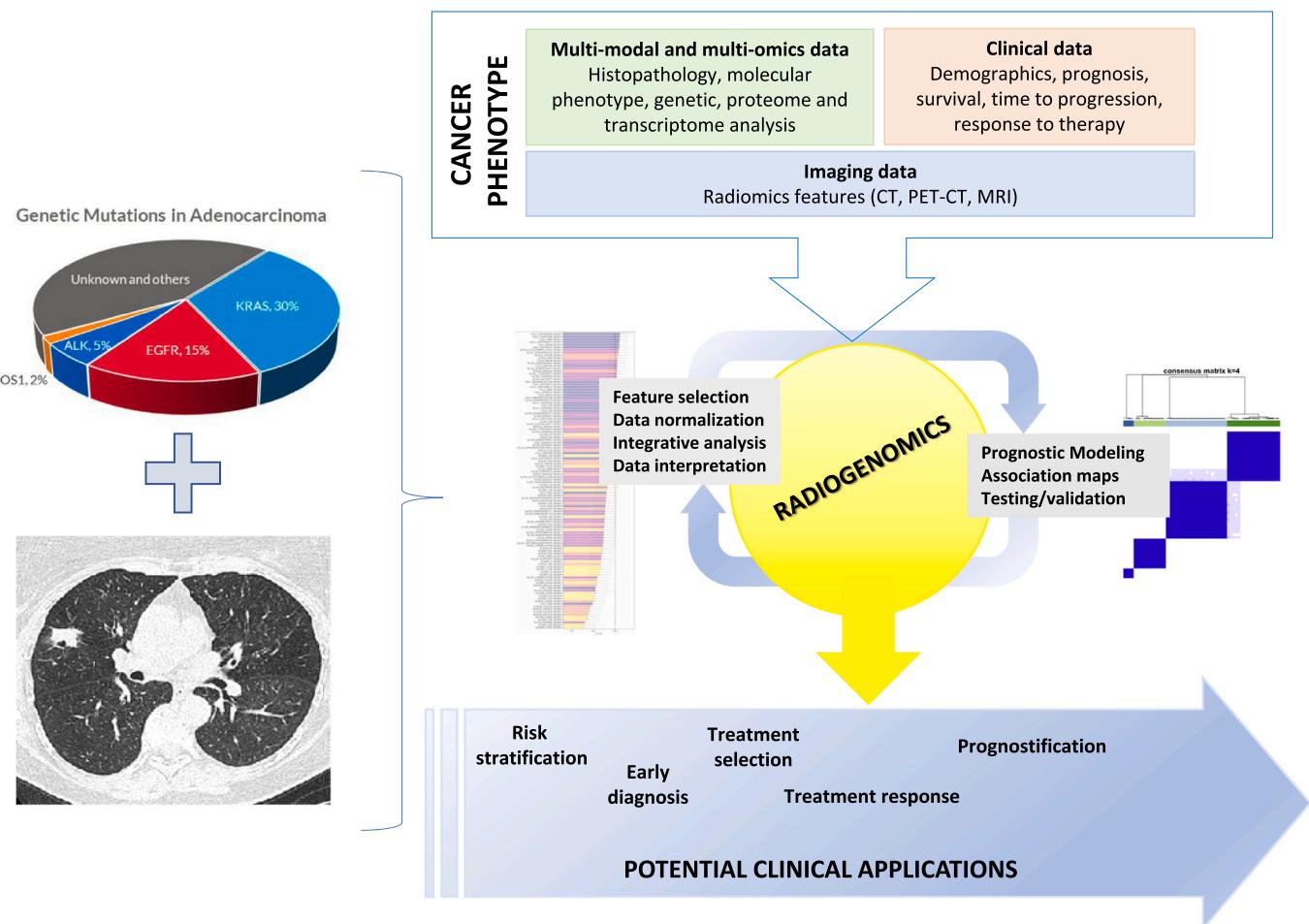
Sustained proliferative signaling is a main hallmark of cancer involving the abnormal activation of several proto-oncogenes. While any component of a multitude of proliferative signaling pathways can be altered, the alteration of tyrosine kinases has been particularly implicated in cancer. Mutations, chromosomal rearrangements, and

amplifications represent the main mechanisms of abnormal tyrosine kinase activation. While amplifications are ligand-dependent phenomena, mutations and chromosomal rearrangements occur in a ligand-independent manner. In the context of evolving tumor heterogeneity, molecular alterations are truncal and thus the earlier we target them the better.<sup>17</sup>

