



Structural and functional radiomics for lung cancer

Guangyao Wu^{1,2,3} · Arthur Jochems¹ · Turkey Refaee^{1,4} · Abdalla Ibrahim^{1,5,6,7} · Chenggong Yan^{1,8} · Sebastian Sanduleanu¹ · Henry C. Woodruff^{1,5} · Philippe Lambin^{1,5}

Received: 23 November 2020 / Accepted: 3 February 2021 / Published online: 11 March 2021
© The Author(s) 2021

Abstract

Introduction Lung cancer ranks second in new cancer cases and first in cancer-related deaths worldwide. Precision medicine is working on altering treatment approaches and improving outcomes in this patient population. Radiological images are a powerful non-invasive tool in the screening and diagnosis of early-stage lung cancer, treatment strategy support, prognosis assessment, and follow-up for advanced-stage lung cancer. Recently, radiological features have evolved from solely semantic to include (handcrafted and deep) radiomic features. Radiomics entails the extraction and analysis of quantitative features from medical images using mathematical and machine learning methods to explore possible ties with biology and clinical outcomes.

Methods Here, we outline the latest applications of both structural and functional radiomics in detection, diagnosis, and prediction of pathology, gene mutation, treatment strategy, follow-up, treatment response evaluation, and prognosis in the field of lung cancer.

Conclusion The major drawbacks of radiomics are the lack of large datasets with high-quality data, standardization of methodology, the black-box nature of deep learning, and reproducibility. The prerequisite for the clinical implementation of radiomics is that these limitations are addressed. Future directions include a safer and more efficient model-training mode, merge multi-modality images, and combined multi-discipline or multi-omics to form “Medomics.”

Keywords Lung cancer · Radiomics · Artificial intelligence · Medical imaging

Introduction

Lung cancer ranks first in cancer-related deaths and second in the new cancer cases in both males and females as reported by the American Cancer Society in 2020 [1]. The 5-year survival rate ranges from 5% for patients with metastatic disease to 57% when

lung cancer is diagnosed and treated at an early stage [2]. Unfortunately, most patients with lung cancer only have mild clinical symptoms at an early stage, but the symptoms appear when the cancer is at an advanced stage [3].

Both the US-based national lung screening trial (NLST) and the Dutch-Belgian lung cancer screening trial (NELSON)

Arthur Jochems and Turkey Refaee; Henry C. Woodruff and Philippe Lambin contributed equally to this work.

This article is part of the Topical Collection on Oncology - Chest

✉ Guangyao Wu
g.wu@maastrichtuniversity.nl

¹ The D-Lab, Department of Precision Medicine, GROW–School for Oncology, Maastricht University Medical Centre+, 6229 Maastricht, The Netherlands

² Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

³ Hubei Province Key Laboratory of Molecular Imaging, Wuhan, China

⁴ Department of Diagnostic Radiology, Faculty of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia

⁵ Department of Radiology and Nuclear Medicine, GROW – School for Oncology, Maastricht University Medical Centre+, Maastricht, The Netherlands

⁶ Division of Nuclear Medicine and Oncological Imaging, Department of Medical Physics, Hospital Center Universitaire De Liege, Liege, Belgium

⁷ Department of Nuclear Medicine and Comprehensive Diagnostic Center Aachen (CDCA), University Hospital RWTH Aachen University, Aachen, Germany

⁸ Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, Guangzhou, China

concluded that the use of low-dose computed tomography (LDCT) for high-risk lung cancer populations decreases lung cancer mortality up to 60% in certain subpopulations, due to early detection and management [4, 5]. A substantial number of early-stage lung cancer patients have been identified using LDCT lung cancer screening [6]. The main goals of precision medicine research in lung cancer could be generally categorized into early-stage detection and/or diagnosis and highly tailored treatment and care in the advanced stage.

The use of radiologic images in medicine has become crucial in clinical practice, for both oncologic and non-oncologic cases. In oncology, medical imaging is used for every aspect of patients' management, including screening, diagnosis, treatment, and prognosis assessment of the disease. Over the past few decades, modern medical imaging has progressed from single structural imaging to combination of functional imaging [7]. Structural imaging refers to methods used to both visualize and evaluate anatomical details, and functional imaging is used to assess the physiology and molecular processes of tissues and organs [8]. Evidence has shown that the relationship between the structure and function of the lung is imperfect, which means that structural and functional images could have some common biomarkers, but certainly have independent biomarkers [9]. However, the correspondence between structural and functional images in the field of lung cancer still remains to be elucidated.

Radiomics refers to the extraction and analysis of quantitative image features from medical images using mathematical and machine learning methods to explore possible ties with biology and clinical outcomes [10]. Radiomic features extracted from structural and functional images (as summarized by Torigian et al. with detailed imaging modalities) can separately reflect the anatomical and functional information of the lesions [11] (Fig. 1). These features, mostly invisible to the unaided eye, have the potential to reduce the workload of clinicians and to increase the quality of diagnosis, prognosis, and treatment. The ultimate goal of radiomics is to build quicker and more reliable clinical decision support systems to assist clinicians rather than replacing them [12].

In this review, we present the development of radiologic features, from semantic and handcrafted radiomic to deep radiomic features, from the perspective of clinicians. In addition, we summarized the latest applications of structural and functional radiomics on early and advanced-stage lung cancer. Furthermore, we address the possible limitations of radiomics and set out future directions with respect to lung cancer.

Development from semantic to radiomic features

Semantic and clinical features

Imaging techniques have been widely used in clinical practice for different purposes in lung cancer. These techniques

include radiography, CT, MRI, and PET/CT. Radiologists analyze the images to detect lesions, and then use a set of qualitative (e.g., shape, location, speculation, and lobulation) and quantitative (e.g., size, volume, density, signal, and standardized uptake values (SUVs)) features to describe and analyze lesions (Fig. 2). Radiologists have been seeking to identify specific signs from images that can be used to determine the pathological type, degree of malignancy, and prognosis of cancer.

However, this approach involves a wealth of experience leading to subjective discrepancies. For this reason, the Fleischner Society has provided a series of glossaries and recommendations for describing and measuring thoracic images [13, 14]. Based on the evidence that semantic CT features have prognostic value, the Lung Imaging Reporting and Data System (Lung-RADS) has been developed to improve the interpretations and understanding of lung cancer screening CT and promote management [15]. Furthermore, the Response Evaluation Criteria in Solid Tumors (RECIST 1.1), a special version (iRECIST) for assessing response to immunotherapy, and a version for PET (PERCIST) have been successfully validated in clinical trials [16–18].

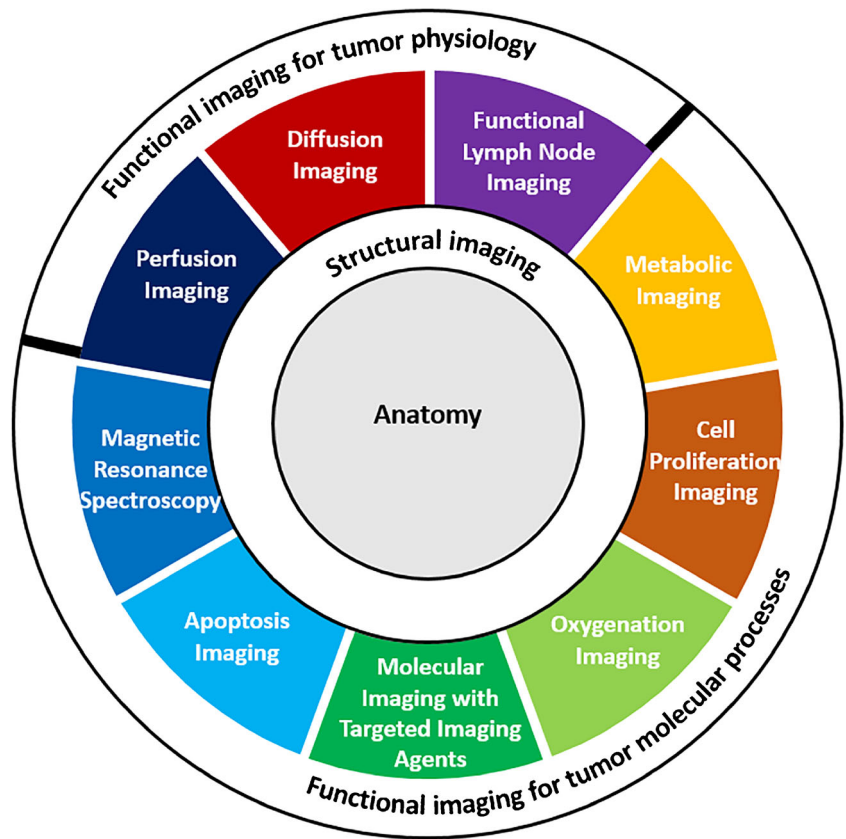
Clinical features have showed additional prognostic utility when combined with semantic CT characteristics, which could be used to construct a clinical model for predicting the risk of malignancy, invasiveness, and poor prognosis of lung cancer [19–21]. A generally recognized and validated clinical model for the malignancy risk of solitary pulmonary nodule is the Brock/PanCan model, which includes both clinical (age, sex, and family history of lung cancer) and semantic CT (emphysema, spiculation, size, type, location, and nodule count) features [22]. However, the subjective nature of these metrics can theoretically restrict the consistency of models based on semantic features [23].

