Radiomic prediction of mutation status based on MR imaging of lung cancer brain metastases

Bihong T. Chen, MD[1],[*], **Taihao Jin, PhD**[1], **Ningrong Ye, MD**[1], **Isa Mambetsariev, BA**[2], **Ebenezer Daniel, PhD**[1], **Tao Wang, MD**[3], **Chi Wah Wong, PhD**[4], **Russell C. Rockne, PhD**[5], **Rivka Colen, MD**[6], **Andrei I. Holodny, MD**[7], **Sagus Sampath, MD**[8], **Ravi Salgia, MD**[2]

[1]Department of Diagnostic Radiology, City of Hope National Medical Center, Duarte, CA, United States

[2]Department of Medical Oncology and Therapeutics Research, City of Hope Comprehensive Cancer Center and Beckman Research Institute, Duarte 91010, CA, USA

[3]Departments of Interventional Radiology, Nanjing First Hospital, Nanjing Medical University, Nanjing, Jiangsu, P.R. China

[4]Center for Informatics, City of Hope National Medical Center, Duarte 91010, CA, United States

[5]Division of Mathematical Oncology, City of Hope National Medical Center, Duarte, CA, United States

[6]Hillman Cancer Center, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania Department of Radiology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

[7]Department of Radiology, Memorial Sloan-Kettering Cancer Center, New York, NY, United States

[8]Department of Radiation Oncology, City of Hope National Medical Center, Duarte, CA, United States

[*]**Correspondence:** Bihong T. Chen, Bechen@coh.org, Department of Diagnostic Radiology, City of Hope Medical Center, 1500 East Duarte Road, Duarte, CA 91010, Phone: 626 218 2318 Fax: 626 930 5451.

[7]·Author Contributions

BTC and RS designed and conducted the study. IM and RS provided the list of lung cancer patients with brain metastases and their mutation status from City of Hope data base. BTC, TJ, NY, TW, IM and RS acquired and evaluated the brain MR imaging data. NY and TW performed tumor segmentation and radiomic feature extraction. BTC and TW reviewed the segmented tumor images for consistency. TJ developed the pipeline for predictive modeling and machine learning. TJ and NY made the figures and tables. TJ and NY performed statistical analysis. BTC, TJ, NY, IM, ED, CWW, RR, TW, RC, AH, SS and RS contributed to interpretation and description of the data. TJ, NY, IM, CWW, RR and RS contributed to the manuscript writing process. BTC prepared the first draft of the entire manuscript and subsequent revisions. All authors approved the final manuscript.

[6]·Conflict of Interest

All authors declared that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

[9]·Supplementary Material

The supplementary Material for this article can be found online.

[10].Data Availability Statement

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

Abstract

Lung cancer metastases comprise most of all brain metastases in adults and most brain metastases are diagnosed by magnetic resonance (MR) scans. The purpose of this study was to conduct an MR imaging-based radiomic analysis of brain metastatic lesions from patients with primary lung cancer to classify mutational status of the metastatic disease. We retrospectively identified lung cancer patients with brain metastases treated at our institution between 2009 and 2017 who underwent genotype testing of their primary lung cancer. Brain MR Images were used for segmentation of enhancing tumors and peritumoral edema, and for radiomic feature extraction. The most relevant radiomic features were identified and used with clinical data to train random forest classifiers to classify the mutation status. Of 110 patients in the study cohort (mean age 57.51 ± 12.32 years; M: F=37:73), 75 had an *EGFR* mutation, 21 had an *ALK* translocation, and 15 had a *KRAS* mutation. One patient had both *ALK* translocation and *EGFR* mutation. Majority of radiomic features most relevant for mutation classification were textural. Model building using both radiomic features and clinical data yielded more accurate classifications than using either alone. For classification of *EGFR*, *ALK*, and *KRAS* mutation status, the model built with both radiomic features and clinical data resulted in area-under-the-curve (AUC) values based on cross-validation of 0.912, 0.915, and 0.985, respectively. Our study demonstrated that MR imaging-based radiomic analysis of brain metastases in patients with primary lung cancer may be used to classify mutation status. This approach may be useful for devising treatment strategies and informing prognosis.

Keywords

Radiomics; Predictive modeling; Mutation; Lung cancer; Brain metastases

1. Introduction

The prognosis of lung cancer patients who develop brain metastasis is poor. Lung cancer is the most common form of cancer to metastasize to the brain [1] with about 7–10% of non-small cell lung cancer (NSCLC) patients presenting with brain metastases upon diagnosis, and 20–40% developing brain metastases later on [2]. Options for the treatment of patients with NSCLC have greatly expanded in the past decade with the advent of targeted therapy. Treatment options are determined according to mutation status. Whereas alterations in epidermal growth factor receptor (*EGFR*), anaplastic lymphoma kinase (*ALK*), ROS proto-oncogene 1 (*ROS-1*), v-Raf murine sarcoma viral oncogene homolog B (*BRAF*), and neurotrophic tropomyosin receptor kinase (*NTRK*) genes can be targeted with FDA-approved drugs, alterations in the Kirsten rat sarcoma virus (*KRAS*) gene—which are estimated to comprise around 27% of lung cancer cases—are not targetable [3-5]. Nevertheless, patients with no actionable alterations may be treated with other therapies such as immunotherapy [6]. The treatment strategy for brain metastases arising from primary lung cancer should similarly be determined according to the genetic mutation status. However, brain metastases are usually small and can be scattered all over the brain, so it is not practical nor always feasible to invasively biopsy or surgically resect the metastases for molecular testing. As a result, most metastatic brain lesions are identified on magnetic

resonance (MR) imaging without pathological tissue confirmation. Therefore, it is prudent to develop non-invasive imaging-based methods to evaluate the mutation status of brain metastases in lung cancer patients [7].

