Published in final edited form as:

Cancer Res. 2017 July 15; 77(14): 3922–3930. doi:10.1158/0008-5472.CAN-17-0122.

Somatic mutations drive distinct imaging phenotypes in lung cancer

Emmanuel Rios Velazquez^{1,*}, Chintan Parmar^{1,*}, Ying Liu^{5,7,*}, Thibaud P. Coroller¹, Gisele Cruz², Olya Stringfield⁵, Zhaoxiang Ye⁷, Mike Makrigiorgos¹, Fiona Fennessy², Raymond H. Mak¹, Robert Gillies⁵, John Quackenbush^{3,4,6}, and Hugo J.W.L. Aerts^{1,2,3}

¹Department of Radiation Oncology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

²Department of Radiology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

³Department of Biostatistics & Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA

⁴Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA

⁵Departments of Cancer Imaging and Metabolism, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA

⁶Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

⁷Department of Radiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin, China; Tiajin's Clinical Research Center for Cancer

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 (AI) 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 computed tomography (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

Corresponding Author: Hugo Aerts, PhD, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Institutes of Medicine – HIM 343, 77 Avenue Louis Pasteur, Boston, MA 02115, P - 617.525.7156, F - 617.582.6037, Hugo_Aerts@dfci.harvard.edu.

*Contributed equally

Conflict of Interests: None

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 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 non-invasively, repeatedly and at low cost.

Keywords

Somatic Mutations; Radiomics; Radiology; Medical Imaging; Imaging-Genomics; Quantitative Imaging; Adenocarcinomas; Artificial Intelligence; Radiogenomics

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 epidermal growth factor receptor (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, i.e. prior to initiation of treatment. Further, 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 (Figure 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 information that biopsies fail to deliver, such as radiographic intra-tumor 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 intra-reader 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 (Figure 1)(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 intra-tumor 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), pathological 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 non-permissive 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 semi-quantitative 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 computed tomography (CT) images and tested these features against the most frequently occurring genetic alterations in the disease, i.e. 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 non-invasively 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 histological 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 were included in our analysis. Clinical stages ranged from stage I to stage 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 of each of the participating centres 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 1 and Supplementary Figure 1. 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. Supplementary Methods 1 a detailed description about the datasets are available.

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. Due to 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 intra-tumor heterogeneity and are calculated in all 3-dimensional directions within the tumor volume, thereby taking the spatial location of each voxel compared to 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 anatomical 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.

Data analysis

First, we investigated the associations between the imaging phenotype and the most common somatic mutations in the integrated dataset (Figure 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(37) (Supplementary Methods 2). 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 Figure 2). 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's $r > 0.90$) 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 Figure 3). Tumor volume and axial diameter were added for comparison. We compared the radiomic features distributions between mutated and non-mutated 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)(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. In order 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. In order 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, Lung-RT and Spore-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, i.e. EGFR, to select a non-redundant and highly informative set of features. Using the top 20 MRMR ranked features, controlling for differences in event ratios, we trained a random forest 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 random forest based clinical multivariate models using five clinical variables, i.e. 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 1). In order to assess the additive effect in prediction, combined models with 20 MRMR ranked radiomic features and 5 clinical variables were built using random forest. 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 cutoff 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, Natick, MA). All other statistical analyses were performed using R (Version 3.0.2).

Results

Genotype-phenotype associations

To investigate genotype-phenotype associations, we compared computed tomography (CT) radiomic features with somatic mutation status in lung adenocarcinoma patients (Figure 1). To incorporate the diversity of genotypic and phenotypic variations between the individual cohorts (see Supplementary Figure 4), we performed an integrated analysis combining four cohorts, totaling 763 patients (Supplementary Table 1). Twenty-six robust and non-redundant radiomic features were included in our analysis. These features were selected in an unsupervised fashion based on test-retest performance and inter-feature correlation, and independent of mutation status or any other outcome (see **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 intra-tumor 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 non-parametric, 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 (Figure 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 co-occurrence 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 indicate that EGFR positive 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 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 to KRAS mutated tumors.

We then compared radiographic features between EGFR and KRAS mutant tumors. We found fourteen significant features all of which were among the sixteen that distinguished

EGFR mutant from EGFR non-mutated tumors. This pronounced difference between EGFR mutated tumors and others is consistent with tumor volumetric analysis. EGFR mutant tumors were smaller than non-mutated tumors ($18.15 \pm 81.7 \text{ cm}^3$ vs $29.7 \pm 61.1 \text{ cm}^3$, FDR $p < 0.05$) while KRAS mutant tumors were more similar in size to non-mutated tumors ($29.6 \pm 62.7 \text{ cm}^3$ vs $25.6 \pm 69.2 \text{ cm}^3$, $p=0.365$).

