Article type: Review

Prediction of oncogene mutation status in non-small cell lung cancer: A systematic review and meta-analysis with a special focus on artificial-intelligence-based methods

Almudena Fuster-Matanzo[†1], Alfonso Picó Peris[†1], Fuensanta Bellvís Bataller[1], Ana Jimenez-Pastor[1], Glen J. Weiss[2], Luis Martí-Bonmatí[3,4], Antonio Lázaro Sánchez[5], Giuseppe L. Banna[6], Alfredo Addeo[7], Ángel Alberich-Bayarri[1]

[1]Quantitative Imaging Biomarkers in Medicine (Quibim), 46021 Valencia, Spain

[2]Quantitative Imaging Biomarkers in Medicine, Quibim, Boston, MA, USA, Boston, MA

[3]Grupo de Investigación Biomédica en Imagen, Instituto de Investigación Sanitaria La Fe, Avenida Fernando Abril Martorell, 106 Torre A planta 7, 46026, Valencia, Spain.

[4]Área Clínica de Imagen Médica, Área Clínica de Imagen Médica, Hospital Universitari i Politècnic La Fe, Avinguda Fernando Abril Martorell, 106 Torre E planta 0, 46026, València, Spain.

[5]

[6]Department of Oncology, Portsmouth Hospitals University NHS Trust, Portsmouth, PO6 3LY, UK; Faculty of Science and Health, School of Pharmacy and Biomedical Sciences, University of Portsmouth, Portsmouth, PO1 2UP, UK

[7]University Hospital Geneva, Geneva, Switzerland.

[†]These authors contributed equally to this work.

Correspondence:

Almudena Fuster-Matanzo

27 Quantitative Imaging Biomarkers in Medicine (Quibim),

28 EDIFICIO EUROPA, Av. d'Aragó, 30, Planta 13

29 46021 Valencia, Spain

30 Mail: almudenafuster@quibim.com

31 Tel: +34 652124031

32

ABSTRACT

Background

In non-small cell lung cancer (NSCLC), alternative strategies to determine patient oncogene mutation status are essential to overcome some of the drawbacks associated with current methods. We aimed to review the use of radiomics alone or in combination with clinical data and to evaluate the performance of artificial intelligence (AI)-based models on the prediction of oncogene mutation status.

Methods

A PRISMA-compliant literature review was conducted. The Medline (via Pubmed), Embase, and Cochrane Library databases were searched for studies published through June 30, 2023 predicting oncogene mutation status in patients with NSCLC using radiomics. Independent meta-analyses evaluating the performance of AI-based models developed with radiomics features or with a combination of radiomics features plus clinical data for the prediction of different oncogenic driver mutations were performed. A meta-regression to analyze the influence of methodological/clinical factors was also conducted.

Results

Out of the 615 studies identified, 89 evaluating models for the prediction of epidermal growth factor-1 (EGFR), anaplastic lymphoma kinase (ALK), and Kirsten rat sarcoma virus (KRAS) mutations were included in the systematic review. A total of 38 met the inclusion criteria for the meta-analyses. The AI algorithms' sensitivity/false positive rate (FPR) in predicting EGFR, ALK, and KRAS mutations using radiomics-based models was 0.753 (95% CI 0.721–0.783)/0.346 (95% CI 0.305–0.390), 0.754 (95% CI 0.639–0.841)/ 0.225 (95% CI 0.163–0.302), and 0.744 (95% CI 0.605–0.846)/0.376 (95% CI 0.274–0.491), respectively. A meta-analysis of combined models was only possible for EGFR mutation, revealing a sensitivity/FPR of 0.800 (95% CI 0.767–0.830)/0.335 (95% CI 0.279–0.396). No statistically significant results were obtained in the meta-regression.

Conclusions

Radiomics-based models may represent valuable non-invasive tools for the determination of oncogene mutation status in NSCLC. Further investigation is required to analyze whether clinical data might boost their performance.

Keywords: radiomics, artificial intelligence, medical imaging, oncogene mutation status, non-small cell lung cancer.

INTRODUCTION

Lung cancer represents the most often diagnosed cancer in both women and men worldwide, ranking first and third, respectively, and remaining the leading cause of cancer death[1]. Non-small cell lung cancer (NSCLC), the most frequent histological subtype, accounts for 80%–85% of cases, being adenocarcinoma the most common subtype (40%–50% of cases). Adenocarcinoma can be further subdivided into distinct molecular subtypes[2]. Indeed, molecular subtyping has become highly relevant in the disease context, as genotype-driven therapy ("targeted therapy") is nowadays the standard of care for a significant subgroup of patients with advanced and metastatic NSCLC[3]. However, traditional methods for determining the molecular genotype, as well as the possible emergence of drug resistance mutations during patient's follow-up, entail invasive biopsies and genetic sequence testing, procedures with multiple number of associated drawbacks including high costs, sampling bias, lack of enough sample, turnaround time, and medical complications[4-6]. Importantly, the overall accessibility of molecular diagnostics and liquid biopsy may be limited for many patients[7], highlighting the need to investigate complementary methods to characterize the oncogene mutation status of lesions.

Radiological imaging represents a potent non-invasive tool for lung cancer, from the screening, diagnosis and staging of the disease to the management, therapeutic planification and follow-up of both early- and advanced-stage cases[8]. Specifically, computed tomography (CT) remains the standard of care for lung cancer visualization, providing excellent morphological and textural information. In recent years, radiomics, the process of extracting and analyzing quantitative features from medical images to investigate potential connections with biology and clinical outcomes, has gained increasing attention for its applicability in several oncological diseases including lung cancer[8]. The application of artificial intelligence (AI) to imaging analyses has enabled important clinical needs to be met. This includes the prognostication of outcomes or the prediction of response to treatment, disease progression, or the mutational and molecular profiling of tumors[9]. In particular, the use of radiomics coupled with AI methods has demonstrated to be a promising non-invasive alternative tool for the prediction of oncogene mutation status in NSCLC[8].

In this systematic review and meta-analysis we aimed to: 1) review the available scientific evidence on the use of imaging-based models and radiomics for the prediction

of the main targetable oncogenic driver alterations in NSCLC, including epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), and Kirsten rat sarcoma virus (KRAS); 2) analyze the overall performance of specifically, AI-based methods, for the prediction of oncogene mutation status; 3) evaluate whether the inclusion of clinical variables in the models improve their performance; 4) evaluate the impact of the available evidence from a clinical perspective.

MATERIAL AND METHODS

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines[10]. The review was registered on PROSPERO before initiation (registration no. CRD42022349809).

Search strategy

A systematic search for eligible publications published through 30 June 2023 was performed in Medline (via Pubmed), Cochrane Library and EMBASE databases using the keywords “Radiomics”, “NSCLC” and “Mutational status”. Further details on the search terms used in each database are provided in **Supplementary Table S1**. There were no limitations on the publishing year, participant age, or nationality. The search was exclusively limited to English-language publications.

Study selection

Literature search and study selection were independently performed by two reviewers (A.F.M. and A.L.S.). To find relevant publications, they reviewed the titles and abstracts. Studies that satisfied the inclusion criteria were then manually assessed for eligibility by full-text screening. Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org) was used as a screening and data extraction tool.

Inclusion criteria

Papers were included in the qualitative synthesis (systematic review) if meeting the following inclusion criteria based on Patient, Index test, Comparator, Reference test, Diagnosis of reference (PIRD) questions: 1) being focused on the ability of radiomics to predict oncogene mutation status in NSCLC; 2) radiomics features were extracted from CT or from F-18 fluoro-deoxy-glucose (FDG)/CT scans; 3) a full text was available; 4) were written in English.

