University of Groningen

Next-Generation Radiogenomics Sequencing for Prediction of EGFR and KRAS Mutation Status in NSCLC Patients Using Multimodal Imaging and Machine Learning Algorithms

Shiri, Isaac; Maleki, Hasan; Hajianfar, Ghasem; Abdollahi, Hamid; Ashrafinia, Saeed; Hatt, Mathieu; Zaidi, Habib; Oveisi, Mehrdad; Rahmim, Arman

Published in:
Molecular Imaging and Biology

DOI:
[10.1007/s11307-020-01487-8](https://doi.org/10.1007/s11307-020-01487-8)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Shiri, I., Maleki, H., Hajianfar, G., Abdollahi, H., Ashrafinia, S., Hatt, M., Zaidi, H., Oveisi, M., & Rahmim, A. (2020). Next-Generation Radiogenomics Sequencing for Prediction of EGFR and KRAS Mutation Status in NSCLC Patients Using Multimodal Imaging and Machine Learning Algorithms. *Molecular Imaging and Biology*, 22(4), 1132-1148. <https://doi.org/10.1007/s11307-020-01487-8>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

RESEARCH ARTICLE

Next-Generation Radiogenomics Sequencing for Prediction of EGFR and KRAS Mutation Status in NSCLC Patients Using Multimodal Imaging and Machine Learning Algorithms

Isaac Shiri,¹ Hasan Maleki,^{2,3} Ghasem Hajianfar,¹ Hamid Abdollahi,^{1,4}
Saeed Ashrafinia,^{5,6} Mathieu Hatt,⁷ Habib Zaidi,^{1,8,9,10} Mehrdad Oveisi,^{2,11}
Arman Rahmim^{5,12,13}

¹Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211, Geneva 4, Switzerland

²Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Science, Tehran, Iran

³Department of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran

⁴Department of Radiologic Sciences and Medical Physics, Faculty of Allied Medicine, Kerman University of Medical Sciences, Kerman, Iran

⁵Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD, USA

⁶Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA

⁷LaTIM, INSERM, UMR 1101, Univ Brest, Brest, France

⁸Geneva University Neurocenter, Geneva University, Geneva, Switzerland

⁹Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands

¹⁰Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark

¹¹Department of Computer Science, University of British Columbia, Vancouver, BC, Canada

¹²Departments of Radiology and Physics, University of British Columbia, Vancouver, BC, Canada

¹³Department of Integrative Oncology, BC Cancer Research Centre, 675 West 10th Ave, Vancouver, BC, V5Z 1L3, Canada

Abstract

Purpose: Considerable progress has been made in the assessment and management of non-small cell lung cancer (NSCLC) patients based on mutation status in the epidermal growth factor receptor (EGFR) and Kirsten rat sarcoma viral oncogene (KRAS). At the same time, NSCLC management through KRAS and EGFR mutation profiling faces challenges. In the present work, we aimed to evaluate a comprehensive radiomics framework that enabled prediction of EGFR and KRAS mutation status in NSCLC patients based on radiomic features from low-dose computed tomography (CT), contrast-enhanced diagnostic quality CT (CTD), and positron emission tomography (PET) imaging modalities and use of machine learning algorithms.

Methods: Our study involved NSCLC patients including 150 PET, low-dose CT, and CTD images. Radiomic features from original and preprocessed (including 64 bin discretizing, Laplacian-of-Gaussian (LOG), and Wavelet) images were extracted. Conventional clinically

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11307-020-01487-8>) contains supplementary material, which is available to authorized users.

Correspondence to: Arman Rahmim; e-mail: arman.rahmim@ubc.ca

used standard uptake value (SUV) parameters and metabolic tumor volume (MTV) were also obtained from PET images. Highly correlated features were pre-eliminated, and false discovery rate (FDR) correction was performed with the resulting q-values reported for univariate analysis. Six feature selection methods and 12 classifiers were then used for multivariate prediction of gene mutation status (provided by polymerase chain reaction (PCR)) in patients. We performed 10-fold cross-validation for model tuning to improve robustness, and our developed models were assessed on an independent validation set with 68 patients (common in all three imaging modalities). The average area under the receiver operator characteristic curve (AUC) was utilized for performance evaluation.

Results: The best predictive power for conventional PET parameters was achieved by SUV_{peak} (AUC 0.69, p value = 0.0002) and MTV (AUC 0.55, p value = 0.0011) for EGFR and KRAS, respectively. Univariate analysis of extracted radiomics features improved AUC performance to 0.75 (q-value 0.003, Short-Run Emphasis feature of GLRLM from LOG preprocessed image of PET with sigma value 1.5) and 0.71 (q-value 0.00005, Large Dependence Low Gray-Level Emphasis feature of GLDM in LOG preprocessed image of CT with sigma value 5) for EGFR and KRAS, respectively. Furthermore, multivariate machine learning-based AUC performances were significantly improved to 0.82 for EGFR (LOG preprocessed image of PET with sigma 3 with variance threshold (VT) feature selector and stochastic gradient descent (SGD) classifier (q-value = 4.86E-05) and 0.83 for KRAS (LOG preprocessed image of CT with sigma 3.5 with select model (SM) feature selector and SGD classifier (q-value = 2.81E-09).

Conclusion: Our work demonstrated that non-invasive and reliable radiomics analysis can be successfully used to predict EGFR and KRAS mutation status in NSCLC patients. We demonstrated that radiomic features extracted from different image-feature sets could be used for EGFR and KRAS mutation status prediction in NSCLC patients and showed improved predictive power relative to conventional image-derived metrics.

Key words: Radiogenomics, PET/CT, Machine learning, NSCLC, KRAS, EGFR

Introduction

Considerable progress has been made recently in the assessment and management of non-small-cell lung cancer (NSCLC) based on mutation status in the epidermal growth factor receptor (EGFR) and Kirsten rat sarcoma viral oncogene (KRAS) genes [1, 2]. Ongoing studies on molecular cancer profiling have revealed that EGFR and KRAS are involved in the occurrence, development, invasion, and metastasis of NSCLC [3]. Moreover, studies have indicated that mutations in KRAS and EGFR are considered as first lines for clinical decision-making in NSCLC treatment and outcome improvement [4]. Furthermore, recent studies have identified that NSCLC patients with mutant KRAS tumors fail to benefit from adjuvant chemotherapy, and their disease does not respond to EGFR inhibitors [5]. In addition, it was reported that outcomes and patterns of failure in genotypic subgroups of NSCLC patients, based on mutations in EGFR or KRAS, can inform the design of future trials integrating targeted therapies [6].

Although KRAS and EGFR mutation profiling is critical in NSCLC management, some studies have raised issues with this approach [7]. First, such mutation status captures only a small degree of tumor heterogeneity and does not provide a complete picture for the assessment of tumor characteristics. Secondly, this method depicts low repeatability and is not feasible for all cases

[8]. Furthermore, the method suffers from invasiveness and patients discomfort [9].