Radiomic analysis, which involves the computerized extraction of data from clinically obtained medical images and it can be used to scrutinize spatially and temporally heterogeneous tumors comprehensively over multiple time points to monitor disease status and treatment-related changes [8]. Conventional imaging evaluation of brain metastatic lesions typically includes only lesion size and location, enhancing characteristics, and the extent of peritumoral edema. By contrast, radiomic analysis of clinically acquired scans can extract highly detailed characteristics regarding tumor texture, shape, and image intensity, which are not discernable to the human eye [9, 10]. Therefore, radiomic analysis is a potentially useful tool that could be used not only to evaluate brain metastases in lung cancer patients but also to identify genetic mutations to guide personalized treatment regimens.

Prior studies have demonstrated the ability of radiomic features to predict the mutation status of driver oncogenes [9, 11, 12]. Most studies in lung cancer were performed in primary lung cancer using computed tomography (CT) images. For example, Aerts and colleagues used features extracted from CT images to identify a prognostic radiomic signature associated with the gene expression profiles of patients with lung cancer and head-and-neck cancer [9]. Gevaert and colleagues developed a signature based on the CT images of primary lung cancer that could predict *EGFR* but not *KRAS* mutation status [11]. Liu and colleagues identified a set of five CTbased features that could be used to predict *EGFR* mutation status [12]. *However, some patients may present with neurological symptoms such as headache leading to brain MRI scans and subsequent imaging diagnosis of brain metastases prior to confirmation of their primary lung cancer. In addition, a complimentary non-invasive imaging-focused method to predict mutation status from brain metastases is desirable than relying solely on the primary lung cancer. Although lung cancer frequently metastasizes to the brain leading to poor prognosis* [1], *prior radiomic studies have mostly focused on the primary lung cancer rather than brain metastases to predict mutation status, leaving a gap in knowledge and a need for further study.*

There is limited information on the radiomic features of brain metastases. Existing radiomic studies of brain metastases have focused on developing models to distinguish between different primary cancers, such as lung cancer versus breast cancer or melanoma or between various lung cancer subtypes such as small cell lung carcinoma and NSCLC [13-16]. However, to the best of our knowledge, no studies have been published using radiomic approaches and predictive modeling to classify the mutation status of brain metastases in lung cancer patients. To address this need, we collected brain magnetic resonance (MR) scans obtained for clinical care of lung cancer patients with brain metastatic lesions, segmented the tumors, extracted radiomic features, and used machine learning algorithms to classify the mutation status from the brain metastases. We hypothesized that radiomic features extracted from MR images of brain metastases could be used to classify *EGFR*, *ALK*, and *KRAS* mutation status in patients with primary lung cancer.

2. Materials and methods

2.1. Patient selection

We obtained data from retrospectively identified consecutive patients with lung cancer treated at City of Hope National Medical Center (Duarte, CA) between 2009 and 2017 under a protocol approved by the Institutional Review Board of our hospital. The inclusion criteria were: diagnosis of lung cancer; confirmation of *EGFR*, *ALK*, or *KRAS* mutation; having brain metastasis; and having brain MR scans for the diagnosis of and prior to treatment for brain metastasis. Demographic and clinical information, including gender, age, race, smoking history, and metastatic sites, were retrieved from electronic medical records (EMR; Table 1). The mutation status of the patients was also retrieved from EMR. All patients in the study underwent genotype testing of their primary lung cancer at the discretion of their clinical providers. Tissue samples were obtained through lung biopsy or lung surgery and tested using several clinically available molecular testing platforms, such as fluorescence in situ hybridization (FISH), immunohistochemistry (IHC), and next-generation sequencing (NGS).

Data for a total of 138 lung cancer patients with brain metastasis with confirmed EGFR, ALK, and/or KRAS alteration status were retrieved from the database of patients treated at City of Hope National Medical Center between 2009 and 2017. Thirteen patients were excluded because their brain lesions were too small (smaller than 5 mm), three patients were excluded because the FLAIR sequence was not available in their brain MR scans, and 12 patients were excluded because the brain MR images were blurry due to head motion during image acquisition. Therefore, a total of 110 patients (mean age 57.51 ± 12.32 years; M: F=37:73; age range: 22–85) were included in the final study cohort (Table 1).

The study was approved by the institutional review board at City of Hope National Medical Center and informed consent was waived due to the retrospective nature of this study. The study was conducted in accordance with the Declaration of Helsinki.

2.2. Brain tumor segmentation

Pre-segmentation image registration was performed with both T1-weighted contrastenhanced (TIC) and T2-weighted fluid-attenuated inversion recovery (FLAIR) images affinely co-registered into the same geometric space using the elastix toolbox [17]. Image transformation and re-slicing were performed using FMRIB Software Library (FSL) scripts.

TIC and FLAIR images were used for segmentation of enhancing tumors and peritumoral edema, respectively. *Most of the brain MRI scans were performed on the same in-house 3T VERIO Siemens scanner (Siemens, Erlangen, Germany). For the T1C sequence, axial T1-weighted three-dimensional (3D) magnetization prepared rapid gradient echo (MPRAGE) imaging data were acquired with the following parameters: repetition time (TR)=1400 ms, echo time (TE)=2.91 millisecond (ms), imaging matrix=256 x 232, flip angle = 12, with a voxel size of 1 x 1 x 1 mm^3 in the axial, coronal, and sagittal planes. MultiHance® (gadobenate dimeglumine) at 0.1 mmol/Kg was administered as an intravenous injection for*

the T1C sequence. For the FLAIR sequence, the imaging parameters included TR=9000, TE=119, slice thickness=4 mm, acquisition matrix=256x204.