Handcrafted radiomic features

Handcrafted radiomic features are the result of mathematical formulas that take image pixel values from within a region of interest (ROI) as an input and output a number, the so-called feature value, a process that can be automated. In order to quantitatively and automatically identify and interpret imaging findings, these handcrafted radiomic features can be used as a machine learning framework to correlate them with the underlying biology and clinical outcomes (Fig. 2) [24, 25]. The classical radiomic workflow involves image acquisition, lesion segmentation, pre-processing, feature extraction, feature selection, modeling, model validation/evaluation, and if the last step proves successful, clinical implementation.

Strict quality management at each step is necessary to ensure applicability and generalization of the developed model, and the first step is no exception as high-quality images are of

Fig. 1 Composition of functional and structural imaging for tumors. Structural imaging refers to techniques, which are used to visualize and analyze the anatomical information of the human body. Functional imaging refers to approaches that are the study of tumor physiology and molecular process



supreme importance in a radiomics study. In the following task, experienced radiologists often manually delineate the images slice by slice to define the ROI. Manual delineation, however, is time consuming and vulnerable to inter-observer variability [26]. Advanced semi-automatic or fully automatic

segmentation can enhance the repeatability of radiomic features [27].

Pre-processing techniques (e.g., resampling, denoising, histogram equalization and image normalization) can be used to minimize the variance among images when various

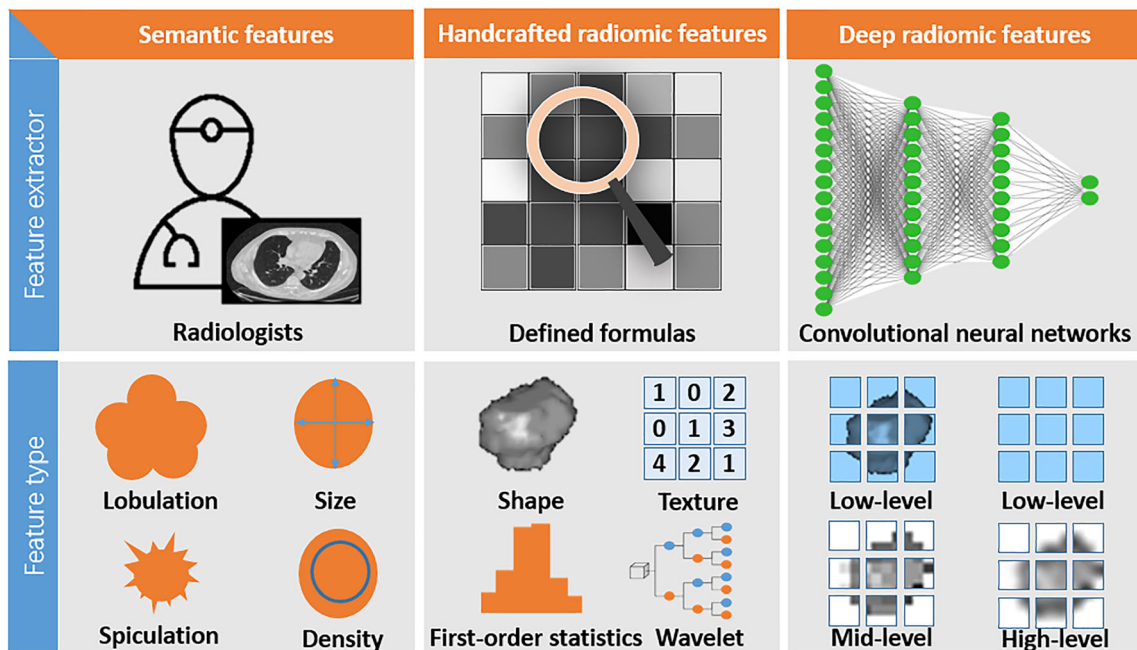


Fig. 2 A comparison of semantic, handcrafted radiomic, and deep radiomic features

scanners or scanning and reconstruction parameters are used [28, 29]. For feature extraction, McNitt-Gray et al. summarized several software or packages that can automatically perform this task, producing first-order statistics, shape, texture, and filter based features [30]. The input format (file type and modality) and the type of features should also be taken into consideration during this processing. Of note, radiomic features are sensitive to differing acquisition and reconstruction parameters employed by the multiple centers, and harmonization in the image domains or feature domains are being explored [31].

The selection of stable, important, and predictive features is another significant step in the radiomics pipeline. There are three methods to perform feature selection: (i) filter methods select features independently from the model, often using statistical tests and correlations of features with the outcome, such as Chi-square, Mann-Whitney, and Pearson's correlation. (ii) wrapper methods generate different subsets of features that are used to train multiple models, which are ranked based on model performance. (iii) embedded methods perform feature selection and model training simultaneously, with the least absolute shrinkage and selection operator (LASSO) method being the most widely published [32]. The choice of feature selection method also depends on which machine learning algorithm is chosen for the next step.

Once the features that will serve as input to the model have been selected, many machine learning algorithms are available to train classification or regression models. These algorithms can be broadly classified into supervised (e.g., logistic regression, support vector machine, random forest, and artificial neural network) and unsupervised (e.g., clustering and auto-encoder) solutions. In short, supervised methods optimize performance by updating the inner model configuration based on the desired outcome, while unsupervised methods rely on patterns in the predictors, without having access to the actual outcome. The composition, consistency, and interpretability of data are considerations that need to be weighed when choosing models. External validation is required to properly assess efficiency and generalization. Many tools from data science are used to evaluate the performance of a trained model, such as the receiver operating characteristic (ROC) or precision recall curve, area under the ROC curve (AUC), concordance index (C-index), confusion matrix, calibration, and decision curve analysis [33]. Multivariate Cox regression and Kaplan-Meier curve are popular tools for survival analysis [34]. When using these metrics, the balance of data, i.e., the number of cases in each class, needs to be carefully considered [35]. There are a number of methods to artificially balance datasets as Fotouhi et al. summarized, especially for classification problems, such as oversampling, undersampling, synthetic minority over-sampling technique (SMOTE), and adaptive synthetic (ADASYN) techniques [36].

Deep radiomic features

Deep learning is a broader family of machine learning methods inspired by our brain's own network of neurons [37]. The convolutional neural network (CNN) is commonly used in the analysis of images and has found its way to the field of medical imaging. The term deep comes from the large number of various types of layers (interconnected "slices" of the network): convolution, pooling, activation, and full connection (Fig. 2). In convolution layers, convolutional kernels slide over the image to automatically extract and select features that can be named deep radiomic features from a tailored and well-trained CNN structure (e.g., AlexNet, VGG, ResNet, Inception, and DenseNet) [38].

Handcrafted and deep radiomic features have many similarities but are also distinct. Deep radiomic features in shallow layers define intensity, shape, and texture details that are similar to handcrafted radiomic features; as the layers get deeper, more and more abstract features are extracted that are learnt from exposure to data, making these features difficult to interpret [39]. This makes a model based on deep radiomic features look like a black box, where the connection between the input and output is not understood. However, if the interpretability of the model is not considered to be of critical relevance, deep radiomic features are an effective supplement for handcrafted radiomic features, since deep learning networks can take whole images as the input, making them independent of ROI segmentation and allowing for added features such as anatomical location. In addition, trained deep learning models and learned features can be adapted to other imaging tasks using transfer learning [40].

