Integrated Systems and Technologies

Cancer Research

Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer

Emmanuel Rios Velazquez[1], Chintan Parmar[1], Ying Liu[2,3], Thibaud P. Coroller[1], Gisele Cruz[4], Olya Stringfield[2], Zhaoxiang Ye[3], Mike Makrigiorgos[1], Fiona Fennessy[4], Raymond H. Mak[1], Robert Gillies[2], John Quackenbush[5,6,7], and Hugo J.W.L. Aerts[1,4,5]

Abstract

Tumors are characterized by somatic mutations that drive biological processes ultimately reflected in tumor phenotype. With regard to radiographic phenotypes, generally unconnected through present understanding to the presence of specific mutations, artificial intelligence methods can automatically quantify phenotypic characters by using predefined, engineered algorithms or automatic deep-learning methods, a process also known as radiomics. Here we demonstrate how imaging phenotypes can be connected to somatic mutations through an integrated analysis of independent datasets of 763 lung adenocarcinoma patients with somatic mutation testing and engineered CT image analytics. We developed radiomic signatures capable of distinguishing between tumor genotypes in a discovery cohort ($n = 353$) and verified them in an independent validation cohort ($n = 352$). All radiomic signatures significantly outperformed conventional radiographic predictors (tumor volume and maximum diameter). We found a radiomic signature related to radiographic heterogeneity that successfully discriminated between $EGFR^+$ and $EGFR^-$ cases (AUC = 0.69). Combining this signature with a clinical model of EGFR status (AUC = 0.70) significantly improved prediction accuracy (AUC = 0.75). The highest performing signature was capable of distinguishing between $EGFR^+$ and $KRAS^+$ tumors (AUC = 0.80) and, when combined with a clinical model (AUC = 0.81), substantially improved its performance (AUC = 0.86). A $KRAS^+/KRAS^-$ radiomic signature also showed significant albeit lower performance (AUC = 0.63) and did not improve the accuracy of a clinical predictor of KRAS status. Our results argue that somatic mutations drive distinct radiographic phenotypes that can be predicted by radiomics. This work has implications for the use of imaging-based biomarkers in the clinic, as applied noninvasively, repeatedly, and at low cost. *Cancer Res; 77(14); 3922–30. ©2017 AACR.*

Introduction

Somatic mutations, alterations in the DNA sequence that can occur during an organism's lifetime, are potential biologic drivers of cancers that, in turn, can accelerate the accumulation of further somatic mutations. It is well known that the location of somatic mutations, for example within specific genes, can influence biological processes involved in the development and progression of tumors, ultimately influencing its phenotype.

With the introduction of genomic profiling in clinical practice, cancer treatment decisions are increasingly based not only on the patient's clinical characteristics and tumor morphology, but also on individual mutational profiles (1, 2). For example, the use of erlotinib and gefitinib, drugs that target specific mutations within the EGFR gene, have resulted in improved outcomes in a subset of lung cancer patients in which those mutations occur (3–5). Although mutational sequencing of biopsies can be informative and has become standard of care in some situations, they typically quantify only a small part of a possibly heterogeneous tumor, and they are often only performed once, that is, prior to initiation of treatment. Furthermore, there are instances in which such screening can be impractical. Repeated tumor sampling, difficult-to-access tissue samples, failure to determine a mutational status due to poor DNA quality, the relative high costs, and long turnaround time can limit the applicability of molecular assays to monitor the cancer progression and its response to treatment (6–8).

Medical imaging is in routine use in oncology for tumor detection, definition of location and extent of disease, treatment planning, and longitudinal response monitoring. Tumor images exhibit strong phenotypic differences between patients (Fig. 1A) that can be used to assess tumor phenotype (including effects of the genotype) and its local microenvironment, and these data can be used in determining potential treatments (9). While imaging cannot replace biopsies, imaging studies can provide additional

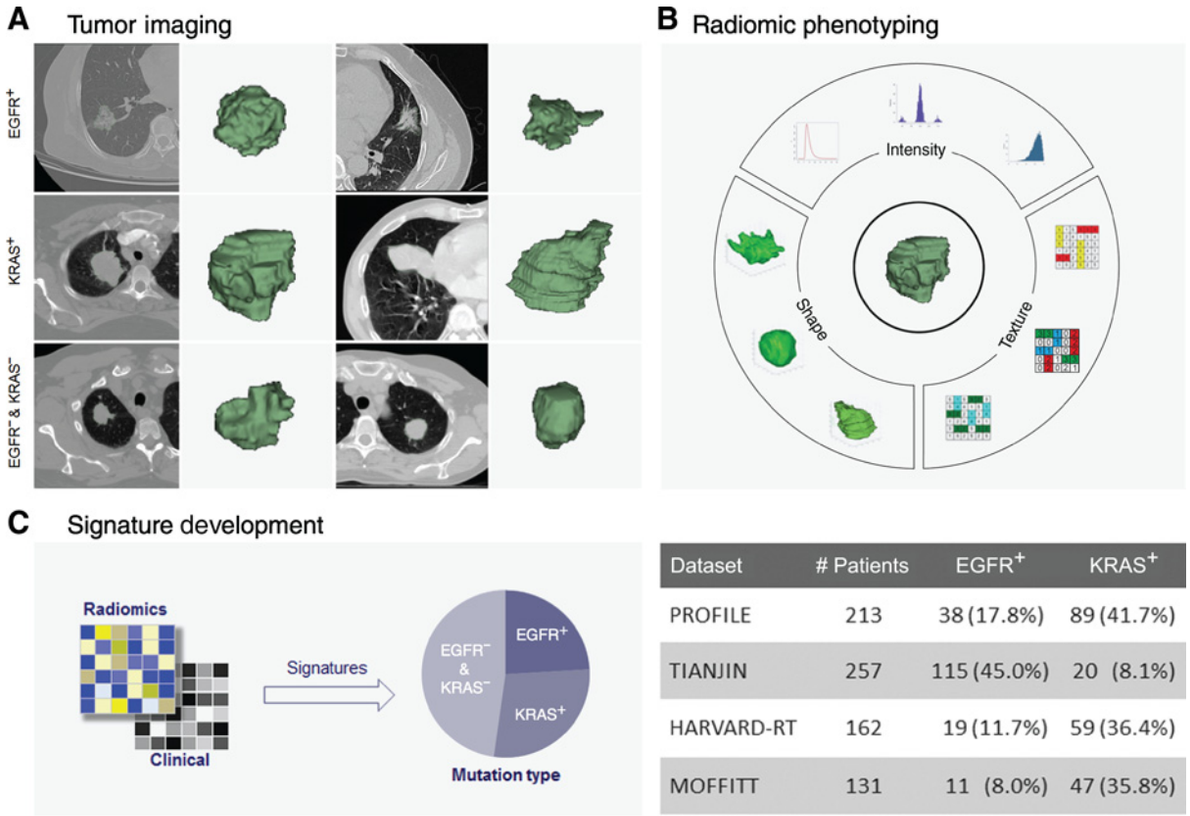
[1]Department of Radiation Oncology Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [2]Departments of Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida. [3]Department of Radiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China. [4]Department of Radiology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [5]Department of Biostatistics & Computational Biology, Dana-Farber Cancer Institute, Boston, Massachusetts. [6]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts. [7]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

