BioMedical Engineering OnLine

REVIEW

Open Access

The potential of MRI radiomics based on extrapulmonary metastases in predicting EGFR mutations: a systematic review and meta-analysis

Linyong Wu[1†], Dayou Wei[1*], Songhua Li[1], Shaofeng Wu[1], Yan Lin[1] and Lifei Chen[1]

[†]Linyong Wu—First author.

*Correspondence: weidayoumm@163.com

[1] Department of Medical Ultrasound, Maoming People's Hospital, Maoming, Guangdong 525011, People's Republic of China

Abstract

Background: Epidermal growth factor receptor (EGFR) gene mutations can lead to distant metastasis in non-small cell lung cancer (NSCLC). When the primary NSCLC lesions are removed or cannot be sampled, the EGFR status of the metastatic lesions are the potential alternative method to reflect EGFR mutations in the primary NSCLC lesions. This review aimed to evaluate the potential of magnetic resonance imaging (MRI) radiomics based on extrapulmonary metastases in predicting EGFR mutations through a systematic reviews and meta-analysis.

Materials and methods: A systematic review of the studies on MRI radiomics based on extrapulmonary metastases in predicting EGFR mutations. The area under the curve (AUC), sensitivity (SNEC), and specificity (SPEC) of each study were separately extracted for comprehensive evaluation of MRI radiomics in predicting EGFR mutations in primary or metastatic NSCLC.

Results: Thirteen studies were ultimately included, with 2369 cases of metastatic NSCLC, including five studies predicting EGFR mutations in primary NSCLC, eight studies predicting EGFR mutations in metastatic NSCL. In terms of EGFR mutations in the primary lesion of NSCLC, the pooled AUC was 0.90, with SENC and SPEC of 0.80 and 0.85, respectively, which seems superior to the radiomics meta-analysis based on NSCLC primary lesions. In terms of EGFR mutations in NSCLC metastases, the pooled AUC was 0.86, with SENC and SEPC of 0.79 and 0.79, respectively, indicating moderate evaluation performance.

Conclusions: MRI radiomics helps to predict the EGFR mutation status in the primary or metastatic lesions of NSCLC, serve as a high-precision supplement to current molecular detection methods.

Keywords: Non-small cell lung cancer, Epidermal growth factor receptor, Magnetic resonance imaging, Metastases, Radiomics, Systematic review

Introduction

Lung cancer is the most common malignant cancer in the world, with incidence rate ranking second and mortality ranking first [1]. According to statistics, among the lung cancer deaths in the United States in 2023, approximately 103,000 cancer cases will be caused by direct smoking [2]. Non-small cell lung cancer (NSCLC) accounts for approximately 80% of lung cancer cases, with adenocarcinoma being the most common. However, about 40% of NSCLC patients experience distant metastasis at initial diagnosis, with the most common locations being the brain, spine, and liver [3–5].

Epidermal growth factor receptor (EGFR), as a tyrosine receptor, is the main driving gene and therapeutic target of NSCLC [6]. The mutation rate of EGFR in East Asian populations was as high as 40% -50%, while in Western populations it was 10%–20% [7]. EGFR gene mutations mainly occur in exons 18–21, and mutations at different loci can lead to different pathways of metastasis. For example, patients with EGFR exon 19 mutations were more likely to experience brain parenchymal metastasis [8], while patients with EGFR exon 21 mutations were more likely to experience liver metastasis [9]. Tyrosine kinase inhibitors (TKI) targeting EGFR have become an effective treatment for improving the prognosis of NSCLC patients with EGFR mutations [10]. However, EGFR mutation detection typically requires invasive tissue or blood biopsy, which has limitations such as difficulty in obtaining primary lesion, sampling bias, high cost, long processing time, and poor representativeness of results [11]. In addition, when the primary lesion is removed or cannot be sampled, metastases may be a potential alternative method to reflect the EGFR status of the primary lesion.

Although magnetic resonance imaging (MRI) is considered an important non-invasive method for evaluating distant metastasis of NSCLC, the potential of traditional MRI images to predict EGFR mutations in the primary or metastatic lesions of NSCLC remains a controversial issue [12]. Radiomics is a computer-assisted approach to breaking through visual information on the basis of traditional medical imaging, quantifying image features similar to high-throughput genes, and further exploring the correlation between features and genes [13]. Through investigation, it was found that radiomics studies based on computed tomography (CT) and positron emission tomography/computed tomography (PET/CT) images had shown great potential in predicting EGFR mutations in the primary lesion of NSCLC. For example, a meta-analysis was conducted based on 14 CT radiomics studies, and it was found that predicting EGFR mutations in primary lesions had a moderate effect, with an area under the curve (AUC) of 0.80 [14]; based on 17 PET/CT radiomic studies, the meta-analysis results of the training and validation cohorts were statistically analyzed, with pooled AUC, sensitivity (SNEC), and specificity (SPEC) of 0.84, 0.76, 0.78, and 0.82, 0.76, 0.75, respectively [15]. However, no meta-analysis based on MRI radiomics had been found to summarize the potential value of radiomics in predicting EGFR mutations in NSCLC; the comprehensive potential of MRI radiomics based on images of extrapulmonary metastases for predicting EFGR mutations in the primary or metastatic lesions of NSCLC had not yet been discovered.

Therefore, this meta-analysis aims to evaluate the accuracy of predicting EGFR mutations in primary or metastatic lesions of NSCLC patients based on MRI radiomics features of extrapulmonary metastatic tumor images, as well as the potential value of evaluating EGFR-TKI treatment response.

Results

Baseline characteristics of literature inclusion

A total of 48 original studies were identified based on keywords. Three studies were based on the same batch of samples for radiomics analysis of intratumoral or peritumoral region features, so only one study with the best summary performance was included. After excluding original studies that met at least one exclusion criterion, 13 studies (12 training cohorts and 15 validation cohorts) were included, with a total of 2369 cases of NSCLC metastases [3–5, 16–25]. According to the different lesion locations of EGFR mutation, there were 5 studies predicting EGFR mutations in primary lung cancer lesions, including 2 NSCLC spinal metastases and 3 NSCLC brain metastases; and 8 studies predicting EGFR mutations in metastatic lung cancer lesions, including 5 cases of NSCLC spinal metastases, 1 case of NSCLC brain metastases, and 2 cases of NSCLC liver metastases. In addition, 2 studies predicting EGFR-ITK response based on metastatic lung cancer lesions, including 1 cases of NSCLC spinal metastases, 1 case of NSCLC brain metastases (Fig. 1).

Quality evaluation

All studies were retrospective, with 6 studies being multicenter studies. 10 studies were based on Siemens scanners, and 3 studies were based on Siemens and GE or GE and Philips scanners. 11 studies were based on ITK-snap software, 1 study was based on 3D slicer software for image segmentation, and 1 study was not mentioned. All studies were based on the open-source PyRadiomics software package for feature extraction, with 10 based on MRI multi sequence extraction and 3 based on MRI single sequence extraction.