Of note, the medical machine-learning field, unlike the conventional clinic or pharmacy, does not have a standardized research methodology. As the field grows the need for standardized methodology increases for study comparability. Methodological recommendations for the prediction or radiomic analysis are helpful in the creation and evaluation of clinical performance [10, 41, 42]. However, given that the field is new and rapidly changing, constant updates and additions to guides are required.

Applications of structural radiomic features in lung cancer

Chest radiography is an initial screening technique for lung cancer due to its low cost and convenience. A deep learning detection algorithm showed a high sensitivity for pulmonary nodules and lung cancer in the NLST and private datasets, and these studies suggested that this technique, as a second reader, could enable radiologists to better detect pulmonary nodules and lung cancer. [43–46]. In addition, Lu et al., developed a model focused on deep learning to classify a population of

smokers with high-risk of lung cancer and the AUC was slightly higher than the eligibility requirements for clinical assessment (0.76 vs. 0.63) [47]. LDCT can detect a small lesion to provide more information than a radiograph. Using deep learning with LDCT, a 90.0% sensitivity for pulmonary nodule detection has been achieved [48].

With the large numbers of small nodules found on LDCT lung cancer screening, an instrument for the classification of malignant nodules would boost clinical management. The performance of a CNN (AUC = 0.90) outperformed the Brock (AUC = 0.87) model for estimating the likelihood of malignant nodules [49]. Garau et al. found that the handcrafted radiomics features model had a higher AUC than the Lung-RADS clinical model (0.86 vs. 0.76) in the external validation [50]. In the case of solitary noncalcified nodules, the radiomic model incorporating perinodular and intranodular features demonstrated improved performance (intranodular radiomic features: AUC of 0.75; combination: AUC of 0.80) in distinguishing adenocarcinomas from granulomas in the non-contrast chest CT [51]. Some investigators, on the other hand, observed that the accuracy of the radiomics model was close to radiologists when contrast CT was used [52]. In addition, MRI radiomics also demonstrated strong success (AUC = 0.88) in the differentiation of lung malignancies and benign lesions [53]. A follow-up scan is a recommended method for the management of accidental pulmonary nodules. The changes of features (delta-radiomic) on the baseline and follow-up scans were able to be used to predict malignancy of nodules, and the dynamic details (e.g., tumor doubling time and growth) can be predicted by baseline scan [54–56].

In regard to the classification of histological subtypes, the combination of handcrafted radiomic and clinical features of a logistic regression nomogram was used to categorize small cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC) with an AUC of 0.94 and an accuracy of 86.2% [57]. In addition, the study selected five handcrafted radiomic features as a signature for differentiating lung squamous cell carcinoma (SCC) from adenocarcinoma with an AUC of 0.89 [58]. Research involving 920 patients showed that both handcraft (AUC of 0.79) and deep (AUC of 0.84) radiomics could attain high performance in distinguishing adenocarcinoma, SCC, and SCLC [59]. Recent studies have documented that a radiomic signature can provide information on the level of Ki-67 expression [60].

The majority type of early lung cancer is adenocarcinoma, in which adenocarcinoma in situ (AIS) or minimally invasive adenocarcinoma (MIA) has a near 100% 5-year survival probability after resection [61]. Recently, several studies have documented that CT-based handcrafted and deep radiomics have been able to determine the invasiveness of lung adenocarcinoma manifesting as sub-solid and solid nodules with AUC of 0.77 to 0.90 [62–64]. The combination of other variables such as clinical-, semantic-, and intraoperatively

pathological features can boost the accuracy of final pathology [65, 66]. In addition, a fusion of intranodular (solid and ground-glass) and perinodular radiomic features can be more predictive than the full gross tumor alone [67, 68]. In response to the specific types of invasive adenocarcinoma, both handcrafted and deep radiomics have shown efficacy in predicting higher invasive levels of solid/micropapillary adenocarcinoma [69–71].

Early-stage lung cancer, adenocarcinoma in particular, spreads through intrusive neighboring lymphovessel, pleura, and air space, which impairs the surgical result and affects the surgical approach. Two-dimensional texture features can individually predict lymphovascular invasion with an AUC of 0.86 [72]. Some researchers proposed that radiomic signature could potentially help to identify the pleural invasion of stage I pulmonary adenocarcinoma [73]. In addition, Zhuo et al. combined radiomic and semantic features (the size of the solid component and mediastinal lymphadenectasis) with an AUC of 0.99 for the prediction of the existence of spread through air space [74]. Intraoperative lymph node status is critical to choose a systematic or selective lymph node dissection. Several studies have shown that handcrafted and deep radiomic features of the intra/peri-tumor can be used as biomarkers to predict lymph node metastases [75–77]. Furthermore, in the case of pleural metastases, radiomic features may have a diagnostic power with AUC of 0.93 [78].

With the development of personalized treatment for lung cancer, the identification of therapeutically actionable mutations (e.g., Epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), programmed cell death 1 ligand, (PD-L1), and v-Raf murine sarcoma viral oncogene homolog B1 (BRAF)) has been a significant premises for an optimal treatment strategy. Thanks to the existence of non-invasive, simple, and low cost of radiomics compared to gene detection, which has demonstrated strong predictive efficacy for the mutation type and can be used as an alternative method [79–82]. In addition to predicting the status of gene mutations, some studies aimed to directly predict the treatment response, such as Immunotherapy, chemotherapy, and radiotherapy [83–86]. Another study has been able to predict biological substrates, such as tumor hypoxia with radiomics [87]. Dercle et al. conducted prospective and retrospective experiments in multicenter clinical trials specifically to predict the susceptibility of lung cancer to nivolumab, docetaxel, and gefitinib with an AUC of 0.77, 0.67, and 0.82, respectively [83]. Radiomics has been proven to be a valuable method in radiotherapy preparation, radiotherapy response, pathologic response to neoadjuvant chemoradiotherapy, and side effects of radiation pneumonitis [84–86].

Long-term prognostic outcomes (e.g., overall survival, disease-free survival, distant metastasis, and local recurrence) after therapy are chronically tracked and are expected to be primarily dependent on therapeutic, pathological, and

histological details. Radiomics can extract useful and unknown image features for predicting prognosis prior to treatment. Choe et al. observed that handcrafted radiomic features provided additional prognostic benefit outside the clinical-pathologic model alone (AUC: 0.78 vs. 0.73) [88]. In addition, radiomics has demonstrated a certain capacity to predict recurrence (AUC of 0.74–0.76) and distant metastasis (AUC of 0.89) [89, 90]. However, most studies have reported mild or even negative prognostic [91–93].

Applications of functional radiomic features in lung cancer

The most common functional imaging for lung cancer is 18F-FDG PET, which can reflect tumor glucose metabolism and capture both metabolic and structural information when paired with CT (PET/CT). Functional imaging has been widely used in clinical practice for diagnosis, staging, genetic mutation estimation, treatment response evaluation, and prognostic assessment.

Some studies have used PET-based radiomics alone to forecast clinical outcomes [94–97]. A research with 210 adenocarcinoma and 186 SCC patients showed that a PET-based radiomics signature could distinguish lung adenocarcinoma from SCC, albeit without external validation [94]. In 264 NSCLC patients, Tau et al. used PET images to train a CNN to predict lymph node metastases (accuracy of 80%) or distant metastases (accuracy of 63%) with moderate performance [95]. Some studies indicated that both PET-based radiomic features provided the prediction of prognostic outcome for lung cancer patients with radiotherapy or chemo-radiotherapy, whereas semiquantitative PET factors were not available [96]. Both PET-based handcrafted and deep features can predict the response of immunotherapy in lung adenocarcinoma [97, 98].

PET/CT provides more precise location information and more detailed surrounding structures than PET alone, meaning that radiomics based on PET/CT is able to extract both functional and structural feature and has a wider application prospect than PET or CT alone. One of the applications of diagnosis is using radiomics to distinguish pulmonary tuberculosis, lymphoma, and other benign lesions from lung cancer [99–101]. Another application of diagnosis based on PET/CT radiomics is to distinguish lung adenocarcinoma from SCC as well as primary from metastatic lung cancer [102, 103]. A recent study with a small sample size (with 91 patients) attempted to use 2 PET and 2 CT features for the identification of growth patterns in early lung invasive adenocarcinoma [104].