Delineation of tumor boundaries was performed in a semi-automated fashion on a slice-by-slice basis using ITK-SNAP, an open-source 3D image analysis software [18]. Tumors smaller than 5 mm in diameter were not segmented, as they were not reliably covered in two consecutive slices, which was the minimal requirement for the 3D approach. The segmented T1C and FLAIR images were reviewed for tumor delineation and consistency by two neuroradiologists (BTC with over 15 years of experience and TW with 8 years of experience). The delineated images of the two segmented tumor phenotypes (enhancing tumor and peritumoral edema) were exported for radiomic analysis. Among the segmented tumors, those smaller than 5x5x5 voxels were excluded from the final analysis due to unreliable feature extraction. For patients with more than 10 tumors, only the 10 largest tumors from each patient were included in the analysis. The entire dataset contained a total of 452 lesions from 110 patients. Fig. 1 presents the schema for brain tumor segmentation, radiomic feature extraction, and predictive modeling.

2.3. Radiomic feature extraction

Radiomic features related to the enhancing tumors and peritumoral edema were extracted from T1C and FLAIR images, respectively, using the corresponding tumor and edema masks. To emphasize specific features and to suppress noise, each image was pre-processed with 6 image filters (Wavelet (HH, HL, LL, HH, LLL), Laplacian of Gaussian (sigma=1, 2, 3), Square, Square Root, Logarithm, Local Binary Pattern, Gradient, and Exponential) before radiomic feature extraction. *Three categories of radiomic features were extracted from each image: (i) 18 intensity-based features (Minimum, Maximum, Mean, etc.); (ii) 13 shape-based features (Volume, Surface Area, Sphericity, etc.); and (iii) 74 textural features, including Gray Level Co-occurrence Matrix (GLCM, 23 features), Gray Level Run Length Matrix (GLRLM, 16 features), Gray Level Size Zone Matrix (GLSZM, 16 features), Neighboring Gray Tone Difference Matrix (NGTDM, 5 features), and Gray Level Dependence Matrix (GLDM, 14). A total of 2786 radiomic features were extracted from each tumor, with equal numbers numbers of features for the enhancing tumor from the T1C sequence and for the peritumoral edema from the FLAIR sequence. The radiomic features were extracted using the open-source python package Pyradiomics (https://github.com/Radiomics/pyradiomics) [19].*

To test the variability of radiomic features and to obtain stable features for the modeling, we randomly selected 20 patients and had their brain tumors independently segmented by the two neuroradiologists in this study (BTC and TW). We then used the interclass correlation coefficient (ICC) test to assess the consistency of the radiomic features between the two neuroradiologists in this study. Using ICC score larger than 0.8 (corrected p value <0.05) Three categories of radiomic features were extracted from each image: (i) 18 intensity-based features (Minimum, Maximum, Mean, etc.); (ii) 13 shape-based features (Volume, Surface Area, Sphericity, etc.); and (iii) 74 textural features, including Gray Level Co-occurrence Matrix (GLCM, 23 features), Gray Level Run Length Matrix (GLRLM, 16 features), Gray Level Size Zone Matrix (GLSZM, 16 features), Neighboring Gray Tone Difference Matrix

(NGTDM, 5 features), and Gray Level Dependence Matrix (GLDM, 14). A total of 2786 radiomic features were extracted from each tumor, with eq03C;0.05) as cut-off, 2520 stable features were obtained. We then performed feature selection with a minimum redundancy and maximum relevance (MRMR) algorithm and identified the 50 most relevant radiomic features from the 2520 radiomic features which were then included in the modeling process.

The pipeline for radiomic feature extraction consisted of 4 steps: (i) skull-stripping, (ii) bias field correction, (iii) image intensity normalization, and (iv) radiomic feature extraction. In the first step, both the Brain Extraction Tool (BET; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET) and Free Surfer (https://surfer.nmr.mgh.harvard.edu/) were used to remove non-brain tissue from the anatomical brain MR images. All skull-stripping results were then checked manually to ensure accuracy. The "N4 ITK Bias Field Correction" module of the nipype framework (https://nipype.readthedocs.io/en/0.12.0/index.html) was used to correct non-uniform intensity caused by field inhomogeneity. An intensity normalization algorithm was applied to standardize the intensity scales across MR images. *We used a python module NyulNormalizer (https://gitlab.com/eferrante/nyul/blob/master/NyulNormalizer.py) which implemented a method for intensity normalization as indicated in the references* [20-22]. This was performed to account for inter-scanner and within-sequence differences in imaging protocols. Finally, radiomic features were extracted using the open-source python package Pyradiomics [19].

2.4. Building predictive models for mutation status

Independent classifiers were built to classify the *EGFR*, *ALK*, and *KRAS* mutation status, as determined by genotype testing of the patients' primary lung cancers. To build each classifier, all 462 lesions from the 110 patients were divided into positive and negative groups. For example, to build the *EGFR* mutation status classifier, the positive group included lesions from all patients who were *EGFR* mutation positive, including the patient with double mutations. The lesions from the remaining patients were in the mutation negative group.