Exclusion criteria

Papers describing studies conducted using MRI scans (not the standard of care for NSCLC patients) or performed in phantom or animal models, or published as case reports, editorials, reviews, poster presentations, letters, editorials, or meeting abstracts were excluded. Papers not on the field of interest were also excluded.

For the quantitative synthesis (meta-analysis), the following additional exclusion criteria were applied: 1) oncogene mutation status was not the primary objective of the paper; 2) were focused on specific mutation subtypes; 3) did not apply AI-based methodologies; 4) developed simultaneous detection models or discriminant models; 5) sensitivity or specificity metrics were not available and could not be calculated; 6) were not comparable with the other articles included (model was developed based on intra- and extra-tumor derived radiomics features); 7) only included models developed with a combination of quantitative features extracted from PET/CT or from PET images (strictly adhering to a clinical perspective, PET scanning equipment is not always available and CT remains the standard of care for NSCLC patients); 8) did not reach a sufficient quality score according to the quality assessment (described below).

Quality assessment

The methodological quality of each study for its possible inclusion in the quantitative assessment was evaluated by using the Checklist for Artificial Intelligence in Medical Imaging (CLAIM)[11]. Classification, image reconstruction, text analysis, and workflow optimization are some of the applications of AI in medical imaging that are addressed by CLAIM, which is modeled after the Standards for Reporting of Diagnostic Accuracy Studies (STARD) guideline[12-15]. CLAIM checklist consists of 42 items divided into the conventional sections included in peer-reviewed scientific articles: title or abstract (1 item), abstract (1 item), introduction (2 items), methods (28 items subdivided into study design [2 items], data [7 items], ground truth [5 items], data partitions [3 items], model [3 items], training [3 items] and evaluation [5 items]), results (5 items subdivided into data [2 items] and model performance [3 items]), discussion (2 items) and other information (3 items). The CLAIM guideline offers a roadmap for writers and reviewers with the intention of fostering clear, open, and verifiable scientific discourse on the use of AI in medical imaging[11].

For our quality assessment, a score was calculated for each paper ([total score, 42 - number of "not applicable" fields in each case]). A cut-off value of at least half of the total score after removing the "not applicable" items was established for the inclusion in the quantitative analysis. Therefore, this cut-off value varied for each study depending on the number of items that were applicable from among the 42 total items included in the CLAIM checklist (e.g., a cut-off value of 19 was established for those studies in which only 38 items of the checklist were applicable). **See Supplementary Table S2**. The assessment of the rigor, quality, and generalizability of the work of all enrolled studies was performed by three reviewers (A.J.P., F.B.B. and A.P.P.).

Data extraction

Data extracted included the following: (1) study details: first author, publication year, research questions, study design; (2) patient details: the source of data acquisition (single-center/multicenter), sample size, smoking history, age, sex, TNM staging, treatment status (naïve or any treatment received prior image acquisition), histological subtype; (3) imaging details: imaging modality, plain or contrast CT; (4) oncogene mutation status-related information: type of mutation, specific subtype of mutation (if available), sequencing method; sequencing kit (5) radiomics details: segmentation software, type of segmentation (manual, automatic, or semi-automatic), radiomics feature extraction software, number of imaging features extracted, number and name of radiomics features included in final models, features selection methods, type of models constructed (machine learning [ML], deep learning [DL], classical statistical model), final classifier used in machine learning models, clinical variables included in the models (if applicable), and models performance. Two independent reviewers (A.F.M. and A.L.S.) completed the initial screening and extracted data from all included studies.

Data analysis

For studies including models based on features extracted from different imaging modalities, only those based on CT scans were included in the quantitative analysis. A bivariate analysis of sensitivity and specificity as proposed by Reitsma et al.[16] was chosen to perform the meta-analyses. This method has the distinct advantage of preserving the two-dimensional nature of the underlying data. It can also produce summary estimates of sensitivity and specificity (false positive rate [FPR, 1-specificity]), recognizing any possible correlation between these two measures. The

method uses a random effect approach in which the values of the sensitivity and FPR estimates are obtained with restricted maximum likelihood. As a complement to the bivariate approach, the summary receiver operating characteristic (sROC) was calculated by converting each pair of sensitivity and specificity into a single measure of accuracy, the diagnostic odds ratio (DOR).

The analyses were carried out by reproducing the confusion matrices of each model presented in the studies, the number of cases and the prevalence of oncogene mutant positive cases. All calculations were performed on the basis of validation cohorts for studies applying a training/validation split method, or on the basis of the total sample when cross-validation was the validation strategy. To ensure homogeneity, calculations were conducted based on internal validation cohort data when external validation was also performed (minority of the cases).

Finally, a meta-regression analysis was performed to measure the possible influence of the following predictors: (1) average age of the cases, (2) manual segmentation vs semi-automatic segmentation vs both procedures (no studies including automatic segmentation approaches met the inclusion criteria for the quantitative analysis), (3) whether the model included only radiomics features or was combined with clinical variables, and (4) whether the model was classified as ML or DL. The heterogeneity in the description of the clinical variables included in the models prevented the inclusion of additional predictors of greatest clinical interest. Only the best model from each study according to its DOR was selected. When the mean/median age was not available due to the heterogeneity among studies when presenting descriptive results, it was inferred from the information obtained. Thus, mean and median values were indistinctly considered; when both values were provided, an average of both was calculated. If mean values were absent, median values were considered and viceversa. If both values were absent from the validation cohort, mean/median age from the total cohort was considered. When this information was not available either, the study was not included in the meta-regression.

All the analyses were performed using R Statistical Software v4.2.2 and the packages mada and tidyverse.

RESULTS

In total 615 articles were obtained according to the search strategy (**Figure 1**). After de-duplication, 397 studies were obtained and screened. According to the inclusion and exclusion criteria, 89 studies were included in the qualitative analysis (systematic review), all of them developing models for the prediction of EGFR, ALK, and/or KRAS. Out of those, 38 were found eligible for the quantitative part of the study (meta-analyses). As detailed in **Supplementary Table S2**, all papers passed this quality check and were therefore included.

Qualitative analysis (systematic review)

Methodological characteristics of the studies

The methodological characteristics of the studies are summarized in **Table 1**. Most of the studies (*n* = 69/ 89) applied exclusively ML algorithms, while this methodology was also used to build comparator models in 10 articles in which DL techniques were the main methodological approach followed. Only three studies exclusively applied DL algorithms, while classical statistical models were used in seven publications. Among the 79 articles applying ML techniques, the most common classifier used was logistic regression (*n* = 38), followed by support vector machine (*n* = 35) and by random forest (*n* = 29). In terms of partitioning strategy, training-validation split was the most frequent technique (*n* = 71). External validation was only performed in a small set of studies (*n* = 9). Regarding imaging techniques, CT was the most frequently used modality (*n* = 61), followed by PET/CT (*n* = 22) and by PET alone (*n* = 4). Additionally, in one study[17] PET/CT scans and contrast-enhanced CT images independently acquired were collected, while in another study[18], PET/CT, CT, and contrast-enhanced diagnostic quality (CTD) images were used. Of the 61 studies conducted with CT scans, 39 included non-contrast-enhanced images, 18 contrast-enhanced images, in two contrast-related information was not specified and in two both contrast- and non-contrast-enhanced scans were included. Regarding tumor segmentation, a manual approach was followed in 48 studies, and automatic and semi-automatic segmentations were applied in two and 31 studies, respectively; three studies applied both methodologies (for verification or a different approach according to the imaging modality used) and five studies did not specify the method utilized for tumor segmentation.