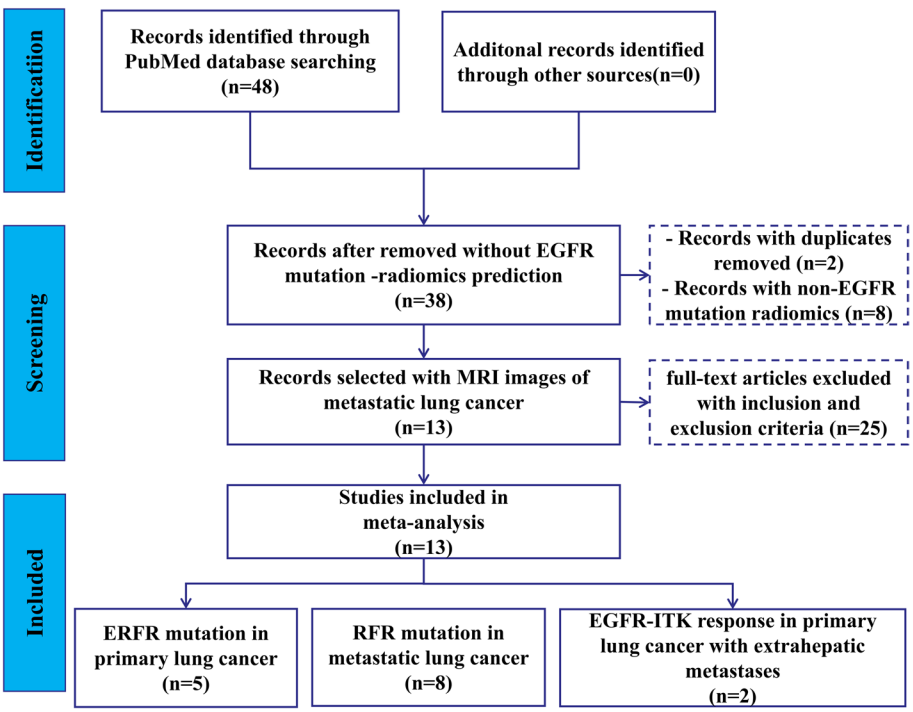

Fig. 1 Flowchart of this study

12 studies included over 1000 features, while only one study included 107 features. All studies were based on multi-strategy feature selection, among which the intra-class correlation coefficient, Mann–Whitney U tests, and LASSO joint strategy were the most common. All studies were based on linear regression for model development. 9 studies developed joint models based on multiple sequence features, while two studies developed joint models based on multiple sequence features and smoking status. 11 studies were based on 10 or fivefold cross-validation to validate model performance, while 2 studies were not mentioned. 9 studies were analyzed based on the features of intratumoral regions, of which 2 studies were based on tumor subregions; 4 studies were analyzed based on the features of intratumoral and peritumoral regions (Table 1).

The RQS score range for 13 studies was 10–17 points, with an average of (13 ± 2) points and an average proportion of 36% (13/36). All studies evaluated study type, image protocols, feature selection, model performance, and model validation; 11 studies used multiple segmentation of images; 4 studies calibrated the statistical data; 5 studies evaluated the clinical utility; one study evaluated the potential of non-radiomic features. All studies had not evaluated biological relevance, cost-effectiveness, and open science. According to the modified QUADS-2 standard, all studies were of high quality (QUADAS-2 ≥ 7). All studies had a lower risk of bias in patient selection tests, reference standard tests, and flow and timing tests. The risk of bias in index testing was higher in 7 studies (54%), while the risk of bias was lower in 6 studies (46%) (Fig. 2).

Comprehensive literature analysis of predicting EGFR mutation in primary NSCLS based on MRI radiomics from extrapulmonary metastases

Table 2 summarizes the basic characteristics of 5 original studies based on MRI radiomics in extrapulmonary metastases for predicting primary NSCLS EGFR mutation. A total of 14 cohorts were included, with 1112 cases of extrapulmonary metastases, including 592 cases of EGFR mutation positive and 520 cases of EGFR mutation negative. The AUC, SENC, and SPEC of all cohorts were between 0.74–0.97, 0.67–0.96, and 0.71–0.97, respectively. The pooled SENC, pooled SPEC, pooled AUC, and sROC curves were used to evaluate the potential of MRI radiomics based on extrapulmonary metastases in predicting primary NSCLS EGFR mutation. The results showed that the pooled SENC and pooled SPEC were 0.80 [95% confidence Interval (CI), 0.74–0.85] and 0.85 (95% CI, 0.80–0.88), respectively. The forest map is shown in Figs. 3A and 4A. There were significant heterogeneity in both pooled SENC ($I^2 = 60.98\%$, $P < 0.01$) and pooled SPEC ($I^2 = 44.95\%$, $P = 0.03$). Through sROC curve analysis, the pooled AUC was 0.90 (95% CI, 0.87–0.92), as shown in Fig. 5A, indicating higher evaluation performance. The existence of publication bias was detected by Deek's funnel plots, which mean that publication bias does not exist ($t = -0.46$, $P = 0.65$, Fig. 6A).

Subgroup analysis of radiomics based on the location and cohort type of extrapulmonary metastases. Based on the location, the pooled SENC, SPEC, and AUC based on brain metastases were 0.84 (95% CI, 0.80–0.88), 0.88 (95% CI, 0.84–0.91), and 0.93, respectively; the pooled SENC, SPEC, and AUC based on spinal metastases were 0.73 (95% CI 0.66–0.79), 0.80 (95% CI 0.73–0.85), and 0.83, respectively. Based on the cohort type, the pooled SENC, SPEC, and AUC based on training cohorts were 0.79 (95% CI 0.74–0.83), 0.87 (95% CI 0.82–0.90), and 0.93, respectively; the pooled SENC,

Table 1 Baseline characteristics of literature inclusion

Author	Type	Scanners	Segmentation	Sequence	Feature software	Feature number	Feature selection	Modeling algorithm	Nomogram	Cross-validation
Jiang et al. [16]	Retrospective	Siemens	ITK-SNAP	T1WI, T2WI, T2FS	Pyradiomics	1967	Mann–Whitney U tests + LASSO	LR	1	10
Fan et al. [17]	Retrospective	Siemens	ITK-SNAP	T1WI, T2FS	Pyradiomics	1595	Mann–Whitney U tests + LASSO	LR	1	10
Ren et al. [25]	Retrospective	Siemens	ITK-SNAP	T1WI, T2WI, T2FS	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO	LR	1	10
Cao et al. [18]	Retrospective	Siemens	ITK-SNAP	T2WI	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO	LR	1	10
Fan et al. [3]	Retrospective	Siemens	ITK-SNAP	T1WI, T2FS	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO	LR	1	10
Zheng et al. [20]	Retrospective	Siemens, GE	ITK-SNAP	T1WI, T2WI, T2FS	Pyradiomics	1470	ICC + PCC + univariate analysis + LASSO	LR	1	–
Fan et al. [19]	Retrospective	Siemens	ITK-SNAP	CE-T1WI	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO	LR	0	5
Fan et al. [21]	Retrospective	Siemens	ITK-SNAP	T2WI, CE-T1WI	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO	LR	1	10
Hou et al. [5]	Retrospective	Siemens	-	CE-T1WI	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO	LR	1	10
Cao et al. [22]	Retrospective	Siemens	ITK-SNAP	T1WI, T2WI, T2FS	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO + mRMR	LR	1	–
Cheng et al. [4]	Retrospective	Siemens	ITK-SNAP	T1WI, T2WI, T2FS	Pyradiomics	1967	ICC + Mann–Whitney U tests + LASSO	LR	1	10
Hou et al. [23]	Retrospective	Siemens, GE	ITK-SNAP	T2WI, CE-T1WI	Pyradiomics	1967	Mann–Whitney U tests + LASSO	LR	0	10
Huang et al. [24]	Retrospective	GE, Philips	3D slicer	T1WI, T2WI, T2FS, ADC, CE-T1WI	Pyradiomics	107	PCC + RFE	LR	1	5

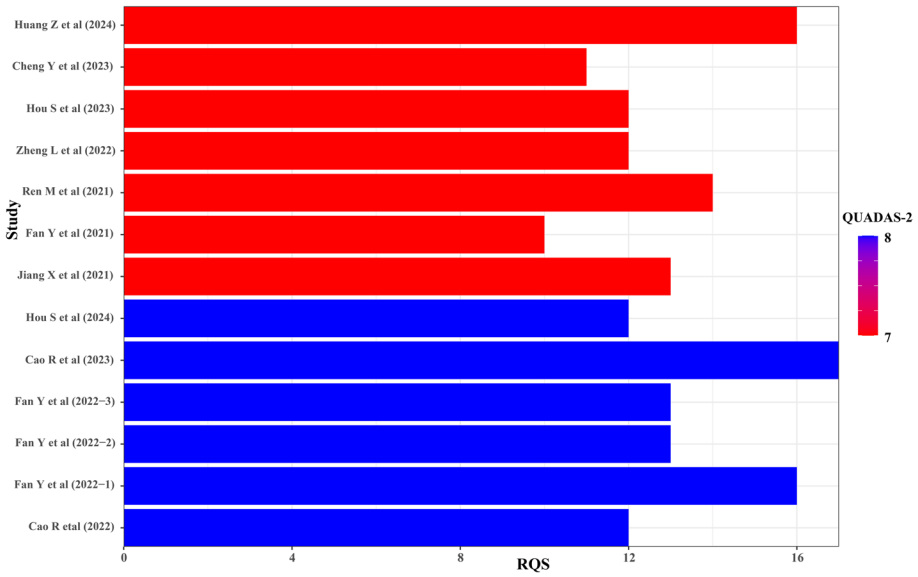

Fig. 2 Methodological quality evaluation of studies based on RQS and QUADAS-2

Table 2 Comprehensive literature analysis of predicting EGFR mutation in primary NSCLC based on MRI radiomics from extrapulmonary metastases