2.5. Feature selection and machine learning algorithm

Classifications were performed using Scikit-Learn software [23]. *Leave one out cross validation (LOOCV) was used to evaluate classification performance* [24]. *Briefly, for a dataset of N samples, the cross validation was performed in N round of training and testing. In each round, one sample in the dataset was successively taken out to test the model that was trained using the rest of the dataset (N-1). During the training process in each round, the 50 most relevant radiomic features were selected from a total of 2520 radiomic features using a minimum redundancy and maximum relevance (MRMR) algorithm* (https://github.com/fbrundu/pymrmr) [25], *and the synthetic minority over-sampling technique (SMOTE) was used to improve learning from imbalanced sample sizes* [26]. Clinical data, including demographic information, additional sites of metastases, and tumor information, were also used for the classification of each mutation. Demographic information included gender (male, female), race (Caucasian, Asian, other), and smoking history (yes, no). Clinical data also included clinical information such as the presence or absence of additional metastasis at 11 sites, including bone, lymph, liver, lung, kidney, pancreas, breast, spinal

cord, mediastinum, pericardium, and pleura. Tumor information included the number of tumors, the volume of the tumor core, and the edema/tumor volume ratio. In total, there were 18 items of clinical data. To select the most suitable machine learning algorithms for classifying mutation status, we tested 30 classifiers implemented in Scikit-Learn and ranked their performances according to the area under curve (AUC) of the receiver operating characteristic curve (ROC). The ROC curves were constructed using the class probability predicted by the classifiers. The area under the curve (AUC) was computed using Scikit-Learn. Other classification performance metrics including accuracy, sensitivity, and specificity were computed based on the prediction results of the classifiers. A total of four classifiers (random forest, extra tree, bagging, and gradient boosting), which were all ensemble learning-based classifiers, ranked among the top three classifiers for at least one of the three predictive models (Table 1S). The random forest classifier was ultimately selected as the most suitable algorithm because it was among the top three classifiers for all three predictive models and this classifier has been widely used in the field [27]. *The random forest classifier was selected based on the performance using radiomic features only. The same classifier was used for building the model using both radiomic features and clinical data. We also made the nomograms based on the top 20 radiomic features selected by the Random Forest Algorithm and the nomograms were included in the* supplementary file (Fig. 1S).

2.6. Statistical analysis

We used the statistical analysis package in the SciPy: Open Source Scientific Tools for Python library (https://www.scipy.org/) to analyze patient and tumor information (Table 1). Analysis of variance (ANOVA) test was used to determine the statistical significance of differences in the means of age among the three patient groups. Fisher's exact test was used to determine the statistical significance of group differences in the distributions of the categorical variables, including gender, race, smoking history, histology, and other metastatic sites. P-values were two-sided and values less than 0.05 were considered statistically significant.

3. RESULTS

3.1. Patient information

Of the 110 patients included in the final study cohort (Table 1), 75 were *EGFR* mutation-positive, 21 were *ALK* translocation-positive, and 15 were KRAS mutation-positive. One patient was positive for both *ALK* and *EGFR* alterations.

There were statistically significant group differences in racial distribution (p<0.05) and smoking history (p<0.001). Pairwise comparisons revealed that the *KRAS* group had a significantly greater percentage of Caucasian patients and fewer Asian patients (p=0.005; Table 1). Pairwise comparisons also revealed that the *KRAS* group had a higher percentage of smokers than the *EGFR* (p=0.0002) and *ALK* (p=0.0019) groups.

We performed comparisons of the demographic data between the mutation positive group and the mutation negative groups for each gene mutation, i.e. EGFR (+) versus EGFR (−),

ALK (+) versus ALK (−), and KRAS (+) versus KRAS (−). We found that the EGFR (+) group had a significantly greater percentage of Asian patients than the EGFR (−) group (p=0.042). The KRAS (+) group in our cohort was significantly older than the KRAS (−) group (p=0.002) and had a higher percentage of smokers than the KRAS (−) group (p= 0.0001) in our cohort.

We also performed feature ranking of the clinical data based on the feature importance assessed by the Random Forest Classifier. For EGFR mutation, the top 5 clinical data in decreasing order were Metastatic Sites (Lung), Metastatic Sites (Lymph Nodes), Age at Diagnosis, Metastatic Sites (Liver) and Smoker_Yes. For ALK mutation, the top 5 clinical data in decreasing order were Age at Diagnosis, Race_Asian, Metastatic Sites (Lung), Metastatic Sites (Lymph Nodes), Race_Caucasian. For KRAS, the top 5 clinical data in decreasing order were Smoker_Yes, Age at Diagnosis, Metastatic Sites (Lung), Metastatic Sites (Lymph Nodes), and Race_Caucasian.

3.2. Classification of mutation status

Classification using 50 most relevant radiomic features alone to classify molecular mutation status was better than using the18 items of clinical data alone, and combining radiomic features and clinical data was better than using either alone (Fig. 2). The AUCs for the classification of *EGFR* mutation status were 0.847, 0.609, and 0.912 when using radiomic features alone, clinical data alone, and radiomic feature and clinical data combined, respectively. The AUCs for the classification of *ALK* alteration status were 0.813, 0.603 and 0.915 when using radiomic features alone, clinical data alone, and radiomic feature and clinical data combined, respectively. The AUCs for the classification of *KRAS* mutation status were 0.938, 0.684, and 0.985 when using radiomic features alone, clinical data alone, and radiomic feature and clinical data combined, respectively.