Furthermore, the prediction of EGFR mutations is a representative example of application of radiomics based on PET/CT. In a study with 248 lung cancer patients without treatment, researchers found that their model for prediction of

EGFR mutations could reach an AUC of 0.87 when combined clinical and radiomic signature [105]. Similar performance has also been reported in another retrospective study [106]. In addition, for patients with EGFR mutation, a deep radiomic score was a non-invasive tool to identify NSCLC patients susceptible to tyrosine kinase or immune checkpoint inhibitors [107].

Yang et al. concluded that a radiomic nomogram based on PET/CT rad-core and clinicopathological features was able to predict the overall survival of NSCLC patients [108]. A study focused on prediction of prognosis after immunotherapy suggested that PET/CT-based radiomics signature could be used before the start of treatment to identify those most likely to benefit from immunotherapy for advanced NSCLC patients [109]. In addition, a multicenter study with 87 early stage NSCLC patients underwent radiotherapy selected one PET and one CT feature to predict local recurrence and reached good performance (100% sensitivity and 96% specificity) [110]. Compared to handcrafted radiomic features, deep radiomic features had a significantly better prognostic value [111].

Dual-energy CT (DECT) provides additional perfusion information of tumor using quantification of iodine enhancement at different phases. In 93 lung cancer patients with examination of DECT, entropy from iodine overlay maps enhanced prediction of overall survival to pathological stage alone (C-index, 0.72 vs. 0.67) [112]. Table 1 summarized the radiomic studies using both structural and functional images.

Limitations and challenges

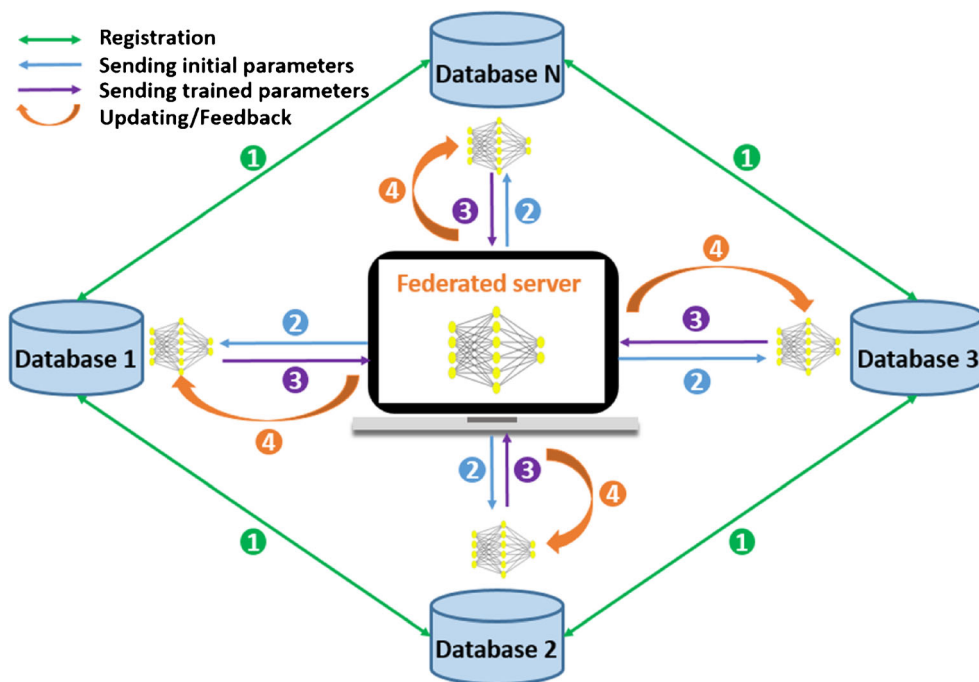
While radiomics has been successfully applied in the quantitative analysis of structural and functional images of lung cancer, certain limitations and obstacles must be faced and resolved before it is implemented in clinical practice. The first is that radiomics is a data-hungry approach. Large, diverse, multicenter, and high-quality data is needed to generalize the results and conclusions of radiomics studies. The creation of vast databases of medical images is currently problematic, mostly because sharing or exchange of data between hospitals and institutes is insufficient and the processing of data collection is time consuming. However, data sharing could pose possible ethical and legal dangers. Furthermore, the imaging data needs to be labeled with correct outcomes in order to be used for training, a process that adds to the cost burden. As the golden standard, the labels are also strongly contingent on the experience of physicians, subjective and complex [113]. Histopathological observations, for example, can be constrained by sampling errors and observer heterogeneity. Unsupervised and self-supervised methods are independent of particular labels, though accuracy and interpretability are

Table 1 Summary of some radiomic studies using both structural and functional images

Study	Study design	Modality	Population	Features type	Features selection	Model algorithm	Type of validation	Outcome
Du et al. (2020)	Retrospective Single-center	PET/CT	77 Tuberculosis 79 Lung cancers	Clinical Semantic Handcrafted radiomic	LASSO	LR	Leave one out Without external validation	Diagnosis
Sibille et al. (2020)	Retrospective Single-center	PET/CT	302 lung cancer 327 lymphoma	Deep radiomic	CNN	CNN	Leave one out Without external validation	Diagnosis
Kang et al. (2019)	Retrospective Single-center	PET/CT	157 malignant 111 benign patients	Clinical Handcrafted radiomic	LASSO	LR	Bootstrapping validation Without external validation	Diagnosis
Han et al. (2020)	Retrospective Single-center	PET/CT	867 adenocarcinomas 552 SCCs	Handcrafted radiomic	Ten feature selection methods	10 ML models and the VGG16	Leave one out Without external validation	Diagnosis
Kirienko et al. (2018)	Retrospective Single-center	PET/CT	534 Lung lesions	Deep radiomic Handcrafted radiomic	–	LDA	Leave one out Without external validation	Primary or metastatic lung lesions
Shao et al. (2020)	Retrospective Single-center	PET/CT	91 GGNs	Semantic Handcrafted radiomic	LASSO	LR	Bootstrapping validation Without external validation	Lepidic or acinar-papillary growth
Zhang et al. (2020)	Retrospective Single-center	PET/CT	248 NSCLCs	Clinical Handcrafted radiomic	LASSO	LR	Leave one out Without external validation	EGFR mutation
Liu et al. (2020)	Retrospective Single-center	PET/CT	148 Adenocarcinomas	Handcrafted radiomic	RF/LR	Xgboost	Leave one out Without external validation	EGFR mutation
Mu et al. (2020)	Retrospective Multi-center	PET/CT	681 NSCLCs	Deep radiomic	CNN	CNN	Leave one out With external validation	EGFR mutation Treatment response
Yang et al. (2020)	Retrospective Single-center	PET/CT	315 NSCLCs	Clinical Handcrafted radiomic	LASSO	LR	Leave one out Without external validation	Survival
Mu et al. (2019)	Retrospective/prospective Single-center	PET/CT	194 Stage IIIB-IV NSCLCs	Clinical Handcrafted radiomic	LASSO	LR	Leave one out With external validation	Survival after immunotherapy
Dissaux et al. (2020)	Retrospective Multi-center	PET/CT	87 Early-stage NSCLCs	Clinical Semantic Handcrafted radiomic	Univariate/ Multivariate analysis	Cox	Leave one out With external validation	Local Recurrence after radiotherapy

SCC, squamous cell carcinoma; NSCLC, non-small cell lung cancer; GGN, ground-glass nodule; LASSO, least absolute shrinkage and selection operator; CNN, convolutional neural network; RF, random forest; LR, logistic regression; ML, machine learning; EGFR, epidermal growth factor receptor