Clinical characteristics of the studies

The 89 studies evaluated in the qualitative synthesis included a total of 32,084 patients with NSCLC. Although most of the studies included >200 patients, in 42 publications, the sample size did not reach this figure and in 11 studies sample size was even lower than 100. All studies were retrospective and mostly unicentric (n = 72); the number of participant centers was not specified in one study[19]. In general, basic clinical and demographic information collected included sex, age, smoking status, TNM stage, histology, and treatment status at the moment of image acquisition, although this information was not available in 13, 9, 22, 34, 22 and 24 studies out of the 89 assessed, respectively. The clinical characteristics of the patients included in the 89 studies are depicted in **Table 2**. The median [range]/ mean ± standard deviation (SD) age of patients was 61.78 [59–64.17] years and 61.71 ± 3.64 years, respectively. In terms of sex, the total population was balanced, with 13,574 females and 14,066 males. The smoking history was available for 23,200 patients, and many were non-smokers (n = 12,813); while smoking history was unknown for 1,146 patients. Out of the 55 studies detailing information about the TNM stage, the majority of them (n = 40) included information about the four stages (I-IV), either provided per group or grouped in stages I-II and stages III-IV. Among the 15 studies that did not include patients of all stages, two studies included only early stage patients (stages I and II)[20, 21], two included patients stage II-IV[22, 23], six included only patients of stages III and IV[24-29] (three of them with a majority of stage IV patients[24, 27, 29]), and five included patients of stages I-III without including the most advanced stage[19, 30-33]. A total of 65 studies included patients with adenocarcinoma: 43 exclusively including this histology subtype and 22 including other NSCLC histology types as well. Finally, in most of the cases (n = 65), images were acquired before patients received any treatment, with two studies also including post-treatment images[34, 35]. In 24 studies, no information on treatment was detailed, although in some of them image acquisition before surgery[31, 36-41], before polymerase chain reaction (PCR)[42], or before pathological diagnosis[43] was detailed as an inclusion criterion. In five studies[19, 44-47], authors specify that patients had not received radiotherapy or chemotherapy, but no information on targeted therapy was provided. Finally, only one study[48] out of the 89 included in the systematic review, which did not meet the inclusion criteria to be considered for the meta-analysis, included patients who had received treatment with tyrosine kinase inhibitors (TKIs).

Quantitative analysis (meta-analysis)

A total of 38 studies met the inclusion criteria for the quantitative assessment (n = 17,066 patients). Three main different meta-analyses including radiomics-based models were conducted: 1) a meta-analysis including studies focused on the detection of EGFR (n = 34 studies)[17, 25-28, 32, 33, 36, 38, 39, 46, 49-71]; 2) a meta-analysis including studies focused on the detection of ALK (n = 3 studies)[72-74]; 3) a meta-analysis including studies focused on the detection of KRAS (n = 4 studies)[47, 50, 54, 62]. In three studies, authors developed models for the detection of both EGFR and KRAS[50, 54, 62]. Furthermore, a separate meta-analysis was conducted for combined models (radiomics features + clinical variables) for the prediction of EGFR (not enough studies for ALK or KRAS mutations). Studies included in all the meta-analyses conducted are summarized in **Supplementary Table S3.** Details on the radiomics features included in the EGFR, ALK, and KRAS models are summarized in **Supplementary Table S4**, **Supplementary Table S5** and **Supplementary Table S6**, respectively. In terms of radiomics variables, models grouped different combinations of first order, shape, gray level co-occurrence matrix (GLCM), gray level size zone matrix (GLSZM), gray level run length matrix (GLRLM), neighboring gray tone difference matrix (NGTDM), and gray level dependence matrix (GLDM) features. Clinical data included sex, smoking history, and/or histological type in the majority of studies.

EGFR

Results of the meta-analysis focused on models built with radiomics features are summarized in **Figure 2**. Note that this meta-analysis also included a study[50] in which predictions were based on features extracted by a multi-channel and multi-task deep learning model with the ability to simultaneously detect EGFR and KRAS oncogene mutations; and consequently, did not include radiomics features (only single-task results for the independent prediction of EGFR and KRAS were considered for the quantitative analysis). A hierarchical sROC curve was plotted for the included 24 studies [17, 25, 27, 32, 36, 39, 46, 49-51, 54, 55, 57-60, 62, 64-66, 68-71] that evaluate the performance of AI algorithms in predicting EGFR mutation status in NSCLC (**Supplementary Figure S1**). Eight studies assessed more than one model[32, 36, 39, 50, 59, 60, 65, 70]. As observed, radiomics-based models exhibited high diagnostic performance in predicting EGFR mutation status with an overall AUC of 0.766. The AI algorithms' sensitivity in determining the EGFR mutation

status varied from 0.362 to 0.948, resulting in an estimate of 0.753 (95% CI 0.721–0.783). The FPR of these algorithms ranged from 0.022 to 0.761, with a estimate of 0.346 (95% CI 0.305–0.390). Detecting a positive case for EGFR mutation was almost six times more likely than not detecting it (DOR = 5.70 [95% CI 4.74–6.81]).

The effect of adding clinical variables to radiomics models or to models including both radiomics and deep features[65] (models including clinical data and radiomic or deep features referred in this work as combined models) in the prediction of EGFR mutation was also analyzed. This meta-analysis included 23 studies[25, 26, 28, 32, 33, 38, 39, 46, 51-53, 56-58, 61-69], of which four of them developed more than one model[32, 39, 56, 67]. Results are depicted in **Figure 3** and sROC curve in **Supplementary Figure S1**. Overall, the performance of combined models slightly improved compared to radiomics models, with an AUC of 0.811 and a sensitivity of 0.800 (95% CI 0.767–0.830; model's sensitivity ranging from 0.523 to 0.944). The FPR resulted similar with a value of 0.335 (95% CI 0.279–0.396; model's FPR ranging from 0.167 to 0.760.). Detecting a positive case for EGFR mutation with combined models was more than eight times more likely than not detecting it (DOR = 8.35 [95% CI 6.77–10.20]).

ALK

The meta-analysis focused on radiomics-based models included three studies[72-74], one of which developed two different models, one based on pre-contrast images and another one on post-contrast images[73]. An overall AUC of 0.831 was obtained for the prediction of ALK aberration, with a sensitivity ranging from 0.682 to 0.825, resulting in an estimate of 0.754 (95% CI 0.639–0.841). The FPR of these algorithms ranged from 0.167 to 0.277, with an estimate of 0.225 (95% CI 0.163–0.302). Detecting a positive case for ALK aberration was 11 times more likely than not detecting it (DOR = 5.70 [95% CI 5.83–19.10]) (**Figure 4** and **Supplementary Figure S2**). Given the lack of enough studies developing combined models, a meta-analysis to assess the effects of adding clinical variables in the prediction of ALK aberration was not possible. The only study[74] that developed a model including age, sex, smoking history, smoking index, clinical stage, distal metastasis and pathological invasiveness of the tumor in combination with conventional CT features and different first order, GLCM, GLSZM, and GLRL radiomics features demonstrated increased performance in predicting ALK aberration of the combined model vs the radiomics-based model, but only in the

primary cohort (AUC , 0.83–0.88, $p = 0.01$), not in the testing cohort (AUC , 0.80–0.88, $p = 0.29$).