Table 2 presents the performance data, including the accuracy, AUC, sensitivity, and specificity, for the classification of *EGFR*, *ALK*, and *KRAS* alteration status using both radiomic features and clinical data. *The 'number of lesions' in* Table 2 *indicated the lesions for each positive mutation group, i.e., EGFR, or ALK, or KRAS. In the classification of each mutation status, the total number of 462 lesions was divided into two classes for each mutation, those positive to the particular mutation versus the rest of the lesions. The total number of lesions was the sum of all lesions (n=462) including the EGFR positive mutation (n=335), ALK positive mutation (n=71) and KRAS positive mutation (n=56).*

The top 20 radiomic features for the classification of *EGFR*, *ALK*, and *KRAS* mutation status with ranking based on random forest feature selection, are presented in the heat maps in Fig. 3 and Table 2S in the supplementary file. The majority of the top 20 radiomic features were textural features.

4. DISCUSSION

In this study, we extracted radiomic features from MR images of brain metastatic lesions and used the radiomic features, as well as clinical data, to build machine learning models for classification of molecular alteration status of the three most common oncogenes in lung

cancer patients, i.e., *EGFR*, *ALK*, and *KRAS*. Our study showed that the MR imaging-based radiomic analysis of brain metastases could potentially serve as a non-invasive technique to classify *EGFR*, *ALK*, and *KRAS* mutations in lung cancer patients.

Our work was novel in that we analyzed MR images of brain metastatic lesions using a radiomic approach to classify the mutation status of the primary lung cancer, whereas prior studies focused on using radiomics to differentiate brain metastases from various primary cancers or from different lung cancer subtypes. For example, Li and colleagues used quantitative radiomic features extracted from MR images of brain metastatic lesions to predict the pathological subtypes of the originating lung cancers [15]. They achieved misclassification rates of 3.1%, 4.3%, 5.8%, and 8.1%, for small cell lung carcinoma, squamous cell carcinoma, adenocarcinoma, and large cell lung carcinoma, respectively. Bérésova and colleagues demonstrated the ability of textural radiomic analysis of MR images to differentiate between brain metastases originating from lung cancer versus breast cancer [13]. Ortiz-Ramon and colleagues used radiomic features from MR images of brain metastatic lesions to predict whether the primary cancer type was lung cancer or melanoma [14]. Their best predictive model (a naïve Bayes algorithm) achieved AUC of 0.947 ± 0.067.

Our work is relevant for clinical care because identifying the mutation status of the common oncogenes may help to determine the prognosis and treatment strategy for patients with lung cancer. On the one hand, patients with common lung cancer driver mutations, such as mutations of *EGFR* and translocations of *ALK*, are particularly prone to developing brain metastases [28]. Therefore, identification of these mutations may help to consult the patients about the likelihood of developing brain metastases and thus poor prognosis. On the other hand, tyrosine kinase inhibitors and ALK inhibitors can significantly increase the progression-free and overall survival of patients who have *EGFR* and *ALK* alterations, respectively [7]. Knowledge of mutation status may help to devise a targeted therapy. Although the mutation status of the patients in our study was obtained from the primary lung cancer, our radiomic study of mutations focusing on the brain metastases should still have merit. First of all, about 10% of NSCLC patients presents with brain metastases upon initial diagnosis of lung cancer [1, 2]. These patients may not have had their molecular genotype testing performed on the biopsy or resection specimen of the primary lung cancer yet. Therefore, a radiomic analysis of their brain metastases may potentially contribute useful information on the mutation status of their primary lung cancer. Second, the mutation status of lung cancer may not always available in clinical practice. It is conceivable that individualized treatment plan with targeted therapy may be devised if the mutation status of the primary lung cancer could be inferred from the radiomic analysis of brain metastases. Third, brain MR scans are non-invasive and can be readily obtained in a much easier way than invasive biopsy or surgery of either the primary lung cancer or the brain metastases. In addition, brain MR scans can be performed at multiple time points throughout the treatment while repeated invasive biopsies or surgical resections might not be practical or feasible even in the event of non-diagnostic biopsies.

We used the leave one out cross validation (LOOCV) method to evaluate classification performance for the following reasons. First, we chose multiple candidates in the first round of model selection based on LOOCV. Our top 3 algorithms for the prediction of each of the

three mutation status are all decision-tree based ensemble learners. Second, the same algorithms were also ranked as top performers in the published research evaluating the performance of machine learning algorithms trained with radiomic features [29, 30]. Third, we intended to choose an approach that would most reliably evaluate the classification performance of each classifier, especially in terms of generalization error [31]. For performance estimation purpose, Molinaro and colleagues have suggested that LOOCV performed well with regard to Mean Square Error and bias. As pointed out by Arlot and colleagues [24], LOOCV was a better risk estimator than the K-fold CV. Zhang and colleagues also pointed out that the 10-fold CV was riskier than LOOCV for predictive performance estimation [32]. Lastly, we chose LOOCV because it used the most probable portion of the dataset to train the algorithms for our modest sample size. These results may provide us the performance metrics of the algorithms closer to their potential performance for a sufficiently larger data size, which may inform on our future studies.