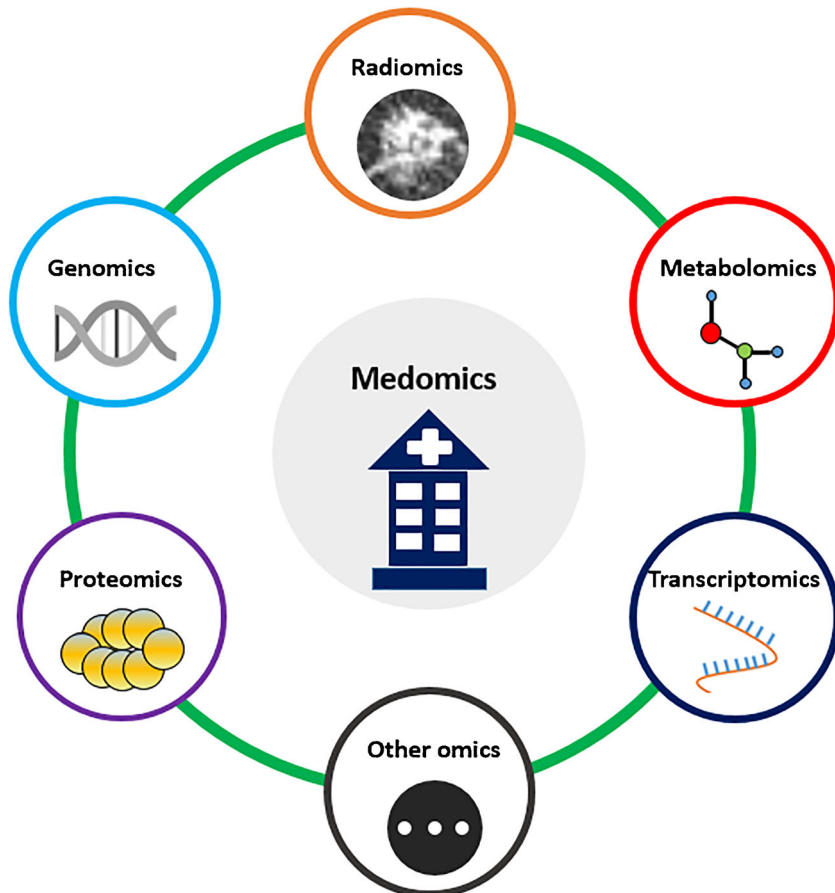
Fig. 3 The pipeline of federated learning, which includes the main four steps: data registration among local databases, sending initial parameters to each local center from federated server, sending trained parameters to federated server from local centers, and federated server aggregates the received parameters to update local model and to give a feedback to local database



sacrificed [114]. Precise segmentation is essential for handcrafted radiomics to select areas of interest for tumors. Manual segmentation is susceptible to inconsistency from

multiple readers and time consuming. While automatic and semi-automatic segmentation has been used to increase objectivity and minimize time costs, there is no norm for guiding or

Fig. 4 A scope of fusion of multi-discipline or multi-omics to form a “Medomics.” Other omics can be included in the Medomics, such as pathomics and lipidomics



assessing segmentation efficiency. Deep radiomics do not need explicit segmentation, although a rough ROI selection is required to conserve computing resources and reduce the effect of noise.

AI algorithms have made great strides in recent years, but the road is long and tortuous. There are plenty of methods to choose from during the process of preprocessing, feature extraction, feature selection, and modeling, though standard and appropriate workflow or guidelines for methodology and evaluation system are still uncommon. Currently, the most effective algorithm is deep learning based on artificial neural network that produce a huge number of computational parameters and requires high-performance computers to provide processing power and hardware support. More powerful and intelligent algorithms will certainly appear in the future, but the timing of the emergence may depend greatly on development of neuroscience and computer science. Data-driven radiomics especially deep radiomic features and models are basically a set of data that transforms into black boxes without intuitive interpretability. These models may have an even higher performance than humans, but they are unacceptable to clinicians who do not grasp how the machine makes decisions and works.

Furthermore, reproducibility is a basic requirement for clinical use. The variation in image acquisition and reconstruction as well as in the radiomics process can influence the reliability of the features, which is why many models do not perform well on independent external validation datasets. Although some harmonization approaches can reduce the batch effects in multicentric studies, the need for a “reference batch” to calibrate the harmonization hampers its use in prospective studies and real-time clinical practice [31]. A phantom study reported many (94%) handcrafted radiomic features were not reproducible and were redundant [115]. In addition, some studies concluded that the reproducibility of handcrafted radiomic features are easily affected by different acquisition and reconstruction parameters [116, 117]. Recently, the Image Biomarker Standardization Initiative (IBSI) assessed the reproducibility of handcrafted radiomic features and found a set of standardized 169 features that are deemed highly reproducible [118].

Most of the published studies were retrospective cohort studies with a limited sample size, and it can only provide low-level clinical evidence to prove the efficacy of radiomics. In order to assess the efficiency of additional prognostic and predictive benefit, radiomic features and models must be compared to typical clinical variables, so that radiomics must match precision and ease of use [119]. Finally, in actual clinical practice, patient conditions and the results of diagnostic imaging are very complex with numerous lesions and comorbidities present, and a prediction model based on a single lesion might not be able to fulfill clinical needs.

Future direction

Federated learning, as a distributed machine-learning framework, can easily solve the dilemma of data silos and make it possible to integrate a model from a local database without sharing data (Fig. 3) [120, 121]. In addition, the knowledge-driven and data-driven approaches will effectively minimize reliance on big data and achieve human-machine cooperation [122].

The implementation of standardized acquisition and reconstruction protocols as well as harmonizing radiomic features will greatly improve repeatability of radiomic features and generalization of radiomic signatures. Standardization and quality evaluation of radiomics methods are crucial in the reproducibility interpretability, generalization, and long-term clinical application for any studies and trials. Elaborate prospective clinical trials with a broad sample size with high-level evidence would be required to validate the utility of radiomics. Automation, ease-of-use, and multitasking can be a one-stop solution to help physicians to accelerate clinical practice and management.

Once existing limitations as mentioned above are overcome, investigators can use more resources to work on multi-modality fusion images and multi-discipline convergence (Fig. 4). More advanced imaging modalities and methods for lung cancer will reflect detailed structural and functional information, and will provide comprehensive and robust radiomic features and increase prospects beyond PET and CT. In addition, at a time when radiomics is taking off, AI technology is also developing rapidly in other disciplines (e.g., genomics, proteomics, and metabolomics). How to integrate radiomics with other omics to form a “Medomics” will be a fascinating avenue of further research and worth pursuing in the future.

Conclusion

In this review, we discussed some of the latest and most impactful radiomics studies for lung cancer in the context of semantic to deep radiomic features, summarized the application of structural and functional radiomics studies, and suggested current limitations and future directions quantitative image analysis in lung cancer. Overall, radiomic approaches focused on both structural and functional images continue to evolve rapidly and are expected to bridge the gap between conventional and precision medicine. In addition, comparing and combining multi-modality functional imaging with structural imaging for lung cancer radiomics should be addressed in the future. While current challenges in data and methodology obstruct the immediate adoption of this approach in clinical practice, radiomics still holds the promise to overcome

these hurdles and to be integrated in the “Medomics” workflows of the future.

Funding Open access funding provided by Maastricht University. This work was funded by ERC advanced grant (ERC-ADG-2015, n° 694812 - Hypoximmuno), European Program H2020 (ImmunoSABR - n° 733008, PREDICT - ITN - n° 766276, CHAIMELEON - n° 952172, EuCanImage - n° 952103), TRANSCAN Joint Transnational Call 2016 (JTC2016 “CLEARLY” - n° UM 2017–8295), China Scholarships Council (n° 201808210318), and Interreg V-A Euregio Meuse-Rhine (“Euradiomics” - n° EMR4). This work was supported by the Dutch Cancer Society (KWF Kankerbestrijding), Project number 12085/2018–2.