To date, seven FDA-approved targeted therapies have been suggested as first-line treatment options for advanced NSCLC harboring *EGFR*, *ALK*, *ROS1*, *BRAF*, *MET* exon 14, *RET*, and *NTRK* alterations.<sup>18</sup> Two other targeted therapies have been suggested as second-line treatment options, against *EGFR* exon 20 mutations and the now druggable *KRAS* mutation.<sup>19,20</sup> The frequency of these alterations varies across different populations; in East Asia, *EGFR* mutations can be found in almost 50% of all lung adenocarcinomas.<sup>21</sup> Most actionable alterations are enriched in never-smokers; however, upfront broad molecular profiling is strongly recommended in all patients with lung adenocarcinoma and in some patients with squamous cell carcinoma, particularly as patients with different alterations may yet show similar pathological and radiological features.<sup>22</sup>

Some of the challenges in designing tyrosine kinase inhibitors are their bioavailability and selectivity against the mutated enzyme while sparing the wild-type form of the enzyme. Most next-generation inhibitors are available in the form of oral administration and can promote robust, profound, and durable response while being safe and well tolerated. However, it is important to note that initial treatment response may not last and cancers might develop resistance and consequently progress.<sup>23</sup>

Finding the right actionable tumor mutations to target is a rate limiting step of targeted therapy. Across all the technologies available for this purpose, a hybrid broad molecular profiling panel based on DNA and RNA next-generation sequencing is the gold standard, being



**Fig. 3.** Overview of radiogenomics workflow using a multi-modal and multi-omics approach.

preferred to polymerase chain reaction (PCR) and fluorescent in situ hybridization (FISH). In addition to identifying actionable tumor mutations, broad panel molecular tumor characterization can also detect other non-actionable tumor mutations that while being non-actionable may yet serve as prognostic and predictive tools. Despite its benefits, however, the major downside of broad panel molecular tumor characterization is that it requires tissue biopsy, and unfortunately, even if the biopsied tissue meets several technical requirements such as fixation and tumor content, it is unlikely to be representative of the entire tumor. The emerging use of liquid biopsy, an alternative to tissue acquisition based in the collection of circulating tumor markers in the peripheral blood, is greatly advantageous in this context as it can more fully account for tumor heterogeneity, although still as a complementary method to broad panel molecular tumor characterization.<sup>24</sup> However, it is important to note that the turnaround time of next-generation sequencing platforms, including that of liquid biopsy, remains a huge caveat.<sup>25</sup> Patients who present with a high burden from symptomatic disease are usually unable to wait for the full molecular results before starting systemic therapy.

#### 4. Radiogenomics in the personalized management of lung cancer patients: potential contributions

Genomics analysis is essential to enable targeted therapy of lung cancer. The gold standard for both histopathologic and genetic phenotyping of lung cancer is tissue samples obtained at surgery or at a diagnostic biopsy. Nevertheless, both surgery and biopsy are invasive procedures that are not without complications. Multiple biopsies may be

needed when the biopsied material is insufficient for histopathologic analysis. On some occasions, biopsies may be forfeited, e.g., cases of inaccessible lung nodules at CT-guided biopsy or patients with advanced emphysema or interstitial lung disease.

Radiogenomics, by combining radiomic and genomic data, is a promising method to non-invasively predict the status of different mutations and could have a decisive role in treatment decision-making. In the context of lung cancer, a heterogeneous disease, radiogenomics could be potentially used as an additional tool to genomic testing, since genetic mutations and expression are assessed based on small tissue samples and may not reflect the heterogeneity of the entire lesion, while this advanced imaging analysis could provide data from the whole tumor.