Author	Year	*N*	EGFR mutation (±)	AUC	TP	FP	FN	TN	Metastases	Cohort
Cao R-T	2022	100	49/51	0.97	47	5	2	46	Brain	Training
Cao R-V1	2022	50	24/26	0.90	20	2	4	24	Brain	Validation 1
Cao R-V2	2022	38	19/19	0.90	17	2	2	17	Brain	Validation 2
Fan Y-T	2022-1	153	90/63	0.95	79	4	11	59	Brain	Training
Fan Y-V1	2022-1	77	45/32	0.88	33	1	12	31	Brain	Validation 1
Fan Y-V2	2022-1	80	40/40	0.90	35	8	5	32	Brain	Validation 2
Fan Y-T	2022-2	106	54/52	0.84	39	8	15	44	Spine	Training
Fan Y-V1	2022-2	54	28/26	0.81	25	7	3	19	Spine	Validation 1
Fan Y-V2	2022-2	32	18/14	0.75	14	4	4	10	Spine	Validation 2
Cao R-T	2023	171	91/80	0.81	61	11	30	69	Spine	Training
Cao R-V1	2023	86	46/40	0.75	32	11	14	29	Spine	Validation 1
Cao R-V2	2023	42	24/18	0.74	17	4	7	14	Spine	Validation 2
Huang Z-T	2024	86	45/41	0.76	34	10	11	31	Brain	Training
Huang Z-V	2024	37	19/18	0.79	14	3	5	15	Brain	Validation

SPEC, and AUC based on validation cohorts were 0.79 (95% CI 0.73–0.83), 0.82 (95% CI 0.76–0.87), and 0.88, respectively.

Comprehensive literature analysis of predicting EGFR mutation in metastatic NSCLS based on MRI radiomics from extrapulmonary metastases

Table 3 summarizes the basic characteristics of 8 original studies based on MRI radiomics in extrapulmonary metastases for predicting metastatic NSCLS EGFR mutation. A total of 18 cohorts were included, with 1198 cases of extrapulmonary metastases, including 657 cases of EGFR mutation positive and 541 cases of EGFR

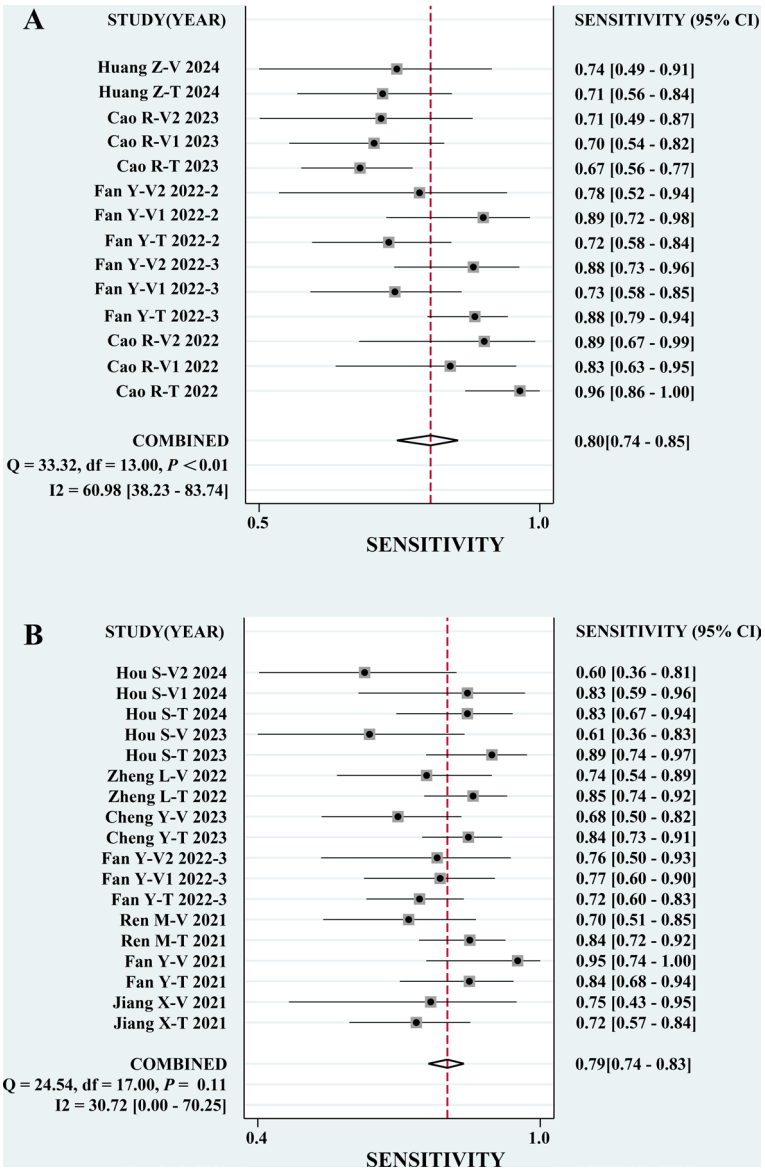

Fig. 3 Forest plots of pooled SENC of predicting EGFR mutations based on MRI radiomics from extrapulmonary metastases. **A** NSCLC primary lesions; **B** NSCLC metastatic lesions

mutation negative. The AUC, SENC, and SPEC of all cohorts were between 0.73–0.90, 0.58–0.95, and 0.61–0.92, respectively. The pooled SENC, pooled SPEC, pooled AUC, sROC curves were used to evaluate the potential of MRI radiomics based on extrapulmonary metastases in predicting primary NSCLS EGFR mutation. The results showed that the pooled SENC and pooled SPEC were 0.79 (95% CI, 0.74–0.83) and 0.79 (95% CI, 0.74–0.83), respectively. The forest map is shown in Figs. 3B and 4B. There were no significant heterogeneity in both pooled SENC ($I^2 = 30.72\%$, $P = 0.11$) and pooled SPEC ($I^2 = 18.23\%$, $P = 0.24$). Through sROC curve analysis, the pooled AUC was 0.86 (95% CI, 0.82–0.89), as shown in Fig. 5B, indicating moderate evaluation performance.