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 (Figure 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 Figure 4). 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 to 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 **Methods** and Supplementary Methods 3). It is noteworthy that a large number of included features (9 out of 20), were common across the three signatures. These included textural features that are sensitive to tumor radiographic heterogeneity, such as gray-level non-uniformity (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 2-4 show the performances of radiomic signatures on the validation cohort. Conventional radiographic predictors, i.e. maximum diameter and tumor volume showed significant, albeit low, performance in distinguishing between EGFR+ and EGFR- tumors ($\text{AUC}=0.61$, $p=5.88 \times 10^{-04}$ and $\text{AUC}=0.60$, $p=8.44 \times 10^{-04}$, respectively). Neither diameter nor volume were able to distinguish between KRAS+ and KRAS- ($\text{AUC}=0.53$, $p=0.44$ and $\text{AUC}=0.52$, $p=0.63$, respectively), or to distinguish between EGFR+ and KRAS+ ($\text{AUC}=0.58$, $p=0.07$ and $\text{AUC}=0.56$, $p=0.12$, respectively).

The radiomic signature we developed showed a significant ability to discriminate between EGFR+ and EGFR- cases ($\text{AUC}=0.69$, $p=5.32 \times 10^{-10}$, Figure 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 ($\text{AUC}=0.70$, $p=1.71 \times 10^{-09}$), 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 ($\text{AUC}=0.75$, $p=8.93 \times 10^{-18}$), 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 4, Supplementary Figure 2).

A KRAS+/KRAS- radiomic signature also showed significant, albeit lower, performance ($AUC=0.63$, $p=5.45\times10^{-05}$), and outperformed maximum diameter ($p=0.02$) and tumor volume ($p=0.01$) (Figure 3B). A clinical predictor of KRAS, that included age, gender, smoking status, race, and clinical stage, status showed a very high performance ($AUC=0.75$, $p=5.86\times10^{-19}$), and was significantly higher than the radiomic model ($p=8.73\times10^{-05}$). 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 4, Supplementary Figure 5).

Finally, we developed a predictive signature to distinguish between EGFR mutated and KRAS mutated tumors (Figure 3C). This radiomic signature had the best overall performance ($AUC=0.80$, $p=1.20\times10^{-20}$), significantly higher than either maximum diameter ($p=1.65\times10^{-05}$) or tumor volume ($p=6.17\times10^{-06}$). Combining this signature with a clinical signature ($AUC=0.81$, $p=1.37\times10^{-21}$), substantially improved its performance ($AUC=0.86$, $p=9.88\times10^{-39}$). This improvement was also significantly higher than either the radiomic or clinical model alone ($p=0.02\times10^{-02}$ and $p=0.02$, respectively). The combined model also had a high specificity (0.87) and accuracy (0.79) (Supplementary Table 4, Supplementary Figure 5).

Discussion

Cancer is characterized by distinct molecular and environmental events that drive tumor development and progression (45,46) (45,46). Radiomic assessment of the tumor phenotype can be used with non-invasive 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 computed tomography (CT) radiomic phenotypes and the most common somatic mutations in lung adenocarcinoma, i.e. 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 based on 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, ¹⁸F- 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-positive and negative tumors. For instance, Wang et al., 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-positive and negative tumors exhibit distinct imaging phenotypes, allowing more sensitive and specific classification of tumor EGFR status. In a recent single-cohort association study, Liu et al. 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 et al, in a single-cohort association study of advanced stage lung adenocarcinomas, found that EGFR positive tumors presented more frequently ground-glass opacity, while KRAS positive 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 iso-volumetric 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 positive and negative 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 3-dimensional scores that are representative of the tumor as a whole, i.e. as an average score for textural features in all directions in the 3-dimensional space. We believe that a detailed analysis of the intra-tumoral 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 intra-tumoral 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 complementary assay to monitor response to therapy. Similarly, if the radiomic signature predicts a different EGFR status compared to tissue assessment, particularly if an EGFR targeted treatment was to be delivered, this could indicate re-biopsy, 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 where biopsies are not possible, and it could also provide complementary information, specially over the course of treatment. It might also be interesting to compare radiomics-based genotype predictions with liquid biopsy-based information since 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 radiological 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 be further investigated in prospective genotype profiling cohorts(33,55).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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

Financial Support: Authors acknowledge financial support from the National Institute of Health (NIH-USA U24CA194354, and NIH-USA U01CA190234).