E. Rios Velazquez, C. Parmar, and Y. Liu contributed equally to this article.

Corresponding Author: Hugo J.W.L. Aerts, Harvard-DFCI, 450 Brookline Ave, Boston, MA 02115. Phone: 617-525-7156; Fax: 617-582-6037; E-mail: Hugo_Aerts@dfci.harvard.edu

doi: 10.1158/0008-5472.CAN-17-0122



Dataset	# Patients	EGFR+	KRAS+
PROFILE	213	38 (17.8%)	89 (41.7%)
TIANJIN	257	115 (45.0%)	20 (8.1%)
HARVARD-RT	162	19 (11.7%)	59 (36.4%)
MOFFITT	131	11 (8.0%)	47 (35.8%)

Figure 1.
Analysis workflow. **A,** Examples of lung adenocarcinomas tumors imaged with CT imaging (left) and segmented in 3D (right). **B,** Quantification of the tumor phenotype using radiomics feature algorithms. **C,** Radiomic and clinical data were used to develop signatures for EGFR and KRAS mutation status from four independent datasets to investigate associations between the radiomic features and somatic mutations in lung adenocarcinomas. Details on the patient and tumor characteristics of these cohorts are shown in Supplementary Table S1 and Supplementary Data S1.

information that biopsies fail to deliver, such as radiographic intratumor heterogeneity, and can do so throughout the course of treatment (10), providing longitudinal information on disease state, evolution, and response to therapy.

The quantification of phenotypic characteristics on medical imaging has classically been performed by (semi)-qualitative assessment of radiologists, characterizing so-called "semantic" features (10). Although some studies have shown association of these measures with clinical outcomes, their use has been limited, primarily because they require substantially more work and have shown strong inter- and intrareader variation.

Advanced image analysis algorithms, originating from the Artificial Intelligence (AI) domain, make it possible to reproducibly quantify imaging phenotypes automatically by extracting a large number of image features. This process of AI-based phenotyping is referred to as "radiomics" and can provide a far more detailed characterization of the phenotype than would be possible by eye (Fig. 1; refs. 9–13). Radiomic methods are either based on predefined engineered features, relying on expert domain knowledge, or on deep learning methods that automatically can learn feature representations from data. Some of these features capture characteristics that are understandable by human observers and often related to semantic features; others capture agnostic characteristics that are generally higher order and filtered metrics. For example, features can capture statistical (or first order) characteristics that measure simple histogram statistics such as average intensity or distribution asymmetry of the image intensity values. Their values are invariant of the position of the voxels or their relationships within the image and can quantify characteristics related to overall intensity (e.g., tumor density on CT). Other features capture textural variations, which quantify spatial relationships of voxels over the image. For example, these can quantify if voxels have similar values (e.g., related to necrosis), or spatial variations (e.g., related to intratumor heterogeneity). Additional features can describe the overall shape and size of the tumor or other properties of the tumor outline such as elongation, sphericity, and compactness.

Radiomic biomarkers have been shown to be associated with several clinical endpoints, including survival (11, 14–16), nodule malignancy (17, 18), pathologic response (19, 20), recurrence, and distant metastasis (21–23), as well as tumor gene expression patterns (11, 14, 21). A natural extension of this observation is that tumor phenotype should be linked to the tumor genotype. Given that somatic mutations affect the ability of cells to grow in otherwise nonpermissive conditions, we decided to test whether these conditions can be quantified by radiomics and if they reflect the underlying mutational landscape, and whether one could use radiomic phenotype to predict tumor genotype. Although, associations between diagnostic imaging features and mutational data have been explored (24–34), most studies suffer from small

cohort sizes, do not include external validation, or have relied on observer-dependent semiquantitative features that make replication difficult. We hypothesized that automated quantitative radiomic feature extraction, applied to a large, heterogeneous cohort, and rigorously validated, could establish the genotype-imaging phenotype linkage.

In this study, we used a cohort of 763 lung adenocarcinoma patients assembled from four institutions to investigate engineered tumor radiomic features extracted from CT images and tested these features against the most frequently occurring genetic alterations in the disease, that is, EGFR and KRAS mutations. We used a discovery cohort of 353 patients and identified multivariate radiomic signatures specific for EGFR and KRAS mutations. We validated the predictive power of these signatures to identify EGFR and KRAS mutations in an independent cohort of 352 patients. Finally, we combined these signatures with clinical parameters to create integrated predictors that, in general, exhibited improved performance. Our results argue for integration of robust, reproducible radiomic signatures into clinical practice as they can be applied noninvasively and provide additional information that can be used to assess mutational status.

Materials and Methods

Datasets

In this study we used the following four independent lung cancer dataset:

- PROFILE: A total of 213 patients with confirmed lung cancer, stages I–IV, were prospectively included, between June 2011 and June 2013, for mass spectrometry genotyping of 471 known mutations in 41 oncogenes and tumor suppressors (PROFILE OncoMap) as described previously (35, 36). Tumor genomic profiling and clinical data of all patients were retrieved from the Clinical and Operational Research Information System of the Dana-Farber Cancer Institute and Brigham and Women's Hospital. All patients with a biopsy performed in the primary tumor were selected. We excluded cases whose histologic type was not of lung origin.
- TIANJIN: A total of 257 surgical patients with lung adenocarcinoma, stages I–IV, with data on EGFR and KRAS mutations were included in our analysis. For all patients, diagnostic CT imaging, and tumor delineations were available. Data on EGFR and KRAS mutations were routinely clinically collected and extracted from the electronic medical records.
- MOFFITT: A cohort of 131 lung cancer patients treated at the Moffitt cancer center was included in our analysis. Clinical stages ranged from stage I–IV. For these patients, mutation status was determined using mass spectrometry in the *KRAS*, *EGFR*, *TP53*, *STK11* genes.
- HARVARD-RT: A cohort of 162 patients with lung adenocarcinoma, stages I–IIIb, treated with radiation oncology at the Dana-Farber Cancer Institute and Brigham and Women's Hospital. Data on EGFR and KRAS mutations were routinely clinically collected and extracted from the electronic medical records.