Declarations

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest P. Lambin reports, outside the submitted work, minority shares in The Medical Cloud Company and grants/sponsored research agreements from Varian medical, Oncoradiomics, ptTheragnostic/DNAmito and, Health Innovation Ventures. P. Lambin received an advisor/presenter fee and/or reimbursement of travel costs/external grant writing fee and/or in kind man-power contribution from Oncoradiomics, BHV, Varian, Elekta, ptTheragnostic and Convert pharmaceuticals. P. Lambin has shares in the company Oncoradiomics SA, Convert pharmaceuticals SA and is co-inventor of two issued patents with royalties on radiomics (PCT/NL2014/050248, PCT/NL2014/050728) licensed to Oncoradiomics and one issued patent on mtDNA (PCT/EP2014/059089) licensed to ptTheragnostic/DNAmito, three non-patented invention (softwares) licensed to ptTheragnostic/DNAmito, Oncoradiomics and Health Innovation Ventures and three non-issues, non-licensed patents on Deep Learning-Radiomics and LSRT (N2024482, N2024889, N2024889). H.C. Woodruff owns shares in Oncoradiomics, outside the submitted work. G. Wu, A. Jochems, T. Refaee, A. Ibrahim, C. Yan, and S. Sanduleanu declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics. 2020. *CA Cancer J Clin.* 2020;70:7–30.
- Howlander N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, et al. SEER cancer statistics review, 1975–2016. National Cancer Institute. 2019. https://seer.cancer.gov/csr/1975_2016/. Accessed 11 Nov 2020.
- Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, et al. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2017;28(suppl_4):iv1–iv21.
- National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011;365:395–409.
- de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med.* 2020;382:503–13.
- National Lung Screening Trial Research Team, Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, et al. Results of initial low-dose computed tomographic screening for lung cancer. *N Engl J Med.* 2013;368:1980–91.
- European Society of Radiology (ESR). Medical imaging in personalised medicine: a white paper of the research committee of the European Society of Radiology (ESR). *Insights Imaging.* 2015;6:141–55.
- Hirsch GV, Bauer CM, Merabet LB. Using structural and functional brain imaging to uncover how the brain adapts to blindness. *Ann Neurosci Psychol.* 2015;2:5.
- Hsia CC, Hyde DM, Ochs M, Weibel ER. ATS/ERS Joint Task Force on Quantitative Assessment of Lung Structure. An official research policy statement of the American Thoracic Society/ European Respiratory Society: standards for quantitative assessment of lung structure. *Am J Respir Crit Care Med.* 2010;181:394–418.
- Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017;14:749–62.
- Torigian DA, Huang SS, Houseni M, Alavi A. Functional imaging of cancer with emphasis on molecular techniques. *CA Cancer J Clin.* 2007;57:206–24.
- Langlotz CP. Will artificial intelligence replace radiologists? *Radiology: Artificial Intelligence.* 2019;1:3.
- Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. *Radiology.* 2008;246:697–722.
- Bankier AA, MacMahon H, Goo JM, Rubin GD, Schaefer-Prokop CM, Naidich DP. Recommendations for measuring pulmonary nodules at CT: a statement from the Fleischner Society. *Radiology.* 2017;285:584–600.
- Pinsky PF, Gierada DS, Black W, Munden R, Nath H, Aberle D, et al. Performance of Lung-RADS in the National Lung Screening Trial: a retrospective assessment. *Ann Intern Med.* 2015;162:485–91.
- Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* 2009;45:228–47.
- Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol.* 2017;18:e143–52.
- O JH, Lodge MA, Wahl RL. Practical PERCIST: A simplified guide to PET response criteria in solid tumors 1.0. *Radiology.* 2016;280:576–84.
- Gould MK, Ananth L, Barnett PG. Veterans Affairs SNAP Cooperative Study Group. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. *Chest.* 2007;131:383–8.

20. Liang J, Xu XQ, Xu H, Yuan M, Zhang W, Shi ZF, et al. Using the CT features to differentiate invasive pulmonary adenocarcinoma from pre-invasive lesion appearing as pure or mixed ground-glass nodules. *Br J Radiol.* 2015;20140811:88.
21. Lin J, Carter CA, McGlynn KA, Zahm SH, Nations JA, Anderson WF, et al. A prognostic model to predict mortality among non-small-cell lung cancer patients in the U.S. Military Health System. *J Thorac Oncol.* 2015;10:1694–702.
22. Zhang Y, Zheng D, Xie J, Li Y, Wang Y, Li C, et al. development and validation of web-based nomograms to precisely predict conditional risk of site-specific recurrence for patients with completely resected non-small cell lung cancer: a multi institutional study. *Chest.* 2018;154:501–11.
23. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18:500–10.
24. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48:441–6.
25. Sanduleanu S, Woodruff HC, de Jong EEC, van Timmeren JE, Jochems A, Dubois L, et al. Tracking tumor biology with radiomics: A systematic review utilizing a radiomics quality score. *Radiother Oncol.* 2018;127:349–60.
26. Vinod SK, Jameson MG, Min M, Holloway LC. Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies. *Radiother Oncol.* 2016;121:169–79.
27. Tingelhoff K, Moral AI, Kunkel ME, Rilk M, Wagner I, Eichhorn KG, et al. Comparison between manual and semi-automatic segmentation of nasal cavity and paranasal sinuses from CT images. *Annu Int Conf IEEE Eng Med Biol Soc.* 2007;2007:5505–8.
28. Um H, Tixier F, Bermudez D, Deasy JO, Young RJ, Veeraraghavan H. Impact of image preprocessing on the scanner dependence of multi-parametric MRI radiomic features and covariate shift in multi-institutional glioblastoma datasets. *Phys Med Biol.* 2019;64:165011.
29. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol Biol Phys.* 2018;102:1143–58.
30. McNitt-Gray M, Napel S, Jaggi A, Mattonen SA, Hadjiiski L, Muzi M, et al. Standardization in quantitative imaging: a multi-center comparison of radiomic features from different software packages on digital reference objects and patient data sets. *Tomography.* 2020;6:118–28.
31. Da-Ano R, Visvikis D, Hatt M. Harmonization strategies for multicenter radiomics investigations. *Phys Med Biol.* 2020. <https://doi.org/10.1088/1361-6560/aba798>.
32. Hira ZM, Gillies DF. A review of feature selection and feature extraction methods applied on microarray data. *Adv Bioinformatics.* 2015;2015:198363.
33. Handelman GS, Kok HK, Chandra RV, Razavi AH, Huang S, Brooks M, et al. Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods. *AJR Am J Roentgenol.* 2019;212:38–43.
34. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. *Br J Cancer.* 2003;89:232–8.
35. Wei Q, Dunbrack RL Jr. The role of balanced training and testing data sets for binary classifiers in bioinformatics. *PLoS One.* 2013;8:e67863.
36. Fotouhi S, Asadi S, Kattan MW. A comprehensive data level analysis for cancer diagnosis on imbalanced data. *J Biomed Inform.* 2019;90:103089.
37. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521:436–44.
38. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging.* 2018;9:611–29.
39. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. *Theranostics.* 2019;9:1303–22.
40. Weiss K, Khoshgoftaar TM, Wang DD. A survey of transfer learning. *J Big Data.* 2016;3:9.
41. Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology.* 2018;286:800–9.
42. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162:W1–73.
43. Yoo H, Kim KH, Singh R, Digumarthy SR, Kalra MK. Validation of a deep learning algorithm for the detection of malignant pulmonary nodules in chest radiographs. *JAMA Netw Open.* 2020;3:e2017135.
44. Sim Y, Chung MJ, Kotter E, Yune S, Kim M, Do S, et al. Deep convolutional neural network-based software improves radiologist detection of malignant lung nodules on chest radiographs. *Radiology.* 2020;294:199–209.
45. Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology.* 2019;290:218–28.
46. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med.* 2019;25:954–61.
47. Lu MT, Raghu VK, Mayrhofer T, Aerts HJWL, Hoffmann U. Deep learning using chest radiographs to identify high-risk smokers for lung cancer screening computed tomography: development and validation of a prediction model. *Ann Intern Med.* 2020. <https://doi.org/10.7326/M20-1868>.
48. Liu C, Hu SC, Wang C, Lafata K, Yin FF. Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data. *Quant Imaging Med Surg.* 2020;10:1917–29.
49. Baldwin DR, Gustafson J, Pickup L, Arteta C, Novotny P, Declerck J, et al. External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. *Thorax.* 2020;75:306–12.
50. Garau N, Paganelli C, Summers P, Choi W, Alam S, Lu W, et al. External validation of radiomics-based predictive models in low-dose CT screening for early lung cancer diagnosis. *Med Phys.* 2020. <https://doi.org/10.1002/mp.14308>.
51. 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.
52. Yang X, He J, Wang J, Li W, Liu C, Gao D, et al. CT-based radiomics signature for differentiating solitary granulomatous nodules from solid lung adenocarcinoma. *Lung Cancer.* 2018;125:109–14.
53. Wang X, Wan Q, Chen H, Li Y, Li X. Classification of pulmonary lesion based on multiparametric MRI: utility of radiomics and comparison of machine learning methods. *Eur Radiol.* 2020;30:4595–605.
54. Huang P, Lin CT, Li Y, Tammemagi MC, Brock MV, Atkar-Khatra S, et al. Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. *Lancet Digit Health.* 2019;1:e353–62.