References

1. Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. *Nat Med.* 2011; 17:297–303. [PubMed: 21383744]
2. Garraway LA, Verweij J, Ballman KV. Precision oncology: an overview. *J Clin Oncol.* 2013; 31:1803–5. jco.ascopubs.org. [PubMed: 23589545]
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. [PubMed: 15886310]
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. [PubMed: 23401433]
5. Chong CR, Janne PA. The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nat Med.* 2013; 19:1389–400. [PubMed: 24202392]
6. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. *Nature.* 2013; 501:355–64. [PubMed: 24048068]
7. Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res.* 2012; 72:4875–82. [PubMed: 23002210]
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. [PubMed: 18325048]
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2015; 151169
10. Aerts, HJWL. The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review. *JAMA Oncol.* 2016. [Internet]Available from: <http://dx.doi.org/10.1001/jamaoncol.2016.2631>
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
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. [PubMed: 22257792]
13. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. *Phys Med Biol.* 2016; 61:R150–66. [PubMed: 27269645]
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. [PubMed: 25220716]
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. [PubMed: 26251068]
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 (NY).* 2017. [Internet]Available from: <http://dx.doi.org/10.1007/s00261-017-1072-0>
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. [PubMed: 27422797]

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. [PubMed: 27903462]
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. [PubMed: 27085484]
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. [PubMed: 25746350]
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. [Internet]Available from: <http://dx.doi.org/10.1016/j.radonc.2016.05.024>
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. [PubMed: 28046060]
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. [PubMed: 24029645]
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. [PubMed: 24516257]
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. [PubMed: 22710433]
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. [PubMed: 24987838]
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. [PubMed: 24885982]
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. [PubMed: 25577516]
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. [Internet]Available from: <http://dx.doi.org/10.1007/s00330-015-3814-0>
31. Liu, Y., Kim, J., Balagurunathan, Y., Li, Q., Garcia, AL., Stringfield, O., et al. Clin Lung Cancer. Elsevier; 2016. Radiomic Features Are Associated With EGFR Mutation Status in Lung Adenocarcinomas. [Internet]Available from: <http://dx.doi.org/10.1016/j.clcc.2016.02.001>
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 (Bellingham).* 2015; 2:041007. medicalimaging.spiedigitallibrary.org/ [PubMed: 26835491]
33. Yoon HJ, Sohn I, Cho JH, Lee HY, Kim JH, Choi YL, et al. Decoding Tumor Phenotypes for ALK, ROS1, and RET Fusions in Lung Adenocarcinoma Using a Radiomics Approach. *Medicine.* 2015; 94:e1753. ncbi.nlm.nih.gov. [PubMed: 26469915]
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.* 2016. [Internet]Available from: <http://dx.doi.org/10.2967/jnumed.116.181826>
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. [PubMed: 25157968]

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 Public Library of Science*. 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. [PubMed: 19561260]
38. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. *J Stat Softw*. 2008; 25(1):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* Blackwell Publishing for the Royal Statistical Society. 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. [PubMed: 21903630]
41. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. *Sci Rep Macmillan Publishers Limited*. 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. [PubMed: 27064691]
45. Oesper L, Satas G, Raphael BJ. Quantifying tumor heterogeneity in whole-genome and whole-exome sequencing data. *Bioinformatics*. 2014; 30:3532–40. [PubMed: 25297070]
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. [PubMed: 24071849]
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. [PubMed: 26712103]
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. [PubMed: 26937803]
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. [PubMed: 27518729]
50. Program, National Cancer Institute Cancer Imaging. Quantitative Imaging for Evaluation of Responses to Cancer Therapies: Quantitative Imaging Network (QIN). 2013. [Internet]Available from: <http://imaging.cancer.gov/programsandresources/specializedinitiatives/qin>
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. [PubMed: 22397650]
52. Querings S, Altmüller 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. [PubMed: 21573178]
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. [PubMed: 23254265]
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; gkw650
55. Brastianos, PK., Horowitz, PM., Santagata, S., Jones, RT., McKenna, A., Getz, G., et al. *Nat Genet*. Nature Publishing Group, a division of Macmillan Publishers Limited; Genomic sequencing of meningiomas identifies oncogenic SMO and AKT1 mutations. All Rights Reserved.; 2013/03// print;45:285–9

Figure 1. Analysis workflow
A) Examples of lung adenocarcinomas tumors imaged with computed tomography (CT) imaging (left) and segmented in 3D (right). **B)** Quantification of the tumor phenotype using radiomics feature algorithms. **C)** Radiomic and clinical data was 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 Supplemental Table 1 and Supplemental Data 1.

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 twenty-six reproducible and variant imaging features of 763 lung adenocarcinoma patients.

Volumetric features were included for comparison. * Indicates 5% false discovery rate (FDR) corrected p-values using a two-sided Wilcoxon test. Note, that many features are significantly different between EGFR+ vs EGFR- and between EGFR+ and KRAS+, although less for KRAS+ vs KRAS-.

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. * 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.