##### 4.1. Prediction of mutation status

In the current literature, a preponderance of radiogenomics and image-based AI studies in lung cancer have focused on predicting *EGFR*, but also *KRAS* and *ALK* mutation status (Table 1). The highest diagnostic performance for *EGFR* mutation was seen in models based on logistic regression which combined radiomics features with clinical variables. In regard to CT-derived radiomics characteristics, the combination of these features with clinical variables such as sex, smoking status, and histological subtype yielded the highest diagnostic performance in the predictive model (validation set area under the curve, AUC = 0.87–0.89).<sup>26,27</sup> Similarly, PET/CT-derived radiomic features in combination with clinical features yielded a higher diagnostic performance compared with the predictive model using PET/CT-derived features alone (validation set AUC = 0.87).<sup>28</sup> Besides logistic regression,

**Table 1**

Radiogenomic studies with independent validation in patients with lung cancer

#	Authors	Year	n	n for validation	Imaging modality	Single vs. multi-center	Histology	Mutation
1	Yamamoto et al. <sup>37</sup>	2014	172	113	CT	Multi-center	NSCLC	ALK
2	Yamamoto et al. <sup>52</sup>	2016	26	166	PET/CT	Single-center	NSCLC	EMT
3	Rios Velazquez et al. <sup>53</sup>	2017	763	352	CT	Multi-center	AC	EGFR and Kras
4	Zhang et al. <sup>27</sup>	2018	180	40	CT	Single-center	NSCLC	EGFR
5	Zhou et al. <sup>54</sup>	2018	113	Not specified (5 public cohorts)	CT	Single-center	NSCLC	Clusters of metagenes
6	Jong et al. <sup>55</sup>	2018	195 <sup>a</sup>	195	CT	Multi-center	NSCLC	EGFR
7	Wang et al. <sup>32</sup>	2019	844	241	CT	Multi-center	NSCLC	EGFR
8	Yang X et al. <sup>56</sup>	2019	467	161	CT	Single-center	AC	EGFR
9	Tu et al. <sup>33</sup>	2019	404	161	CT	Single-center	NSCLC	EGFR
10	Jia et al. <sup>30</sup>	2019	503	158	CT	Single-center	AC and SCC	EGFR
11	Zhang et al. <sup>31</sup>	2020	914	205	CT	Multi-center	AC	EGFR
12	Song et al. <sup>40</sup>	2020	342	101 and 96	CT	Multi-center	NSCLC	EGFR
13	Mu et al. <sup>57</sup>	2020	616	187 and 65	PET/CT	Multi-center	AC and SCC	EGFR
14	Lu et al. <sup>26</sup>	2020	104	Not specified	CT	Single-center	AC and SCC	EGFR
15	Lu et al. <sup>58</sup>	2020	228	123	CT	Multi-center	AC and SCC	EGFR
16	Zhang et al. <sup>29</sup>	2021	420	126	CT	Single-center	ACC	EGFR

AC, adenocarcinoma; ALK, anaplastic lymphoma kinase; CT, computed tomography; EMT, epithelial-to-mesenchymal transition; EGFR, epidermal growth factor receptor; Kras, Kirsten rat sarcoma viral oncogene; NSCLC, non-small cell lung cancer; PET, positron emission tomography; SCC, squamous cell carcinoma.

<sup>a</sup> A radiomics signature from a previous article was validated using two independent cohorts.

different machine learning models including support vector machines, decision trees, and random forest classifiers<sup>29,30</sup> and deep learning models like convolutional neural networks<sup>31,32</sup> have been built, also with excellent results (validation set AUC = 0.81–0.84).

However, other studies have shown less robust AI models for the prediction of EGFR mutation status. The models with the worst prediction performances used qualitative CT features such as spiculations, subsolid appearance and pleural indentation, combined or not with clinical variables with machine learning analysis (AUC = 0.62–0.64).<sup>32,33</sup> These results suggest the superiority of radiomics in comparison with qualitative CT analysis to predict EGFR mutation status, although more studies directly comparing these two different approaches in the same group of patients are needed.