As a recently developed paradigm of advanced medical image quantification, radiomics has garnered significant interest given its cost-effectiveness and reliability to characterize tumor heterogeneity and has enabled improved assessment of therapy response and prediction of molecular pathways [10–12]. Accumulating evidence has identified several radiomic features extracted from computed tomography (CT), magnetic resonance (MR), or positron emission tomography (PET) images as highly correlated with genomic parameters in several cancers [8, 13, 14]. For NSCLC patients, radiomics studies have shown that several CT image features can predict mutation status in EGFR and KRAS [8]. For example, features such as size, edge, lucency, and homogeneity extracted from CT images could identify EGFR mutation status [15]. Also, some diffusion-weighted (DW) MR image-intensity histogram features, including mean, skewness, and 10th and 90th percentiles, have been shown to predict EGFR mutation in lung adenocarcinoma [16]. On the other hand, Velazquez et al. developed radiomics models based on CT image features and clinical parameters to distinguish between $EGFR^-$ and $EGFR^+$ and $KRAS^+$ and $KRAS^-$ [17]. Liu et al. also evaluated the ability of CT image features to predict EGFR mutation status in 298 surgically

resected peripheral lung adenocarcinomas in an Asian cohort of patients and built a high-performance predictive model by using multiple logistic regression algorithms [15]. Zhang et al. also developed a radiogenomic model based on CT image features to predict EGFR and KRAS mutations in lung adenocarcinoma patients [18]. Wang et al. [19] and Zhao et al. [20] recently reported on a deep learning and radiomics study on EGFR prediction using CT images. Their radiomics study includes one feature selector (recursive feature elimination, ICC reproducibility test) and classifier (random forest, Logistic regression). They concluded that deep learning can be a non-invasive tool for the prediction of EGFR mutation status. Li et al. [21] investigated EGFR mutation status using PET SUV biomarkers as well as radiomics features. They reported the feasibility of quantitative measurement of tumor phenotype using radiomics. More recently, Pinheiro et al. [22] presented a radiomic study using semantic and radiomic features on EGFR and KRAS prediction using only CT image features. They concluded that lung semantic features in conjunction with tumor-specific semantic features can improve the prediction power of gene mutation status while the radiomic features did not provide good prediction power.

Incorporating radiomics for a prediction study requires a multi-step process that involves reliable statistical analyses, such as feature selection and classification, to reduce overfitting and to build robust predictive or prognostic models [23]. A number of machine learning (ML) algorithms can provide robust means to identify a subset of features to combine into a multi-parametric model [24]. Although several ML algorithms, alone or in combination, have been used in radiomics analysis for feature selection and classification, there is no “one fits all” approach as performance of various ML workflows has been shown to depend on application and/or type of data [25–28]. Previous studies have tested cross-combination of different ML approaches and have suggested distinct ML algorithms that depict high performance for feature selection and classification [24, 25].

In the present study, we aimed to evaluate a comprehensive multimodal (diagnostic CT, low-dose CT, and PET modality) radiomics, univariate analysis, and machine learning framework to predict EGFR and KRAS mutation status in NSCLC patients.

Materials and Methods

Radiomics Analysis

Our radiomics analysis included seven steps as shown in Fig. 1 and elaborated next.

Data Collection This study was conducted on 211 NSCLC patients with available imaging and genomic data.

Imaging studies for all patients were obtained from The Cancer Imaging Archive (TCIA) [29–32] and included contrast-enhanced diagnostic quality CT (CTD) and PET/CT (*i.e.*, low-dose CT (CT) as used for PET attenuation correction and the PET image). The clinical characteristics of all patients are presented in Table 1 while the inclusion and exclusion criteria are presented in Fig. 2. Detailed information about image acquisition and generation were presented in previous works [29–32].

Imaging

All images were acquired with two vendors of imaging systems, including GE Healthcare and Siemens Healthcare. For CTD, 73 % of data were acquired on a GE Discovery CT750HD and 27 % on a Siemens. Slice thicknesses for CTD were 0.625–3 mm (median 1.5 mm) and an x-ray tube current of 124–699 mA (mean 220 mA) at 80–140 kVp (mean 120 kVp), and spiral pitch factor of 0.9–1. A GE Discovery D690 PET/CT and a GE Discovery PET/CT scanner were used for PET/CT scanning. ^{18}F -FDG injected activity and uptake time were 138.90–572.25 MBq (mean 309.26 MBq) and 23.08–128.90 min (mean 66.58 min), respectively. Each bed position acquisition time was 1–5-min, where iterative Ordered Subset Expectation Maximization (OSEM) algorithm was used for PET image reconstruction. Slice thicknesses for low-dose CT were set to 3–5 mm, mA range 36–400, kVp 120–140, and spiral pitch factor 0.9–1. Due to certain problems, including image noise, artifacts, absence of images for some sequences, and image mis-segmentation, some patients were excluded. All in all, 186 PET and 175 low-dose CT and CTD images were segmented.

Genomics

Tumor samples were excised and cut by a surgeon with slice thickness of 3–5 mm and were frozen for 30 min. For single nucleotide mutation detection, SNaPshot technology based on dideoxy single-base extension of oligonucleotide primers was performed after multiplex polymerase chain reaction (PCR) [33]. To detect EGFR mutations, Exons including 18, 19, 20, and 21 were assessed. In addition, Exon 2 positions 12 and 13 were examined with amino acid substitution for missense KRAS mutations.

Image Preprocessing Prior to feature extraction, all images were preprocessed with interpolation (sitkBSpline algorithm, B-spline of order 3 interpolation) to isotropic voxel spacing to be rotationally invariant, for feature extraction to allow comparison between image data from different samples and scanner and center. Re-sampling to $1 \times 1 \times 1$, $1 \times 1 \times 1$, and $3 \times 3 \times 3$ mm³ was performed for CTD, CT, and PET images respectively. Laplacian of Gaussian (LOG), wavelet decomposition (WAV), and

Fig. 1. The radiogenomics framework employed in the present study.

discretized into 64 bins (BIN64) preprocessing were performed to generate different set of features. For LOG filter, different sigma values were used to extract fine, medium, and coarse features; specifically, they ranged

from 0.5 to 5 with 0.5 steps [34]. Wavelet filtering yields eight decompositions per level (all possible combinations of applying either a high- or a low-pass filter in each of the three dimensions, including HHH, HHL, HLH, HLL,