The Institutional Review Boards (IRB) of each of the participating centers approved the studies: Profile and Harvard-RT (Dana-Farber/Harvard Cancer Center IRB, Boston, MA), Tianjin (Tianjin Medical University IRB, Tianjin, China) and Moffitt (IRB Moffitt Cancer Center, Tampa, FL). All research was performed according to the International Ethical Guidelines for Biomedical Research Involving Human Subjects (CIOMS). Details on the patient's and tumor characteristics as well as data available are provided on Supplementary Table S1 and Supplementary Fig. S1. Analysis of this study was performed under an institutional review board within the Consented Research Data Repository of the Dana-Farber/Harvard Cancer Center. Written informed consent for all patients in the MOFFITT and TIANJIN datasets was obtained. Informed consent was waived for the PROFILE and HARVARD datasets due to the retrospective nature of our analyses, according to local IRB protocols. A detailed description about the datasets is available in Supplementary Methods S1.

Radiomic quantification

The tumor imaging phenotype was described using a set of quantitative radiomic features extracted from the segmented tumor regions on the CT scans. Briefly, CT images and tumor contours were imported into 3D-Slicer in NRRD format. Because of the differences in pixel spacing and slice thickness, the images and tumor contours were subsequently normalized to isometric voxels (3 mm) using a cubic interpolation. Next, feature extraction was performed using an in-house developed Radiomics plug-in for 3D-Slicer. All features have been described in detail previously (11, 21).

Features were grouped as follows: (i) tumor intensity features: these include first-order statistics, calculated from the histogram of all tumor voxel intensity values. (ii) Textural features: these quantify intratumor heterogeneity and are calculated in all three-dimensional directions within the tumor volume, thereby taking the spatial location of each voxel compared with the surrounding voxels into account. The size-zone matrix was used to quantify regional heterogeneity. This matrix allows characterization of arrangements of voxels within the tumors, therefore describing tumor regional heterogeneity. (iii) Shape features: metrics of the three-dimensional shape and size of the tumor. (iv) Wavelet features: features in groups I and II are extracted after applying a series of wavelet transforms to the CT images. The wavelet transform decomposes the original image into low and high frequencies, thereby focusing the features on different frequency ranges within the tumor volume. (v) Laplacian of Gaussian features: these are textural features extracted after in-plane filtration using a Laplacian of Gaussian spatial band-pass filter. This filter highlights textural and anatomic patterns of different width depending of the spatial scale of the filter. By modifying the filter width, fine, medium, and coarse textures can be highlighted and textural features are subsequently calculated.

Statistical analysis

First, we investigated the associations between the imaging phenotype and the most common somatic mutations in the integrated dataset (Fig. 1). We used an unsupervised two-step feature selection methodology. First, we used the RIDER NSCLC test-retest dataset ($n = 31$) to assess stability of the radiomic features (Supplementary Methods S2; ref. 37). For each patient, we extracted radiomic features from the test and re-test scans. The intra-class correlation coefficient (ICC) was used to determine the stability of the features (Supplementary Fig. S2). Features with an ICC lower than 0.8 were excluded from the analysis. In a second step, we performed a principal component (PCA)-based analysis

(38) to identify the highest correlated features (Pearson $r > 0.90$) to the principal components that describe at least 90% of the variance in the radiomic data. This resulted in a selection of 26 variance-retaining features (Supplementary Fig. S3). Tumor volume and axial diameter were added for comparison. We compared the radiomic features distributions between mutated and nonmutated cases for each gene using a two-sided Wilcoxon test. To correct for multiple comparisons, we adjusted *P* values by the false discovery rate (FDR = 5%) procedure according to Benjamini and Hochberg (39).

The ability to predict the mutational status of the radiomic features was assessed by the area under the curve (AUC) of the receiver operator characteristic (ROC) as implemented in the survcomp R package (Version 1.12.0; ref. 40). Significance of AUCs was determined using the "noether" method implemented in the R survcomp package.

For the multivariate analysis, we used a temporal split (median scan acquisition date) to divide each of the four cohorts into training and validation sets. All training cohorts were combined into an integrated discovery cohort to identify radiomic signatures for EGFR and KRAS mutations while the validation cohorts were combined into an integrated validation dataset. To statistically compare radiomics and clinical multivariate models, we excluded all the samples with any missing clinical (stage, gender, smoking status, age, or race) or mutation (EGFR or KRAS status) information from each of the four cohorts before the temporal split. This exclusion resulted in total 257, 186, 142, and 120 samples in TIANJIN, PROFILE, LUNG-RT, and SPORE-MOFFITT cohorts, respectively, and hence, 353 patients were used for discovery and 352 patients for independent validation. To compare the two positive mutations ($EGFR^+$ and $KRAS^+$), we also excluded the wild-type cases ($EGFR^-$ and $KRAS^-$) from each cohort before the temporal split. This further reduction resulted in total 136, 114, 78, and 53 samples in TIANJIN, PROFILE, HARVARD-RT, and MOFFITT cohorts, respectively, and hence, 190 patients were used for discovery and 191 patients for independent validation. Radiomic signatures to predict mutation status were built in the integrated discovery cohort, by minimum redundancy maximum relevance (MRMR) feature selection. MRMR has been shown previously to be a stable feature selection algorithm for radiomics (41). The MRMR algorithm (42) was applied on all radiomic features with respect to a given mutational status, that is, EGFR, to select a nonredundant and highly informative set of features. Using the top 20 MRMR ranked features, controlling for differences in event ratios, we trained a random forest (RF) classifier on the discovery cohort. RF classifiers have also shown stability and high accuracy on radiomics analyses (15, 41, 43, 44). RF models were built for radiomic features and mutation data on the discovery cohort and their performance was evaluated on the validation cohort; therefore, none of the models were overfitted because models were trained only on the discovery cohort. The prediction performance was assessed using area under receiver operator characteristics curve (AUC). We also built RF-based clinical multivariate models using five clinical variables, that is, tumor stage, gender, smoking status, age, and race. Distributions of these clinical variables across the four cohorts can be obtained from the patient characteristics table (Supplementary Table S1). To assess the additive effect in prediction, combined models with 20 MRMR-ranked radiomic features and 5 clinical variables were built using RF. Prediction performance (AUC) of these different models were statistically compared using one-sided *t* test as implemented in the R package survcomp. To compute additional prediction measures (e.g., sensitivity, specificity, accuracy, NPV and PPV), we used the event (mutation) ratios of the discovery cohort as a probability threshold and obtained a corresponding cut-off point on the ROC curves. These prediction measures were computed using the R package pROC. MRMR feature selection was implemented using the MATLAB toolbox FEAST (42) and Matlab (Version R2012b, The Mathworks). All other statistical analyses were performed using R (Version 3.0.2).