In the case of the radiomic models, a source of variability between studies and a potential cause of reduced performance to predict EGFR mutation status could be related to the radiomic features themselves which are included in the algorithm. The number of features extracted and selected to build the predictive model can vary widely between studies,<sup>31,34</sup> and the type of features found to be significantly predictive for a specific outcome are also usually different among different studies. The lack of uniformity in radiomics-extracted features, among other obstacles, currently hinders the development of a universal radiogenomics model that could potentially be used across different institutions. Other aspects that may alter the predictive ability of the radiomic model could be inherent to the imaging acquisition protocol. It has been reported that the performance of the radiomics signature is higher for contrast-enhanced images than for non-contrast images, irrespective of the machine learning method used for the predictive model.<sup>35</sup> The diagnostic performance of radiogenomics still needs to be improved before this image-based mutation identifier can be used consistently in daily practice to select patients for genotype-directed therapy of lung cancer.

Although currently there is not much evidence in the literature, radiomics-based models could also be useful to non-invasively predict ALK mutation<sup>36,37</sup> and PD-L1 expression levels,<sup>38,39</sup> which would also have an impact in the clinical setting. Central tumoral location, absence of pleural tail and large pleural effusion are distinct CT characteristics selected by random forest models that have been shown to be a strong predictor for ALK positive status.<sup>37</sup> However, the performance of predictive radiomic models for ALK mutation (validation set AUC = 0.68–0.88)<sup>36</sup> and PD-L1 expression level >50% (validation set AUC = 0.67–0.91)<sup>38,39</sup> has been reported to be highly variable, so these preliminary results entail further study.

#### 4.2. Prediction of treatment response and other applications

The benefits of combining radiomic and genomic data go beyond the mere prediction of mutation status; this information could also be integrated in more complex models to predict treatment response in order to select those patients who would be more likely to benefit from genotype-directed therapy, in addition to stratifying patients according to survival metrics to predict disease prognosis, as has already been reported for EGFR mutation.<sup>40–42</sup> In the last few years there has been an increasing interest in the radiomic analysis of the lung parenchyma surrounding the vicinity of tumoral lesions. Some studies have reported that this perinodular tissue might harbor a different microenvironment from healthy parenchyma that could be used in radiomic models to distinguish benign from malignant lesions<sup>43</sup> or to predict response to chemotherapy,<sup>44</sup> survival outcomes,<sup>45</sup> or tumoral spread through air spaces in lung cancer.<sup>46</sup> Further applications of AI predictive models could be directed to predict mediastinal lymph node invasion and distant metastases.<sup>47,48</sup> Recent investigations have also explored the utility of incorporating nucleic acids sequencing data to radiogenomics models. A multi-center study developed a radiogenomics signature which combined CT-extracted features and RNA sequencing data to assess CD8 tumor infiltration and thus predict response to immunotherapy.<sup>49</sup> Another promising application of radiogenomics is the detection of patients at risk of progression during treatment based on clonal heterogeneity by means of a multi-modal model integrating radiomics and DNA sequencing data from liquid biopsy.<sup>50,51</sup> Considering the range of possibilities that can emerge from this innovative strategy, further studies are needed.

#### 5. Challenges and future perspectives

Radiogenomics in lung cancer has the potential to add value to patient care toward personalized medicine; however, several improvements are needed before it can be clinically implemented and accepted among members of the multidisciplinary team involved in the management of lung cancer. Currently, the majority of radiogenomics studies are single-center studies with variable study methodologies and statistical models and which employ manual lesion segmentation. Automatic segmentation and extraction of the quantitative imaging features is of key importance, e.g., via deep learning. Deep learning strategies can also be used to reduce bias in feature and model selection, and to reduce the scale of the images, consequently reducing computational burden. Currently, the main limitation of deep learning is the need for a large amount of data with clinically relevant and reliable

endpoints. Research regarding ethical regulations is a barrier and limitation to acquire such data. Thus, new regulations need to be created to allow the building of large datasets and to facilitate multi-institutional collaborations. Finally, integrated models using multi-omics sources of data as well as investment in technology to integrate them into current practice systems are needed.