55. Yoon HJ, Park H, Lee HY, Sohn I, Ahn J, Lee SH. Prediction of tumor doubling time of lung adenocarcinoma using radiomic margin characteristics. *Thorac Cancer*. 2020;11:2600–9.
56. Huang C, Lv W, Zhou C, Mao L, Xu Q, Li X, et al. Discrimination between transient and persistent subsolid pulmonary nodules on baseline CT using deep transfer learning. *Eur Radiol*. 2020. <https://doi.org/10.1007/s00330-020-07071-6>.
57. Liu S, Liu S, Zhang C, Yu H, Liu X, Hu Y, et al. Exploratory study of a CT Radiomics model for the classification of small cell lung cancer and non-small-cell lung cancer. *Front Oncol*. 2020;10:1268.
58. Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. *Eur Radiol*. 2018;28:2772–8.
59. Guo Y, Song Q, Jiang M, Guo Y, Xu P, Zhang Y, et al. Histological subtypes classification of lung cancers on ct images using 3d deep learning and radiomics. *Acad Radiol*. 2020;S1076-6332(20):30360–3.
60. Gu Q, Feng Z, Liang Q, Li M, Deng J, Ma M, et al. Machine learning-based radiomics strategy for prediction of cell proliferation in non-small cell lung cancer. *Eur J Radiol*. 2019;118:32–7.
61. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, 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.
62. Xia X, Gong J, Hao W, Yang T, Lin Y, Wang S, et al. Comparison and fusion of deep learning and radiomics features of ground-glass nodules to predict the invasiveness risk of stage-i lung adenocarcinomas in CT scan. *Front Oncol*. 2020;10:418.
63. Sun Y, Li C, Jin L, Gao P, Zhao W, Ma W, et al. Radiomics for lung adenocarcinoma manifesting as pure ground-glass nodules: invasive prediction. *Eur Radiol*. 2020;30:3650–9.
64. Wang X, Li Q, Cai J, Wang W, Xu P, Zhang Y, et al. Predicting the invasiveness of lung adenocarcinomas appearing as ground-glass nodule on CT scan using multi-task learning and deep radiomics. *Transl Lung Cancer Res*. 2020;9:1397–406.
65. Wu G, Woodruff HC, Sanduleanu S, Refaee T, Jochems A, Leijenaar R, et al. Preoperative CT-based radiomics combined with intraoperative frozen section is predictive of invasive adenocarcinoma in pulmonary nodules: a multicenter study. *Eur Radiol*. 2020;30:2680–91.
66. Wang B, Tang Y, Chen Y, Hamal P, Zhu Y, Wang T, et al. Joint use of the radiomics method and frozen sections should be considered in the prediction of the final classification of peripheral lung adenocarcinoma manifesting as ground-glass nodules. *Lung Cancer*. 2020;139:103–10.
67. Wu G, Woodruff HC, Shen J, Refaee T, Sanduleanu S, Ibrahim A, et al. diagnosis of invasive lung adenocarcinoma based on chest CT radiomic features of part-solid pulmonary nodules: a multicenter study. *Radiology*. 2020;297:451–8.
68. Wu L, Gao C, Xiang P, Zheng S, Pang P, Xu M. CT-imaging based analysis of invasive lung adenocarcinoma presenting as ground glass nodules using peri- and intra-nodular radiomic features. *Front Oncol*. 2020;10:838.
69. Ding H, Xia W, Zhang L, Mao Q, Cao B, Zhao Y, et al. CT-based deep learning model for invasiveness classification and micropapillary pattern prediction within lung adenocarcinoma. *Front Oncol*. 2020;10:1186.
70. Park S, Lee SM, Noh HN, Hwang HJ, Kim S, Do KH, et al. Differentiation of predominant subtypes of lung adenocarcinoma using a quantitative radiomics approach on CT. *Eur Radiol*. 2020;30:4883–92.
71. Wang X, Zhang L, Yang X, Tang L, Zhao J, Chen G, et al. Deep learning combined with radiomics may optimize the prediction in differentiating high-grade lung adenocarcinomas in ground glass opacity lesions on CT scans. *Eur J Radiol*. 2020;129:109150. <https://doi.org/10.1016/j.ejrad.2020.109150>.
72. Yang G, Nie P, Zhao L, Guo J, Xue W, Yan L, et al. 2D and 3D texture analysis to predict lymphovascular invasion in lung adenocarcinoma. *Eur J Radiol*. 2020;129:109111. <https://doi.org/10.1016/j.ejrad.2020.109111>.
73. Yuan M, Liu JY, Zhang T, Zhang YD, Li H, Yu TF. Prognostic impact of the findings on thin-section computed tomography in stage I lung adenocarcinoma with visceral pleural invasion. *Sci Rep*. 2018;8:4743.
74. Zhuo Y, Feng M, Yang S, Zhou L, Ge D, Lu S, et al. Radiomics nomograms of tumors and peritumoral regions for the preoperative prediction of spread through air spaces in lung adenocarcinoma. *Transl Oncol*. 2020;100820:13.
75. Zhao X, Wang X, Xia W, Li Q, Zhou L, Li Q, et al. A cross-modal 3D deep learning for accurate lymph node metastasis prediction in clinical stage T1 lung adenocarcinoma. *Lung Cancer*. 2020;145:10–7.
76. Yang M, She Y, Deng J, Wang T, Ren Y, Su H, 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.
77. He L, Huang Y, Yan L, Zheng J, Liang C, Liu Z. Radiomics-based predictive risk score: A scoring system for preoperatively predicting risk of lymph node metastasis in patients with resectable non-small cell lung cancer. *Chin J Cancer Res*. 2019;31:641–52.
78. Yang M, Ren Y, She Y, Xie D, Sun X, Shi J, et al. Imaging phenotype using radiomics to predict dry pleural dissemination in non-small cell lung cancer. *Ann Transl Med*. 2019;7:259.
79. 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;1800986:53.
80. Zhu Y, Liu YL, Feng Y, Yang XY, Zhang J, Chang DD, et al. A CT-derived deep neural network predicts for programmed death ligand-1 expression status in advanced lung adenocarcinomas. *Ann Transl Med*. 2020;8:930.
81. Padole A, Singh R, Zhang EW, Mendoza DP, Dagogo-Jack I, Kalra MK, 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.
82. 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.
83. Derclé L, Fronheiser M, Lu L, Du S, Hayes W, Leung DK, et al. Identification of non-small cell lung cancer sensitive to systemic cancer therapies using radiomics. *Clin Cancer Res*. 2020;26:2151–62.
84. Bi N, Wang J, Zhang T, Chen X, Xia W, Miao J, et al. Deep learning improved clinical target volume contouring quality and efficiency for postoperative radiation therapy in non-small cell lung cancer. *Front Oncol*. 2019;9:1192.
85. Khorrami M, Jain P, Bera K, Alilou M, Thawani R, Patil P, et al. Predicting pathologic response to neoadjuvant chemoradiation in resectable stage III non-small cell lung cancer patients using computed tomography radiomic features. *Lung Cancer*. 2019;135:1–9.
86. Liang B, Tian Y, Chen X, Yan H, Yan L, Zhang T, et al. Prediction of radiation pneumonitis with dose distribution: a convolutional neural network (CNN) based model. *Front Oncol*. 2020;9:1500.
87. Sanduleanu S, Jochems A, Upadhaya TT, Even AJG, Leijenaar RTH, Dankers FJWM, et al. Non-invasive imaging prediction of tumor hypoxia: a novel developed and externally validated CT