Table 1. Clinical characteristics of patients

Characteristic		PET	CT and CTD	Validation PET, CT, CTD
Patient no.		82	82	68
Age, mean (SD)		68.6 (9.11)	69.2 (8.85)	69.1 (8.8)
Weight (Kg), mean (SD)		77.0 (18.5)	77.8 (18.6)	77.6 (18.6)
Sex (%)	Male	66.7	66.3	67.1
	Female	33.3	33.7	32.9
Smoking status (%)	Nonsmoker	21	21.7	21.5
	Former	61.3	61.7	61.2
	Current	17.7	16.6	17.3
Histology (%)	Adenocarcinoma	79.6	82.3	81.4
	Squamous cell carcinoma	18.3	16	16.8
	NOS (not otherwise specified)	2.1	1.7	1.8
T stage (%)	T1	40.6	43.65	43.95
	T2	35.1	33.45	33.31
	T3	15.8	14.55	14.35
	T4	8.5	8.35	8.39
N stage (%)	N0	71.87	70.07	71.07
	N1	13.27	13.57	12.97
	N2	14.86	16.36	15.96
M stage (%)	M0	92.3	88.3	91.2
	M1	7.7	11.7	8.8
AJCC Staging (%)	I	53.86	55.30	55.81
	II	23.28	21.00	21.13
	III	16.78	17.00	16.93
	IV	6.08	6.70	6.13
Histopathological Grade (%)	G1	19.50	19.90	19.31
	G2	43.70	43.90	44.43
	G3	21.60	20.40	20.53
	other	15.20	15.80	15.73
Lymph vascular invasion (%)	Absent	78.50	78.55	79.35
	Present	21.50	21.45	20.65
EGFR mutation status (%)	Mutant	22.50	25.10	24.54
	Wild type	77.50	74.90	75.46
KRAS mutation status (%)	Mutant	23.80	24.60	22.75
	Wild type	76.20	75.40	77.25
PET SUV	SUVmax	7.5 (6.7)		7.1 (6.3)
	SULmax	6.1 (5.2)		5.6 (4.9)
	SUVpeak	9.6 (6.5)		8.9 (6.3)
	SULpeak	7.4 (4.9)		6.8 (4.8)
Volume (cc) mean (SD)		19.15 (48.23)		19.25 (40.18)
Recurrence (%)		17.7	17.7	17.4
Time to last follow-up, mean (SD)		26.9 (21.7)	27.0 (22.1)	26.8 (22.2)
Death (%)		21	20	20.4

LHH, LHL, LLH, and LLL). Preprocessing steps (including discretization, LOG, and wavelet) were performed on all intensity, histogram, and textural features.

Image Segmentation All PET image segmentations were performed using OSIRIX® [35]. Lesions were delineated manually on PET images. CT images were segmented via automatic region growing using 3D-Slicer [36] and were edited and verified by an experienced radiologist. In total, we provide the analysis of 150 PET, low-dose CT, and CTD segmentations.

Feature Extraction In the next step, several features from different feature classes were extracted. The classes include first-order statistics (19 FOS features), shape-based (16 Shape features), gray-level co-occurrence matrix (GLCM 23

features), gray-level run length matrix (16 GLRLM features), gray-level size zone matrix (16 GLSZM features), neighboring gray tone difference matrix (5 NGTDM features), and gray-level dependence matrix (14 GLDM features). Image feature extraction was performed using the open-source python library PyRadiomics [37]. Radiomic features calculated by this package are in compliance with feature definitions described by the image biomarker standardization initiative (IBSI) that ensures the harmonization and reproducibility of calculated radiomic features, and thus facilitates the reproducibility of this study [37, 38]. More details about the feature classes are provided in Table 2.

In addition to the radiomic features, we also extracted conventional clinical PET biomarkers including metabolic tumor volume (MTV) and standard uptake values (SUV_{max} , SUV_{peak}), including normalization to lean body mass (SUL_{max} , SUL_{peak}).

Fig. 2. The flow chart of inclusion and exclusion criteria of patient to this study.

Feature Selection We implemented six different feature selection methods in our framework and compared their performances (Table 3). These feature selection techniques, including filter-based method (select K Best (SKB), variance threshold (VT), and select percentile (SP)), embedded-based method (select from Model (SM)), and combine methods (variance threshold and select from model (VT-SM) and variance threshold and select K best (VT-KB)).

Classification (Multivariate) We implemented and compared 12 classifiers (Table 4, details of each classifier are provided in supplemental Table 1). Different methods of classification, including generalized linear models (logistic regression and stochastic gradient descent (SGD)), naive Bayes models (naive Bayes and Gaussian naive Bayes), nearest neighbor's model (k-nearest neighbors), decision trees model (C5.0), quadratic discriminant analysis model (QDA), support vector machines (SVC), supervised neural network models (multilayer perceptron), and ensemble learning methods (adaptive boost, bagging and random forest) were used in this study.

Analysis

Univariate Analysis

For univariate analysis, each feature value was normalized to obtain Z-scores followed by student t test statistical analysis for comparison. A p value of < 0.05 was used as a criterion for statistically significant differences. Spearman correlation between features was performed to eliminate highly correlated features while the false discovery rate (FDR) correction was assessed and reported along with the q -value (FDR adjusted p value). Statistical analysis was performed using R 3.5.1 software (using “pROC” and “stats” packages). We also studied the prediction performance of conventional clinical image-derived PET metrics (SUV_{max} , SUV_{peak} , SUL_{max} , SUL_{peak} , and MTV).

Multivariate Analysis

All our analyses, including feature selection and classification, were performed using an in-house developed python framework in open-source python library Scikit-Learn [39]. Cross-validation

Table 2. Radiomic features extracted from the different imaging modalities

First-order statistics (FOS)	Gray-level co-occurrence matrix (GLCM)	Gray-level run length matrix (GLRLM)
1. Energy 2. Total energy 3. Entropy 4. Minimum 5. 10th percentile 6. 90th percentile 7. Maximum 8. Mean 9. Median 10. Interquartile range 11. Range 12. Mean absolute deviation (MAD) 13. Robust mean absolute deviation (rMAD) 14. Root mean squared (RMS) 15. Standard deviation 16. Skewness 17. Kurtosis 18. Variance 19. Uniformity	1. Autocorrelation 2. Joint average 3. Cluster prominence 4. Cluster shade 5. Cluster tendency 6. Contrast 7. Correlation 8. Difference average 9. Difference entropy 10. Difference variance 11. Joint energy 12. Joint entropy 13. Informal measure of correlation (IMC) 1 14. Informal measure of correlation (IMC) 2 15. Inverse difference moment (IDM) 16. Inverse Difference Moment Normalized (IDMN) 17. Inverse difference (ID) 18. Inverse Difference Normalized (IDN) 19. Inverse variance 20. Maximum probability 21. Sum average 22. Sum entropy 23. Sum of squares Gray-Level Size Zone Matrix (GLSZM) 1. Small Area Emphasis (SAE) 2. Large Area Emphasis (LAE) 3. Gray-Level Non-Uniformity (GLN) 4. Gray-Level Non-Uniformity Normalized (GLNN) 5. Size-Zone Non-Uniformity (SZN) 6. Size-Zone Non-Uniformity Normalized (SZNN) 7. Zone Percentage (ZP) 8. Gray-level variance (GLV) 9. Zone variance (ZV) 10. Zone entropy (ZE) 11. Low Gray-Level Zone Emphasis (LGLZE) 12. High Gray-Level Zone Emphasis (HGLZE) 13. Small Area Low Gray-Level Emphasis (SALGLE) 14. Small Area High Gray-Level Emphasis (SAHGLE) 15. Large Area Low Gray-Level Emphasis (LALGLE) 16. Large Area High Gray-Level Emphasis (LAHGLE)	1. Short-run emphasis (SRE) 2. Long-run emphasis (LRE) 3. Gray-Level Non-Uniformity (GLN) 4. Gray-Level Non-Uniformity Normalized (GLNN) 5. Run Length Non-Uniformity (RLN) 6. Run Length Non-Uniformity Normalized (RLNN) 7. Run percentage (RP) 8. Gray-level variance (GLV) 9. Run variance (RV) 10. Run entropy (RE) 11. Low Gray-Level Run Emphasis (LGLRE) 12. High Gray-Level Run Emphasis (HGLRE) 13. Short-Run Low Gray-Level Emphasis (SRLGLE) 14. Short-Run High Gray-Level Emphasis (SRHGLE) 15. Long-Run Low Gray-Level Emphasis (LRLGLE) 16. Long-Run High Gray-Level Emphasis (LRHGLE) Gray-level dependence matrix (GLDM) 1. Small Dependence Emphasis (SDE) 2. Large Dependence Emphasis (LDE) 3. Gray-Level Non-Uniformity (GLN) 4. Dependence Non-Uniformity (DN) 5. Dependence Non-Uniformity Normalized (DNN) 6. Gray-level variance (GLV) 7. Dependence variance (DV) 8. Dependence entropy (DE) 9. Low Gray-Level Emphasis (LGLE) 10. High Gray-Level Emphasis (HGLE) 11. Small Dependence Low Gray-Level Emphasis (SDLGLE) 12. Small Dependence High Gray-Level Emphasis (SDHGLE) 13. Large Dependence Low Gray-Level Emphasis (LDLGLE) 14. Large Dependence High Gray-Level Emphasis (LDHGLE) Neighboring gray tone difference matrix (NGTDM) 1. Coarseness 2. Contrast 3. Busyness 4. Complexity 5. Strength
Shape and morphological features 1. Volume 2. Surface area 3. Surface area to volume ratio 4. Sphericity 5. Compactness 1 6. Compactness 2 7. Spherical disproportion 8. Maximum 3D diameter 9. Maximum 2D diameter (slice) 10. Maximum 2D diameter (column) 11. Maximum 2D diameter (row) 12. Major axis 13. Minor axis 14. Least axis 15. Elongation 16. Flatness		