## 6. Conclusions

The integration of radiogenomics data into lung cancer management may represent a new paradigm in this field and has the potential to revolutionize management of patients with lung cancer in the future. The use of radiogenomics as a predictive tool remains investigational and the standardization and improved robustness of radiogenomics models are still needed for this approach to be effectively integrated into the clinical practice. Familiarity with the principles, potential applications, and limitations of radiogenomics in the personalized management of patients with lung cancer is essential to identify clinically meaningful correlations.

## Funding

This research was funded by the National Institutes of Health/National Cancer Institute Cancer Center Support Grant P30 CA008748.

The funding source was not involved in the design, interpretation of data or in the writing process of this paper.

## Declaration of competing interest

None.

## References

- Vijay P, McIntyre AB, Mason CE, Greenfield JP, Li S. Clinical genomics: challenges and opportunities. *Crit Rev Eukaryot Gene Expr* 2016;26:97–113. <https://doi.org/10.1615/CritRevEukaryotGeneExpr.2016015724>.
- Kim G, Kim J, Cha H, Park WY, Ahn JS, Ahn MJ, et al. Metabolic radiogenomics in lung cancer: associations between FDG PET image features and oncogenic signaling pathway alterations. *Sci Rep* 2020;10:13231. <https://doi.org/10.1038/s41598-020-70168-x>.
- Shui L, Ren H, Yang X, Li J, Chen Z, Yi C, et al. The era of radiogenomics in precision medicine: an emerging approach to support diagnosis, treatment decisions, and prognostication in oncology. *Front Oncol* 2020;10:570465. <https://doi.org/10.3389/fonc.2020.570465>.
- Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp* 2018;2:36. <https://doi.org/10.1186/s41747-018-0068-z>.
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;278:563–77. <https://doi.org/10.1148/radiol.2015151169>.
- Lee G, Lee HY, Park H, Schiebler ML, van Beek EJR, Ohno Y, 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. <https://doi.org/10.1016/j.ejrad.2016.09.005>.
- Hassani C, Vaghese BA, Nieva J, Duddalwar V. Radiomics in pulmonary lesion imaging. *Am J Roentgenol* 2019;212:497–504. <https://doi.org/10.2214/AJR.18.20623>.
- Thawani R, McLane M, Beig N, Ghose S, Prasanna P, Velcheti V, et al. Radiomics and radiogenomics in lung cancer: a review for the clinician. *Lung Cancer* 2018;115:34–41. <https://doi.org/10.1016/j.lungcan.2017.10.015>.
- Mu W, Tunali I, Gray JE, Qi J, Schabath MB, Gillies RJ. Radiomics of 18 F-FDG PET/CT images predicts clinical benefit of advanced NSCLC patients to checkpoint blockade immunotherapy. *Eur J Nucl Med Mol Imaging* 2020;47:1168–82. <https://doi.org/10.1007/s00259-019-04625-9>.
- Liu Y, Wu M, Zhang Y, Luo Y, He S, Wang Y, et al. Imaging biomarkers to predict and evaluate the effectiveness of immunotherapy in advanced non-small-cell lung cancer. *Front Oncol* 2021;11:657615. <https://doi.org/10.3389/fonc.2021.657615>.
