

RESEARCH ARTICLE

Open Access



MR-based radiomics predictive modelling of EGFR mutation and HER2 overexpression in metastatic brain adenocarcinoma: a two-centre study

Yanran Li^{1†}, Yong Jin^{2†}, Yunling Wang¹, Wenya Liu¹, Wenxiao Jia¹ and Jian Wang^{1*}

Abstract

Objectives Magnetic resonance (MR)-based radiomics features of brain metastases are utilised to predict epidermal growth factor receptor (EGFR) mutation and human epidermal growth factor receptor 2 (HER2) overexpression in adenocarcinoma, with the aim to identify the most predictive MR sequence.

Methods A retrospective inclusion of 268 individuals with brain metastases from adenocarcinoma across two institutions was conducted. Utilising T1-weighted imaging (T1 contrast-enhanced [T1-CE]) and T2 fluid-attenuated inversion recovery (T2-FLAIR) sequences, 1,409 radiomics features were extracted. These sequences were randomly divided into training and test sets at a 7:3 ratio. The selection of relevant features was done using the least absolute shrinkage selection operator, and the training cohort's support vector classifier model was employed to generate the predictive model. The performance of the radiomics features was evaluated using a separate test set.

Results For contrast-enhanced T1-CE cohorts, the radiomics features based on 19 selected characteristics exhibited excellent discrimination. No significant differences in age, sex, and time to metastasis were observed between the groups with EGFR mutations or HER2+ and those with wild-type EGFR or HER2 ($p > 0.05$). Radiomics feature analysis for T1-CE revealed an area under the curve (AUC) of 0.98, classification accuracy of 0.93, sensitivity of 0.92, and specificity of 0.93 in the training cohort. In the test set, the AUC was 0.82. The 19 radiomics features for the T2-FLAIR sequence showed AUCs of 0.86 in the training set and 0.70 in the test set.

Conclusions This study developed a T1-CE signature that could serve as a non-invasive adjunctive tool to determine the presence of EGFR mutations and HER2+ status in adenocarcinoma, aiding in the direction of treatment plans.

Clinical relevance statement We propose radiomics features based on T1-CE brain MR sequences that are both evidence-based and non-invasive. These can be employed to guide clinical treatment planning in patients with brain metastases from adenocarcinoma.

Keywords T1-CE brain MR sequences, Adenocarcinoma, Brain metastasis, Radiomics, Gene expression

[†]Yanran Li and Yong Jin contributed equally to this work.

*Correspondence:

Jian Wang
jianwang1634@163.com

¹Department of Radiology, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830054, China

²Department of Radiology, Changzhi People's Hospital, Changzhi 046000, Shanxi Province, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Cell surface receptor tyrosine kinases, such as the epidermal growth factor receptor (EGFR) and human epidermal growth factor receptor 2 (HER2), dimerise with other HER family receptors to transmit growth signals [1]. Many common cancers often involve the activation, overexpression, or mutation of EGFR and HER2. When EGFR and HER2 are overexpressed in tumour cells, these cells multiply rapidly [2]. Numerous solid tumours have been found to harbour mutant forms of EGFR, with these mutations being associated with tumour progression, drug resistance, and patient survival [3]. In certain individuals with various solid tumours, HER2 is overexpressed [4], substantially predicting poorer overall survival. Recent studies have identified HER2 kinase domain mutations in 4% of all primary lung cancer tumours and 10% of those with adenocarcinoma histology [5]. As a result, timely determination of EGFR mutations and HER2 overexpression is crucial for predicting therapeutic response and determining specific treatment plans.

The currently accepted method for determining genetic status is biopsy. However, its highly invasive nature often precludes the collection of tissue samples from the main tumour or brain metastases. The size of the lesion, degree of intratumoral cystic necrosis, degree of paratumoral oedema, and degree of imaging signal are all characteristics of brain metastasis that may be effectively reflected by magnetic resonance (MR) imaging [6, 7]. Nonetheless, MR imaging is unable to predict alterations at the molecular level. Therefore, precise, non-invasive quantitative approaches are required to provide a real-time complement to histologic evaluation and represent intratumoral heterogeneity.

Existing brain metastasis radiomic research has primarily focused on building models to differentiate between various primary malignancies or lung cancer subtypes [8–11]. However, there has been limited information available on the radiomic characteristics of brain metastases. Brain metastases might offer a genetic opportunity to select patients for specific treatments.

This study explores the potential of radiomics in analysing EGFR and HER2 gene statuses in patients with brain metastases. It aims to create a relevant model using radiomics for this purpose and verify the model's validity. Such a model could be clinically guided by radiomics for treatment planning in patients with brain metastases from adenocarcinoma.

Methods and materials

Study content

This study aims to utilise T1 contrast-enhanced (T1-CE) and T2 fluid-attenuated inversion recovery (T2-FLAIR) radiomic features to predict EGFR mutations and HER2 overexpression in adenocarcinoma in a personalised

manner. Radiological features were derived from segmented regions of the MRI images. The collected data were divided into training and test sets according to the experimental design and were subsequently used for training and independent validation of the model, respectively. The most valuable radiological features were selected using the least absolute shrinkage and selection operator (LASSO) technique. The prediction model was developed using the support vector classifier (SVC) method. A detailed description and analysis workflow can be seen in Fig. 1.

Participants

Between January 2015 and January 2022, this two-centre retrospective cohort study was authorised by our hospital's ethics committee. Patients were initially diagnosed with brain metastases resulting from cancer using MR imaging and genetic testing. Pathology confirmed the diagnosis of all primary lesions. All MR-confirmed metastases were reported independently by nuclear medicine specialists and radiologists with at least five years of board certification.