(CV) was applied to models, and tuning of models was performed on testing set of 10-fold CV (sample is randomly partitioned into 10 equal size subsamples. Of the 10 subsamples, a single subsample is retained as the test data, while the remaining nine subsamples are used as training data, repeated for 10 times). This process was performed 20 times to get the highest stable results. Furthermore, model evaluation was performed on independent validation sets (68 patients). The predictive power of all features was investigated using the area under the receiver operator characteristic (ROC) curve (AUC). A heatmap was generated to compare the different models. Mean $\pm$ standard deviation (SD) of

AUC of all classifier and feature selection algorithms were calculated and depicted. The cross-combination of feature selection and classification methods were depicted as a heatmap (mean AUC value in cross-validation).

Results

Univariate Analysis

EGFR The results of EGFR mutation status prediction after elimination of high correlated features in each image

Table 3. Feature selector algorithm characteristics

Feature selection method	Feature selection model	Abbreviation	Algorithm of action
Filter-based method	Select K best	SKB	This method scores the features using a chi square function and then “removes all but the k highest scoring features.
	Variance threshold	VT	It removes all features whose variance does not meet some threshold. By default, it removes all zero-variance features, <i>i.e.</i> , features that have the same value in all sample. In this method, we used 0.8 threshold.
	Select percentile	SP	It removes all but a user-specified highest scoring percentage of features using common univariate statistical tests for each feature: false positive rate or family wise error
Embedded-based method	Select from model	SM	It is a meta-transformer that can be used along with any estimator that has a coefficient or feature importance’s attribute after fitting. The features are considered unimportant and removed, if the corresponding coefficient or feature importance’s values are below the provided threshold parameter. Apart from specifying the threshold numerically, there are built-in heuristics for finding a threshold using a string argument. Available heuristics are “mean”, “median” and float multiples of these like “0.1*mean”.
Combine methods	Variance threshold and select from model	VT-SM	Combine of variance threshold and select from model
	Variance threshold and select K best	VT-KB	Combine of variance threshold and select K best

set and FDR correction are presented in supplemental Table 2 (only q-value < 0.05). About 90 features from all image datasets had significant correlation with EGFR mutation status (q-value < 0.05) with mean AUC of 0.67 ± 0.05 (0.52–0.75). The highest prediction performance was achieved by Short-Run Emphasis feature of GLRLM (AUC: 0.75, q-value: 0.003) and Run Variance feature of GLRLM (AUC: 0.75, q-value: 0.016) from LOG preprocessed PET image with sigma values of 1.5 and 1, respectively. In addition, Small Area Emphasis feature and Size Zone Non Uniformity Normalized from GLSZM of CTD dataset which preprocessed by LLL of wavelet and LOG with sigma 2, respectively (AUC 0.75, q-value 0.001). More detail of univariate analysis in EGFR mutation status prediction including AUC, *p* value, and q-value are presented in supplementary figures 1–3, respectively.

KRAS The results of KRAS mutation status prediction after elimination of highly correlated features in each image sets and FDR correction are presented in supplemental Table 3 (only q-values < 0.05). Fourteen features from all image datasets were significantly correlated with KRAS mutation status (q-value < 0.05) with mean AUC of 0.67 ± 0.02 (0.53–0.71). The Large Dependence Low Gray-Level Emphasis from GLDM in LOG (sigma: 5) preprocessed image of CTD with AUC 0.71 (q-value 0.00005) had highest performances. In addition, the Minimum feature from first order of LOG preprocessed in sigma 4 (AUC: 0.71, q-value 0.013), 4.5 (AUC 0.71, q-value 0.01), and 5 (AUC 0.70, q-value 0.014) showed high performances in KRAS mutation status prediction. More details about univariate analysis in KRAS mutation status

prediction, including AUC, *p* value and q-value were presented in supplementary figures 4–6, respectively.

Shape Features of EGFR and KRAS

Details of EGFR and KRAS mutation status prediction using shape features, including AUC, *p* value, and q-value, are presented in supplementary figures 7–9, respectively. The Surface Volume Ratio feature from PET images (AUC 0.60, q-value 0.28) and Flatness from CT images (AUC: 0.67, q-value: 0.11) had highest prediction power in EGFR and KRAS, respectively; however, the differences were not statistically significant.

Conventional PET Metrics

Our univariate analysis on conventional PET metrics for EGFR mutation status prediction showed that SUV_{peak} with AUC = 0.69 (*p* value = 0.0002) had the highest performance. The results for SUL_{peak} , SUV_{max} , SUL_{max} , and MTV were 0.60 (*p* value = 0.0003), 0.59 (*p* value < 0.0001), 0.55 (*p* value < 0.0001), and 0.56 (*p* value = 0.0004), respectively. Our results involving conventional features for KRAS mutation status prediction were all poor: MTV (AUC 0.55, *p* value = 0.0011), SUV_{max} (AUC 0.52, *p* value < 0.0001), SUL_{max} (AUC 0.53, *p* value < 0.0001), SUV_{peak} (AUC 0.52, *p* value = 0.0013), and SUL_{peak} (AUC 0.51, *p* value = 0.0008).