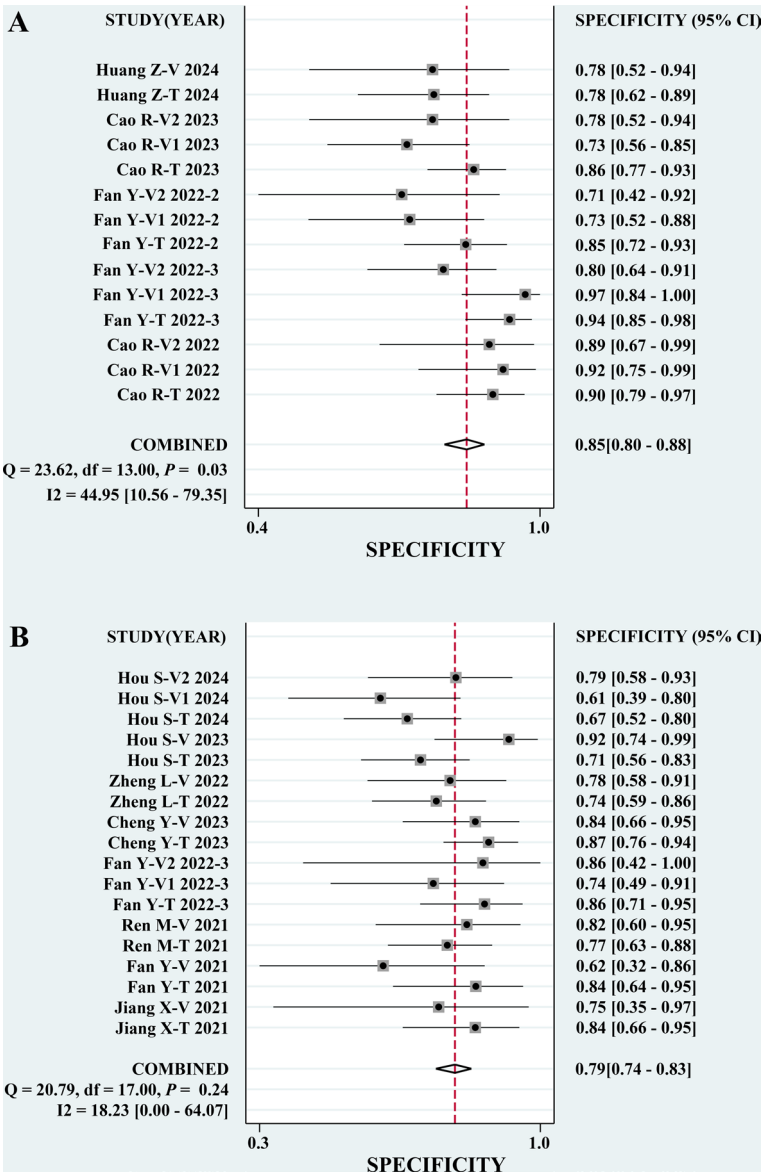

Fig. 4 Forest plots of pooled SPEC of predicting EGFR mutations based on MRI radiomics from extrapulmonary metastases. **A** NSCLC primary lesions; **B** NSCLC metastatic lesions

The existence of publication bias was detected by Deek's funnel plots, which mean that publication bias does not exist ($t = -1.80$, $P = 0.09$, Fig. 6B).

Based on the location, the pooled SENC, SPEC, and AUC based on liver metastases were 0.78 (95% CI, 0.70–0.85), 0.73 (95% CI 0.65–0.79), and 0.83, respectively; the pooled SENC, SPEC, and AUC based on spinal metastases were 0.78 (95% CI 0.74–0.82), 0.82 (95% CI 0.77–0.86), and 0.88, respectively. Based on the cohort type, the pooled SENC, SPEC, and AUC based on training cohorts were 0.81 (95% CI 0.77–0.85), 0.78 (95% CI 0.74–0.83), and 0.88, respectively; the pooled SENC, SPEC, and AUC based on validation cohorts were 0.73 (95% CI 0.67–0.79), 0.78 (95% CI 0.71–0.83), and 0.83, respectively.

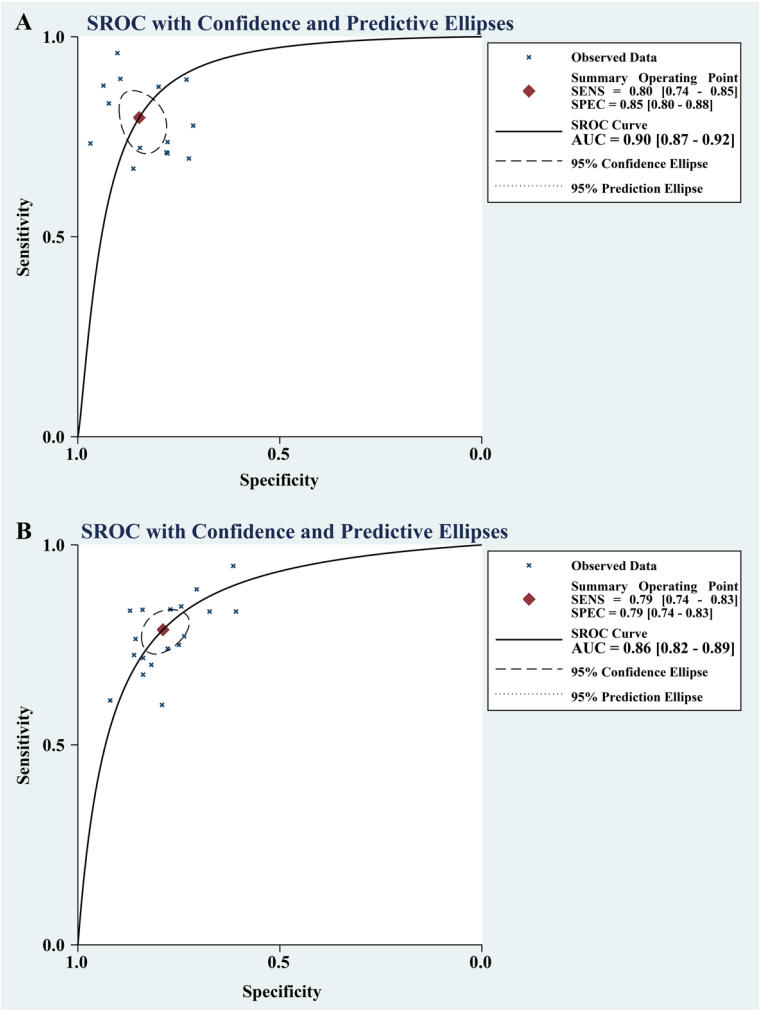

Fig. 5 sROC of predicting EGFR mutations based on MRI radiomics from extrapulmonary metastases. **A** NSCLC primary lesions; **B** NSCLC metastatic lesions

Comprehensive literature analysis of predicting EGFR-ITK response based on MRI radiomics from extrapulmonary metastases

Table 4 summarized the basic characteristics of 2 original studies for predicting EGFR-ITK response based on MRI radiomics from extrapulmonary metastases. A total of 5 cohorts were included, with 285 cases of extrapulmonary metastases, including 138 cases of EGFR-ITK response positive and 147 cases of EGFR-ITK response negative. The AUC, SENC, and SPEC of all cohorts were between 0.76–0.87, 0.71–0.88, and 0.63–0.88, respectively. The pooled SENC, pooled SPEC, pooled AUC, sROC curves were used to evaluate the potential of MRI radiomics based on extrapulmonary metastases in predicting primary NSCLS EGFR mutation. The results showed that the pooled SENC and pooled SPEC were 0.77 (95% CI 0.67–0.84) and 0.80 (95% CI 0.69–0.87), respectively. There were no significant heterogeneity in both pooled SENC ($I^2 = 0.00\%$, $P = 0.53$) and pooled SPEC ($I^2 = 51.08\%$, $P = 0.09$). Through sROC curve analysis, the pooled AUC was 0.84 (95% CI 0.81–0.87), indicating moderate evaluation performance. The existence of publication bias was detected

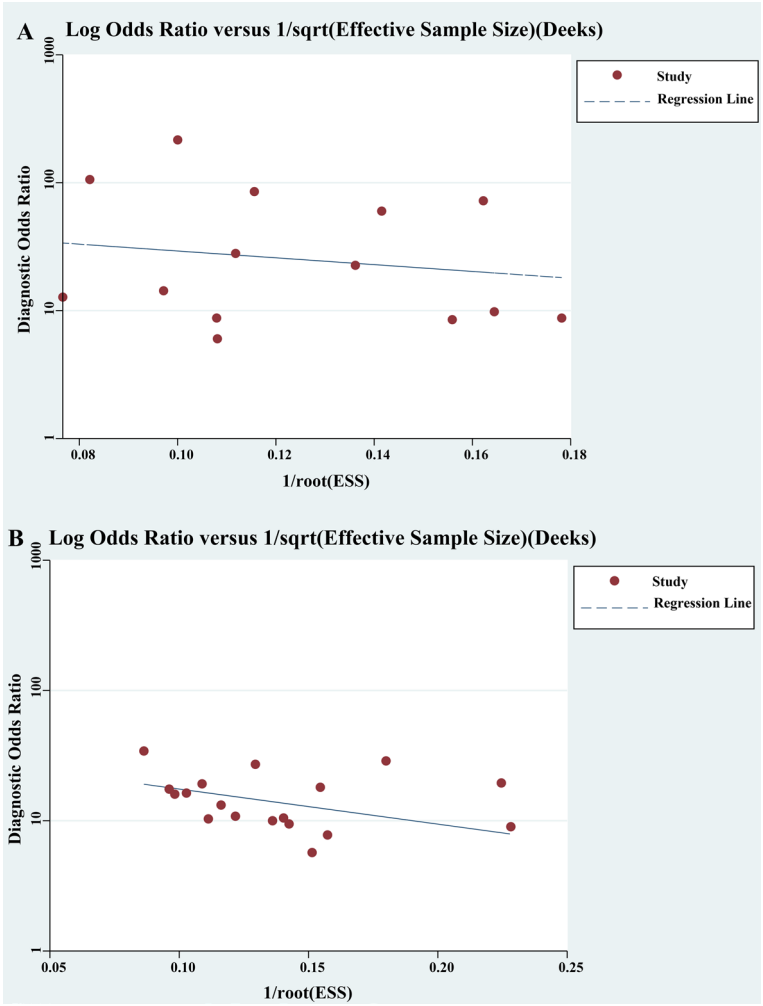

Fig. 6 Begg's funnel plots for the publication bias test of predicting EGFR mutations based on MRI radiomics from extrapulmonary metastases. **A** NSCLC primary lesions; **B** NSCLC metastatic lesions