To be included in the study, patients had to meet the following criteria: (1) primary tumour pathologically confirmed as adenocarcinoma by surgery or biopsy, (2) presence of brain metastases, (3) brain metastases confirmed by pathology, and (4) assessment using MR imaging. The exclusion criteria were as follows: (1) various types of original malignancies, (2) absence of genetic testing, (3) patients who had received systemic therapy, and (4) lack of follow-up data.

EGFR mutation status and overexpressed HER2

All adenocarcinoma patients underwent an evaluation of their EGFR status using next-generation sequencing technologies.

The HER2 status of all patients with adenocarcinoma was assessed using immunohistochemistry (IHC) or fluorescence in situ hybridisation (FISH) [12]. Two independent pathologists reviewed the patients' IHC results at the time of the HER2 status assessment. Patients were considered negative for HER2 if IHC results were 0/1+ and positive if IHC results were 3+. HER2 status was further determined using FISH results when IHC results were 2+ [12, 13].

MR image acquisition

Brain MR images were obtained from patients with adenocarcinoma at their initial diagnosis of brain metastases. At the First Affiliated Hospital of Xinjiang Medical University, patients underwent evaluation using a 3.0 T MR scanner (MAGNETOM Skyra, Siemens Healthineers, USA) and a 3.0 T MR scanner (Discovery MR750, GE Healthcare, USA). At Changzhi People's Hospital,

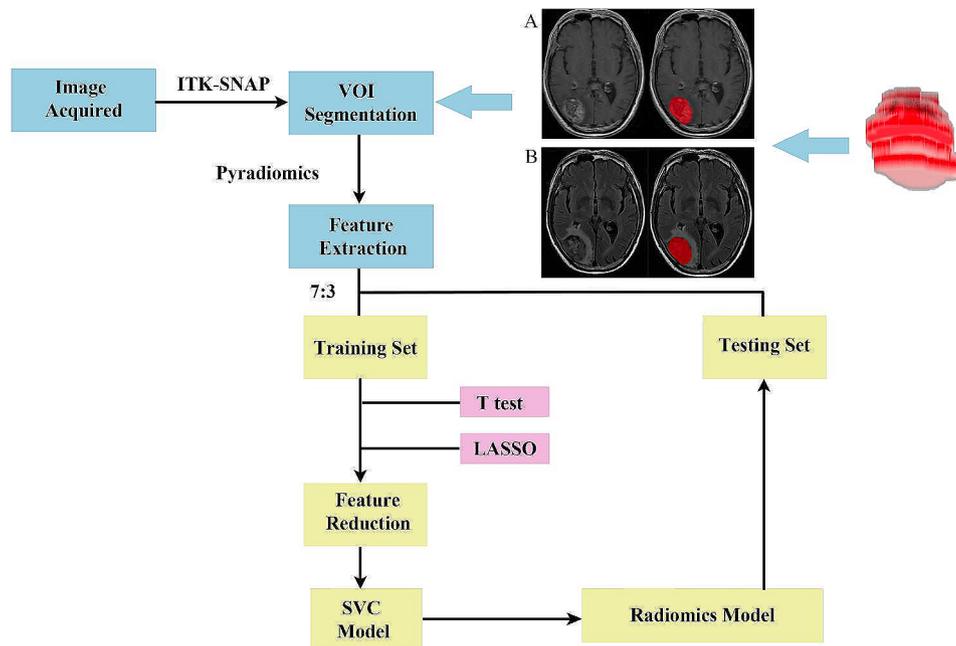


Fig. 1 Workflow for radiomics modeling and analysis including segmentation of metastases and examples from MR images

a 1.5 T MR scanner and a 3.0 T MR scanner (MAGNETOM Verio, Siemens Healthineers, USA) were employed. The T1-CE and T2-FLAIR sequences, which are commonly used for detecting brain metastases, were chosen for the extraction of image features [7, 14–17]. The image matrix size varied from 240×240 to 256×256 pixels, with slice thickness ranging from 1 mm to 6 mm. Additional scanning parameters for these sequences are presented in Supplementary Table 1.

Image segmentation

Regions of interest (ROIs) on the MR images of the aforementioned sequences were delineated using ITK-SNAP [18] around the metastases. To minimise the impact of clustering, only one ROI per image was selected. Initially, ROIs were created around each axial section of the metastatic profile to examine tumour heterogeneity. Subsequently, areas of oedema, haemorrhage, cystic changes, and necrosis were excluded from the images. Representative sections of the metastases for each sequence were chosen. The red spots in Fig. 1 indicate the axial distribution of the metastases.

Radiomics features extraction

Before feature extraction, several preprocessing steps were undertaken to enhance texture recognition. First, the bin width was set to 25. Second, Z-score normalisation was applied to the MR images. Finally, cubic interpolation was utilised to resample the ROIs isotropically to an in-plane resolution of $3 \times 3 \times 3$ mm, ensuring

the consistency of proportions and orientation in the acquired 3D features [19].

Prior to radiomic feature extraction, each image underwent preprocessing with eight image filters to highlight specific details and reduce noise. These filters, comprising Wavelet (HH, HL, LL, LHH, LLL), Laplacian of Gaussian ($\sigma=1, 2, 3$), Square, Square Root, Logarithm, Local Binary Pattern, Gradient, and Exponential, were used to evaluate three different categories of radiomic features in each image. A total of 1,409 radiomic features were extracted per tumour, with an equivalent number of features generated from both T1-CE and T2-FLAIR sequences. The radiomic features were extracted using the open-source Python program Pyradiomics (<https://github.com/Radiomics/pyradiomics>). Various radiomic features were utilised for this extraction process.