KRAS

Four studies met the inclusion criteria for the meta-analysis assessing models for KRAS mutation prediction[47, 50, 54, 62], among which, three of them also developed models for EGFR mutation prediction[50, 54, 62]. KRAS/EGFR models were independently built except in one study, in which a multi-channel multi-task DL model for the prediction of both KRAS and EGFR mutations was developed[50]. However, and according to the inclusion criteria, only single-task metrics were considered for the quantitative analysis despite the multi-channel version displayed the highest performance for the simultaneous detection of both oncogenic driver mutations. Results of the meta-analysis evaluating radiomics-based models are shown in **Figure 4** and **Supplementary Figure S3**. KRAS mutation was predicted with an overall AUC of 0.732 and a sensitivity of 0.744 (95% CI 0.605–0.846; model's sensitivity ranging from 0.641 to 0.875. The FPR was 0.376 (95% CI 0.274–0.491; model's FPR ranging from 0.259 to 0.468). Detecting a positive case for KRAS mutation with radiomics-based models was more than five times more likely than not detecting it (DOR = 8.35 [95% CI 1.98–11.70]). Like ALK, the lack of enough KRAS studies made it impossible to perform a meta-analysis analyzing combined models. Only Ríos Velázquez et al.[62] built a model including age, sex, smoking status, race, and clinical stage together with radiomics features that performed similar to the radiomics model (AUC = 0.69 [95% CI: 0.63– 0.75] vs AUC = 0.63 [95% CI: 0.57– 0.69]) and worse than a model developed only with clinical data AUC = 0.75 [95% CI: 0.69–0.80].

Meta-regression and subgroup analysis

The possible effects of different predictors on the predictive performance of the models was evaluated for EGFR mutation (not enough studies were available for ALK or KRAS mutations). Neither age, nor the use of contrast, nor the type of segmentation (manual/semi-automatic/automatic), nor the model (radiomics/combined), nor the AI methodology (machine learning/deep learning), yielded statistically significant results (**Supplementary Table S7**).

DISCUSSION

At present, molecular testing performed on biopsied tissue remains the gold standard for diagnosis and genotyping in advanced NSCLC[75, 76]. However, given the associated limitations and inconveniences, such as the lack of enough tissue for successful testing[77, 78], or the long turnaround times[76], there is a need to validate and incorporate new procedures into routine clinical practice. In recent years, liquid biopsy has emerged as a promising alternative in NSCLC, especially in clinical scenarios[78]. Likewise, radiomics have shown encouraging results in prognosis and prediction in this setting[79]. In general, both methodologies possess great potential, since they are both simple, straightforward to do, and repeatable at patient follow-up visits, which makes it possible to gather important data about the type of tumor, its aggressiveness, its progression, and its response to therapy[80]. Radiomics has the additional advantage of only requiring medical images and capturing patient-level and tissue-level heterogeneity, such as CT scans in lung cancer, that are usually acquired as part of the patient's standard journey, representing an affordable methodology both in terms of resources and costs. It is important that new techniques are properly validated to facilitate their standardization, prior to incorporation into the routine clinical workflow.

To our knowledge, this is the first systematic review and meta-analysis that analyzes the performance and applicability of different imaging-based models for the prediction of three of the most common oncogene mutations—EGFR, ALK and KRAS—in NSCLC from a clinical perspective and with a special focus on AI methodologies. So far, results were only available for EGFR studies and did not take clinical aspects into account[81]. Thus, the results of our different meta-analyses demonstrate that AI-based models developed with CT-derived radiomics features showed good performance in predicting EGFR, ALK, and KRAS mutations with a sensitivity of 0.753 [95% CI (0.721–0.783)], 0.754 [95% CI (0.639–0.841)] and 0.744 [95% CI (0.605–0.846)], respectively. Whether the inclusion of clinical variables increase models' performance cannot be concluded from our results, although we believe that increasing the number of studies would probably confirm the trends observed in our quantitative analysis of EGFR mutation.

Our outcomes point to radiomics as a candidate screening tool for oncogene mutation status determination. We especially focused on CT-based models, aiming to obtain

conclusions as applicable as possible to the standard clinical workflow since CT remains the most utilized imaging tool in NSCLC[82]. From our work, we conclude that in addition to additional validation of our findings that future studies should be conducted that consider the following important aspects. Firstly, a minimum sample size should be guaranteed to ensure the reliability of the results obtained with AI-based models[83, 84]. In both our systematic review and meta-analyses, more than half of the studies were conducted in >200 patients (n = 46/89 and n = 23/38 [n = 20/34 for EGFR, n = 2/3 for ALK and n = 3/4 for KRAS]), but still a sizable number had small sample sizes, which definitely limited the relevance of their conclusions. Multicentric designs would be also desirable to get more solid conclusions, an approach that few studies followed (n = 16/89 in the systematic review and n = 10/39 in the meta-analysis [n = 10/34 for EGFR, n = 0/3 for ALK and n = 3/4 for KRAS]). Secondly, including independent cohorts for external validations would reinforce the results, leading to more robust and reproducible models. Out of the 89 studies included in the qualitative analysis, only 9 used external cohorts for validation[43, 46, 60, 63, 67, 68, 85-87], of which five were included in the EGFR meta-analysis[46, 60, 63, 67, 68]. Finally, it is important that patient populations reflect clinical practice. Thus, considering the potential applicability of the models for diagnostic purposes, studies should be conducted in treatment naïve populations to avoid possible therapy-related confounding effects, an inclusion criterion mostly applied in the studies evaluated in this work, but still missing in some of them. Additionally, studies should be carried out preferably in stage III-IV NSCLC patients (especially in those at stage IV, for whom clinical guidelines recommend molecular testing[75, 76]). As demonstrated in this work, most of the studies published so far do not provide information on TNM stage or include patients from all stages. Despite the heterogeneity of the studies evaluated, we believe that the evidence provided is enough as to demonstrate the potential of radiomics in oncogene mutation status determination. Thus, AI-based models using radiomics extracted from CT scans could be effective non-invasive screening tools to detect targetable driver mutations in NSCLC with good sensitivity and moderate specificity. These tools would not be intended to replace gold standard techniques, such as PCR or next-generation sequencing, but to allow for the potential earlier identification of ideal candidates to be genetically tested, saving time, costs, and samples. Consequently, a high sensitivity would ensure the identification of oncogene mutation positive patients for whom laboratory-based testing would be subsequently confirmed.

When the influence of different factors on the prediction of EGFR mutation was evaluated, no statistically significant results were obtained, probably due to the limited number of studies included and the presence of missing data. However, some of those factors might play an essential role and should be considered when developing accurate models to be potentially implemented into clinical practice. Indeed, some of the studies included in our qualitative analysis analyzed the impact of different methodological aspects on the performance of the models. For example, Huang et al.[35] demonstrated that interobserver variability in tumor segmentation affects the use of radiomics to predict oncogene mutation status, which suggests that automatic or semi-automatic models might be more suitable. In the study by Shiri et al.[88], the application to radiomics features of ComBat harmonization improved the performance of the models toward more successful prediction of EGFR and KRAS mutations. Likewise, other authors have pointed to the impact of the experimental settings on the robustness of radiomics features[89], or the influence of CT slice thickness on the predictive performance of radiomics-based models[31]. It is also worth mentioning the relevance of using a particular AI methodology. Although we found no differences in the EGFR mutation predictive performance between ML and DL methods, most likely due to the limited number of available DL-based studies, the latter might offer some advantages over the former. Thus, while in radiomics analysis a process of lesion segmentation and subsequent feature extraction is required, which introduces certain degree of variability and can be a high time-consuming task, DL models only required a bounding box of the lesion, greatly reducing this effect. On the other hand, DL models, and in particular end-to-end convolutional neural network (CNN) models, such those developed in most of the DL studies included in our work[37, 42, 43, 50, 58, 68, 85, 86, 90, 91], are generally more complex in terms of the number of parameters, allowing to solve more complicated problems than traditional ML models. Considering available evidence, it seems reasonable to think that methodologic approaches should be carefully revised when validation studies are designed and conducted.