Multivariate Machine Learning Radiomic Models

EGFR Figure 3 shows a heatmap of EGFR mutation status prediction results with different combinations of feature

Table 4. Classifier methods and abbreviations

Machine learning methods	Machine learning algorithm	Abbreviation	Algorithm of action
Support vector classifier	Support vector machine	SVM	Data items are plotted as a point in n-dimensional space (n is number of features) with the value of each feature being the value of a particular coordinate and classification is performed by finding the hyper-plane that differentiate the two classes.
K-nearest neighbors	K-nearest neighbors	KNN	As a non-parametric, lazy learning algorithm, KNN algorithm is based on feature similarity and an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k-nearest neighbors.
Decision tree	Decision Tree (C 5.0)	DT	As a decision support tool, uses a tree-like graph of decisions and their possible consequences, including chance event outcomes, resource costs, and utility, builds a model of decisions.
Quadratic discriminant analysis	Quadratic discriminant analysis	QDA	QDA separate measurements of two classes of objects with a quadratic decision boundary, generated by fitting class conditional densities to the data. It is a more general version of the linear classifier.
Supervised neural network models	Multilayer perceptron	MLP	MLP is a class of feed forward artificial neural network which optimizes the log-loss function using stochastic gradient descent (SGD).
Generalized linear methods	Stochastic gradient descent	SGD	Works based on the differentiable objective function optimization which is a stochastic approximation of gradient descent optimization. This classifier implements a plain SGD learning routine which supports different loss functions and penalties for classification.
	Logistic regression	LREG	The probabilities describing the possible outcomes of a single trial are modeled using a logistic function.
Naive Bayes methods	Naive Bayes	NB	As a probabilistic classifier, works based on applying Bayes' theorem with strong (naive) independence assumptions between the features.
	Gaussian naive Bayes	GNB	It is a generalization of Naive Bayes Networks, which are a special case of probabilistic networks that allows treating continuous variables. This method computes conditional class probabilities and then predicts the most probable class of a vector of training data X, according to sample data D.
Ensemble learning methods	Random forest	RF	It is an ensemble learning approach which is essentially a collection of decision trees, where each tree is slightly different from the others. This algorithm reduces the amount of overfitting by averaging the result of all decision trees.
	Adaptive boosting	AB	This ensemble learning method, consists of very simple base classifiers (weak learners). This approach focuses on the training samples that are hard to classify, that is, to let the weak learners subsequently learn from misclassified training samples to improve the performance of the ensemble.
	Bagging	BAG	BAG fits base classifiers each on random subsets of the original dataset and then aggregate their individual predictions (by voting) to form a final prediction. The base estimator to fit on random subsets of the dataset is a decision tree model.

selection, classifiers, and image sets. Here, we observe a wide range of performance from 0.5 to 0.82. We further observe that the combination of LOG preprocessing of PET image sets with a sigma of 3.5 (PET_LOG_3.5S) with VT feature selector and SGD classifier (PET_LOG_3.5S + VT + SGD, AUC 0.82) had the highest predictive performance, followed by PET image sets preprocessed by LHH of wavelet (PET_W_LHH) with VT_SM feature selector and MLP classifier (PET_W_LHH + VT_SM + MLP, AUC 0.81), LOG preprocessing of PET image sets with a sigma of 4.0 (PET_LOG_4.0S) with SM feature selector and SGD classifier (PET_LOG_4.0S + SM + SGD, AUC 0.81), PET_W_LHH image set with SM feature selector and SGD classifier (PET_W_LHH + SM + SGD, AUC 0.80).

Our results for EGFR mutation status prediction based on feature selection methods and image sets are depicted in supplemental Tables 4 and 5 (mean $\pm$ SD and Min-Max). It can be deduced from these results that selection performance ranges from 0.51 to 0.82, and the combination of VT_SM

feature selection with PET_W_LHH image sets had the highest performance (AUC 0.73 ± 0.074 and 0.58–0.81), followed by SM feature selection with PET_W_LHH image set (AUC 0.72 ± 0.079 and 0.5–0.8), VT feature selection with PET_W_LHH image set (AUC: 0.71 ± 0.065 and 0.6–0.79), and SKB feature selection with PET_W_LHH image set (AUC 0.7 ± 0.06 and 0.6–0.78). Figures 4 and 5a indicate the box plot of EGFR mutation status prediction based on feature selection methods.

Supplemental Tables 6 and 7 (mean $\pm$ SD and Min-Max) show the results regarding EGFR mutation status prediction based on classifiers and image sets. According to these results, classifier performance has a range from 0.50 to 0.82 and combination of RF classifier with PET_W_LHH image set had the highest performance (AUC: 0.76 ± 0.024 and 0.73–0.79), followed by BNB classifier with PET image set which discretized into 64 Bin (PET_BIN64) (AUC 0.75 ± 0.0096 and 0.74–0.76) and SVM classifier with PET_W_LHH image set (AUC 0.75 ± 0.051 and 0.65–

Fig. 5. a AUC box plot of different feature selector in independent validation set of EGFR. **b** AUC box plot of different classifier in independent validation set of EGFR.

0.78) and AB classifier with PET_BIN64 image set (AUC 0.74 ± 0.018 and $0.72-0.77$). Figure 5 b shows the box plot of EGFR mutation status prediction based on classifier methods.

KRAS Figure 4 shows a heatmap of KRAS mutation status prediction for different combinations of feature selection, classifier, and image sets. The model performance has a wide variability, ranging from 0.5 to 0.83, similar to EGFR models. According to this figure, combination methods with the highest predictive performance are: LOG preprocessing of CT image sets with a sigma of 3.5 with SM feature selector and SGD classifier (CT_LOG_3.5S + SM + SGD, AUC: 0.83) and CT_LOG_5 S image sets with VT_SM feature selector SGD classifier (CT_LOG_5 S + VT_SM +

SGD, AUC: 0.82). These models are followed by CTD_LOG_4 S image set with VT feature selector and SGD classifier (CTD_LOG_4S + VT + SGD, AUC 0.80), PET_LOG_0 S image set with SM feature selector and SGD classifier (PET_LOG_0.5S + SM + SGD, AUC 0.80) and PET image sets preprocessed by HHH of wavelet with SKB feature selector and BNB classifier (PET_W_HHH + SKB + BNB, AUC: 0.80).

Supplemental Tables 8 and 9 show KRAS mutation status prediction results based on feature selection methods and image sets. In these results, feature selection performance ranges from 0.5 to 0.83, and the combination of SM feature selection with PET_LOG_4.5S image set had the highest performance (AUC 0.66 ± 0.093 and $0.51-0.79$), followed by VT_SM feature selection with CT_LOG_3.5S image set (AUC 0.66 ± 0.079 and $0.55-0.8$), SM feature selection with

Fig. 6. a AUC box plot of different feature selector in independent validation set of KRAS. **b** AUC box plot of different classifier in independent validation set of KRAS.

CT_LOG_3.5S image set (AUC 0.65 ± 0.11 and 0.5–0.83), and SM feature selection with PET_BIN64 image set (AUC 0.64 ± 0.058 and 0.51–0.71). Figure 6 a shows the box plot of KRAS mutation status prediction based on feature selection methods.

KRAS mutation status prediction results based on classifier and image sets is presented in supplemental Tables 10 and 11 (mean $\pm$ SD and Min-Max). Here, the classifier performance ranges from 0.5 to 0.83 and the combination of SVM classifier with CT_LOG_3.5S image set had the highest performance (AUC 0.79 ± 0.021 and 0.76–0.8), followed by BNB classifier PET_W_HHH image set (AUC 0.75 ± 0.1 and 0.6–0.81) and SGD classifier with CTD_LOG_4 S image set (AUC: 0.71 ± 0.08 and 0.56–0.81). Figure 6 b indicates the box plot of KRAS mutation status prediction based on feature selection methods.