- and FDG-PET-based radiomic signatures. *Radiother Oncol.* 2020;S0167-8140(20):30852–5. <https://doi.org/10.1016/j.radonc.2020.10.016>.
88. Choe J, Lee SM, Do KH, Kim S, Choi S, Lee JG, et al. Outcome prediction in resectable lung adenocarcinoma patients: value of CT radiomics. *Eur Radiol.* 2020;30:4952–63.
 89. Khorrani M, Bera K, Leo P, Vaidya P, Patil P, Thawani R, et al. Stable and discriminating radiomic predictor of recurrence in early stage non-small cell lung cancer: Multi-site study. *Lung Cancer.* 2020;142:90–7.
 90. Zhou H, Dong D, Chen B, Fang M, Cheng Y, Gan Y, et al. Diagnosis of distant metastasis of lung cancer: based on clinical and radiomic features. *Transl Oncol.* 2018;11:31–6.
 91. Kim H, Goo JM, Lee KH, Kim YT, Park CM. Preoperative CT-based deep learning model for predicting disease-free survival in patients with lung adenocarcinomas. *Radiology.* 2020;296:216–24.
 92. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology.* 2016;28:947–57.
 93. van Timmeren JE, van Elmpt W, Leijenaar RTH, Reymen B, Monshouwer R, Bussink J, et al. Longitudinal radiomics of cone-beam CT images from non-small cell lung cancer patients: Evaluation of the added prognostic value for overall survival and locoregional recurrence. *Radiother Oncol.* 2019;136:78–85.
 94. Hyun SH, Ahn MS, Koh YW, Lee SJ. A machine-learning approach using PET-based radiomics to predict the histological subtypes of lung cancer. *Clin Nucl Med.* 2019. <https://doi.org/10.1097/RLU.0000000000002810>.
 95. Tau N, Stundzia A, Yasufuku K, Hussey D, Metser U. Convolutional neural networks in predicting nodal and distant metastatic potential of newly diagnosed non-small cell lung cancer on FDG PET images. *AJR Am J Roentgenol.* 2020;215:192–7.
 96. Arshad MA, Thornton A, Lu H, Tam H, Wallitt K, Rodgers N, et al. Discovery of pre-therapy 2-deoxy-2-18F-fluoro-D-glucose positron emission tomography-based radiomics classifiers of survival outcome in non-small-cell lung cancer patients. *Eur J Nucl Med Mol Imaging.* 2019;46:455–66.
 97. Valentinuzzi D, Vrankar M, Boc N, Ahac V, Zupancic Z, Unk M, et al. [18F] FDG PET immunotherapy radiomics signature (iRADIOMICS) predicts response of non-small-cell lung cancer patients treated with pembrolizumab. *Radiol Oncol.* 2020;54:285–94.
 98. Park C, Na KJ, Choi H, Ock CY, Ha S, Kim M, et al. Tumor immune profiles noninvasively estimated by FDG PET with deep learning correlate with immunotherapy response in lung adenocarcinoma. *Theranostics.* 2020;10:10838–48.
 99. Du D, Gu J, Chen X, Lv W, Feng Q, Rahmim A, et al. Integration of PET/CT radiomics and semantic features for differentiation between active pulmonary tuberculosis and lung cancer. *Mol Imaging Biol.* 2020. <https://doi.org/10.1007/s11307-020-01550-4>.
 100. Sibille L, Seifert R, Avramovic N, Vehren T, Spottiswoode B, Zuehlsdorff S, et al. 18F-FDG PET/CT uptake classification in lymphoma and lung cancer by using deep convolutional neural networks. *Radiology.* 2020;294:445–52.
 101. Kang F, Mu W, Gong J, Wang S, Li G, Li G, et al. Integrating manual diagnosis into radiomics for reducing the false positive rate of 18F-FDG PET/CT diagnosis in patients with suspected lung cancer. *Eur J Nucl Med Mol Imaging.* 2019;46:2770–9.
 102. Han Y, Ma Y, Wu Z, Zhang F, Zheng D, Liu X, et al. Histologic subtype classification of non-small cell lung cancer using PET/CT images. *Eur J Nucl Med Mol Imaging.* 2020. <https://doi.org/10.1007/s00259-020-04771-5>.
 103. Kirienko M, Cozzi L, Rossi A, Voulaz E, Antunovic L, Fogliata 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.
 104. Shao X, Niu R, Shao X, Jiang Z, Wang Y. Value of 18F-FDG PET/CT-based radiomics model to distinguish the growth patterns of early invasive lung adenocarcinoma manifesting as ground-glass opacity nodules. *EJNMMI Res.* 2020;10:80.
 105. Zhang J, Zhao X, Zhao Y, Zhang J, Zhang Z, Wang J, et al. Value of pre-therapy 18F-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.
 106. Liu Q, Sun D, Li N, Kim J, Feng D, Huang G, et al. Predicting EGFR mutation subtypes in lung adenocarcinoma using 18F-FDG PET/CT radiomic features. *Transl Lung Cancer Res.* 2020;9:549–62.
 107. Mu W, Jiang L, Zhang J, Shi Y, Gray JE, Tunali I, et al. Non-invasive decision support for NSCLC treatment using PET/CT radiomics. *Nat Commun.* 2020;11:5228.
 108. Yang B, Zhong J, Zhong J, Ma L, Li A, Ji H, et al. Development and validation of a radiomics nomogram based on 18F-fluorodeoxyglucose positron emission tomography/computed tomography and clinicopathological factors to predict the survival outcomes of patients with non-small cell lung cancer. *Front Oncol.* 2020;10:1042.
 109. Mu W, Tunali I, Gray JE, Qi J, Schabath MB, Gillies RJ. Radiomics of 18F-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.
 110. Dissaux G, Visvikis D, Da-Ano R, Pradier O, Chajon E, Barillot I, et al. Pretreatment 18F-FDG PET/CT radiomics predict local recurrence in patients treated with stereotactic body radiotherapy for early-stage non-small cell lung cancer: a multicentric study. *J Nucl Med.* 2020;61:814–20.
 111. Afshar P, Mohammadi A, Tyrrell PN, Cheung P, Sigiuk A, Plataniotis KN, et al. DRTOP: deep learning-based radiomics for the time-to-event outcome prediction in lung cancer. *Sci Rep.* 2020;10:12366.
 112. Choe J, Lee SM, Do KH, Lee JB, Lee SM, Lee JG, et al. Prognostic value of radiomic analysis of iodine overlay maps from dual-energy computed tomography in patients with resectable lung cancer. *Eur Radiol.* 2019;29:915–23.
 113. Peck M, Moffat D, Latham B, Badrick T. Review of diagnostic error in anatomical pathology and the role and value of second opinions in error prevention. *J Clin Pathol.* 2018;71:995–1000.
 114. Manning CD, Clark K, Hewitt J, Khandelwal U, Levy O. Emergent linguistic structure in artificial neural networks trained by self-supervision. *Proc Natl Acad Sci U S A.* 2020;201907367.
 115. Berenguer R, Pastor-Juan MDR, Canales-Vázquez J, Castro-García M, Villas MV, Mansilla Legorburu F, et al. Radiomics of CT features may be nonreproducible and redundant: influence of CT acquisition parameters. *Radiology.* 2018;288:407–15.
 116. Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol.* 2015;50:757–65.
 117. Zhovannik I, Bussink J, Traverso A, Shi Z, Kalendralis P, Wee L, et al. Learning from scanners: Bias reduction and feature correction in radiomics. *Clin Transl Radiat Oncol.* 2019;19:33–8.
 118. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology.* 2020;295:328–38.
 119. Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: The need for safeguards. *Radiother Oncol.* 2019;130:2–9.

120. Jochems A, Deist TM, van Soest J, Eble M, Bulens P, Coucke P, et al. Distributed learning: Developing a predictive model based on data from multiple hospitals without data leaving the hospital - A real life proof of concept. *Radiother Oncol.* 2016;121:459–67.
121. Jochems A, Deist TM, El Naqa I, Kessler M, Mayo C, Reeves J, et al. Developing and validating a survival prediction model for nsccl patients through distributed learning across 3 countries. *Int J Radiat Oncol Biol Phys.* 2017;99:344–52.
122. Sun J, Hu J, Luo D, Markatou M, Wang F, Edabollahi S, et al. Combining knowledge and data driven insights for identifying risk factors using electronic health records. *AMIA Annu Symp Proc.* 2012;2012:901–10.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.