Results

Genotype–phenotype associations

To investigate genotype–phenotype associations, we compared CT radiomic features with somatic mutation status in lung adenocarcinoma patients (Fig. 1). To incorporate the diversity of genotypic and phenotypic variations between the individual cohorts (see Supplementary Fig. S4), we performed an integrated analysis combining four cohorts, totaling 763 patients (Supplementary Table S1). Twenty-six robust and nonredundant radiomic features were included in our analysis. These features were selected in an unsupervised fashion-based on test–retest performance and interfeature correlation, and independent of mutation status or any other outcome (see Materials and Methods). Selected features included intensity histogram metrics, shape, and texture features, with or without wavelet or Laplacian-of-Gaussian filters. These features are capable to quantify a panel of phenotypic characteristics, such as intratumor homogeneity and heterogeneity, tumor density, and spherical disproportion, describing tumor roundness.

We then investigated the association of selected features with the most frequent somatic mutations in lung adenocarcinoma, KRAS (28.2%, 215 of 763) and EGFR (24%, 183 of 763). Using a nonparametric, two-sided Wilcoxon test on the integrated cohort and correcting for multiple testing (5% FDR), we separately compared EGFR-mutated and KRAS-mutated tumors to cases without EGFR or KRAS mutations, respectively.

We found sixteen radiomic features to be significantly associated with EGFR mutations and ten features associated with KRAS mutations (Fig. 2). Significant features were reported as overrepresented (+) or underrepresented (−), indicating the relative feature representation.

For EGFR-mutated tumors, we found the Homogeneity and Inverse Variance radiomic features to be underrepresented, whereas Sum Entropy and Short Run Emphasis were overrepresented. Homogeneity is sensitive to the number of unique discrete values in the images such that the fewer unique values that are accessed, the more homogeneous the image. Inverse variance assesses variations in intensity of voxels close to each other and therefore quantifies another aspect of homogeneity. Sum Entropy is the entropy of the cooccurrence matrix and therefore quantifies complexity. Short Run Emphasis is a run length feature that indicates successive voxels have similar intensity values. Together, the representation of these features indicates that $EGFR^+$ tumors are more likely to be heterogeneous.

In contrast, Sum Entropy was underrepresented for KRAS-mutated tumors, indicating that KRAS mutants are more homogeneous. Total Energy is associated with a LLL wavelet

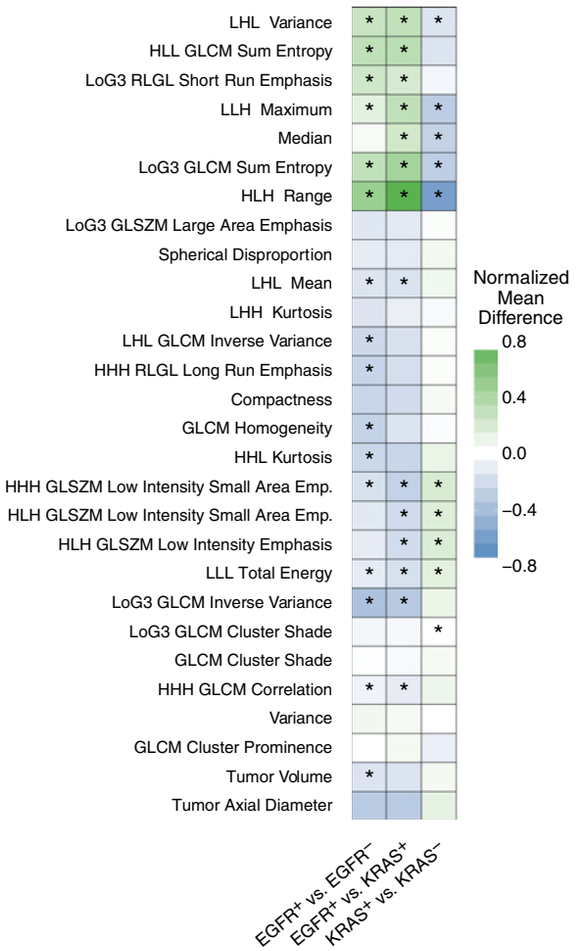


Figure 2.

Somatic genotype-imaging phenotype associations by comparing radiomic feature distributions between mutation subtypes. Heatmap shows the normalized mean difference of radiomic feature distributions for 26 reproducible and variant imaging features of 763 lung adenocarcinoma patients. Volumetric features were included for comparison. *, 5% FDR-corrected *P* values using a two-sided Wilcoxon test. Note that many features are significantly different between $EGFR^+$ versus $EGFR^-$ and between $EGFR^+$ and $KRAS^+$, although less for $KRAS^+$ versus $KRAS^-$.

filter, which is a low pass filter that enhances the general information of the image while reducing its noise. Total Energy was underrepresented in EGFR-mutated tumors and overrepresented in KRAS-mutated tumors so that EGFR-mutated tumors had a quantitatively lower Total Energy metric, compared with KRAS-mutated tumors.

We then compared radiographic features between EGFR- and KRAS-mutant tumors. We found 14 significant features, all of which were among the 16 that distinguished EGFR-mutant from EGFR nonmutated tumors. This pronounced difference between EGFR-mutated tumors and others is consistent with tumor volumetric analysis. EGFR-mutant tumors were smaller than nonmutated tumors (18.15 ± 81.7 cm^3 vs. 29.7 ± 61.1 cm^3, FDR $P < 0.05$) while KRAS-mutant tumors were more similar in size to nonmutated tumors (29.6 ± 62.7 cm^3 vs. 25.6 ± 69.2 cm^3, $P = 0.365$).