Furthermore, supplemental Table 12 compares (q-value, *t* test) on predictive performance (AUC) of univariate vs. multivariate analysis approaches for EGFR and KRAS mutation status for different image sets. The data show that multivariate analysis significantly improved prediction power against univariate analysis, underscoring the importance of utilizing machine learning methods for improved performance.

Discussion

Accurate detection of EGFR/KRAS mutation status in NSCLCs allows improved selection of patients for effective therapeutic strategies [3]. Although bio-techniques provide intra- and inter-tumor EGFR/KRAS mutation status in

patients, they are not entirely feasible and suffer from limitations [40]. In the present study, we developed a comprehensive radiomics framework that calculates radiomic features from low-dose CT, diagnostic CT, and PET images to predict EGFR and KRAS mutation status using univariate and multivariate ML algorithms in NSCLCs patients. In our radiogenomics study, several radiomics models were found predictive for EGFR and KRAS mutation status, and as such are suitable for non-invasive cost-effective assessment.

In recent years, the field of radiogenomics—correlation of imaging features with genomic parameters—has garnered significant interest, as a non-invasive assessment framework to tailor therapy based on imaging biomarkers and underlying biological pathways [8, 41, 42]. In our work, we observed a wide range of predictive model performances for different radiomics features, feature selection, and classification algorithms and for different imaging modalities. For EGFR and KRAS mutation status prediction, univariate analysis of radiomic features and cross-combination of imaging/validation /feature selection/classification as multivariate analysis methods resulted in a wide range of performance (AUC 0.5~0.83). The best predictive power of conventional PET parameters was achieved by SUV_{peak} (AUC: 0.69, p value = 0.0002) and MTV (AUC 0.55, p value = 0.0011) for EGFR and KRAS. Univariate analysis of extracted radiomic features improved prediction power up to AUC 0.75 (q -value 0.003, Short-Run Emphasis feature of GLRLM from LOG preprocessed PET images with a sigma value of 1.5) and AUC 0.71 (q -value 0.00005, the Large Dependence Low Gray-Level Emphasis from GLDM in LOG preprocessed CTD images with a sigma value of 5) for EGFR and KRAS, respectively. Furthermore, the machine learning algorithm improved the prediction power of gene mutation status up to AUC 0.82 for EGFR (LOG preprocessed PET image set with a sigma of 3 with VT feature selector and SGD classifier) and AUC 0.83 for KRAS (CT image set with a sigma of 3.5 with SM feature selector and SGD classifier). Using radiomics feature and machine learning algorithm outperformed conventional methods in prediction of EGFR and KRAS gene mutation status in NSCLC patients.

The role of feature selection is more prominent in cross-combination of feature selection and image sets for EGFR mutation status prediction. Random forest (RF) and support vector machine (SVM) methods proved to yield a higher performance in EGFR and KRAS prediction. For classification, the selection of an appropriate algorithm has a more prominent role. Zhao et al. [19] developed a conventional radiomics model and 3D deep learning system to predict EGFR-mutant pulmonary adenocarcinoma from CT images, the best-reported AUC was 75.8 using deep learning. However, in the present study, the EGFR status achieved an AUC 0.82 which shows the role of combining different image sets, feature selector, and classifier. Parmer et al. [24] compared 14 feature selection and 12 classification methods

in terms of performance and stability and found that the Wilcoxon rank-sum test and random forest (RF) had the highest prognostic performance for feature selection and classification, respectively, while demonstrating high stability against data perturbation. Moreover, Zhang et al. [43] built a 54-cross-combination ML algorithm framework including six feature selections and nine classification methods for survival prediction of advanced nasopharyngeal carcinoma. They found that using RF for both feature selection and classification had the highest prognostic performance, followed by RF + AB and sure independence screening (SIS) + linear support vector machines (SVMs). Abdollahi et al. [44] studied different combinations of feature selector and classifier for prediction of intensity-modulated radiation therapy response, Gleason score, and stage in prostate cancer and showed that different tasks of classification can be achieved by different combinations of feature selector and classifiers. In addition, we compared the predictive power of conventional PET parameters and radiomic features. In our radiogenomics study, we showed that radiomics features and machine learning-based radiomic models are more predictive than conventional parameters. Du et al. [45] provided a cross-combination of six features selection and seven classifiers for PET/CT radiomics-based differentiation between recurrence and inflammation in nasopharyngeal carcinoma. They reported the most accurate and reliable machine learning methods through the different combinations of feature selectors and classifier methods. They suggested that parameter tuning can result in higher performance; however, at the same time, it may cause overfitting. In our current study, we have fine-tuned the model hyperparameters and tested them on external validation set to prevent overfitting. Wang et al. [19] and Zhao et al. [20] recently performed a radiomics study using deep learning for EGFR prediction using CT images. In the radiomic study reported by Wang et al. [19], radiomic features extraction was performed using PyRadiomics with eight features selected using recursive feature elimination imported into a random forest classifier. They reported AUCs of 0.64 and 0.81 on external validation sets for radiomics and deep learning models. In Zhao et al. [20] study, 401 features were selected using an Intra Class Correlation (ICC) test that were subsequently input to logistic regression to construct the radiomics model. They reported AUCs of 0.68 and 0.75 for radiomics and deep learning models. Tu et al. [46] more recently developed a radiomics model based on CT image features and selected the features based on an ICC test and then predicted EGFR mutation using logistic regression models. They reported an AUC of 0.76 for their radiomics signature and an AUC of 0.798 for their integrated model (radiomic + clinical). In comparison with their imaging modality and machine learning algorithms, our current study resulted in 0.75 and 0.78 as the highest AUCs using low-dose CT and CTD images, SM and SKB as features selector, and DT and SGD classifiers, respectively. The random forest and logistics

regression classifier resulted in 0.71 and 0.75 AUCs in our study using CT images. An AUC of 0.82 was achieved *via* LOG preprocessed PET images, the VT feature selector, and the SGD classifier. Owing to the lack of one-fits-all models in machine learning, different combinations of imaging modalities, feature selectors, and classifiers can provide different results. In the current study, we considered a wide variety of imaging modalities, feature selectors, and classifiers to provide a high-performance model for EGFR and KRAS mutation prediction.

Rizzo *et al.* [47] developed a radiogenomic model based on CT radiological hand-crafted features to predict EGFR and KRAS mutations in NSCLC patients. In their study, 20 radiological features were extracted by two radiologists, with the analysis performed on an external validation set. They reported AUCs of 0.82 and 0.60 for prediction of EGFR and KRAS, respectively. A recent study performed by Li *et al.* [21] reported EGFR mutation status using conventional PET/CT SUV and radiomic features. For conventional PET SUV features, they reported AUCs of 0.621 (p value = 0.026), 0.624 (p value = 0.023), 0.615 (p value = 0.035), and 0.597 (p value = 0.074) for SUV_{max} , SUV_{mean} , SUV_{peak} , and TLG, respectively. In their radiomics analysis using the LASSO feature selector and an ensemble classifier, they reported AUCs of 0.667, 0.789, and 0.805 for CT, PET, and PET/CT images radiomic model and 0.686, 0.774, and 0.822 for CT, PET, and PET/CT radiomics + clinical model, respectively. In our current study for EGFR prediction, the conventional SUV PET features provided AUCs of 0.69 (p value = 0.0002), 0.60 (p value = 0.0003), 0.59 (p value < 0.0001), 0.55 (p value < 0.0001), and 0.56 (p value = 0.0004) for SUV_{peak} , SUL_{peak} , SUV_{max} , SUL_{max} , and MTV, respectively. For ensemble learning in our current study, the combination of VT feature selector with AB classifier in LOG preprocessed CTD images results in an AUC of 0.73. Combination of SM feature selector with AB classifier in LLL wavelet preprocessed low-dose CT image results in an AUC of 0.73, while the combination of VT feature selector with RF classifier in LHH wavelet preprocessed PET images resulted in AUC of 0.79. The best radiomics model for the prediction of EGFR and KRAS was achieved *via* the PET LOG preprocessing method, the VT features selector, the SGD classifier (PET_LOG_3.5S + VT + SGD, AUC 0.82), and the LOG preprocessing of CT images, the SM feature selector, and the SGD model (CT_LOG_3.5S + SM + SGD, AUC 0.83), respectively.