Establishment of radiomics features

Data from Institutions 1 and 2 were combined and subsequently divided into training (70%) and test (30%) sets after manually segmenting and extracting features from all ROIs for each participant. A PyRadiomics-based pipeline, previously tested, was used to extract radiomics features and conduct feature selection. Among the radiomics features retrieved, there were several duplicated, unstable, and irrelevant imaging characteristics. Feature selection methods were used to identify and select the most informative, stable radiomics features. This approach also helps prevent overfitting. The final feature set was established using the LASSO method and recursive feature elimination.

Model building

A radiomics model was constructed using a linear SVC based on these two MR image sequences after feature selection. To address the data imbalance, the synthetic minority oversampling technique algorithm with default parameters was applied [20]. This created minority instances along a line connecting a minority instance to its nearest neighbours.

Interobserver reproducibility evaluation

In our radiomics inquiry, ROI segmentation repeatability and radiomics feature extraction reproducibility are integral to the interobserver reproducibility process. The repeatability of the ROI segmentation was tested through multi-reader segmentation. For this purpose, two independent radiologists manually delineated ROIs in four MR sequences using the same annotation program, ITK-SNAP (<http://www.itksnap.org>). For reproducible radiomics feature extraction, we employed PyRadiomics, an open-source program, to compute the radiomics features and to perform image preprocessing. The feature extraction process was guided by a.yaml file, allowing separate individuals to replicate the extraction of radiomics features outlined in our study using the same file.

Please include the supplier's information.

Statistical analysis

SPSS (version 22.0) was utilised for the statistical analysis. To explore group differences, we applied chi-squared tests to categorical variables and t-tests for continuous data. The model's effectiveness was assessed using the receiver operating characteristic (ROC) curve and the area under the curve (AUC). Sensitivity and specificity metrics were calculated using a threshold of 0.5. A p-value of less than 0.05 was deemed statistically significant.

Results

Patients' information

Between January 2015 and January 2022, we conducted a retrospective analysis of 513 patients with brain metastasis from two institutions (Institution 1=272, Institution 2=241). In Institution 1, 81 patients and in Institution 2, 69 patients with brain metastasis were excluded as these metastases did not originate from adenocarcinoma. Additionally, 52 patients in Institution 1 and 43 in Institution 2 were excluded due to a lack of genetic testing on these brain metastases. Ultimately, 268 patients with documented T1-CE sequences in hospital archives were enrolled across both institutions (Fig. 2). The gene expression features of these 268 patients are summarised in Table 1. Overall, there were no significant differences