Predictive radiomic signatures for EGFR and KRAS mutation status

To evaluate the value of radiomic data to predict EGFR and KRAS mutation status, we developed and independently validated radiomic signatures, and compared their performance with clinical models (Fig. 1C). To incorporate the diversity of genotypic and phenotypic variations across the datasets into the signature development, we divided each of the four cohorts into an independent training set ($n = 353$) and an independent validation set ($n = 352$; Supplementary Fig. S4). Three radiomic signatures were developed for classifying: (i) between $EGFR^+$ and $EGFR^-$, (ii) between $KRAS^+$ and $KRAS^-$, and (iii) between $EGFR^+$ and $KRAS^+$. Furthermore, we also developed clinical models incorporating age, gender, smoking status, race, and clinical stage, to classify between these three groups. The performances of the radiomic and clinical signatures were compared with each other, and to conventional radiographic parameters used in clinical settings (axial diameter and volume of the tumor).

Each of the three signatures used twenty radiomic features (see Materials and Methods; Supplementary Methods S3). It is noteworthy that a large number of included features (9 of 20), were common across the three signatures. These included textural features that are sensitive to tumor radiographic heterogeneity, such as gray-level nonuniformity (GLNU) and low intensity small area emphasis (LISAE), both sensitive to complex patterns or high variation, as well as cluster prominence (CP) and inverse difference moment (IDM), which emphasize voxel pattern from close range intensity (e.g., smooth transition between voxel intensity), and are related to radiographic homogeneity.

Figure 3 and Supplementary Tables S2–S4 show the performances of radiomic signatures on the validation cohort. Conventional radiographic predictors, that is, maximum diameter and tumor volume showed significant, albeit low, performance in distinguishing between $EGFR^+$ and $EGFR^-$ tumors (AUC = 0.61, $P = 5.88 \times 10^{-04}$ and AUC = 0.60, $P = 8.44 \times 10^{-04}$, respectively). Neither diameter nor volume were able to distinguish between $KRAS^+$ and $KRAS^-$ (AUC = 0.53, $P = 0.44$ and AUC = 52, $P = 0.63$, respectively), or to distinguish between $EGFR^+$ and $KRAS^+$ (AUC = 0.58, $P = 0.07$ and AUC = 0.56, $P = 0.12$, respectively).

The radiomic signature we developed showed a significant ability to discriminate between $EGFR^+$ and $EGFR^-$ cases (AUC = 0.69, $P = 5.32 \times 10^{-10}$, Fig. 3A). This signature significantly outperformed axial diameter ($P < 0.03$) and tumor volume ($P < 0.02$). A clinical model of EGFR status, including age, gender, smoking status, race, and clinical stage, also showed high performance (AUC = 0.70, $P = 1.71 \times 10^{-09}$), and was similar to the performance of the radiomic signature ($P = 0.46$). We tested whether the radiomic and clinical signatures were complementary by creating a combined predictor and identified a combined signature that improved the accuracy (AUC = 0.75, $P = 8.93 \times 10^{-18}$), significantly better than the radiomic ($P = 0.05 \times 10^{-02}$) and clinical ($P = 0.03$) signatures alone. The complementary power of the two signatures was supported by a higher sensitivity for the radiomic signature and greater specificity for the clinical signature alone, with the combined signature having both high sensitivity and specificity (Supplementary Table S4; Supplementary Fig. S2).

A $KRAS^+/KRAS^-$ radiomic signature also showed significant, albeit lower, performance (AUC = 0.63, $P = 5.45 \times 10^{-05}$), and outperformed maximum diameter ($P = 0.02$) and tumor

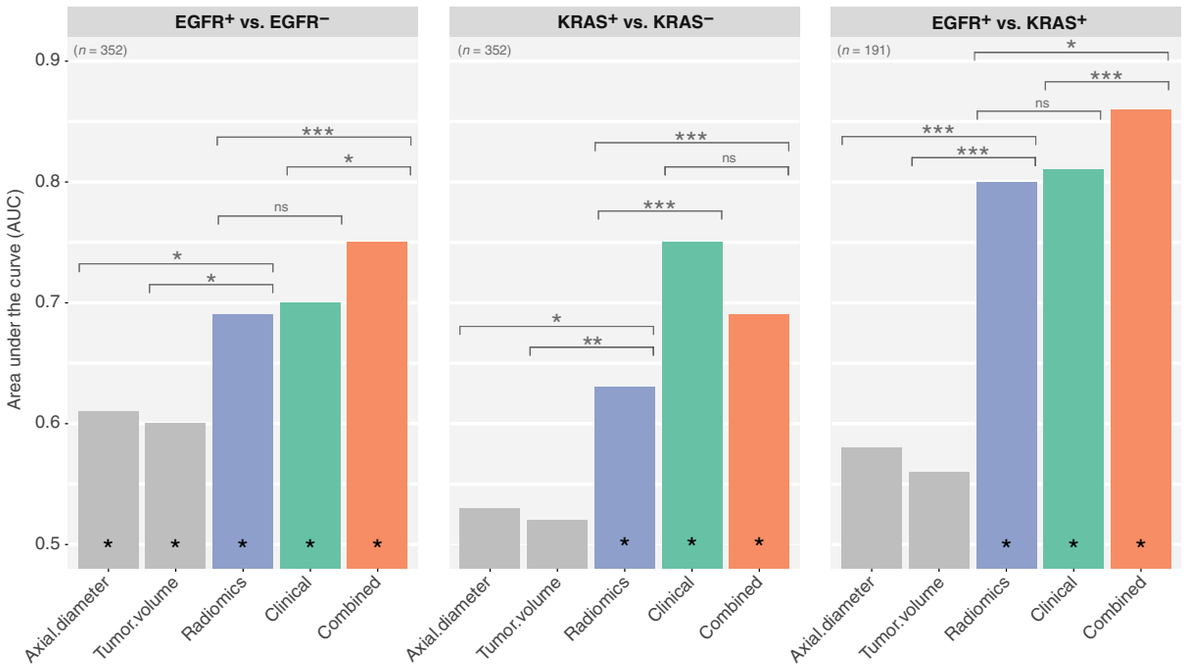


Figure 3.