Despite considerable progress in recently reported studies using radiomics to determine mutation status [12, 29, 33], the performance, reliability, and reproducibility of this approach have not reached acceptable standards for clinical practice. Most studies in the field have used radiomic features extracted from lung CT images. The highest reported AUCs for predicting *EGFR*, *ALK*, and *KRAS* mutation status based on lung CT images were 0.89, 0.649, and 0.667, respectively [11, 12, 29, 33]. To develop more reliable and accurate predictive models, it is necessary to evaluate radiomic features extracted from various imaging sources including both the primary cancer and the metastatic lesions. In addition, it is also important to optimize modeling process by using a combination of clinical, pathological and imaging features, and by testing a wide range of machine learning algorithms. The performance of our predictive models based on radiomic features extracted from brain MR images compared favorably to the published results using data extracted from lung CT images [11, 12, 29, 33]. Importantly, our study cohort was similar to those used in previously published studies, enabling a reasonable comparison of our results to others. For example, there was a significantly higher percentage of smokers in our *KRAS* mutation-positive group than in our *EGFR* and *ALK* alteration-positive groups, which was similar to the study by Sweis and colleagues [34]. Our models for classifying *EGFR*, *ALK*, and *KRAS* mutation status achieved AUC values of 0.917, 0.916, and 0.986, respectively. The modeling algorithms used in our models for feature selection (MRMR) and machine learning (random forest algorithm) were determined to be the top performing algorithms in a study by Parmar and colleagues, who systematically compared the prognostic performances of various combinations of feature selections and machine learning algorithms [30]. Our study suggests that combining radiomic features extracted from MR images with clinical data will improve predictive modeling performance and reliability.

There were several limitations to this study. First, this was a retrospective study based on MR imaging of brain metastases in lung cancer patients identified through a database search of patients treated at our institution over a 9-year interval. Some confounding variables, such as the imaging protocols, the MR scanners used, and treatment regimen, could not be properly controlled for. For example, we were not able to control the variation in contrast administration (e.g., route and timing of contrast injection) or the timing of image acquisition after contrast administration. Furthermore, the patients in our study cohort

received a wide range of treatments, as lung cancer treatment has evolved significantly over the last nine years, and we did not have the statistical power to control for the effects of different treatment regimens on the predictive modeling of mutation status. Second, although our study might be one of the larger retrospective studies performed in brain metastases using a radiomic method, with a total of 110 patients, this sample size is still modest. This limitation was compounded by the fact that the three mutation groups were unevenly distributed. Whereas there were 75 patients in the *EGFR* mutation-positive group, there were only 21 in the *ALK* mutation-positive group and 15 in the *KRAS* mutation-positive group. To address this issue, we used the SMOTE to improve learning based on imbalanced sample sizes [26]. Third, genetic mutation status was obtained from the primary lung cancer through biopsy or surgical resection because there was no tissue confirmation for most of the brain metastatic lesions for which the radiomic features were extracted. Therefore, we could not detect or account for potentially rare cases in which the primary lung cancer and brain metastases from the same patient may have different mutations. Lastly, the current study was performed on data collected from a single institution with no validation of external cohorts, which limited its generalizability. Our study approach needs to be validated in the setting of a multi-center study with larger cohorts across different institutions. Nevertheless, the results reported here may serve as preliminary data to support future prospective studies using machine learning algorithms to classify mutation status of patients with brain metastases. To the best of our knowledge, our study is the first to use MR imaging-based radiomic analysis of brain metastatic lesions and machine learning algorithms to classify mutation status of patients with primary lung cancer.

In summary, our study showed that a radiomic approach capturing the critical radiological features of brain metastases from patients with lung cancer could be used to differentiate the three common mutations, i.e., *EGFR*, *ALK* and *KRAS*. The detection of mutation status using this approach may be useful for informing treatment strategies and prognosis for lung cancer patients with brain metastases. Future studies should be performed with a larger sample size and external cohorts to validate our results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

11.   REFERENCES

[1]. Nolan C, Deangelis LM. Overview of metastatic disease of the central nervous system. Handb Clin Neurol 2018;149:3–23. [PubMed: 29307359]

[2]. Ali A, Goffin JR, Arnold A, Ellis PM. Survival of patients with non-small-cell lung cancer after a diagnosis of brain metastases. Curr Oncol 2013;20(4):e300–6. [PubMed: 23904768]

[3]. Soria JC, Ohe Y, Vansteenkiste J, Reungwetwattana T, Chewaskulyong B, Lee KH, et al. Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. N Engl J Med 2018;378(2):113–25. [PubMed: 29151359]

[4]. Consortium APG. AACR Project GENIE: Powering Precision Medicine through an International Consortium. Cancer Discov 2017;7(8):818–31. [PubMed: 28572459]

[5]. Mak KS, Gainor JF, Niemierko A, Oh KS, Willers H, Choi NC, et al. Significance of targeted therapy and genetic alterations in EGFR, ALK, or KRAS on survival in patients with non-small cell lung cancer treated with radiotherapy for brain metastases. Neuro Oncol 2015;17(2):296–302. [PubMed: 25053852]

[6]. Brahmer J, Reckamp KL, Baas P, Crino L, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med 2015;373(2):123–35. [PubMed: 26028407]

[7]. Di Lorenzo R, Ahluwalia MS. Targeted therapy of brain metastases: latest evidence and clinical implications. Ther Adv Med Oncol 2017;9(12):781–96. [PubMed: 29449898]

[8]. Kuo MD, Jamshidi N. Behind the numbers: Decoding molecular phenotypes with radiogenomics--guiding principles and technical considerations. Radiology 2014;270(2):320–5. [PubMed: 24471381]

[9]. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature communications 2014;5:4006.

[10]. 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(4):441–6. [PubMed: 22257792]

[11]. Gevaert O, Echegaray S, Khuong A, Hoang CD, Shrager JB, Jensen KC, et al. Predictive radiogenomics modeling of EGFR mutation status in lung cancer. Sci Rep 2017;7:41674. [PubMed: 28139704]

[12]. 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(5):441–8 e6. [PubMed: 27017476]

[13]. Béresová M, Larroza A, Arana E, Varga J, Balkay L, Moratal D. 2D and 3D texture analysis to differentiate brain metastases on MR images: proceed with caution. Magnetic Resonance Materials in Physics, Biology and Medicine 2018;31(2):285–94.