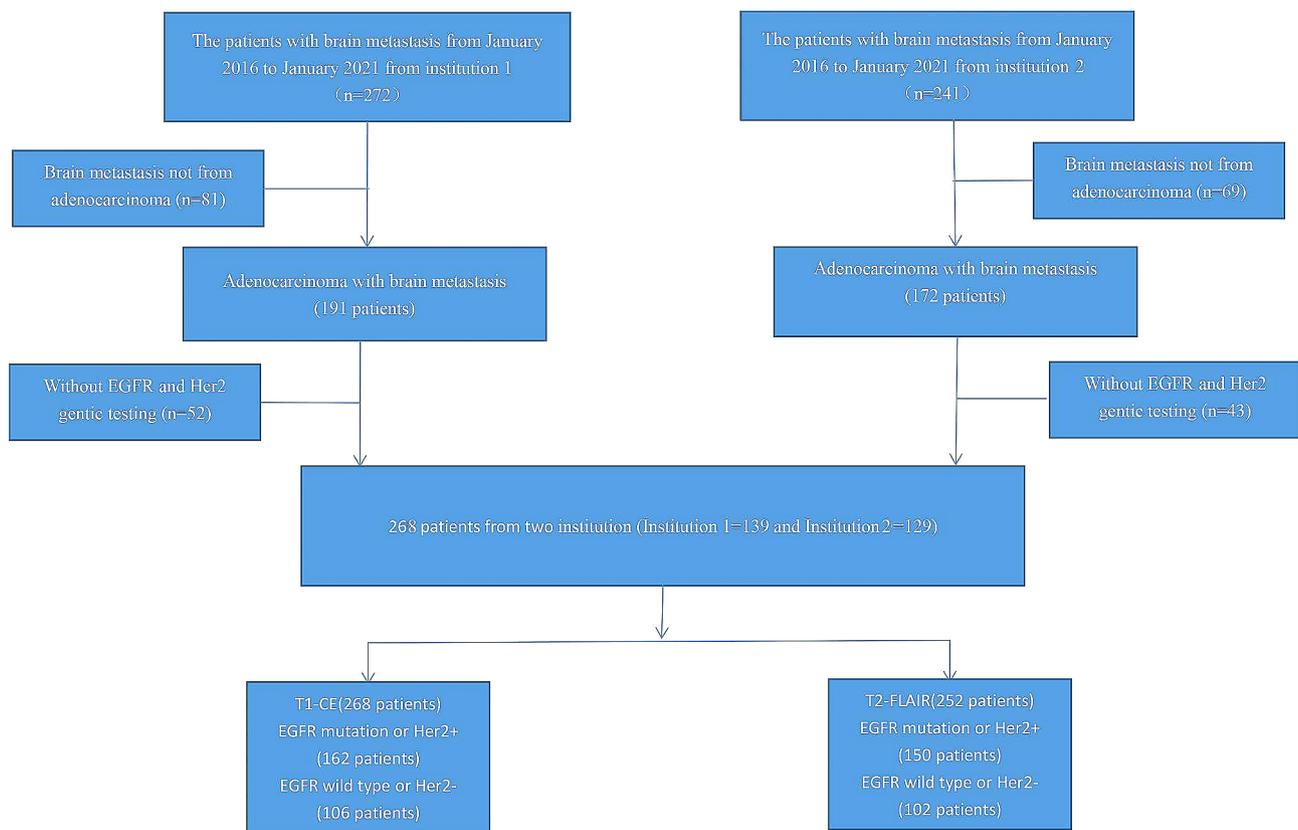


Fig. 2 The flow chart of patient enrollment

Table 1 Pathological diagnosis of primary tumors

Pathological diagnosis of primary focus	Gene expression				Total
	EGFR mutation	HER2+	EGFR wild-type	HER2-	
Lung adenocarcinoma	105 (39.2%)		56 (20.9%)		161 (60.1%)
Breast adenocarcinoma	8 (3%)	24 (9%)		22 (8.2%)	54 (20.1%)
Gastric adenocarcinoma		9 (3.4%)		7 (2.6%)	16 (6%)
Adenocarcinoma of colon			6 (2.2%)	8 (3%)	14 (5.2%)
Undefined adenocarcinoma	16 (6%)		7 (2.6%)		23 (8.6%)
Total	129 (48.1%)	33 (12.3%)	69 (25.8%)	37 (13.8%)	268

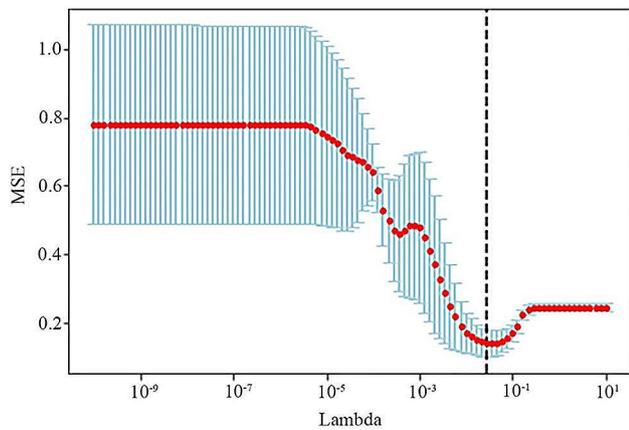


Fig. 3 The MSE versus Lambda plot for T1-CE

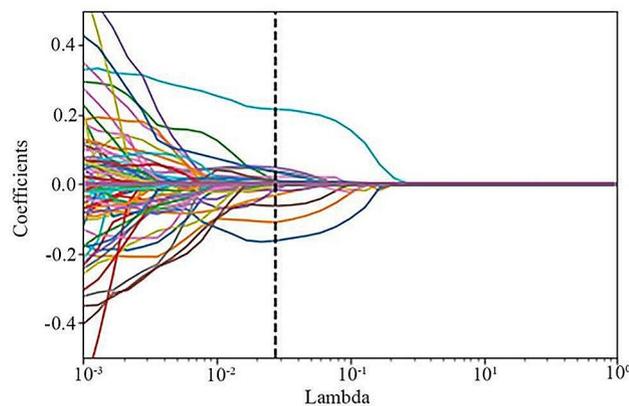


Fig. 4 The LASSO coefficients versus Lambda for T1-CE

in age, sex, and time to metastasis between the groups with EGFR mutant or HER2+ and those with wild-type EGFR or HER2- ($p > 0.05$).

Selection of features and creation of radiomics features

For a more robust feature selection, we employed a combination of cross-validation and grid search techniques.

Table 2 Performance of each MR sequence’s radiomics signature in the training group and the testing cohort

Metrics	Training cohort		Testing cohort	
	T1-CE	T2-FLAIR	T1-CE	T2-FLAIR
AUC	0.98	0.86	0.82	0.70
ACC	0.93	0.74	0.70	0.62
SPE	0.93	0.67	0.69	0.67
SEN	0.92	0.83	0.73	0.57

AUC area under the curve, ACC classification accuracy, SEN sensitivity, SPE specificity

The search range consisted of 100 evenly spaced values within the range $[-10, 1]$ (lambda). After examining these values, 0.027826 emerged as the optimal lambda value for T1-CE. Both the mean squared error (MSE) versus lambda plot and the LASSO coefficients versus lambda plot were provided. Figure 3 illustrates the MSE vs. lambda plot, whereas Fig. 4 displays the LASSO coefficients versus the lambda plot.

A total of 19 out of 1,409 considered radiomics features were ultimately selected for the T1-contrast sequence. The chosen radiomics features for the T1-contrast sequence included original_firstorder_InterquartileRange, original_firstorder_Range, original_glszm_SmallAreaEmphasis, original_glszm_SmallAreaLowGrayLevelEmphasis, wavelet-LLH_firstorder_Maximum, wavelet-LLH_gldm_ClusterProminence, wavelet-LLH_gldm_ClusterShade, wavelet-LLH_glszm_SmallAreaEmphasis, wavelet-LLH_ngtdm_Contrast, wavelet-LHL_firstorder_Mean, wavelet-LHH_firstorder_RobustMeanAbsoluteDeviation, wavelet-LHH_gldm_LongRunHighGrayLevelEmphasis, wavelet-LHH_glszm_SizeZoneNonUniformityNormalized, wavelet-HLH_glszm_SizeZoneNonUniformityNormalized, wavelet-HLH_ngtdm_Complexity, wavelet-HHL_gldm_HighGrayLevelEmphasis, lbp-2D_firstorder_MeanAbsoluteDeviation, logarithm_glszm_SmallAreaEmphasis, and logarithm_glszm_SmallAreaLowGrayLevelEmphasis.

Radiomics features performance

After feature selection, a radiomics feature set was created using an SVC model to differentiate between mutant EGFR and HER2+ and wild-type EGFR and HER2. For the MR sequences, a radiomics feature set was developed using the training cohort, and the test set was employed to evaluate the feature set’s performance. Its complete performance is detailed in Table 2.