Radiomic signatures to predict somatic mutations. Performance of radiomic signatures on the validation cohort for each mutational status classification on an independent validation dataset ($n = 353$). For comparison, conventional radiographic parameters (axial diameter and tumor volume) were included. Clinical models, including age, gender, smoking status, race, and clinical stage, were developed for each classification. Asterisk (*) on bottom of bars indicates that the performance of a model is significantly better than random. Furthermore, the segments indicate if a signature is significantly higher than another (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant). Note, that overall the radiomic signatures significantly outperform radiographic volumetric predictors.

volume ($P = 0.01$; Fig. 3B). A clinical predictor of KRAS, that included age, gender, smoking status, race, and clinical stage, status showed a very high performance (AUC = 0.75, $P = 5.86 \times 10^{-19}$), and was significantly higher than the radiomic model ($P = 8.73 \times 10^{-05}$). Combining both signatures reduced the overall performance (AUC = 0.69). In fact, the accuracy was higher for the clinical signature (0.66) than for either the radiomic (0.56) or combined (0.60) signature (Supplementary Table S4; Supplementary Fig. S5).

Finally, we developed a predictive signature to distinguish between EGFR-mutated and KRAS-mutated tumors (Fig. 3C). This radiomic signature had the best overall performance (AUC = 0.80, $P = 1.20 \times 10^{-20}$), significantly higher than either maximum diameter ($P = 1.65 \times 10^{-05}$) or tumor volume ($P = 6.17 \times 10^{-06}$). Combining this signature with a clinical signature (AUC = 0.81, $P = 1.37 \times 10^{-21}$), substantially improved its performance (AUC = 0.86, $P = 9.88 \times 10^{-39}$). This improvement was also significantly higher than either the radiomic or clinical model alone ($P = 0.02 \times 10^{-02}$ and $P = 0.02$, respectively). The combined model also had a high specificity (0.87) and accuracy (0.79) (Supplementary Table S4; Supplementary Fig. S5).

Discussion

Cancer is characterized by distinct molecular and environmental events that drive tumor development and progression (45, 46). Radiomic assessment of the tumor phenotype can be used with noninvasive images that are collected routinely in the clinic throughout the course of care. Our driving hypothesis is that the tumor phenotype, measured quantitatively through radiomics, should reflect the tumor genotype. We investigated the association between CT radiomic phenotypes and the most common somatic mutations in lung adenocarcinoma, that is, EGFR and KRAS mutations. We performed an integrated analysis of four large independent cohorts of lung adenocarcinoma patients for whom clinical, imaging, and mutational profiling data were available. We applied a stringent statistical design with independent training and validation cohorts in a large number of patients, to ensure validity of results. We found that EGFR tumors were more likely to be heterogeneous, with a smaller volume, and presenting an overall lower density on CT images. On the other hand, KRAS tumors could not be discriminated on the basis of volumetric information; however, they were more likely to be homogeneous.

We developed radiomic signatures predictive of mutational status in a discovery cohort of 353 patients that showed strong predictive performance in a validation cohort of 352 patients. We found radiomic signatures were predictive of EGFR mutations (AUC = 0.69) and were able to reliably distinguish between $EGFR^+$ and $KRAS^+$ (AUC = 0.80) tumors. The $EGFR^+/EGFR^-$ radiomic signature complemented the predictive value of a signature based on clinical factors (AUC = 0.75).

Few studies have investigated associations between the tumor imaging phenotype and the underlying molecular landscape (24–31, 33). These studies generally had small sample sizes, used subjective observer-dependent imaging descriptors, and did not perform robust external validation. For example, associations

have been reported between textural imaging features and KRAS mutations in combined imaging modalities (CT, ^{18}F- FDG PET and dynamic contrast-enhanced CT) in colorectal cancer (25), and using CT in NSCLC (27, 47), but within small cohorts (<100 patients), and, more importantly, without external validation, which is a critical component of radiomic analyses (9, 10). Our analysis of radiomic features extracted from CT images found that they had relatively weak predictive power for KRAS, a result that we validated in an independent validation set, indicating a generally similar imaging phenotype for KRAS$^+$ and KRAS$^-$ tumors. For instance, Wang and colleagues, found a single semantic CT feature, tumor spiculation, to be associated with the presence of KRAS mutations in a single cohort of stage I lung adenocarcinoma patients (47). We found, however, that EGFR$^+$ and EGFR$^-$ tumors exhibit distinct imaging phenotypes, allowing more sensitive and specific classification of tumor EGFR status. In a recent single cohort association study, Liu and colleagues found 16 semantic CT annotations to be associated with EGFR status in a cohort of NSCLC patients. EGFR mutations were associated with tumor size, with ground glass opacity, and enhancement heterogeneity. Their study, however, lacked validation of their multivariate models and did not include multiple testing corrections (48). Similarly, Park and colleagues, in a single cohort association study of advanced stage lung adenocarcinomas, found that EGFR$^+$ tumors presented more frequently ground-glass opacity, while KRAS$^+$ cases were more likely solid tumors (49).

This is the first study to evaluate associations between somatic mutations and radiomic features in a large cohort of lung adenocarcinoma patients, and to validate its findings in an independent patient cohort. Our findings suggest the need for studies in other cancers to learn more general rules for mapping clinical and imaging features to the tumor's mutational status in single tissues and across tissues, and to understand whether temporal changes in tumor morphology can be associated with changes in tumor mutational status.

Despite strong results, our study has a number of limitations. We used diagnostic CT scans as performed routinely in the clinic with heterogeneous scanning protocols. CT images were acquired using scanners manufactured by different companies, with a range of image reconstruction algorithms, different slice thicknesses, with and without contrast, and using different dosages. We normalized all images to isovolumetric voxels to reduce the effect of different slice thicknesses. However, despite these factors potentially adding noise to the data, we were able to identify a strong signal predictive of EGFR mutations. It is conceivable that heterogeneity in CT protocols could have obscured more subtle differences in the phenotype of KRAS$^+$ and KRAS$^-$ tumors. Further optimization and standardization of imaging data is an important aspect for the introduction of imaging-based biomarkers. Several groups, such as the Quantitative Imaging Network (50), investigate these issues by implementing feature standardization efforts such as segmentation challenges, radiomic feature definition standardizations, and developing open source, publicly available analyses platforms.