- Horvat N, Bates DDB, Petkovska I. Novel imaging techniques of rectal cancer: what do radiomics and radiogenomics have to offer? A literature review. *Abdom Radiol* 2019;44:3764–74. <https://doi.org/10.1007/s00261-019-02042-y>.
- Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT texture analysis: definitions, applications, biologic correlates, and challenges. *Radiographics* 2017;37:1483–503. <https://doi.org/10.1148/rg.2017170056>.
- Giger ML. Machine learning in medical imaging. *J Am Coll Radiol* 2018;15:512–20. <https://doi.org/10.1016/j.jacr.2017.12.028>.
- Suzuki K. Overview of deep learning in medical imaging. *Radiol Phys Technol* 2017;10:257–73. <https://doi.org/10.1007/s12194-017-0406-5>.
- Larue RT, Defraene G, De Ruysscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol* 2017;90:20160665. <https://doi.org/10.1259/bjr.20160665>.
- Wang H, Ding MX, Zhai ZH. Dynamics of DNA synthesis in IBRV (infectious bovine rhinotracheitis virus) infected cells and the morphogenesis of the viruses. *Wei Sheng Wu Xue Bao* 1985;25:306–12.
- Saab S, Zalzale H, Rahal Z, Khalifeh Y, Sinjab A, Kadara H. Insights into lung cancer immune-based biology, prevention, and treatment. *Front Immunol* 2020;11:159. <https://doi.org/10.3389/fimmu.2020.00159>.
- Jakubek Y, Lang W, Vattathil S, Garcia M, Xu L, Huang L, et al. Genomic landscape established by allelic imbalance in the cancerization field of a normal appearing airway. *Cancer Res* 2016;76:3676–83. <https://doi.org/10.1158/0008-5472.CAN-15-3064>.
- Vyse S, Huang PH. Targeting EGFR exon 20 insertion mutations in non-small cell lung cancer. *Signal Transduct Target Ther* 2019;4:5. <https://doi.org/10.1038/s41392-019-0038-9>.
- Friedlaender A, Drilon A, Weiss GJ, Banna GL, Addeo A. KRAS as a druggable target in NSCLC: rising like a phoenix after decades of development failures. *Cancer Treat Rev* 2020;85:101978. <https://doi.org/10.1016/j.ctrv.2020.101978>.
- Shi Y, Au JS, Thongprasert S, Srivivasan S, Tsai CM, Khoa MT, et al. A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). *J Thorac Oncol* 2014;9:154–62. <https://doi.org/10.1097/JTO.0000000000000033>.
- Network CGAR. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014;511:543–50. <https://doi.org/10.1038/nature13385>.
- Herbrink M, Nuijen B, Schellens JH, Beijnen JH. Variability in bioavailability of small molecular tyrosine kinase inhibitors. *Cancer Treat Rev* 2015;41:412–22. <https://doi.org/10.1016/j.ctrv.2015.03.005>.
- Revelo AE, Martin A, Velasquez R, Kulandaisamy PC, Bustamante J, Keshishyan S, et al. Liquid biopsy for lung cancers: an update on recent developments. *Ann Transl Med* 2019;7:349. <https://doi.org/10.21037/atm.2019.03.28>.
- Osoegawa K, Montero-Martín G, Mallempati KC, Bauer M, Milius RP, Maiers M, et al. Challenges for the standardized reporting of NGS HLA genotyping: surveying gaps between clinical and research laboratories. *Hum Immunol* 2021. <https://doi.org/10.1016/j.humimm.2021.08.011>. S0198-8859(21)00210-X.
- Lu X, Li M, Zhang H, Hua S, Meng F, Yang H, et al. A novel radiomic nomogram for predicting epidermal growth factor receptor mutation in peripheral lung adenocarcinoma. *Phys Med Biol* 2020;65:055012. <https://doi.org/10.1088/1361-6560/ab6f98>.