by Deek's funnel plots, which mean that publication bias does not exist ($t = -1.70$, $P = 0.19$).

Clinical utility

Fagan's analysis indicated that MRI radiomics, based on extrapulmonary metastasis, predicted a post-detection probability of EGFR mutations to be 57% and 48%, respectively, in primary and metastatic lesions of NSCLC (Fig. 7).

Meta-analysis investigation of NSCLC EGFR mutations based on artificial intelligence

The meta-analysis of NSCLC EGFR mutations based on AI is presented in Table 5. A total of 3 meta-analyses were found [14, 15, 26], including 35 (CT, PET/CT, MRI), 17 (PET/CT), and 14 (CT) radiomics studies, with pooled AUC of 0.79, 0.84 (0.82), and 0.80, respectively. The pooled AUC of this review was 0.90, which seemed superior to the AI meta-analysis based on NSCLC primary lesions.

Table 3 Comprehensive literature analysis of predicting EGFR mutation in metastatic NSCLC based on MRI radiomics from extrapulmonary metastases

Author	Year	*N*	EGFR mutation (±)	AUC	TP	FP	FN	TN	Metastases	Cohort
Jiang X-T	2021	77	46/31	0.86	33	5	13	26	Spine	Training
Jiang X-V	2021	20	12/8	0.77	9	2	3	6	Spine	Validation
Fan Y-T	2021	62	37/25	0.88	31	4	6	21	Spine	Training
Fan Y-V	2021	32	19/13	0.78	18	5	1	8	Spine	Validation
Ren M-T	2021	110	62/48	0.87	52	11	10	37	Spine	Training
Ren M-V	2021	52	30/22	0.78	21	4	9	18	Spine	Validation
Fan Y-T	2022-3	105	69/36	0.85	50	5	19	31	Spine	Training
Fan Y-V1	2022-3	54	35/19	0.78	27	5	8	14	Spine	Validation 1
Fan Y-V2	2022-3	24	17/7	0.81	13	1	4	6	Spine	Validation 2
Cheng Y-T	2023	135	73/62	0.90	61	8	12	54	Spine	Training
Cheng Y-V	2023	68	37/31	0.78	25	5	12	26	Spine	Validation
Zheng L-T	2022	108	65/43	0.85	55	11	10	32	Brain	Training
Zheng L-V	2022	54	27/27	0.81	20	6	7	21	Brain	Validation
Hou S-T	2023	87	36/51	0.82	32	15	4	36	Liver	Training
Hou S-V	2023	43	18/25	0.79	11	2	7	23	Liver	Validation
Hou S-T	2024	82	36/46	0.84	30	15	6	31	Liver	Training
Hou S-V1	2024	41	18/23	0.77	15	9	3	14	Liver	Validation 1
Hou S-V2	2024	44	20/24	0.73	12	5	8	19	Liver	Validation 2

Table 4 Comprehensive literature analysis of predicting EGFR-ITK response based on MRI radiomics from extrapulmonary metastases

Author	Year	*N*	EGFR-ITK response (±)	AUC	TP	FP	FN	TN	Metastases	Cohort
Fan Y-T	2022-1	89	46/43	0.87	33	5	13	38	Brain	Training
Fan Y-V1	2022-1	46	24/22	0.79	17	4	7	18	Brain	Validation 1
Fan Y-V2	2022-1	40	20/20	0.80	15	4	5	16	Brain	Validation 2
Cheng Y-T	2023	73	32/41	0.80	28	15	4	26	Spine	Training
Cheng Y-V	2023	37	16/21	0.76	12	4	4	17	Spine	Validation

Discussion

EGFR mutation can lead to distant metastasis of NSCLC and is also an important site for targeted treatment of advanced NSCLC. Non-invasive acquisition of EGFR mutation status helps in the development of targeted treatment plans. To the best of our knowledge, this review is the first meta-analysis of the potential of MRI radiomics based on extrapulmonary metastases to predict EGFR mutations in the primary or metastatic lesions of NSCLC. The meta-analysis results showed satisfactory diagnostic accuracy. In terms of EGFR mutations in the primary lesion of NSCLC, the pooled AUC was 0.90, with SENC and SEPC of 0.80 and 0.85, respectively, which seems superior to the radiomics meta-analysis based on NSCLC primary lesions. In terms of EGFR mutations in NSCLC metastases, the pooled AUC was 0.86, with SENC and SEPC of 0.79 and 0.79, respectively, indicating moderate evaluation performance. Furthermore, in the evaluation of targeted therapy for NSCLC with distant metastasis, the pooled

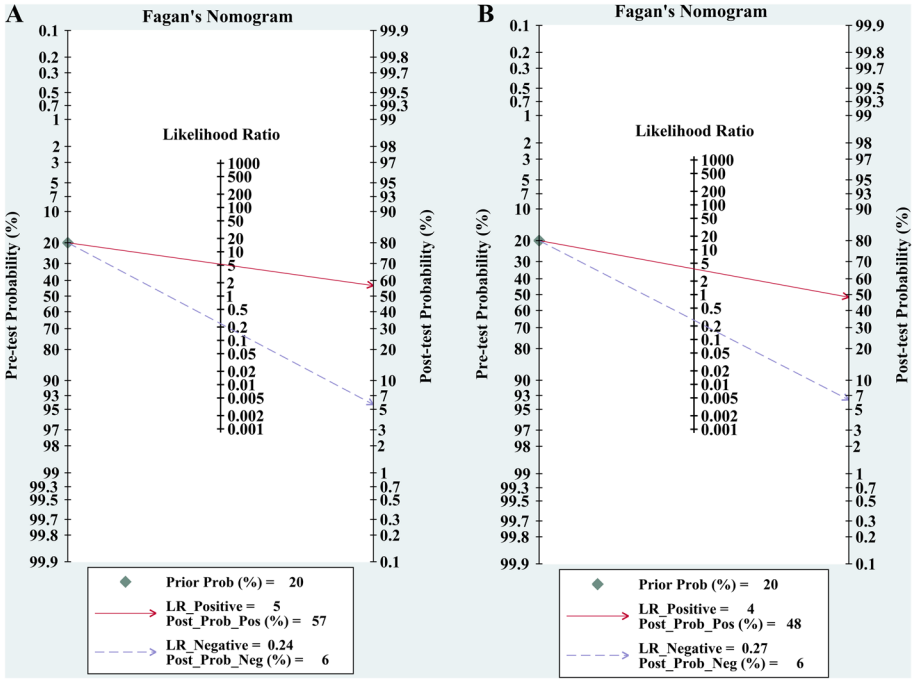

Fig. 7 Fagan nomogram for the elucidation of post-test probabilities with a pre-test probability. **A** NSCLC primary lesions; **B** NSCLC metastatic lesions

Table 5 Meta-analysis investigation of EGFR mutation in primary NSCLC based on artificial intelligence

Author	Year	Study number	Cohort	Image	AUC	SENC	SEPC
Nguyen et al	2023	35	–	CT, PET/CT, MRI	0.79	0.72 (0.67–0.76)	0.73 (0.69–0.78)
Felfli et al	2023	14	–	CT	0.8	–	–
Ma et al	2024	17	Training	PET/CT	0.84	0.76 (0.70–0.81)	0.78 (0.74–0.82)
Ma et al	2024	10	Validation	PET/CT	0.82	0.76 (0.67–0.83)	0.75 (0.68–0.80)
This review	2024	5	–	MRI	0.90	0.80 (0.74–0.85)	0.85 (0.80–0.88)

AUC was 0.84, with SENC and SEPC of 0.77 and 0.80, respectively, showing moderate performance. However, lower RQS evaluations indicated that methodological quality assessment remained a significant challenge in current AI transformation applications, highlighting the importance of developing standardized guidelines. Despite the limitations of MRI radiomics studies, it had great potential in predicting EGFR mutations and evaluating targeted therapies in the primary or metastatic lesions of NSCLC.