Radiomic features report three-dimensional scores that are representative of the tumor as a whole, that is, as an average score for textural features in all directions in the three-dimensional space. We believe that a detailed analysis of the intratumoral tumor heterogeneity, by compartment mappings, particularly if paired with multiple biopsies at distinct geographical locations, such as investigated in renal cell carcinomas (7, 51), may allow us to further elucidate intratumoral genotype–phenotype relationships. Localized and temporal tumor heterogeneity represents a challenge for repeated tissue sampling for genomic assays (6); thus, imaging signatures may be useful as a surrogate for genomic assays when a biopsy is not possible, or as a complementary assay to monitor response to therapy. Similarly, if the radiomic signature predicts a different EGFR status compared with tissue assessment, particularly if an EGFR-targeted treatment was to be delivered, this could indicate rebiopsy, to rule out sampling error or misdiagnosis (52). Furthermore, quantitative radiomic analysis can add in the evaluation of treatment response in EGFR-mutant lung adenocarcinoma patients, treated with EGFR tyrosine kinase inhibitors, beyond tumor volume assessments and RECIST criteria (53). This, however, still needs to be evaluated in clinical data.

There are fundamental differences between radiology and pathology-based tests. Radiology can capture the phenotype at a macroscopic level (millimeter resolution), and not a microscopic level, as provided by histopathology and required for a detailed quantification of underlying biological processes. An advantage of imaging is that it easily can sample the complete disease burden sequentially over time. Therefore, imaging-based biomarkers could potentially be applied in clinical situations were biopsy-based assays are not possible, and it could also provide complementary information, especially over the course of treatment. It might also be interesting to compare radiomics-based genotype predictions with liquid biopsy–based information, as both of these approaches can reflect the overall tumor load as opposed to the partial sampling provided by biopsies (54).

Nevertheless, this study lays important groundwork for establishing radiomics as an important adjunct approach to existing clinical predictors of disease status and therefore treatment protocol. Ongoing prospective data collection projects such as PROFILE at the Dana-Farber/Brigham and Women's Cancer Center, are collecting comprehensive mutational data on almost all lung cancer patients receiving treatment (35) and these data are linked to comprehensive clinical records including radiologic scans collections. Such initiatives are needed to further validate imaging-based predictors that can be useful for clinical application.

In summary, we demonstrated an association between the imaging phenotype captured with a radiomic signature and EGFR-mutant tumors, in four independent cohorts of lung adenocarcinomas. This association may have clinical impact in selecting patients for targeted therapies. Imaging phenotype associations with other molecular subtypes of NSCLC should to be further investigated in prospective genotype profiling cohorts (33, 55).

Disclosure of Potential Conflicts of Interest

R.J. Gillies has ownership interest (including patents) and is a consultant/advisory board member for Health Myne. J. Quackenbush is a founder and board chair at Genospace, has ownership interest (including patents) in Genospace, and is a consultant/advisory board member for Caris Life Sciences and Shpera, LLC. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: E. Rios Velazquez, C. Parmar, Y. Liu, T.P. Coroller, Z. Ye, R.H. Mak, R.J. Gillies, J. Quackenbush, H. Aerts

Development of methodology: E. Rios Velazquez, C. Parmar, O. Stringfield, R.H. Mak, J. Quackenbush, H. Aerts

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): E. Rios Velazquez, Y. Liu, T.P. Coroller, O. Stringfield, Z. Ye, F.M.M. Fennessy, R.H. Mak, H. Aerts
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): E. Rios Velazquez, C. Parmar, Y. Liu, T.P. Coroller, F.M.M. Fennessy, R.H. Mak, R.J. Gillies, H. Aerts
Writing, review, and/or revision of the manuscript: E. Rios Velazquez, C. Parmar, Y. Liu, T.P. Coroller, O. Stringfield, Z. Ye, G.M. Makrigiorgos, F.M.M. Fennessy, R.H. Mak, R.J. Gillies, J. Quackenbush, H. Aerts
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y. Liu, R.H. Mak, R.J. Gillies, H. Aerts
Study supervision: E. Rios Velazquez, Z. Ye, H. Aerts
Other (collection of data and analysis of radiology imaging studies): G. Cruz

Acknowledgments

The authors would like to thank the PROFILE team of the Dana-Farber Cancer Institute for their help with somatic mutation testing.

Grant Support

This work was financially supported by NIH (NIH-USA U24CA194354 and NIH-USA U01CA190234).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 17, 2017; revised March 13, 2017; accepted May 22, 2017; published OnlineFirst May 31, 2017.