[14]. Ortiz-Ramon R, Larroza A, Arana E, Moratal D. A radiomics evaluation of 2D and 3D MRI texture features to classify brain metastases from lung cancer and melanoma. Conf Proc IEEE Eng Med Biol Soc 2017;2017:493–6.

[15]. Li Z, Mao Y, Li H, Yu G, Wan H, Li B. Differentiating brain metastases from different pathological types of lung cancers using texture analysis of T1 postcontrast MR. Magn Reson Med 2016;76(5):1410–9. [PubMed: 26621795]

[16]. Nardone V, Tini P, Biondi M, Sebaste L, Vanzi E, De Otto G, et al. Prognostic Value of MR Imaging Texture Analysis in Brain Non-Small Cell Lung Cancer Oligo-Metastases Undergoing Stereotactic Irradiation. Cureus 2016;8(4):e584. [PubMed: 27226944]

[17]. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2010;29(1):196–205. [PubMed: 19923044]

[18]. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31(3):1116–28. [PubMed: 16545965]

[19]. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res 2017;77(21):e104–e7. [PubMed: 29092951]

[20]. Nyul LG, Udupa JK. On standardizing the MR image intensity scale. Magn Reson Med 1999;42(6):1072–81. [PubMed: 10571928]

[21]. Shah M, Xiao Y, Subbanna N, Francis S, Arnold DL, Collins DL, et al. Evaluating intensity normalization on MRIs of human brain with multiple sclerosis. Med Image Anal 2011;15(2):267–82. [PubMed: 21233004]

[22]. Shinohara RT, Sweeney EM, Goldsmith J, Shiee N, Mateen FJ, Calabresi PA, et al. Statistical normalization techniques for magnetic resonance imaging. Neuroimage Clin 2014;6:9–19. [PubMed: 25379412]

[23]. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. Journal of machine learning research 2011;12(Oct):2825–30.

[24]. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statist Surv 2010;4:40–79.

[25]. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of maxdependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 2005;27(8):1226–38. [PubMed: 16119262]

[26]. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research 2002;16:321–57.

[27]. Sarica A, Cerasa A, Quattrone A. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review. Front Aging Neurosci 2017;9:329. [PubMed: 29056906]

[28]. Rangachari D, Yamaguchi N, VanderLaan PA, Folch E, Mahadevan A, Floyd SR, et al. Brain metastases in patients with EGFR-mutated or ALK-rearranged non-small-cell lung cancers. Lung Cancer 2015;88(1):108–11. [PubMed: 25682925]

[29]. Zhang L, Chen B, Liu X, Song J, Fang M, Hu C, et al. Quantitative Biomarkers for Prediction of Epidermal Growth Factor Receptor Mutation in Non-Small Cell Lung Cancer. Transl Oncol 2018;11(1):94–101. [PubMed: 29216508]

[30]. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJ. Machine Learning methods for Quantitative Radiomic Biomarkers. Sci Rep 2015;5:13087. [PubMed: 26278466]

[31]. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. Bioinformatics 2005;21(15):3301–7. [PubMed: 15905277]

[32]. Zhang Y, Yang Y. Cross-validation for selecting a model selection procedure. Journal of Econometrics 2015;187(1):95–112.

[33]. 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 2016;26(1):32–42. [PubMed: 25956936]

[34]. Sweis RF, Thomas S, Bank B, Fishkin P, Mooney C, Salgia R. Concurrent EGFR Mutation and ALK Translocation in Non-Small Cell Lung Cancer. Cureus 2016;8(2):e513. [PubMed: 27026837]
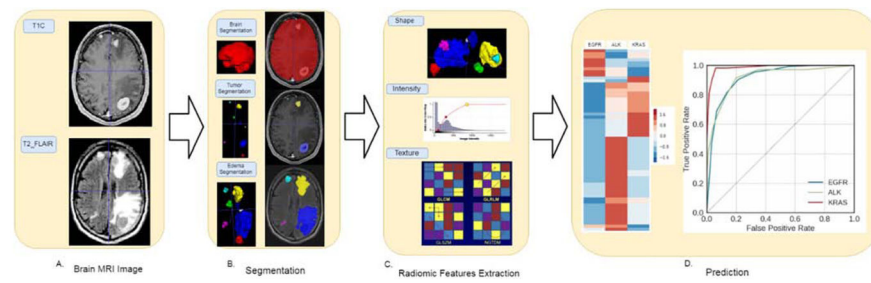
Fig. 1.