AI, such as, radiomics, deep learning, and pathomics, have been widely used for predicting EGFR mutations in primary lung cancer. For example, the multiphase CT radiomics models based on preoperative non-enhanced and enhanced image features had been validated to perform well in predicting the EGFR mutation status of 424 NSCLC cases [27]; deep learning models based on PET/CT demonstrated high accuracy in predicting EGFR mutation status from NSCLC patient cohorts from different centers [28]. The development of deep learning networks based on H&E images could serve as a high-precision supplement to current molecular detection methods and

provide treatment opportunities for NSCLC patients with limited available samples [29]. As far as we know, CT and PET/CT are the most explored medical images in AI, and their potential value in carrying image information has been comprehensively evaluated through meta-analysis. However, the meta-analysis based on CT radiomics only analyzed the pooled AUC, lacking a comprehensive analysis of SNEC and SPEC; a meta-analysis based on PET/CT radiomics lacked a comparative analysis of a single PET radiomics. In addition, the application of MRI in the evaluation of primary lesions in NSCLC is still in its early stages, with more applications in the evaluation of metastatic lesions in NSCLC, especially in the brain, spine, and liver. When the primary lesion is removed or cannot be sampled, the EGFR status of the metastatic lesions can reflect the EGFR mutation status of the primary lesion. Therefore, MRI radiomics based on extrapulmonary metastases have gradually been applied to predict the EGFR mutation status of primary or metastatic lesions. For example, the combination model developed based on radiomics features of the entire tumor, tumor activity area, and peritumoral edema area performed well in predicting EGFR mutations in NSCLC primary lesions, with the SENC of 0.75 [18]; the combined model developed based on radiomics and deep learning features had been validated to achieve medium to high performance in predicting EGFR mutations in NSCLC metastases [16]. However, the conclusions of a single study still cannot represent the potential clinical value of MRI radiology, and a meta-analysis is still needed to integrate all published studies in order to obtain more reliable conclusions.

All steps in AI medical studies are interrelated, and the main steps include: imaging acquisition and protocol, image segmentation, feature extraction, model development and evaluation, quality control, and interpretability analysis (Fig. 8). Any error in any stage can lead to the accumulation and transmission of errors, resulting in unreliable results and inability to achieve repeatability and validation [30]. In imaging acquisition and protocol, the variations in MRI imaging are primarily attributed to machine types and parameter settings, which can be addressed through image normalization methods such as Z-score and max-min. Although normalization methods are frequently applied in this context, the analysis of differences before and after normalization remains worthy. Image segmentation encompasses manual, semi-automatic, and fully automatic methods, with manual segmentation currently being the predominant approach,

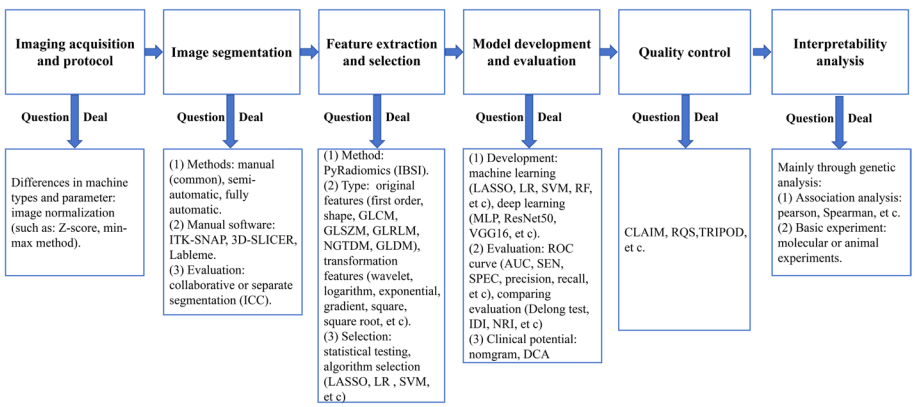

Fig. 8 The main steps and quality control interpretation of radiomics

largely accomplished through software such as ITK-SNAP, 3D-SLICER, and Labelme. The effectiveness of segmentation primarily relies on collaborative segmentation by multiple individuals or individual segmentation followed by assessment using ICC, which involves segmentation by different physicians at the same time point and by the same physician at different time points. Feature extraction primarily depends on the Pyradiomics package, with the rationalization of feature formulation being subject to the quality control of the Image Biomarker Standardization Initiative (IBSI). Feature types include original features and transformation features. Original features encompass first order, shape, gray-level co-occurrence matrix (GLCM), gray-level dependence matrix (GLDM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), and neighboring gray tone difference matrix (NGTDM). Transformation features include wavelet, logarithm, exponential, gradient, square, square root, etc. Interestingly, wavelet transformation features, with a relatively larger number compared to other features, seem to be more frequently selected for model development. Whether this is due to the importance of wavelet features or their quantity remains worthy. Feature selection often employs multi-strategy approaches, including testing and algorithm selection. Model development primarily utilizes machine learning or deep learning algorithms. However, most studies only employ one or a few algorithms, raising the question of which algorithm is best suited for a particular feature—whether a feature performs well in only one algorithm or across multiple algorithms. Model evaluation predominantly mentions AUC in many studies, but comparisons between algorithms are not solely based on intuitive AUC comparisons; they can be further facilitated through methods such as the Delong test, net reclassification index (NRI), and integrated discrimination improvement (IDI). Many studies adopt decision curve analysis (DCA) as an assessment of clinical potential, although its substantial value still warrants further analysis. AI medical studies needs to establish strict methodological quality evaluation standards and reporting guidelines in order to accelerate clinical translation. This meta-analysis evaluated the methodological quality of the studies based on the RQS and QUADAS-2 methods. The average RQS level of this meta-analysis was 13 points, ranging from 10 to 17 points, indicating that the quality of MRI radiomics was not very high, similar to the results of other meta-analyses [14]. The main reason was that all studies did not disclose biological significance, threshold setting, cost, open science, multiple timepoints, and were all retrospective designs. However, based on the modified QUADS-2 method, all studies had scores greater than or equal to 7, and disclosed patient selection, reference criteria, flow and timing. However, due to internal validation used in six studies, the risk of index testing bias was high. Based on another perspective of methodological quality analysis, this meta-analysis included 13 studies, of which 50% were designed based on multi centers, 77% of image acquisition were based on Siemens scanners, 77% of image sequences were based on multi sequences, 85% of image segmentation were based on ITK-snap software, 100% of feature extraction were based on PyRadiomics software package, 92% of feature selection were based on multi strategies (ICC, Mann Whitney U tests, and LASSO), 100% of model development were based on logistic linear regression, 85% of model development were based on cross-validation, 69% of feature sources were based on intratumoral regions, and 62% of metastases were located in the spine. The study differences come from: (1) data sources: multicenter research makes

the conclusions more stable; (2) the difference in information carried by MRI image sequences: the complementary information of multiple sequences leads to higher prediction results; (3) the MRI features come from different sources: the overall intra tumor regional features represent the average features, ignoring the heterogeneity within the tumor, while the sub regional and extra regional features of the tumor explain the heterogeneity, resulting in higher prediction results [31]; and (4) differences in the location of metastases. The interpretability of radiomics features is currently an urgent issue that needs to be addressed. As of now, the biological significance of radiomics is interpreted based on radiogenomics [32]. For example, it had been found that clustering based on PET features was associated with the cell cycle and WNT signaling pathways in lung adenocarcinoma, as well as with the cell cycle, p53, and WNT in lung squamous cell carcinoma [33]. Multiple correlations between CT radiomics features and representative genes of typical molecular pathways had been identified, allowing for non-invasive identification of the molecular nature of lung cancer [34]. Currently, the interpretability of radiomics features is still being explored in terms of correlations with clinical samples, lacking validation through molecular and animal experiments. And as revealed by EGFR investigations, the causal relationship between extrapulmonary MRI radiomics features and EGFR remains an unstudied area.