Pinheiro *et al.* [22] reported on a radiomics study on the same public datasets used in our current study for EGFR and KRAS prediction using only CT image features (radiomics and semantic radiological features) on a 80 %/20 % training and test setup (116 patients for EGFR and 114 for KRAS). They extracted wavelet and LOG preprocessed features using PyRadiomics from 3D segmented CT images, and subsequently used principal component analysis (PCA) followed by t-Distributed Stochastic Neighbor Embedding (t-SNE) to reduce the extracted features. Subsequently, an

Extreme Gradient Boosting (XGBoost) classifier was used to classify KRAS and EGFR gene mutation status. They reported AUCs of 0.57 ± 0.12 and 0.50 ± 0.01 using radiomics model and 0.74 ± 0.08 and 0.50 ± 0.07 for semantic features for EGFR and KRAS prediction, respectively. In our current study, the CT radiomics feature model arrived at AUCs of 0.78 and 0.83 by SGD_SM and SGD_SM models for EGFR and KRAS prediction in 68 external validation sets, respectively.

Radiomics/genomics studies suffer from several challenges, and a robust framework for clinical decision making is highly desired [41, 48, 49]. As an approved guideline, standardization efforts (IBSI in particular) [38] have sought to address the challenge of reproducing and validating reported findings by comparing and standardizing definitions and implementation of several image feature sets between participating institutions. The IBSI also provides image biomarker nomenclature and definitions, benchmark datasets, and benchmark values to verify image processing and image biomarker calculations, as well as reporting guidelines, for high-throughput image analysis [38]. In this study, we followed the IBSI protocol to resample all images to a constant voxel size to reduce feature sensitivity to variable image generation parameters [38], also ensuring the robustness and reproducibility of the study using publicly available datasets and open-source feature extraction and machine learning library [50]. Having an image dataset generated using varying acquisition [51, 52], reconstruction [53], processing, and segmentation [54] results produce inconsistencies in feature evaluation. Future studies are required to assess the effect of other sources of variation in scans, including but not limited to post-injection scan time window, administered dosage, reconstruction parameters (different inherent spatial resolution, PSF kernel, iterations, subsets, time-of-flight (TOF) setting, *etc.*). Combining the radiomic features and clinical information would also be suggested to investigate the possibility of predicted model improvement. The significance of our study is in addressing convolutional biomarkers *vs.* radiomics analysis with various ML algorithms on three different imaging modalities.

The limited size of a dataset is a limiting factor in radiomic studies, as the analysis with many radiomic features and only a few imaging data is prone to overfitting. We tuned our models using 10-fold cross-validation to reduce the sensitivity of our results to input data and repeated for 20 times to make our results more reliable, and all model evaluations were performed on 68 patient's independent validation set which was not used in the training process. Our presented results were developed using a standardized radiomics analysis workflow that showed strong and significant predictability of gene mutations. Specifically, in the feature selection phase, we used ML algorithms to reduce the number of features to decrease the dimensionality and reduce overfitting.

Conclusion

We demonstrated that prediction of EGFR and KRAS mutation status in NSCLC patients can be significantly improved *via* robust and systematic radiomics analysis. We evaluated different machine learning methods to find optimal methods for radiogenomics analyses. Overall, we showed that radiomic features extracted from different imaging modalities can be used for effective prediction of EGFR and KRAS mutation status, having significantly higher predictive power relative to conventional imaging biomarkers.

Funding Information This study was funded by “Rajaie Cardiovascular Medical and Research Centre” and the Swiss National Science Foundation under Grant SNSF 320030_176052.

Compliance with Ethical Standards

Conflict of Interest

The authors declare that they have no conflict of interest.

Ethical Approval

For this type of study formal consent is not required.

Informed Consent

For this type of study formal consent is not required.

References

- Roberts PJ, Stinchcombe TE, Der CJ, Socinski MA (2010) Personalized medicine in non-small-cell lung cancer: is KRAS a useful marker in selecting patients for epidermal growth factor receptor-targeted therapy? *J Clin Oncol* 28:4769–4777
- Ludovini V, Bianconi F, Pistola L, Chiari R, Minotti V, Colella R, Giuffrida D, Tofanetti FR, Siggillino A, Flacco A, Baldelli E, Iacono D, Mameli MG, Cavaliere A, Crinò L (2011) Phosphoinositide-3-kinase catalytic alpha and KRAS mutations are important predictors of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in patients with advanced non-small cell lung cancer. *J Thorac Oncol* 6:707–715
- Linardou H, Dahabreh IJ, Kanaklopiti D, Siannis F, Bafaloukos D, Kosmidis P, Papadimitriou CA, Murray S (2008) Assessment of somatic k-RAS mutations as a mechanism associated with resistance to EGFR-targeted agents: a systematic review and meta-analysis of studies in advanced non-small-cell lung cancer and metastatic colorectal cancer. *Lancet Oncol* 9:962–972
- Eberhard DA, Johnson BE, Amler LC, Goddard AD, Heldens SL, Herbst RS, Ince WL, Jänne PA, Januario T, Johnson DH, Klein P, Miller VA, Ostland MA, Ramies DA, Sebisanovic D, Stinson JA, Zhang YR, Seshagiri S, Hillan KJ (2005) Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol* 23:5900–5909
- Riely GJ, Marks J, Pao W (2009) KRAS mutations in non-small cell lung cancer. *Ann Am Thorac Soc* 6:201–205
- Mak R, Hermann G, Aerts H et al (2016) Outcomes by EGFR, KRAS and ALK genotype after combined modality therapy for locally advanced non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 96:S156
- Sun X, Xiao Z, Chen G et al (2018) A PET imaging approach for determining EGFR mutation status for improved lung cancer patient management. *Sci Transl Med* 10:eaa8840
- Thawani R, McLane M, Beig N et al (2017) Radiomics and radiogenomics in lung cancer: a review for the clinician. *Lung Cancer* 115:34–41
- Incoronato M, Aiello M, Infante T et al (2017) Radiogenomic analysis of oncological data: a technical survey. *Int J Mol Sci* 18:805
- Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout R, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ (2012) Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 48:441–446
- Gillies RJ, Kinahan PE, Hricak H (2015) Radiomics: images are more than pictures, they are data. *Radiology* 278:563–577
- Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D (2017) Characterization of PET/CT images using texture analysis: the past, the present ... any future? *Eur J Nucl Med Mol Imaging* 44:151–165
- Zinn PO, Majadan B, Sathyan P et al (2011) Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. *PLoS One* 6:e25451
- Kang J, Rancati T, Lee S et al (2018) Machine learning and radiogenomics: lessons learned and future directions. *Front Oncol* 8:228
- Liu Y, Kim J, Balagurunathan Y et al (2016) Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. *Clin Lung Cancer* 17:441–448.e446
- Lin YC, Lin G, Hong JH, Lin YP, Chen FH, Ng SH, Wang CC (2017) Diffusion radiomics analysis of intratumoral heterogeneity in a murine prostate cancer model following radiotherapy: Pixelwise correlation with histology. *J Magn Reson Imaging* 46:483–489
- Velazquez ER, Parmar C, Liu Y et al (2017) Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res* 77:3922–3930
- Zhang L, Chen B, Liu X et al (2018) Quantitative biomarkers for prediction of epidermal growth factor receptor mutation in non-small cell lung cancer. *Transl Oncol* 11:94–101
- Wang S, Shi J, Ye Z et al (2019) Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* 53:1800986
- Zhao W, Yang J, Ni B et al (2019) Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. In: *Cancer medicine*
- Li X, Yin G, Zhang Y, et al. (2019) Predictive power of a radiomic signature based on 18F-FDG PET/CT images for EGFR mutational status in NSCLC. *Frontiers in Oncology* 9
- Pinheiro G, Pereira T, Dias C, et al. (2019) Identifying relationships between imaging phenotypes and lung cancer-related mutation status: EGFR and KRAS. *bioRxiv*:794123
- Lambin P, Leijenaar RT, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE, Walsh S (2017) Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 14:749–762
- Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJ (2015) Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 5:13087
- Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJ (2015) Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. *Front Oncol* 5:272
- Leger S, Zwanenburg A, Pilz K et al (2017) A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep* 7:13206
- Deist TM, Dankers FJ, Valdes G et al (2018) Machine learning algorithms for outcome prediction in (chemo) radiotherapy: an empirical comparison of classifiers. *Med Phys*