- Zhang L, Chen B, Liu X, Song J, Fang M, Hu C, et al. Quantitative biomarkers for prediction of epidermal growth factor receptor mutation in non-small cell lung cancer. *Transl Oncol* 2018;11:94–101. <https://doi.org/10.1016/j.tranon.2017.10.012>.
- Zhang J, Zhao X, Zhao Y, Zhang Z, Wang J, Wang Y, et al. Value of pre-therapy 18 F-FDG PET/CT radiomics in predicting EGFR mutation status in patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging* 2020;47:1137–46. <https://doi.org/10.1007/s00259-019-04592-1>.
- Zhang G, Cao Y, Zhang J, Ren J, Zhao Z, Zhang X, et al. Predicting EGFR mutation status in lung adenocarcinoma: development and validation of a computed tomography-based radiomics signature. *Am J Cancer Res* 2021;11:546–60.
- Jia TY, Xiong JF, Li XY, Yu W, Xu ZY, Cai XW, et al. Identifying EGFR mutations in lung adenocarcinoma by noninvasive imaging using radiomics features and random forest modeling. *Eur Radiol* 2019;29:4742–50. <https://doi.org/10.1007/s00330-019-06024-y>.
- Zhang B, Qi S, Pan X, Li C, Yao Y, Qian W, et al. Deep CNN model using CT radiomics feature mapping recognizes EGFR gene mutation status of lung adenocarcinoma. *Front Oncol* 2020;10:598721. <https://doi.org/10.3389/fonc.2020.598721>.
- Wang S, Shi J, Ye Z, Dong D, Yu D, Zhou M, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* 2019;53. <https://doi.org/10.1183/13993003.00986-2018>.
- Tu W, Sun G, Fan L, Wang Y, Xia Y, Guan Y, et al. Radiomics signature: a potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. *Lung Cancer* 2019;132:28–35. <https://doi.org/10.1016/j.lungcan.2019.03.025>.
- Zhang M, Bao Y, Rui W, Shangguan C, Liu J, Xu J, et al. Performance of 18 F-FDG PET/CT radiomics for predicting EGFR mutation status in patients with non-small cell lung cancer. *Front Oncol* 2020;10:568857. <https://doi.org/10.3389/fonc.2020.568857>.
- Hong D, Xu K, Zhang L, Wan X, Guo Y. Radiomics signature as a predictive factor for EGFR mutations in advanced lung adenocarcinoma. *Front Oncol* 2020;10:28. <https://doi.org/10.3389/fonc.2020.00028>.
- Song L, Zhu Z, Mao L, Li X, Han W, Du H, et al. Clinical, conventional CT and radiomic feature-based machine learning models for predicting ALK rearrangement status in lung adenocarcinoma patients. *Front Oncol* 2020;10:369. <https://doi.org/10.3389/fonc.2020.00369>.
- Yamamoto S, Korn RL, Oklu R, Migdal C, Gotway MB, Weiss GJ, et al. ALK molecular phenotype in non-small cell lung cancer: CT radiogenomic characterization. *Radiology* 2014;272:568–76. <https://doi.org/10.1148/radiol.14140789>.
- Yoon J, Suh YJ, Han K, Cho H, Lee HJ, Hur J, et al. Utility of CT radiomics for prediction of PD-L1 expression in advanced lung adenocarcinomas. *Thorac Cancer* 2020;11:993–1004. <https://doi.org/10.1111/1759-7714.13352>.
- Jiang M, Sun D, Guo Y, Xiao J, Wang L, Yao X. Assessing PD-L1 expression level by radiomic features from PET/CT in nonsmall cell lung cancer patients: an initial result. *Acad Radiol* 2020;27:171–9. <https://doi.org/10.1016/j.acra.2019.04.016>.