Schema for brain tumor segmentation and radiomic feature extraction. **(A)** Representative images of post contrast T1-weighted (T1C) and fluid-attenuated inversion recovery (FLAIR) images. **(B)** Segmented enhancing tumor and peritumoral edema masks (top panel) overlapping the T1C images (middle panel) and FLAIR images (bottom panel). **(C)** Illustrations of the three categories of radiomic features extracted from the images, including shape, intensity, and texture. GLCM: Gray Level Co-occurrence Matrix; GLRLM: Gray Level Run Length Matrix; GLSZM: Gray Level Size Zone Matrix; NGTDM: Neighboring Gray Tone Difference Matrix. **(D)** The left panel indicates the top 50 radiomic features used for classifying each of the three types of mutation status. The right panel shows the receiver operating characteristic curves for classifying the three types of mutation status.
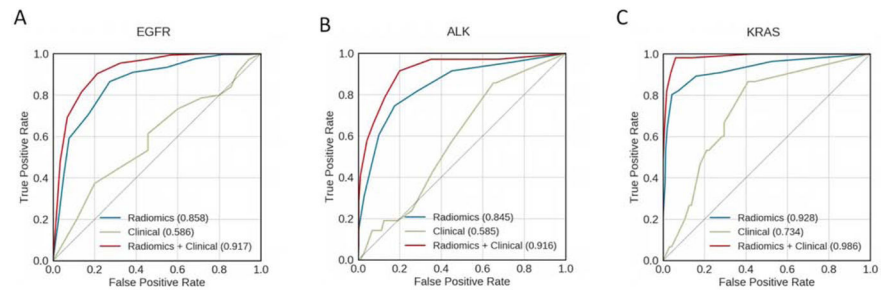
Fig. 2.
Receiver operating characteristic curves for classifying **(A)** *EGFR*, **(B)** *ALK*, and **(C)** *KRAS*
mutation status. Curves are shown for models using clinical data only (green), radiomic
features only (blue), and a combination of both clinical data and radiomic features (red). The
areas under the receiver operating characteristic curves (AUCs) are indicated in each panel.
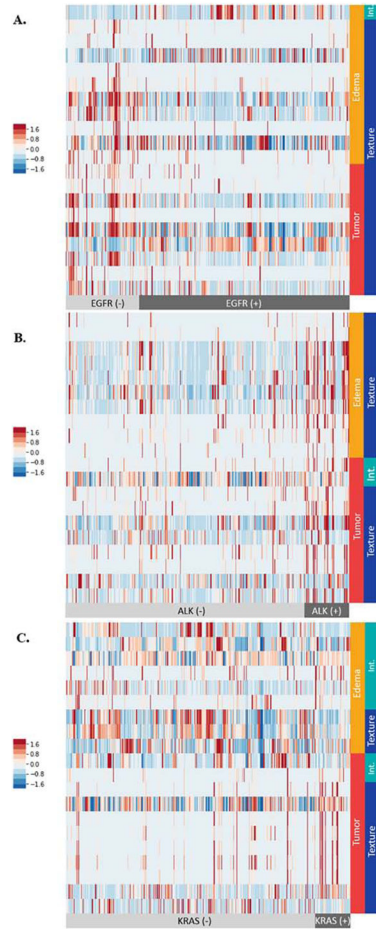
Fig. 3.
Heat maps showing the top 20 radiomic features for classifying **(A)** *EGFR*, **(B)** *ALK*, and **(C)** *KRAS* mutation status. Each row represents one of the top 20 radiomic features determined by random forest feature selection. Each column represents a lesion. The mutation status of the patients is indicted by light gray (mutation negative) and black (mutation positive) color bars at the bottom of each panel. The tumor delineations from which the radiomic features were extracted are indicated by vertical red (enhancing tumor) and orange (peritumoral edema) color bars in the right side of each panel. The categories of radiomic features are indicated by vertical blue (textural) and green (intensity-based, abbreviated: Int.) color bars. The normalized values for each radiomic feature are indicated according to the color scale on the left side of each panel. The detailed description for each of the top 20 radiomic features for each panel is listed in Table 2S in the supplementary file.

Table 1

Demographic and clinical information for the study cohort

	EGFR (+) N = 75	*ALK* (+) N = 21	*KRAS* (+) N = 15	p
Gender				
Male	24 (32%)	8 (38.10%)	5 (33.33%)	0.83
Female	51 (68%)	13 (61.90%)	10 (66.67%)	
Age				
Mean ± SD	57.43 ± 12.09	53.81 ± 14.79	63.67 ± 6.40	0.09
Race				
Asian	35 (46.67%)	7 (33.33%)	1 (6.67%)	
Caucasian	34 (45.33%)	13(61.90%)	11 (73.33%)	0.016
Other *	6 (8%)	1 (0.04%)	3 (20%)	
History of Smoking				
Yes	20 (26.67%)	5 (23.80%)	12 (80%)	<0.001
No	55 (73.33%)	16 (76.19%)	3 (20%)	
Histology				
Adenocarcinoma	71 (94.67%)	21 (100%)	13 (86.67%)	0.50
Other **	4 (5.33%)	0 (0%)	2 (13.33%)	
Other Metastatic Sites				
Bones	45 (60%)	9 (34.62%)	4 (26.67%)	
Lymph Nodes	27 (36%)	4 (15.38%)	3 (20%)	
Lung	26 (34.67%)	4 (15.38%)	1 (6.67%)	0.99
Liver	22 (29.33%)	5 (19.23%)	2 (13.33%)	
Pleura	13 (17.33%)	2 (7.69%)	1 (6.67%)	
Other ***	9 (12%)	0 (0%)	0 (0%)	

*
 American Indian or Alaska Native, African American, Native Hawaiian, or Pacific Islander

**
 Squamous cell lung carcinoma, adenosquamous cell lung cancer, lung carcinosarcoma, non-small cell lung carcinoma (NOS)

***
 Kidneys, mediastinum, pancreas, pericardium

Table 2

Performance data for *EGFR*, *ALK*, and *KRAS* mutation classification models using both clinical data and radiomic features.

Mutation	Accuracy	AUC[*]	Sensitivity	Specificity	Number of lesions
EGFR	77.7%	0.912	73.1%	90.6%	335
ALK	86.7%	0.915	76.1%	88.7%	71
KRAS	96.7%	0.985	89.3%	97.7%	56

[*] AUC: area under the receiver operating characteristic curve

Note: Number of lesions indicates the lesions for each positive gene mutation, i.e., EGFR, or ALK, or KRAS. In the classification of each mutation status, the total number of 462 lesions were divided into two classes for each mutation, those positive to the particular mutation versus the rest of the lesions.