28. Hajianfar G, Shiri I, Maleki H, Oveisi N, Haghparast A, Abdollahi H, Oveisi M (2019) Noninvasive O6 methylguanine-DNA methyltransferase status prediction in glioblastoma multiforme cancer using magnetic resonance imaging radiomics features: univariate and multivariate radiogenomics analysis. *World Neurosurg* 132:e140–e161
29. Prior FW, Clark K, Commean P, et al. (2013) TCIA: an information resource to enable open science [abstract]. 1282-1285P
30. Gevaert O, Xu J, Hoang CD et al (2012) Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results. *Radiology* 264:387–396
31. Clark K, Vendt B, Smith K et al (2013) The cancer imaging archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 26:1045–1057
32. Bakr S, Gevaert O, Echegaray S et al (2018) A radiogenomic dataset of non-small cell lung cancer. *Sci Data* 5:180202
33. Asano H, Toyooka S, Tokumo M, Ichimura K, Aoe K, Ito S, Tsukuda K, Ouchida M, Aoe M, Katayama H, Hiraki A, Sugi K, Kiura K, Date H, Shimizu N (2006) Detection of EGFR gene mutation in lung cancer by mutant-enriched polymerase chain reaction assay. *Clin Cancer Res* 12:43–48
34. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K (2012) Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *Eur Radiol* 22:796–802
35. Rosset A, Spadola L, Ratib O (2004) OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging* 17:205–216
36. Pieper S, Halle M, Kikinis R (2004) 3D slicer [abstract]. 632-635P
37. Van Griethuysen JJ, Fedorov A, Parmar C et al (2017) Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 77:e104–e107
38. Zwanenburg A, Leger S, Vallières M, Löck S (2016) Image biomarker standardisation initiative. *arXiv preprint arXiv:161207003*
39. Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: machine learning in python. *J Mach Learn Res* 12:2825–2830
40. Simon R, Roychowdhury S (2013) Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov* 12:358
41. Wu J, Tha KK, Xing L, Li R (2018) Radiomics and radiogenomics for precision radiotherapy. *J Radiat Res* 59:i25–i31
42. Shiri I, Maleki H, Hajianfar G, et al. (2018) PET/CT Radiomic sequencer for prediction of EGFR and KRAS mutation status in NSCLC patients [abstract]. 1-4P
43. Zhang B, He X, Ouyang F et al (2017) Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. *Cancer Lett* 403:21–27
44. Abdollahi H, Mofid B, Shiri I et al (2019) Machine learning-based radiomic models to predict intensity-modulated radiation therapy response, Gleason score and stage in prostate cancer. *La radiologia medica* 124:555–567
45. Du D, Feng H, Lv W et al (2019) Machine learning methods for optimal radiomics-based differentiation between recurrence and inflammation: application to nasopharyngeal carcinoma post-therapy PET/CT images. *Molecular imaging and biology*:1–9
46. Tu W, Sun G, Fan L, Wang Y, Xia Y, Guan Y, Li Q, Zhang D, Liu S, Li Z (2019) Radiomics signature: a potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. *Lung Cancer* 132:28–35
47. Rizzo S, Raimondi S, de Jong EE, van Elmp W, de Piano F, Petrella F, Bagnardi V, Jochems A, Bellomi M, Dingemans AM, Lambin P (2019) Genomics of non-small cell lung cancer (NSCLC): association between CT-based imaging features and EGFR and K-RAS mutations in 122 patients—an external validation. *Eur J Radiol* 110:148–155
48. Reuzé S, Schernberg A, Orlhac F et al (2018) Radiomics in nuclear medicine applied to radiation therapy: methods, pitfalls and challenges. *Int J Radiat Oncol Biol Phys*
49. Aerts HJ (2016) The potential of radiomic-based phenotyping in precision medicine: a review. *JAMA Oncol* 2:1636–1642
50. Pinker K, Shitano F, Sala E, Do RK, Young RJ, Wibmer AG, Hricak H, Sutton EJ, Morris EA (2018) Background, current role, and potential applications of radiogenomics. *J Magn Reson Imaging* 47:604–620
51. Leijenaar RT, Carvalho S, Velazquez ER, van Elmp W, Parmar C, Hoekstra OS, Hoekstra CJ, Boellaard R, Dekker AL, Gillies RJ, Aerts HJ, Lambin P (2013) Stability of FDG-PET Radiomics features: an integrated analysis of test-retest and inter-observer variability. *Acta Oncol* 52:1391–1397
52. Abdollahi H, Shiri I, Heydari M (2019) Medical imaging technologists in radiomics era: an Alice in wonderland problem. *Iran J Public Health* 48:184
53. Shiri I, Rahmim A, Ghaffarian P, Geramifar P, Abdollahi H, Bitarafan-Rajabi A (2017) The impact of image reconstruction settings on 18F-FDG PET radiomic features: multi-scanner phantom and patient studies. *Eur Radiol* 27:4498–4509
54. Lu L, Lv W, Jiang J, Ma J, Feng Q, Rahmim A, Chen W (2016) Robustness of radiomic features in [11 C] choline and [18 F] FDG PET/CT imaging of nasopharyngeal carcinoma: impact of segmentation and discretization. *Mol Imaging Biol* 18:935–945

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