40. Song J, Wang L, Ng NN, Zhao M, Shi J, Wu N, et al. Development and validation of a machine learning model to explore tyrosine kinase inhibitor response in patients with stage IV EGFR variant-positive non-small cell lung cancer. *JAMA Netw Open* 2020;3:e2030442. <https://doi.org/10.1001/jamanetworkopen.2020.30442>.
41. Kim KH, Kim J, Park H, Kim H, Lee SH, Sohn I, et al. Parallel comparison and combining effect of radiomic and emerging genomic data for prognostic stratification of non-small cell lung carcinoma patients. *Thorac Cancer* 2020;11: 2542–51. <https://doi.org/10.1111/1759-7714.13568>.
42. Chen YH, Wang TF, Chu SC, Lin CB, Wang LY, Lu KH, et al. Incorporating radiomic feature of pretreatment 18F-FDG PET improves survival stratification in patients with EGFR-mutated lung adenocarcinoma. *PLoS One* 2020;15:e0244502. <https://doi.org/10.1371/journal.pone.0244502>.
43. Beig N, Khorrami M, Alilou M, Prasanna P, Braman N, Orooji M, et al. Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas. *Radiology* 2019;290:783–92.
44. Khorrami M, Khunger M, Zagouras A, Patil P, Thawani R, Bera K, et al. Combination of peri- and intratumoral radiomic features on baseline CT scans predicts response to chemotherapy in lung adenocarcinoma. *Radiol Artif Intell* 2019;1:180012.
45. Pérez-Morales J, Tunali I, Stringfield O, Eschrich SA, Balagurunathan Y, Gillies RJ, et al. Peritumoral and intratumoral radiomic features predict survival outcomes among patients diagnosed in lung cancer screening. *Sci Rep* 2020;10:10528.
46. Liao G, Huang L, Wu S, Zhang P, Xie D, Yao L, et al. Preoperative CT-based peritumoral and tumoral radiomic features prediction for tumor spread through air spaces in clinical stage I lung adenocarcinoma. *Lung Cancer* 2021;163:87–95.
47. Gao X, Chu C, Li Y, Lu P, Wang W, Liu W, et al. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from (18)F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol* 2015;84:312–7. <https://doi.org/10.1016/j.ejrad.2014.11.006>.
48. Xu X, Huang L, Chen J, Wen J, Liu D, Cao J, et al. Application of radiomics signature captured from pretreatment thoracic CT to predict brain metastases in stage III/IV ALK-positive non-small cell lung cancer patients. *J Thorac Dis* 2019;11:4516–28. <https://doi.org/10.21037/jtd.2019.11.01>.
49. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, 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. [https://doi.org/10.1016/S1470-2045\(18\)30413-3](https://doi.org/10.1016/S1470-2045(18)30413-3).
50. Lafata KJ, Corradetti MN, Gao J, Jacobs CD, Weng J, Chang Y, et al. Radiogenomic analysis of locally advanced lung cancer based on CT imaging and intratreatment changes in cell-free DNA. *Radiol Imaging Cancer* 2021;3:e200157. <https://doi.org/10.1148/rycan.2021200157>.
51. Cucchiara F, Del Re M, Valleggi S, Romei C, Petrini I, Lucchesi M, et al. Integrating liquid biopsy and radiomics to monitor clonal heterogeneity of EGFR-positive non-small cell lung cancer. *Front Oncol* 2020;10:593831. <https://doi.org/10.3389/fonc.2020.593831>.
52. Yamamoto S, Huang D, Du L, Korn RL, Jamshidi N, Burnette BL, et al. Radiogenomic analysis demonstrates associations between 18F-Fluoro-2-deoxyglucose PET, prognosis, and epithelial-mesenchymal transition in non-small cell lung cancer. *Radiology* 2016;280:261–70. <https://doi.org/10.1148/radiol.2016160259>.
53. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res* 2017;77: 3922–30. <https://doi.org/10.1158/0008-5472.CAN-17-0122>.
54. Zhou M, Leung A, Echegaray S, Gentles A, Shrager JB, Jensen KC, et al. Non-small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. *Radiology* 2018;286:307–15. <https://doi.org/10.1148/radiol.2017161845>.
55. Jong EEC, van Elmpt W, Rizzo S, Colarieti A, Spitaleri G, Leijenaar RTH, et al. Applicability of a prognostic CT-based radiomic signature model trained on stage I–III non-small cell lung cancer in stage IV non-small cell lung cancer. *Lung Cancer* 2018;124:6–11. <https://doi.org/10.1016/j.lungcan.2018.07.023>.
56. Yang X, Dong X, Wang J, Li W, Gu Z, Gao D, et al. Computed tomography-based radiomics signature: a potential indicator of epidermal growth factor receptor mutation in pulmonary adenocarcinoma appearing as a subsolid nodule. *Oncologist* 2019;24:e1156–64. <https://doi.org/10.1634/theoncologist.2018-0706>.
57. Mu W, Jiang L, Zhang JY, Shi Y, Gray JE, Tunali I, et al. Non-invasive decision support for NSCLC treatment using PET/CT radiomics. *Nat Commun* 2020;11:5228. <https://doi.org/10.1038/s41467-020-19116-x>.
58. Lu L, Sun SH, Yang HEL, Guo P, Schwartz LH, et al. Radiomics prediction of EGFR status in lung cancer - our experience in using multiple feature extractors and the cancer imaging archive data. *Tomography* 2020;6:223–30. <https://doi.org/10.18383/j.tom.2020.00017>.