The radiomics signature of T1-CE demonstrated the best performance in the training group. Analysis of radiomics profiles for T1-CE revealed an AUC of 0.98, classification accuracy (ACC) of 0.93, sensitivity (SEN) of 0.92, and specificity (SPE) of 0.93. This performance was also evaluated on a separate test set, where the T1-CE radiomics signature showed an AUC of 0.82, ACC

of 0.70, SEN of 0.73, and SPE of 0.69, indicating strong performance.

In contrast, the T2-FLAIR sequence exhibited substantially weaker performance. The radiomics signature of T2-FLAIR in the training group achieved an AUC of 0.86, ACC of 0.74, SEN of 0.83, and SPE of 0.67. In the test set, the AUC of the T2-FLAIR was 0.70. Figure 5 illustrates the ROC curves for these two MR sequences. Additionally, the AUC scores associated with different radiomics interpolators are presented in Supplementary Fig. 1.

Discussion

At present, invasive biopsy or surgical resection of brain metastases is typically required to identify gene mutation status in clinical practice. However, these invasive examinations may lead to serious complications, such as postoperative bleeding and cerebrospinal fluid leakage. Furthermore, it is challenging to obtain accurate gene mutation statuses for patients who cannot tolerate surgery [21]. The imageomics model created in this study demonstrates high accuracy in predicting gene mutations in adenocarcinoma brain metastases and possesses substantial predictive power. Consequently, this study provides a potent non-invasive tool for predicting EGFR gene mutation status and HER2 overexpression in adenocarcinoma brain metastases, thereby greatly enhancing patient prognosis. Additionally, prior studies have primarily focused on the prediction of gene mutations in lung cancer brain metastases, with the research scope being somewhat limited. Most of these studies target

EGFR mutations, with fewer investigations into other gene mutations [22]. Imageomics studies predicting EGFR mutation status in patients with lung cancer typically rely on CT or MRI images of primary lung lesions [23]. In contrast, most MRI imageomics studies of brain metastases aim to distinguish between different lung cancer subtypes or primary sites [24], with a limited number focused on predicting EGFR mutation status in lung cancer. Sui et al. [25] predicted the EGFR mutation status in primary lung adenocarcinoma based on MRI imaging omics characteristics of brain metastases, exploring the optimal MRI sequence for EGFR mutation prediction. Their results indicated that imaging omics characteristics based on multiple combined MRI sequences (enhanced T1WI, FLAIR, and DWI) could serve as a non-invasive auxiliary tool for predicting EGFR mutation status in lung adenocarcinoma. Jiang et al. [26] established a 3D-T1WI enhanced imaging omics model of brain metastases from non-small cell lung cancer to assess its diagnostic efficacy for EGFR mutation status in these patients. The findings showed that 3D-T1WI enhanced imaging omics of brain metastases from non-small cell lung cancer had high diagnostic efficacy and net clinical benefit in determining EGFR mutation status. This study's establishment of an imageomics model to predict gene mutation in brain metastases and the identification of effective MRI diagnostic sequences enrich the research in this field and provide a novel diagnostic method that can accurately predict EGFR mutation/HER2+ status in adenocarcinoma.

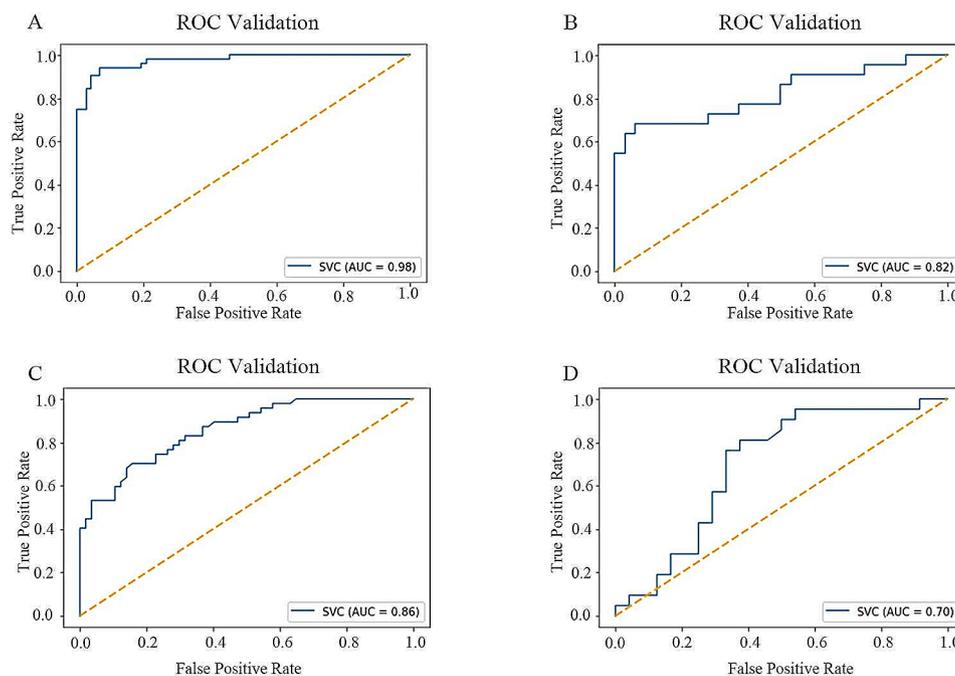


Fig. 5 ROC curves in the training and testing cohort for various MR sequences. A. T1-CE training cohort. B. T1-CE testing cohort. C. T2-FLAIR training cohort. D. T2-FLAIR testing cohort

Ahn utilised radiomics features to identify EGFR mutations in the treatment of primary lung cancer and brain metastases, achieving a prediction AUC of 86.81 [27]. We examined 268 patients using the T1-CE sequence and 252 using the T2-FLAIR sequence. Higher rates of EGFR mutation were observed in women than in men, with the majority of changes being prevalent [28, 29]. Clinical factors such as age, sex, smoking, and drinking cannot effectively predict tumour mutation status. Meanwhile, multiple studies have demonstrated that classic imaging characteristics, including breast density [30], a spiculated mass [31], and microcalcifications [32] on mammograms, are strongly correlated with HER2 status. However, the predictive capability of these characteristics for HER2 status is limited.