Our study has also some limitations, mainly derived from the limitations of the publications included. Thus, it is based on retrospective studies displaying great heterogeneity in terms of methodology and patient clinical characteristics, which clearly hamper the impact of our conclusions. Additionally, the limited available evidence for ALK and KRAS mutations, makes it difficult to draw solid conclusions. Despite this,

our work gathers the most up-to-date and complete evidence (all models developed in each of the studies were analyzed) on imaging-based models for the prediction of three of the most important oncogene mutations in NSCLC, following a clinical approach and a special focus on AI models. Our exhaustive review and meta-analyses are intended to provide solid evidence for future research in the field.

In conclusion, radiomics-based models offer a useful and non-invasive method for determining the status of EGFR mutations in NSCLC and seem to retain similar predictive value for ALK and KRAS mutations. Additionally, although the inclusion of clinical variables tends to increase the performance of the models, further validation is required.

ACKNOWLEDGEMENTS

CONFLICTS OF INTEREST

GJW reports personal fees from Quibim related to this work. He is a former employee of SOTIO Biotech Inc., and reports personal fees from Imaging Endpoints II, MiRanostics Consulting, Gossamer Bio, International Genomics Consortium, Angiex, Genomic Health, Oncacare, Rafael Pharmaceuticals, Roche, Immunocore, Kymera, and SPARC-all outside this submitted work; has ownership interest in MiRanostics Consulting, Exact Sciences, Moderna, Agenus, Aurinia Pharmaceuticals, and Circulogene-outside the submitted work; and has issued patents- all outside the submitted work. The remaining authors declare no conflicts of interest.

FUNDING

REFERENCES

1 Sung H, Ferlay J, Siegel RL et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 Cancers in 185 countries. *CA Cancer J Clin*. 2021; 71(3): 209-249.

2 Osmani L, Askin F, Gabrielson E et al. Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): moving from targeted therapy to immunotherapy. *Semin Cancer Biol*. 2018; 52(Pt 1): 103-109.

3 König D, Savic Prince S, Rothschild SI. Targeted therapy in advanced and metastatic non-small cell lung cancer. An update on treatment of the most important actionable oncogenic driver alterations. *Cancers (Basel)*. 2021; 13(4).

4 Chiu YW, Kao YH, Simoff MJ et al. Costs of biopsy and complications in patients with lung cancer. *Clinicoecon Outcomes Res*. 2021; 13: 191-200.

5 Manicone M, Poggiana C, Facchinetti A et al. Critical issues in the clinical application of liquid biopsy in non-small cell lung cancer. *J Thorac Dis*. 2017; 9(Suppl 13): S1346-s1358.

6 Young M SR. Percutaneous lung lesion biopsy. [Updated 2023 Jun 19]. *StatPearls [Internet]*. . Treasure Island (FL): StatPearls Publishing.

7 Di Capua D, Bracken-Clarke D, Ronan K et al. The liquid biopsy for lung cancer: state of the art, limitations and future developments. *Cancers (Basel)*. 2021; 13(16).

8 Wu G, Jochems A, Refaee T et al. Structural and functional radiomics for lung cancer. *Eur J Nucl Med Mol Imaging*. 2021; 48(12): 3961-3974.

9 Bera K, Braman N, Gupta A et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol*. 2022; 19(2): 132-146.

10 Page MJ, McKenzie JE, Bossuyt PM et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021; 372: n71.

11 Mongan J, Moy L, Kahn CE. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. *Radiol Artif Intell*. 2020; 2(2): e200029.

12 Bossuyt PM, Reitsma JB. The STARD initiative. *Lancet*. 2003; 361(9351): 71.

13 Bossuyt PM, Reitsma JB, Bruns DE et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Radiology*. 2003; 226(1): 24-28.

14 Bossuyt PM, Reitsma JB, Bruns DE et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology*. 2015; 277(3): 826-832.

15 Cohen JF, Korevaar DA, Altman DG et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016; 6(11): e012799.

16 Reitsma JB, Glas AS, Rutjes AW et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005; 58(10): 982-990.

17 Nair JKR, Saeed UA, McDougall CC et al. Radiogenomic models using machine learning techniques to predict EGFR mutations in non-small cell lung cancer. *Can Assoc Radiol J*. 2021; 72(1): 109-119.

18 Shiri I, Maleki H, Hajianfar G et al. Next-generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using

multimodal imaging and machine learning algorithms. *Mol Imaging Biol*. 2020; 22(4): 1132-1148.

19 Dang Y, Wang R, Qian K et al. Clinical and radiological predictors of epidermal growth factor receptor mutation in nonsmall cell lung cancer. *J Appl Clin Med Phys*. 2021; 22(1): 271-280.

20 Omura K, Murakami Y, Hashimoto K et al. Detection of EGFR mutations in early-stage lung adenocarcinoma by machine learning-based radiomics. *Transl Cancer Res*. 2023; 12(4): 837-847.

21 Wang X, Kong C, Xu W et al. Decoding tumor mutation burden and driver mutations in early stage lung adenocarcinoma using CT-based radiomics signature. *Thorac Cancer*. 2019; 10(10): 1904-1912.

22 Liu Q, Sun D, Li N et al. Predicting EGFR mutation subtypes in lung adenocarcinoma using (18)F-FDG PET/CT radiomic features. *Transl Lung Cancer Res*. 2020; 9(3): 549-562.

23 Yang L, Xu P, Li M et al. PET/CT radiomic features: a potential biomarker for EGFR mutation status and survival outcome prediction in NSCLC patients treated with TKIs. *Front Oncol*. 2022; 12: 894323.

24 Hong D, Xu K, Zhang L et al. Radiomics signature as a predictive factor for EGFR mutations in advanced lung adenocarcinoma. *Front Oncol*. 2020; 10: 28.

25 Lu J, Ji X, Wang L et al. Machine learning-based radiomics for prediction of epidermal growth factor receptor mutations in lung adenocarcinoma. *Dis Markers*. 2022; 2022: 2056837.

26 Wu S, Shen G, Mao J et al. CT radiomics in predicting EGFR mutation in non-small cell lung cancer: a single institutional study. *Front Oncol*. 2020; 10: 542957.

27 Yang C, Chen W, Gong G et al. Application of CT radiomics features to predict the EGFR mutation status and therapeutic sensitivity to TKIs of advanced lung adenocarcinoma. *Transl Cancer Res*. 2020; 9(11): 6683-6690.

28 Zhang L, Chen B, Liu X 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.

29 Zhu Y, Guo YB, Xu D et al. A computed tomography (CT)-derived radiomics approach for predicting primary co-mutations involving TP53 and epidermal growth factor receptor (EGFR) in patients with advanced lung adenocarcinomas (LUAD). *Ann Transl Med*. 2021; 9(7): 545.

30 Aide N, Weyts K, Lasnon C. Prediction of the presence of targetable molecular alteration(s) with clinico-metabolic (18) F-FDG PET radiomics in non-Asian lung adenocarcinoma patients. *Diagnostics (Basel)*. 2022; 12(10).