To the best of our knowledge, compared to CT (AUC = 0.80) and PET/CT (AUC = 0.84), the meta-analysis of MRI radiomics based on extrapulmonary metastases appeared to perform better in predicting EGFR mutations in the primary lesion of NSCLC (AUC = 0.9). According to subgroup analysis, it was found that the predictive performance was still above 0.8 depending on the location of different metastatic foci and cohort types. In terms of pooled SENC and SEPC evaluation, there was low heterogeneity, however, compared to meta-analyses of other EGFR mutations, heterogeneity was relatively low. In evidence-based medicine, publication bias may significantly affected the results of meta-analysis and may lead to misleading conclusions. Based on Begg's funnel plots for evaluating publication bias in included studies, it was found that there was no statistically significant publication bias. When the primary lesions are removed or cannot be sampled, EGFR mutations based on metastatic lesions can serve as a substitute for the primary lesion situation. The meta-analysis of MRI radiomics based on extrapulmonary metastases also demonstrated excellent potential in predicting EGFR mutations of NSCLC metastases (AUC = 0.86). However, there was no significant heterogeneity in the pooled SENC and SEPC. Based on Begg's Funnel plots for evaluating publication bias in included studies, it was also found that there was no statistically significant publication bias. In summary, it seems that radiomics analysis based on extrapulmonary metastases carries more EGFR mutation information, which may be related to differences in imaging principles.

This meta-analysis also has limitations. Firstly, the analyzed population was all Chinese, and the generalization ability of the conclusion was insufficient. Secondly, all included studies were retrospective, emphasizing the necessity of prospective studies. Thirdly, both RQS and QUADAS-2 had evaluation limitations and their interpretations were controversial. Finally, the number of studies included was small, which did not exclude the possibility of excessive exclusion during the literature screening process.

Conclusions

In summary, although there are some limitations, the prospects of MRI radiomics in predicting genetic information changes of EGFR gene in NSCLC primary and metastatic lesions cannot be denied, which is more in line with the requirements of precision medicine. However, the overall methodological quality of MRI radiomics still needs to be improved. In addition, the causal relationship between the information carried by MRI and the genetic information of genes still needs to be verified through scientific prospective experiments to promote the clinical translation of artificial intelligence.

Methods

Literature search program

A systematic search was conducted on original studies published before March 1, 2024 in the PubMed, Embase, and Web of Science database using keywords "Radiomics", "Lung", "EGFR", "EGFR-TKI", "Metastasis", and "Metastases".

Two reviewers with more than 3 years of experience in oncological imaging diagnosis independently reviewed the original study of preoperative radiomics, including the study abstract and full text. When disputes arose among reviewers, the final decision should be made by reviewers with more than 5 years of experience in oncological imaging diagnosis.

Literature screening criteria

The inclusion criteria were as follows: (1) radiomics studies; (2) confirmed by pathological or imaging examination as metastatic lung cancer; (3) preoperative MRI examination with metastatic lung cancer; (4) EGFR mutation/EGFR-ITK response data; (5) SENC and SPEC indicators could be directly/indirectly extracted from the full text.

The exclusion criteria were as follows: (1) comments, meta-analyses, case reports, guidelines or errata, repeated studies; (2) postoperative radiomics studies; (3) preoperative anti-tumor treatment; (4) deep learning or non-radiomics studies of EGFR mutation/EGFR-ITK response.

Literature data extraction

The literature data were extracted from the original studies: (1) basic characteristics (author, publication year, and study design); (2) cohort characteristics (cohort type, cohort sample size, and cohort EGFR mutation/EGFR-ITK response); (3) image characteristics (image segmentation software, radiomics software, feature selection strategy, and model algorithms); (4) evaluation indicators (AUC, SENC, and SPEC). The number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) were calculated according to the SENC and SPEC in each study report [35]. If there were two or more models based on the same cohort in a study, the model with higher performance was included.

Quality evaluation

The Radiomics Quality Score (RQS) was used to evaluate radiomics quality, which was an important tool to measure the rigor of artificial intelligence (AI) study [36,

37]. RQS included 16 evaluation indexes, including image acquisition, image preprocessing, validation, performance evaluation, practicality, open science. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) was used to evaluate the risk of bias at the study level, including bias in fields such as: (1) patient selection, (2) index test, (3) reference standard, and (4) flow and timing. The risk of bias in each field was classified as low (score = 2), high (score = 1), or unclear (score = 0). A modified version of QUADAS-2 proposed by Sollini et al. and verified by Bedrikovetski et al. was used [38, 39]. RQS, QUADAS-2, and phases classification were developed by two reviewers with more than 3 years of experience in oncological imaging diagnostics independently.

Statistical analysis

Stata and MetaDiSc software were used for summary analysis and plotting [40]. The Cochrane diagnostic test and I^2 statistic are used to evaluate heterogeneity between studies, with I^2 values greater than 50% indicating high heterogeneity [41]. Deek's funnel plots were used to assess whether the analysis was subject to publication bias [42]. The summary receiver operating characteristic (sROC) curve demonstrated the predictive potential of radiomics studies. Clinical practicality was evaluated based on probability after testing and the Fagan plots were created [43]. $P < 0.05$ was considered statistically significant.

Abbreviations

EGFR	Epidermal growth factor receptor
NSCLC	Non-small cell lung cancer
MRI	Magnetic resonance imaging
AUC	Area under the curve
SENC	Sensitivity
SPEC	Specificity
TKI	Tyrosine kinase inhibitors
CT	Computed tomography
PET/CT	Positron emission tomography/computed tomography
TP	True positives
TN	True negatives
FP	False positives
FN	False negatives
RQS	Radiomics Quality Score
AI	Artificial intelligence
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies
sROC	Summary receiver operating characteristic
IBSI	Image Biomarker Standardization Initiative
GLCM	Gray-level co-occurrence matrix
GLDM	Gray-level dependence matrix
GLRLM	Gray-level run length matrix
GLSZM	Gray-level size zone matrix
NGTDM	Neighboring gray tone difference matrix
NRI	Net reclassification index
IDI	Integrated discrimination improvement
DCA	Decision curve analysis

Acknowledgements

Not applicable.

Author contributions

Study design: Linyong Wu and Dayou Wei. Literature search and study selection: Linyong Wu and Songhua Li. Data extraction and quality assessment: Yan Li, Shaofeng Wu and Lifei Chen. Statistical analysis: Linyong Wu and Dayou Wei. Study supervision: Linyong Wu and Dayou Wei. Editing and review of the manuscript: all authors. All authors contributed to the article and approved the submitted version.

Funding
This study was funded by the Science and Technology Planning Project of Maoming City (2024109).