In establishing HER2 status, challenges include the scarcity of sufficient tissue samples for genetic testing, the high cost of genetic detection, and the extended duration required to obtain genetic testing results. To address these challenges, our study measured heterogeneity within MR images by exploring the potential correlation between the spatial distribution of voxel grey levels and the shape of the underlying tissue. These methods may provide additional insights into the adenocarcinoma phenotype and genotype, thereby improving the prediction of related gene statuses.

Nineteen key characteristics, all statistically significant, were selected from 1,409 candidate features extracted from normalised MR imaging data. We then developed a radiomics signature using SVC regression on the selected variables, and its efficacy was evaluated in the test set. This proposed radiomics signature demonstrates high accuracy in differentiating between mutant EGFR or HER2+ and normal EGFR or HER2-.

Previous research indicates that radiomics analysis is effective in determining the status of EGFR mutations [33–35]. Furthermore, the radiomics signature of T2-FLAIR provided an AUC of 0.7, ACC of 0.62, SEN of 0.57, and SPE of 0.67. These findings suggest that using T1-CE data, the proposed radiomics signature can predict EGFR mutation/HER2+ status non-invasively. However, the T2-FLAIR sequence exhibits limited discriminatory ability for classifying EGFR mutation status.

Associating radiomics properties with clinical imaging is an emerging field that enhances molecular biology relevance. The correlation between radiomics traits and molecular characteristics may reflect gene expression variations within and across tumours. Tumours enriched in genetic mutations during development lead to alterations in downstream signalling pathways, causing widespread biological responses and driving processes such as cell proliferation and inflammation. Extracting a broad array of quantitative imaging metrics, proven effective in other studies, can translate these changes

into interpretable data. Segal et al. were the first to demonstrate the use of radiogenomic maps, reconstructing tumour gene expression patterns using 28 imaging characteristics [36]. Koay et al. posited that quantifiable imaging parameters could reveal genetic and pathological tumour heterogeneity. They observed that patients with prominent tumours showed substantially less stroma, more mesenchymal features, and a higher likelihood of multiple signalling pathway alterations [37]. Grossmann et al. [38] suggested that radiomics facilitates the non-invasive assessment of adenocarcinoma molecular and clinical traits, hypothesising a link between radiomic imaging properties and signalling pathways. These studies demonstrate that image phenotypes can be used for non-invasive assessments with therapeutic implications and may indicate the activity of different biological pathways.

In our study, we established a connection between the resilience of the radiomics signature, its enhanced performance, and the tumour's molecular biological characteristics. Lung cancer development and progression are influenced by the transmembrane protein EGFR, which has cytoplasmic kinase activity. EGFR promotes cell proliferation upon binding to specific ligands [39]. Aberrant EGFR expression, such as gene amplification, overexpression of EGFR ligands, and mutations, all contribute to tumour growth and metastasis [3]. Experimental models using animal models with tumour cell xenografts have shown that blocking EGFR increases tumour cell mortality, reduces angiogenic factor production, and prolongs survival [3].

It has been demonstrated that HER2, through heterodimerisation with EGFR and other HER family members, activates a wide array of intracellular signalling pathways. Notably, HER2 appears to be the preferred heterodimerisation partner for other members of the HER family [40]. When HER2 is overexpressed, a substantial number of HER2 homodimers and heterodimers are formed. This aberrant downstream signalling is partly mediated by the constitutive activation of the PI3-kinase/Akt and ras/mitogen-activated protein kinase pathways. These pathways inhibit apoptosis, promote cell motility, and support cell growth and survival. Transfecting HER2 into human tumour cell lines has been shown to create a more aggressive tumour cell phenotype [41].

However, there are several limitations to this research. Since the amount of segmentation determines the radiomics characteristic, segmentation is initially one of the more challenging aspects of radiomics. To ensure the consistency of ROI delineation, we aim to establish a deep-learning segmentation model based on large datasets in future work. Second, the information was obtained retrospectively, introducing a potential selective bias that cannot be eliminated. The participants in this study

also originated from two separate institutions, each with its own set of protocols and scanners, resulting in notable variations in patients' characteristics between the two centres. A normalisation strategy was employed to mitigate data discrepancies across multiple centres and enhance the stability of features and various models.

Conclusion

In conclusion, we have proposed a radiomics signature derived from T1-CE brain MR sequences that are both reliable and non-invasive. Our brain metastasis model has the potential to more accurately predict EGFR mutation/HER2+status in patients with adenocarcinoma compared with models predominantly focused on primary lesions. This could substantially assist in directing individualised treatment regimens. While larger prospective studies are necessary to validate our findings, our initial results indicate that radiomics could serve as a non-invasive tool for tumour genotyping through the examination of metastases.