31 Li Y, Lu L, Xiao M et al. CT slice thickness and convolution Kernel affect performance of a radiomic model for predicting egfr status in non-small cell lung cancer: a preliminary study. *Sci Rep*. 2018; 8(1): 17913.

595 32 Zhu H, Song Y, Huang Z et al. Accurate prediction of epidermal growth factor
596 receptor mutation status in early-stage lung adenocarcinoma, using radiomics
597 and clinical features. *Asia Pac J Clin Oncol*. 2022; 18(6): 586-594.

598 33 Dong Y, Jiang Z, Li C et al. Development and validation of novel radiomics-
599 based nomograms for the prediction of EGFR mutations and Ki-67 proliferation
600 index in non-small cell lung cancer. *Quant Imaging Med Surg*. 2022; 12(5):
601 2658-2671.

602 34 Aerts HJ, Grossmann P, Tan Y et al. Defining a adiomic response phenotype: a
603 pilot study using targeted therapy in NSCLC. *Sci Rep*. 2016; 6: 33860.

604 35 Huang Q, Lu L, Dercle L et al. Interobserver variability in tumor contouring
605 affects the use of radiomics to predict mutational status. *J Med Imaging*
606 *(Bellingham)*. 2018; 5(1): 011005.

607 36 Feng Y, Song F, Zhang P et al. Prediction of EGFR mutation status in non-small
608 cell lung cancer based on ensemble learning. *Front Pharmacol*. 2022; 13:
609 897597.

610 37 Huang X, Sun Y, Tan M et al. Three-dimensional convolutional neural network-
611 based prediction of epidermal growth factor receptor expression status in
612 patients with non-small cell lung cancer. *Front Oncol*. 2022; 12: 772770.

613 38 Jia TY, Xiong JF, Li XY et al. Identifying EGFR mutations in lung
614 adenocarcinoma by noninvasive imaging using radiomics features and random
615 forest modeling. *Eur Radiol*. 2019; 29(9): 4742-4750.

616 39 Liu G, Xu Z, Ge Y et al. 3D radiomics predicts EGFR mutation, exon-19
617 deletion and exon-21 L858R mutation in lung adenocarcinoma. *Transl Lung*
618 *Cancer Res*. 2020; 9(4): 1212-1224.

619 40 Mei D, Luo Y, Wang Y et al. CT texture analysis of lung adenocarcinoma: can
620 Radiomic features be surrogate biomarkers for EGFR mutation statuses. *Cancer*
621 *Imaging*. 2018; 18(1): 52.

622 41 Zhang T, Xu Z, Liu G et al. Simultaneous identification of
623 EGFR,KRAS,ERBB2, and TP53 mutations in patients with non-small cell lung
624 cancer by machine learning-derived three-dimensional radiomics. *Cancers*
625 *(Basel)*. 2021; 13(8).

626 42 Song J, Ding C, Huang Q et al. Deep learning predicts epidermal growth factor
627 receptor mutation subtypes in lung adenocarcinoma. *Med Phys*. 2021; 48(12):
628 7891-7899.

629 43 Wang C, Xu X, Shao J et al. Deep learning to predict EGFR mutation and PD-
630 L1 expression status in non-small-cell lung cancer on computed tomography
631 images. *J Oncol*. 2021; 2021: 5499385.

632 44 Huang W, Wang J, Wang H et al. PET/CT based egfr mutation status
633 classification of NSCLC using deep learning features and radiomics features.
634 *Front Pharmacol*. 2022; 13: 898529.

635 45 Kawazoe Y, Shiinoki T, Fujimoto K et al. Investigation of the combination of
636 intratumoral and peritumoral radiomic signatures for predicting epidermal

growth factor receptor mutation in lung adenocarcinoma. *J Appl Clin Med Phys*. 2023; 24(6): e13980.

46 Lu X, Li M, Zhang H et al. A novel radiomic nomogram for predicting epidermal growth factor receptor mutation in peripheral lung adenocarcinoma. *Phys Med Biol*. 2020; 65(5): 055012.

47 Wang J, Lv X, Huang W et al. Establishment and optimization of radiomics algorithms for prediction of KRAS gene mutation by integration of NSCLC gene mutation mutual exclusion information. *Front Pharmacol*. 2022; 13: 862581.

48 Yang X, Fang C, Li C et al. Can CT radiomics detect acquired T790M mutation and predict prognosis in advanced lung adenocarcinoma with progression after first- or second-generation EGFR TKIs? *Front Oncol*. 2022; 12: 904983.

49 Chang C, Zhou S, Yu H et al. A clinically practical radiomics-clinical combined model based on PET/CT data and nomogram predicts EGFR mutation in lung adenocarcinoma. *Eur Radiol*. 2021; 31(8): 6259-6268.

50 Dong Y, Hou L, Yang W et al. Multi-channel multi-task deep learning for predicting EGFR and KRAS mutations of non-small cell lung cancer on CT images. *Quant Imaging Med Surg*. 2021; 11(6): 2354-2375.

51 Gao J, Niu R, Shi Y et al. The predictive value of [(18)F]FDG PET/CT radiomics combined with clinical features for EGFR mutation status in different clinical staging of lung adenocarcinoma. *EJNMMI Res*. 2023; 13(1): 26.

52 Huo JW, Luo TY, Diao L et al. Using combined CT-clinical radiomics models to identify epidermal growth factor receptor mutation subtypes in lung adenocarcinoma. *Front Oncol*. 2022; 12: 846589.

53 Jiang M, Yang P, Li J et al. Computed tomography-based radiomics quantification predicts epidermal growth factor receptor mutation status and efficacy of first-line targeted therapy in lung adenocarcinoma. *Front Oncol*. 2022; 12: 985284.

54 Le NQK, Kha QH, Nguyen VH et al. Machine learning-based radiomics signatures for EGFR and KRAS mutations prediction in non-small-cell lung cancer. *Int J Mol Sci*. 2021; 22(17).

55 Li S, Li Y, Zhao M et al. Combination of (18)F-fluorodeoxyglucose PET/CT radiomics and clinical features for predicting epidermal growth factor receptor mutations in lung adenocarcinoma. *Korean J Radiol*. 2022; 23(9): 921-930.

56 Li S, Luo T, Ding C et al. Detailed identification of epidermal growth factor receptor mutations in lung adenocarcinoma: Combining radiomics with machine learning. *Med Phys*. 2020; 47(8): 3458-3466.

57 Li X, Yin G, Zhang Y et al. Predictive power of a radiomic signature based on (18)F-FDG PET/CT images for EGFR mutational status in NSCLC. *Front Oncol*. 2019; 9: 1062.

58 Li XY, Xiong JF, Jia TY et al. Detection of epithelial growth factor receptor (EGFR) mutations on CT images of patients with lung adenocarcinoma using

radiomics and/or multi-level residual convolutionary neural networks. *J Thorac Dis*. 2018; 10(12): 6624-6635.

59 Liu Y, Zhou J, Wu J et al. Development and validation of machine learning models to predict epidermal growth factor receptor mutation in non-small cell lung cancer: a multi-center retrospective radiomics study. *Cancer Control*. 2022; 29: 10732748221092926.

60 Ninomiya K, Arimura H, Chan WY et al. Robust radiogenomics approach to the identification of EGFR mutations among patients with NSCLC from three different countries using topologically invariant Betti numbers. *PLoS One*. 2021; 16(1): e0244354.