Availability of data and materials
No datasets were generated or analyzed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 3 July 2024 Accepted: 6 January 2025
Published online: 17 January 2025

References

1. Siegel RL, Miller KD, Wagle NS, et al. Cancer statistics, 2023. CA Cancer J Clin. 2023;73(1):17–48.
2. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. Nat Rev Clin Oncol. 2023;20(9):624–39.
3. Fan Y, Zhao Z, Wang X, et al. Radiomics for prediction of response to EGFR-TKI based on metastasis/brain parenchyma (M/BP)-interface. Radiol Med. 2022;127(12):1342–54.
4. Cheng Y, Wang H, Yuan W, et al. Combined radiomics of primary tumour and bone metastasis improve the prediction of EGFR mutation status and response to EGFR-TKI therapy for NSCLC. Phys Med. 2023;116: 103177.
5. Hou S, Fan Y, Wang X, et al. Radiomics for detection of the EGFR mutation in liver metastatic NSCLC. Acad Radiol. 2023;30(6):1039–46.
6. da Cunha SG, Shepherd FA, Tsao MS. EGFR mutations and lung cancer. Annu Rev Pathol. 2011;6:49–69.
7. Midha A, Dearden S, McCormack R. EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). Am J Cancer Res. 2015;5(9):2892–911.
8. Li H, Cao J, Zhang X, et al. Correlation between status of epidermal growth factor receptor mutation and distant metastases of lung adenocarcinoma upon initial diagnosis based on 1063 patients in China. Clin Exp Metastasis. 2017;34(1):63–71.
9. Doebele RC, Lu X, Sumey C, et al. Oncogene status predicts patterns of metastatic spread in treatment-naive nonsmall cell lung cancer. Cancer. 2012;118(18):4502–11.
10. Gu Y, Xu Y, Zhuang H, et al. Value and significance of brain radiation therapy during first-line EGFR-TKI treatment in lung adenocarcinoma with EGFR sensitive mutation and synchronous brain metastasis: appropriate timing and technique. Thorac Cancer. 2021;12(23):3157–68.
11. Song J, Shi J, Dong D, et al. A new approach to predict progression-free survival in stage IV EGFR-mutant NSCLC patients with EGFR-TKI therapy. Clin Cancer Res. 2018;24(15):3583–92.
12. Farris JC, Hughes RT, Razavian NB, et al. Brain metastasis incidence and patterns of presentation after definitive treatment of locally advanced non-small cell lung cancer: a potential argument for brain magnetic resonance imaging surveillance. Adv Radiat Oncol. 2022;8(3): 101058 (**Published 2022 Oct 23**).
13. Qi Y, Zhao T, Han M. The application of radiomics in predicting gene mutations in cancer. Eur Radiol. 2022;32(6):4014–24.
14. Felfli M, Liu Y, Zerka F, et al. Systematic review, meta-analysis and radiomics quality score assessment of CT radiomics-based models predicting tumor EGFR mutation status in patients with non-small-cell lung cancer. Int J Mol Sci. 2023;24(14):11433.
15. Ma N, Yang W, Wang Q, et al. Predictive value of 18F-FDG PET/CT radiomics for EGFR mutation status in non-small cell lung cancer: a systematic review and meta-analysis. Front Oncol. 2024;14:1281572.
16. Jiang X, Ren M, Shuang X, et al. Multiparametric MRI-based radiomics approaches for preoperative prediction of EGFR mutation status in spinal bone metastases in patients with lung adenocarcinoma. J Magn Reson Imaging. 2021;54(2):497–507.
17. Fan Y, Dong Y, Yang H, et al. Subregional radiomics analysis for the detection of the EGFR mutation on thoracic spinal metastases from lung cancer. Phys Med Biol. 2021. https://doi.org/10.1088/1361-6560/ac2ea7.
18. Cao R, Pang Z, Wang X, et al. Radiomics evaluates the EGFR mutation status from the brain metastasis: a multi-center study. Phys Med Biol. 2022. https://doi.org/10.1088/1361-6560/ac7192.
19. Fan Y, Dong Y, Wang H, et al. Development and externally validate MRI-based nomogram to assess EGFR and T790M mutations in patients with metastatic lung adenocarcinoma. Eur Radiol. 2022;32(10):6739–51.
20. Zheng L, Xie H, Luo X, et al. Radiomic signatures for predicting EGFR mutation status in lung cancer brain metastases. Front Oncol. 2022;12: 931812.
21. Fan Y, Dong Y, Sun X, et al. Development and validation of MRI-based radiomics signatures as new markers for preoperative assessment of EGFR mutation and subtypes from bone metastases. BMC Cancer. 2022;22(1):889.
22. Cao R, Chen H, Wang H, et al. Comprehensive analysis of prediction of the EGFR mutation and subtypes based on the spinal metastasis from primary lung adenocarcinoma. Front Oncol. 2023;13:1154327.
23. Hou S, Wang H, Wang X, et al. Tumor-liver interface in MRI of liver metastasis enables prediction of EGFR mutation in patients with lung cancer: a proof-of-concept study. Med Phys. 2024;51(2):1083–91.

24. Huang Z, Tu X, Yu T, Zhan Z, Lin Q, Huang X. Peritumoural MRI radiomics signature of brain metastases can predict epidermal growth factor receptor mutation status in lung adenocarcinoma. Clin Radiol. 2024;79(2):e305–16.
25. Ren M, Yang H, Lai Q, et al. MRI-based radiomics analysis for predicting the EGFR mutation based on thoracic spinal metastases in lung adenocarcinoma patients. Med Phys. 2021;48(9):5142–51.
26. Nguyen HS, Ho DKN, Nguyen NN, et al. Predicting EGFR mutation status in non-small cell lung cancer using artificial intelligence: a systematic review and meta-analysis. Acad Radiol. 2024;31(2):660–83.
27. Zhang G, Man Q, Shang L, et al. Using multi-phase CT radiomics features to predict EGFR mutation status in lung adenocarcinoma patients. Acad Radiol. 2024. https://doi.org/10.1016/j.acra.2023.12.024.
28. Mu W, Jiang L, Zhang J, et al. Non-invasive decision support for NSCLC treatment using PET/CT radiomics. Nat Commun. 2020;11(1):5228.
29. Zhao D, Zhao Y, He S, et al. High accuracy epidermal growth factor receptor mutation prediction via histopathological deep learning. BMC Pulm Med. 2023;23(1):244.
30. Huang EP, O'Connor JPB, McShane LM, et al. Criteria for the translation of radiomics into clinically useful tests. Nat Rev Clin Oncol. 2023;20(2):69–82.
31. Tabassum M, Suman AA, et al. Radiomics and machine learning in brain tumors and their habitat: a systematic review. Cancers (Basel). 2023;15(15):3845.
32. Aguado-Barrera ME, Sosa-Fajardo P, Gómez-Caamaño A, et al. Radiogenomics in lung cancer: where are we? Lung Cancer. 2023;176:56–74.
33. Kim G, Kim J, Cha H, et al. Metabolic radiogenomics in lung cancer: associations between FDG PET image features and oncogenic signaling pathway alterations. Sci Rep. 2020;10(1):13231.
34. Zhou M, Leung A, Echegaray S, et al. Non-small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. Radiology. 2018;286(1):307–15.
35. Bhatti KM, Khanzada ZS, Kuzman M, Ali SM, Iftikhar SY, Small P. Diagnostic performance of artificial intelligence-based models for the detection of early esophageal cancers in Barret's esophagus: a meta-analysis of patient-based studies. Cureus. 2021;13(6): e15447.
36. Spadarella G, Calareso G, Garanzini E, Ugga L, Cuocolo A, Cuocolo R. MRI based radiomics in nasopharyngeal cancer: systematic review and perspectives using radiomic quality score (RQS) assessment. Eur J Radiol. 2021;140: 109744.
37. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749–62.
38. Sollini M, Antunovic L, Chiti A, et al. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. Eur J Nucl Med Mol Imaging. 2019;46(13):2656–72.
39. Bedrikovetski S, Dudi-Venkata NN, Kroon HM, et al. Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. BMC Cancer. 2021;21(1):1058.
40. Huang H, Wang FF, Luo S, et al. Diagnostic performance of radiomics using machine learning algorithms to predict MGMT promoter methylation status in glioma patients: a meta-analysis. Diagn Interv Radiol. 2021;27(6):716–24.
41. Oh KE, Vasandani N, Anwar A. Radiomics to differentiate malignant and benign breast lesions: a systematic review and diagnostic test accuracy meta-analysis. Cureus. 2023;15(11): e49015.
42. Shen J, Ye Z, Xie H, et al. The relationship between *Helicobacter pylori* infection and recurrent aphthous stomatitis: a systematic review and meta-analysis. Clin Oral Investig. 2023;27(11):6345–56.
43. Zhu F, Yang C, Zou J, et al. The classification of benign and malignant lung nodules based on CT radiomics: a systematic review, quality score assessment, and meta-analysis. Acta Radiol. 2023;64(12):3074–84.

Publisher's Note