Abbreviations

ACC	Accuracy
AUC	Area Under Curves
CRC	Colorectal Cancer
EGFR	Epidermal Growth Factor Receptor
GLCM	Gray Level Co-occurrence Matrix
GLRLM	Gray Level Dependence Matrix
GLSZM	Gray Level Size Zone Matrix
GLDM	Gray Level Dependence Matrix
HER2	Human Epidermal Growth Factor Receptor 2
LASSO	Least Absolute Shrinkage and Selection Operator
linear SVC	Linear Support Vector Classifier
MAP	Mitogen-Activated Protein
NGS	Next-Generation Sequencing
NGTDM	Neighboring Gray Tone Difference Matrix
NSCLC	Non-Small-Cell Lung Cancer
OS	Overall Survival
ROI	Region of Interest
SCCHN	Squamous Cell Carcinoma of the Head and Neck
SEN	Sensitivity
SMOTE	Synthetic Minority Oversampling Technique
SPE	Specificity
T1-CE	Contrast-Enhanced T1-Weighted Imaging
T2-FLAIR	T2 Fluid-Attenuated Inversion Recovery
TKs	Tyrosine Kinases

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40644-024-00709-4>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

For this research or the publishing of this paper, no money or support was given.

Author contributions

Li YR was responsible for the conception and design of the article. Liu WY and Jia WX conducted the research implementation and feasibility analysis. Wang YL and Wang J were responsible for the implementation of guidance

and data collection. Jin Y conducted the statistical analysis, result analysis and interpretation and wrote the paper. Li YR revised the paper in English. Li YR and Wang J revised the paper. Wang J is in charge of managing and supervising the article's quality control and review processes.

Funding

There was no financial support for this study.

Data availability

All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

The Helsinki Declaration was followed while conducting this study. The First Affiliated Hospital of Xinjiang Medical University and Changzhi People's Hospital's Ethics Committees both approved this study. A formal informed consent form was signed by each participant.

Consent for publication

The manuscript is not submitted for publication or consideration elsewhere.

Competing interests

The authors declare that they have no competing interests.

Received: 23 November 2023 / Accepted: 11 May 2024

Published online: 21 May 2024

References

- Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med*. 2004;350(21):2129–39. <https://doi.org/10.1056/NEJMoa040938>.
- Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science*. 2004;304(5676):1497–500. <https://doi.org/10.1126/science.1099314>.
- Baselga J. Why the epidermal growth factor receptor? The rationale for cancer therapy. *Oncologist*. 2002;7(Suppl 4):2–8. https://doi.org/10.1634/theoncologist.7-suppl_4-2.
- Iqbal N, Iqbal N. Human epidermal growth factor receptor 2 (HER2) in cancers: overexpression and therapeutic implications. *Mol Biol Int*. 2014;2014:852748. <https://doi.org/10.1155/2014/852748>.
- Stephens P, Hunter C, Bignell G, Edkins S, Davies H, Teague J, et al. Lung cancer: intragenic ERBB2 kinase mutations in tumors. *Nature*. 2004;431(7008):525–6. <https://doi.org/10.1038/431525b>.
- Wu SG, Rao MY, Zhou J, Lin Q, Wang ZJ, Chen YX, et al. Distribution of metastatic disease in the brain in relation to the hippocampus: a retrospective single-center analysis of 6064 metastases in 632 patients. *Oncotarget*. 2015;6(41):44030–6. <https://doi.org/10.18632/oncotarget.5828>.
- Takano K, Kinoshita M, Takagaki M, Sakai M, Tateishi S, Achiha T, et al. Different spatial distributions of brain metastases from lung cancer by histological subtype and mutation status of epidermal growth factor receptor. *Neuro Oncol*. 2016;18(5):716–24. <https://doi.org/10.1093/neuonc/nov266>.
- 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. *MAGMA*. 2018;31(2):285–94. <https://doi.org/10.1007/s10334-017-0653-9>.
- 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. *Annu Int Conf IEEE Eng Med Biol Soc*. 2017;2017:493–6. <https://doi.org/10.1109/embc.2017.8036869>.
- 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 post-contrast MR. *Magn Reson Med*. 2016;76(5):1410–9. <https://doi.org/10.1002/mrm.26029>.
- 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. <https://doi.org/10.7759/cureus.584>.

12. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *J Clin Oncol*. 2018;36(20):2105–22. <https://doi.org/10.1200/jco.2018.77.8738>.
13. Hanna WM, Slodkowska E, Lu FI, Nafisi H, Nofech-Mozes S. Comparative Analysis of Human Epidermal Growth Factor Receptor 2 testing in breast Cancer according to 2007 and 2013 American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations. *J Clin Oncol*. 2017;35(26):3039–45. <https://doi.org/10.1200/jco.2016.70.5319>.
14. Nyúl LG, Udupa JK. On standardizing the MR image intensity scale. *Magn Reson Med*. 1999;42(6):1072–81. [https://onlinelibrary.wiley.com/doi/abs/10.1002/\(SICI\)1522-2594\(199912\)42:6<3C1072::AID-MRM11>3E3.0.CO;2-M](https://onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1522-2594(199912)42:6<3C1072::AID-MRM11>3E3.0.CO;2-M).
15. 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. <https://doi.org/10.1016/j.media.2010.12.003>.
16. 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. <https://doi.org/10.1016/j.nicl.2014.08.008>.
17. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. *Front Neuroinform*. 2014;8:14. <https://doi.org/10.3389/fninf.2014.00014>.
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. <https://doi.org/10.1016/j.neuroimage.2006.01.015>.
19. Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. *Med Image Anal*. 2014;18(1):176–96. <https://doi.org/10.1016/j.media.2013.10.005>.
20. Nakamura M, Kajiwara Y, Otsuka A, Kimura H. LVQ-SMOT - learning Vector Quantization based Synthetic Minority over-sampling technique for biomedical data. *BioData Min*. 2013;6(1):16. <https://doi.org/10.1186/1756-0381-6-16>.
21. Li B, Peng Y, Qin W, Xiao F, Li H, Chen J, et al. Prediction of EGFR gene mutation in lung cancer brain metastases based on MRI imaging omics model. *Magn Reson Imaging*. 2024;15(03):86–92.
22. Lin Z, Zuo M. Research progress of radiomics in the diagnosis of brain metastases in lung cancer. *Mod Oncol*. 2024;32(02):383–6. <https://doi.org/10.3969/j.issn.1672-4992.2024.02.034>.
23. Rossi G, Barabino E, Fedeli A, Ficarra G, Coco S, Russo A, et al. Radiomic Detection of EGFR mutations in NSCLC. *Cancer Res*. 2021;81:724–31. <https://doi.org/10.1158/0008-5472.CAN-20-0999>.
24. Chen Q, Zou L, Chen Q, Zhao H, Wang L, Ge Y, et al. Predictive efficacy of image morphology and Radiomics features combined with clinical indicators for NSCLC gene mutations. *J Clin Radiol*. 2021;40:485–90.
25. Sui L, Wang J, Ren J, Cai J, Zhang Y, Yin X. Radiogenomics models based on MRI of Brain Metastasis for Predicting EGFR Mutation Status in Lung Adenocarcinoma. *J Clin Radiol*. 2023;42(01):20–5. <https://doi.org/10.13437/j.cnki.jcr.2023.01.033>.
26. Jiang Y, Zhang J, Chen M, Duan S, Xu X. Predicting EGFR gene mutation status in brain metastases of non-small cell lung cancer based on MRI 3D-T1WI enhanced imaging omics model. *J Clin Radiol*. 2022;41(07):1212–6. <https://doi.org/10.13437/j.cnki.jcr.2022.07.003>.
27. Ahn SJ, Kwon H, Yang JJ, Park M, Cha YJ, Suh SH, Lee JM. Contrast-enhanced T1-weighted image radiomics of brain metastases may predict EGFR mutation status in primary lung cancer. *Sci Rep*. 2020;10(1):8905. <https://doi.org/10.1038/s41598-020-65470-7>.
28. Yotsukura M, Yasuda H, Shigenobu T, Kaseda K, Masai K, Hayashi Y, et al. Clinical and pathological characteristics of EGFR mutation in operable early-stage lung adenocarcinoma. *Lung Cancer*. 2017;109:45–51. <https://doi.org/10.1016/j.lungcan.2017.04.014>.
29. 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–e4486. <https://doi.org/10.1016/j.clcc.2016.02.001>.
30. Li E, Guida JL, Tian Y, Sung H, Koka H, Li M, et al. Associations between mammographic density and tumor characteristics in Chinese women with breast cancer. *Breast Cancer Res Treat*. 2019;177(2):527–36. <https://doi.org/10.1007/s10549-019-05325-6>.
31. Jiang L, Ma T, Moran MS, Kong X, Li X, Hafft BG, et al. Mammographic features are associated with clinicopathological characteristics in invasive breast cancer. *Anticancer Res*. 2011;31(6):2327–34.
32. Farrokh D, Boloursaz S, Homai F. Relationship among mammographic findings with histopathologic type of breast cancer and human epidermal growth factor receptor 2 (HER2) in young women. *Electron Physician*. 2017;9(5):4300–5. <https://doi.org/10.19082/4300>.
33. Mei D, Luo Y, Wang Y, Gong J. CT texture analysis of lung adenocarcinoma: can Radiomic features be surrogate biomarkers for EGFR mutation statuses. *Cancer Imaging*. 2018;18(1):52. <https://doi.org/10.1186/s40644-018-0184-2>.
34. Digumarthy SR, Padole AM, Gullo RL, Sequist LV, Kalra MK. Can CT radiomic analysis in NSCLC predict histology and EGFR mutation status? *Medicine*. 2019;98(1):e13963. <https://doi.org/10.1097/md.00000000000013963>.
35. Tu W, Sun G, Fan L, Wang Y, Xia Y, Guan Y, et al. Radiomics signature: a potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. *Lung Cancer*. 2019;132:28–35. <https://doi.org/10.1016/j.lungcan.2019.03.025>.
36. Zhou J, Tan H, Li W, Liu Z, Wu Y, Bai Y, et al. Radiomics Signatures Based on Multiparametric MRI for the preoperative prediction of the HER2 status of patients with breast Cancer. *Acad Radiol*. 2021;28(10):1352–60. <https://doi.org/10.1016/j.acra.2020.05.040>.
37. Koay EJ, Lee Y, Cristini V, Lowengrub JS, Kang Y, Lucas FAS, et al. A visually apparent and quantifiable CT imaging feature identifies Biophysical subtypes of pancreatic ductal adenocarcinoma. *Clin Cancer Res*. 2018;24(23):5883–94. <https://doi.org/10.1158/1078-0432.Ccr-17-3668>.
38. Grossmann P, Stringfield O, El-Hachem N, Bui MM, Velazquez ER, Parmar C, et al. Defining the biological basis of radiomic phenotypes in lung cancer. *ELife*. 2017;6. <https://doi.org/10.7554/eLife.23421>.
39. Siegelin MD, Borczuk AC. Epidermal growth factor receptor mutations in lung adenocarcinoma. *Lab Invest*. 2014;94(2):129–37. <https://doi.org/10.1038/labinvest.2013.147>.
40. Rubin I, Yarden Y. The basic biology of HER2. *Annals Oncology: Official J Eur Soc Med Oncol*. 2001;12(Suppl 1):S3–8. https://doi.org/10.1093/annonc/12.suppl_1.s3.
41. Neve RM, Lane HA, Hynes NE. The role of overexpressed HER2 in transformation. *Ann Oncol*. 2001;12(Suppl 1):S9–13. https://doi.org/10.1093/annonc/12.suppl_1.s9.

Publisher's Note

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