61 Ninomiya K, Arimura H, Tanaka K et al. Three-dimensional topological radiogenomics of epidermal growth factor receptor Del19 and L858R mutation subtypes on computed tomography images of lung cancer patients. *Comput Methods Programs Biomed*. 2023; 236: 107544.

62 Rios Velazquez E, Parmar C, Liu Y et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res*. 2017; 77(14): 3922-3930.

63 Rossi G, Barabino E, Fedeli A et al. Radiomic detection of EGFR mutations in NSCLC. *Cancer Res*. 2021; 81(3): 724-731.

64 Tu W, Sun G, Fan L 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.

65 Wang C, Ma J, Shao J et al. Predicting EGFR and PD-L1 status in NSCLC patients using multitask ai system based on CT images. *Front Immunol*. 2022; 13: 813072.

66 Weng Q, Hui J, Wang H et al. Radiomic feature-based nomogram: a novel technique to predict EGFR-activating mutations for EGFR tyrosin kinase inhibitor therapy. *Front Oncol*. 2021; 11: 590937.

67 Yang X, Liu M, Ren Y et al. Using contrast-enhanced CT and non-contrast-enhanced CT to predict EGFR mutation status in NSCLC patients-a radiomics nomogram analysis. *Eur Radiol*. 2022; 32(4): 2693-2703.

68 Zhang B, Qi S, Pan X et al. Deep CNN model using CT radiomics feature mapping recognizes EGFR gene mutation status of lung adenocarcinoma. *Front Oncol*. 2020; 10: 598721.

69 Zhang G, Cao Y, Zhang J et al. Predicting EGFR mutation status in lung adenocarcinoma: development and validation of a computed tomography-based radiomics signature. *Am J Cancer Res*. 2021; 11(2): 546-560.

70 Zhang M, Bao Y, Rui W et al. Performance of (18)F-FDG PET/CT radiomics for predicting EGFR mutation status in patients with non-small cell lung cancer. *Front Oncol*. 2020; 10: 568857.

71 Zhao HY, Su YX, Zhang LH et al. Prediction model based on 18F-FDG PET/CT radiomic features and clinical factors of EGFR mutations in lung adenocarcinoma. *Neoplasma*. 2022; 69(1): 233-241.

72 Chang C, Sun X, Wang G et al. A machine learning model based on PET/CT radiomics and clinical characteristics predicts ALK rearrangement status in lung adenocarcinoma. *Front Oncol*. 2021; 11: 603882.

73 Ma DN, Gao XY, Dan YB et al. Evaluating solid lung adenocarcinoma anaplastic lymphoma kinase gene rearrangement using noninvasive radiomics biomarkers. *Onco Targets Ther*. 2020; 13: 6927-6935.

74 Song L, Zhu Z, Mao L et al. Clinical, conventional ct and radiomic feature-based machine learning models for predicting ALK rearrangement status in lung adenocarcinoma patients. *Front Oncol*. 2020; 10: 369.

75 Hendriks LE, Kerr KM, Menis J et al. Oncogene-addicted metastatic non-small-cell lung cancer: ESMO clinical practice guideline for diagnosis, treatment and follow-up. *Ann Oncol*. 2023; 34(4): 339-357.

76 Ettinger DS, Wood DE, Aisner DL et al. Non-Small Cell Lung Cancer, Version 3.2022, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2022; 20(5): 497-530.

77 Ferry-Galow KV, Datta V, Makhlouf HR et al. What can be done to improve research biopsy quality in oncology clinical trials? *J Oncol Pract*. 2018; 14(11): Jop1800092.

78 Gutierrez ME, Choi K, Lanman RB et al. Genomic profiling of advanced non-small cell lung cancer in community settings: gaps and opportunities. *Clin Lung Cancer*. 2017; 18(6): 651-659.

79 Fornacon-Wood I, Faivre-Finn C, O'Connor JPB et al. Radiomics as a personalized medicine tool in lung cancer: separating the hope from the hype. *Lung Cancer*. 2020; 146: 197-208.

80 Neri E, Del Re M, Paiar F et al. Radiomics and liquid biopsy in oncology: the holons of systems medicine.

Insights Imaging. 2018; 9(6): 915-924.

81 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*. 2023.

82 Steinert HC. PET and PET-CT of lung cancer. *Methods Mol Biol*. 2011; 727: 33-51.

83 Alwosheel A, van Cranenburgh S, Chorus CG. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *J Choice Model*. 2018; 28: 167-182.

84 Rajput D, Wang W-J, Chen C-C. Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*. 2023; 24(1): 48.

85 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.

86 Zhao W, Yang J, Ni B et al. Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. *Cancer Med*. 2019; 8(7): 3532-3543.

87 Zuo Y, Liu Q, Li N et al. Optimal (18)F-FDG PET/CT radiomics model development for predicting EGFR mutation status and prognosis in lung adenocarcinoma: a multicentric study. *Front Oncol*. 2023; 13: 1173355.

88 Shiri I, Amini M, Nazari M et al. Impact of feature harmonization on radiogenomics analysis: prediction of EGFR and KRAS mutations from non-small cell lung cancer PET/CT images. *Comput Biol Med*. 2022; 142: 105230.

89 Yip SSF, Parmar C, Kim J et al. Impact of experimental design on PET radiomics in predicting somatic mutation status. *Eur J Radiol*. 2017; 97: 8-15.

90 Shao J, Ma J, Zhang S et al. Radiogenomic system for non-invasive identification of multiple actionable mutations and PD-L1 expression in non-small cell lung cancer based on ct images. *Cancers (Basel)*. 2022; 14(19).

91 Xiao Z, Cai H, Wang Y et al. Deep learning for predicting epidermal growth factor receptor mutations of non-small cell lung cancer on PET/CT images. *Quant Imaging Med Surg*. 2023; 13(3): 1286-1299.

92 Agüloğlu N, Aksu A, Akyol M et al. Importance of pretreatment 18F-FDG PET/CT texture analysis in predicting EGFR and ALK mutation in patients with non-small cell lung cancer. *Nuklearmedizin*. 2022; 61(6): 433-439.

93 Agazzi GM, Ravanelli M, Roca E et al. CT texture analysis for prediction of EGFR mutational status and ALK rearrangement in patients with non-small cell lung cancer. *Radiol Med*. 2021; 126(6): 786-794.

94 Chen W, Hua Y, Mao D et al. A computed tomography-derived radiomics approach for predicting uncommon EGFR mutation in patients with NSCLC. *Front Oncol*. 2021; 11: 722106.

95 Chen Q, Li Y, Cheng Q et al. EGFR mutation status and subtypes predicted by CT-based 3D radiomic features in lung adenocarcinoma. *Onco Targets Ther*. 2022; 15: 597-608.

96 Choe J, Lee SM, Kim W et al. CT radiomics-based prediction of anaplastic lymphoma kinase and epidermal growth factor receptor mutations in lung adenocarcinoma. *Eur J Radiol*. 2021; 139: 109710.

97 Digumarthy SR, Padole AM, Gullo RL et al. Can CT radiomic analysis in NSCLC predict histology and EGFR mutation status? *Medicine (Baltimore)*. 2019; 98(1): e13963.

98 Hao P, Deng BY, Huang CT et al. Predicting anaplastic lymphoma kinase rearrangement status in patients with non-small cell lung cancer using a machine learning algorithm that combines clinical features and CT images. *Front Oncol*. 2022; 12: 994285.

99 He R, Yang X, Li T et al. A machine learning-based predictive model of epidermal growth factor mutations in lung adenocarcinomas. *Cancers (Basel)*. 2022; 14(19).