References

1. Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. Nat Med 2011;17:297–303.
2. Garraway LA, Verweij J, Ballman KV. Precision oncology: an overview. J Clin Oncol 2013;31:1803–5.
3. Janne PA, Engelman JA, Johnson BE. Epidermal growth factor receptor mutations in non-small-cell lung cancer: implications for treatment and tumor biology. J Clin Oncol 2005;23:3227–34.
4. Li T, Kung HJ, Mack PC, Gandara DR. Genotyping and genomic profiling of non-small-cell lung cancer: implications for current and future therapies. J Clin Oncol 2013;31:1039–49.
5. Chong CR, Janne PA. The quest to overcome resistance to EGFR-targeted therapies in cancer. Nat Med 2013;19:1389–400.
6. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. Nature 2013;501:355–64.
7. Swanton C. Intratumor heterogeneity: evolution through space and time. Cancer Res 2012;72:4875–82.
8. Taniguchi K, Okami J, Kodama K, Higashiyama M, Kato K. Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. Cancer Sci 2008;99:929–35.
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563–77.
10. Aerts HJWL. The potential of radiomic-based phenotyping in precision medicine: a review. JAMA Oncol 2016;2:1636–42.
11. Aerts H, Rios Velazquez E, Leijenaar R, Parmar C, Grossmann P, Cavalho S, et al. Decoding the tumor phenotype by non-invasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
12. 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.
13. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. Phys Med Biol 2016;61:R150–66.
14. Fried DV, Tucker SL, Zhou S, Liao Z, Mawlawi O, Ibbott G, et al. Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2014;90:834–42.
15. Parmar C, Leijenaar RTH, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. Sci Rep 2015;5:11044.
16. Zhou Y, He L, Huang Y, Chen S, Wu P, Ye W, et al. CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma. Abdom Radiol 2017;42:1695–1704.
17. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting malignant nodules from screening CT scans. J Thorac Oncol 2016;11:2120–8.
18. Schabath M, Balagurunathan Y, Dmitry G, LAwrence H, Samuel H, Stringfield O, et al. Radiomics of lung cancer. J Thorac Oncol 2016;11:S5–6.
19. Coroller TP, Agrawal V, Huynh E, Narayan V, Lee SW, Mak RH, et al. Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC. J Thorac Oncol 2017;12:467–76.
20. Coroller TP, Agrawal V, Narayan V, Hou Y, Grossmann P, Lee SW, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. Radiother Oncol 2016;119:480–6.
21. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol 2015;114:345–50.
22. Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. Radiother Oncol 2016;120:258–66.
23. Huynh E, Coroller TP, Narayan V, Agrawal V, Romano J, Franco I, et al. Associations of radiomic data extracted from static and respiratory-gated CT scans with disease recurrence in lung cancer patients treated with SBRT. PLoS One 2017;12:e0169172.
24. Karlo CA, Di Paolo PL, Chaim J, Hakimi AA, Ostrovnaya I, Russo P, et al. Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations. Radiology 2014;270:464–71.
25. Miles KA, Ganeshan B, Rodriguez-Justo M, Goh VJ, Ziauddin Z, Engledow A, et al. Multifunctional imaging signature for V-KI-RAS2 Kirsten rat sarcoma viral oncogene homolog (KRAS) mutations in colorectal cancer. J Nucl Med 2014;55:386–91.
26. Nair VS, Gevaert O, Davidzon G, Napel S, Graves EE, Hoang CD, et al. Prognostic PET 18F-FDG uptake imaging features are associated with major oncogenomic alterations in patients with resected non-small cell lung cancer. Cancer Res 2012;72:3725–34.
27. Weiss GJ, Ganeshan B, Miles KA, Campbell DH, Cheung PY, Frank S, et al. Noninvasive image texture analysis differentiates K-ras mutation from pan-wildtype NSCLC and is prognostic. PLoS One 2014;9:e100244.
28. Yamamoto S, Korn RL, Oklu R, Migdal C, Gotway MB, Weiss GJ, et al. ALK molecular phenotype in non-small cell lung cancer: CT radiogenomic characterization. Radiology 2014;272:568–76.
29. Zhou JY, Zheng J, Yu ZF, Xiao WB, Zhao J, Sun K, et al. Comparative analysis of clinicoradiologic characteristics of lung adenocarcinomas with ALK rearrangements or EGFR mutations. Eur Radiol 2015;25:1257–66.
30. Rizzo S, Petrella F, Buscarino V, De Maria F, Raimondi S, Barberis M, et al. CT radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer. Eur Radiol 2015;26:32–42.
31. Liu Y, Kim J, Balagurunathan Y, Li Q, Garcia AL, Stringfield O, et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. Clin Lung Cancer 2016;17:441–8.e6.
32. Guo W, Li H, Zhu Y, Lan L, Yang S, Drukker K, et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and genomics data. J Med Imaging 2015;2:041007.
33. Yoon HJ, Sohn I, Cho JH, Lee HY, Kim J-H, Choi Y-L, et al. Decoding tumor phenotypes for ALK, ROS1, and RET fusions in lung adenocarcinoma using a radiomics approach. Medicine 2015;94:e1753.
34. Yip SS, Kim J, Coroller T, Parmar C, Rios Velazquez E, Huynh E, et al. Associations between somatic mutations and metabolic imaging phenotypes in non-small cell lung cancer. J Nucl Med 2017;58:569–76.
35. MacConaill LE, Garcia E, Shivdasani P, Ducar M, Adusumilli R, Breneiser M, et al. Prospective enterprise-level molecular genotyping of a cohort of cancer patients. J Mol Diagn 2014;16:660–72.
36. MacConaill LE, Campbell CD, Kehoe SM, Bass AJ, Hatton C, Niu L, et al. Profiling critical cancer gene mutations in clinical tumor samples. PLoS One 2009;4:e7887.

37. Zhao B, James LP, Moskowitz CS, Guo P, Ginsberg MS, Lefkowitz RA, et al. Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. Radiology 2009;252: 263–72.
38. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. J Stat Softw 2008;25:1–18.
39. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995;57:289–300.
40. Schroder MS, Culhane AC, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. Bioinformatics 2011;27:3206–8.
41. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for quantitative radiomic biomarkers. Sci Rep 2015;5: 13087.
42. Brown G, Pocock A, Zhao M-J, Luján M. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. J Mach Learn Res 2012;13:27–66.
43. Breiman L. Random forests. Mach Learn 2001;45:5–32.
44. Wu W, Parmar C, Grossmann P, Quackenbush J, Lambin P, Bussink J, et al. Exploratory study to identify radiomics classifiers for lung cancer histology. Front Oncol 2016;6:71.
45. Oesper L, Satas G, Raphael BJ. Quantifying tumor heterogeneity in whole-genome and whole-exome sequencing data. Bioinformatics 2014;30: 3532–40.
46. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113–20.
47. Wang H, Schabath MB, Liu Y, Stringfield O, Balagurunathan Y, Heine JJ, et al. Association between computed tomographic features and kirsten rat sarcoma viral oncogene mutations in patients with stage I lung adenocarcinoma and their prognostic value. Clin Lung Cancer 2016;17:271–8.
48. Liu Y, Kim J, Qu F, Liu S, Wang H, Balagurunathan Y, et al. CT features associated with epidermal growth factor receptor mutation status in patients with lung adenocarcinoma. Radiology 2016;280:271–80.
49. Park J, Kobayashi Y, Urayama KY, Yamaura H, Yatabe Y, Hida T. Imaging characteristics of driver mutations in EGFR, KRAS, and ALK among treatment-naïve patients with advanced lung adenocarcinoma. PLoS One 2016;11:e0161081.
50. National Cancer Institute Cancer Imaging Program. Quantitative imaging for evaluation of responses to cancer therapies: quantitative imaging network (QIN) [Internet]; 2013. Available from: https://imaging.cancer.gov/programs_resources/specialized_initiatives/qin.htm.
51. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883–92.
52. Querings S, Altmuller J, Ansen S, Zander T, Seidel D, Gabler F, et al. Benchmarking of mutation diagnostics in clinical lung cancer specimens. PLoS One 2011;6:e19601.
53. Nishino M, Cardarella S, Dahlberg SE, Jackman DM, Ramaiya NH, Hatabu H, et al. Radiographic assessment and therapeutic decisions at RECIST progression in EGFR-mutant NSCLC treated with EGFR tyrosine kinase inhibitors. Lung Cancer 2013;79:283–8.
54. Song C, Chen S, Yibin L, Rachel F, Alexander M, Harvey M, et al. Elimination of unaltered DNA in mixed clinical samples via nuclease-assisted minor-allele enrichment. Nucleic Acids Res 2016;44: e146.
55. Brastianos PK, Horowitz PM, Santagata S, Jones RT, McKenna A, Getz G, et al. Genomic sequencing of meningiomas identifies oncogenic SMO and AKT1 mutations. Nat Genet 2013;45:285–9.