100 Hou D, Li W, Wang S et al. Different clinicopathologic and computed tomography imaging characteristics of primary and acquired EGFR T790M mutations in patients with non-small-cell lung cancer. *Cancer Manag Res*. 2021; 13: 6389-6401.

101 Jiang M, Zhang Y, Xu J et al. Assessing EGFR gene mutation status in non-small cell lung cancer with imaging features from PET/CT. *Nucl Med Commun*. 2019; 40(8): 842-849.

102 Kawazoe Y, Shiinoki T, Fujimoto K et al. Comparison of the radiomics-based predictive models using machine learning and nomogram for epidermal growth factor receptor mutation status and subtypes in lung adenocarcinoma. *Phys Eng Sci Med*. 2023; 46(1): 395-403.

103 Koyasu S, Nishio M, Isoda H et al. Usefulness of gradient tree boosting for predicting histological subtype and EGFR mutation status of non-small cell lung cancer on (18)F FDG-PET/CT. *Ann Nucl Med*. 2020; 34(1): 49-57.

104 Li S, Ding C, Zhang H et al. Radiomics for the prediction of EGFR mutation subtypes in non-small cell lung cancer. *Med Phys*. 2019; 46(10): 4545-4552.

105 Li H, Gao C, Sun Y et al. Radiomics analysis to enhance precise identification of epidermal growth factor receptor mutation based on positron emission tomography images of lung cancer patients. *J Biomed Nanotechnol*. 2021; 17(4): 691-702.

106 Liu Y, Kim J, Balagurunathan Y et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. *Clin Lung Cancer*. 2016; 17(5): 441-448.e446.

107 Liu Z, Zhang T, Lin L et al. Applications of radiomics-based analysis pipeline for predicting epidermal growth factor receptor mutation status. *Biomed Eng Online*. 2023; 22(1): 17.

108 Lu L, Sun SH, Yang H et al. Radiomics prediction of EGFR status in lung cancer-our experience in using multiple feature extractors and The Cancer Imaging Archive Data. *Tomography*. 2020; 6(2): 223-230.

109 Ruan D, Fang J, Teng X. Efficient 18F-Fluorodeoxyglucose positron emission tomography/computed tomography-based machine learning model for predicting epidermal growth factor receptor mutations in non-small cell lung cancer. *Q J Nucl Med Mol Imaging*. 2022.

110 Trivizakis E, Souglakos J, Karantanas A et al. Deep radiotranscriptomics of non-small cell lung carcinoma for assessing molecular and histology subtypes with a data-driven analysis. *Diagnostics (Basel)*. 2021; 11(12).

111 Yamazaki M, Yagi T, Tominaga M et al. Role of intratumoral and peritumoral CT radiomics for the prediction of EGFR gene mutation in primary lung cancer. *Br J Radiol*. 2022; 95(1140): 20220374.

112 Yang B, Ji HS, Zhou CS et al. (18)F-fluorodeoxyglucose positron emission tomography/computed tomography-based radiomic features for prediction of epidermal growth factor receptor mutation status and prognosis in patients with lung adenocarcinoma. *Transl Lung Cancer Res*. 2020; 9(3): 563-574.

113 Zhang J, Zhao X, Zhao Y et al. Value of pre-therapy (18)F-FDG PET/CT radiomics in predicting EGFR mutation status in patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging*. 2020; 47(5): 1137-1146.

114 Zhang T, Liu Z, Lin L et al. Detection of the gene mutation of epidermal growth factor receptor in lung adenocarcinoma by radiomic features from a small amount of PET data. *Nucl Med Commun*. 2023.

115 Zhao W, Wu Y, Xu Y et al. The potential of radiomics nomogram in non-invasively prediction of epidermal growth factor receptor mutation status and subtypes in lung adenocarcinoma. *Front Oncol*. 2019; 9: 1485.

Table 1. Methodological characteristics of the studies (*N* = 89) included in the systematic review. For those studies with the same name for the first author and published the same year, a hashtag was added to unequivocally indicate those that were included in the different meta-analyses and consequently, that are represented in the forest plots.

Author-Year	Imaging modality	Contrast-CT*	Tumor segmentation	Model	Classifier (ML)	Datasets			Partition strategy
						Training	Validation	Test	
Agüloğlu et al. 2022[92]	PET/CT	Non-contrast CT	Semi-automatic	ML	RF NB KNN DT SVM LR	133	56	–	Training-Validation split
Aerts et al. 2016[34]	CT	Non-contrast CT	Semi-automatic	Classical statistical model	–	–	–	–	–
Agazzi et al. 2021[93]	CT	Contrast-enhanced	Manual	ML	GBM	104	67	–	Training-Validation split

Aide et al. 2022[30]	PET	–	Manual	ML	LASSO	87	22	–	Training-Validation split
Chang et al. 2021#[49]	PET/CT	Non-contrast CT	Manual	ML	LASSO	409	174	–	Training-Validation split
Chang et al. 2021##[72]	PET/CT	Non-contrast CT	Manual	ML	LR†	367	159	–	Training-Validation split
Chen et al. 2021[94]	CT	Non-contrast CT	Manual	ML	SVM	179	44	–	Training-Validation split
Chen et al. 2022[95]	CT	Non-contrast CT	Semi-automatic	ML	LASSO	176	57	–	Training-Validation split
Choe et al. 2021[96]	CT	Contrast-enhanced	Semi-automatic	ML	LR	349	154	–	Training-Validation split
Dang et al. 2021[19]	CT	Non-contrast CT	Semi-automatic	ML	LASSO	88	30	–	Training-Validation split
Digumarthy et	CT	Contrast-enhanced	Not specified	Classical	–	–	–	–	–

al. 2019[97]				statistical model					
Dong et al. 2022[33]	CT	Non-contrast CT	Not specified	ML	LR	87	45	–	Training-Validation split
Dong et al. 2021[50]	CT	Non-contrast CT	Manual	DL ML	RF	363	162	–	Training-Validation split
Feng et al. 2022[36]	CT	Non-contrast CT	Manual	ML	RF XGBoost LR SVM	151	–	–	Training-Validation split
Gao et al. 2023[51]	PET/CT	Non-contrast CT	Semi-automatic	ML	LR RF SVM	404	111	–	Training-Validation split
Hao et al. 2022[98]	CT	Non-contrast CT	Manual	ML	SVM XGBoost AdaBoost LBP DT	154	39	–	Training-Validation split

					LR				
He et al. 2022[99]	CT	Non-contrast CT	Semi-automatic	ML	RF KNN LGBM SVM	–	–	–	Training-Validation split
Hong et al. 2020[24]	CT	Contrast-enhanced	Manual	ML	NBC KNN RF SVM DT LR	140	61	–	Training-Validation split
Huang et al. 2018[35]	CT	Non-contrast CT	Semi-automatic	Classical statistical model	–	–	–	–	–
Huang et al. 2022[44]	PET/CT	Non-contrast CT	Manual	DL ML	LR	138	57	–	Training-Validation split
Huang et al. 2022[37]	CT	Non-contrast CT	Manual	DL	LR	770	304	–	Training-Validation split

				ML					
Huo et al. 2022[52]	CT	Contrast-enhanced	Manual	ML	GBT	487	121	–	Training-Validation split
Hou et al. 2021[100]	CT	Contrast-enhanced	Semi-automatic	Classical statistical model	–	144	62	–	Training-Validation split
Jia et al. 2019[38]	CT	Non-contrast CT	Semi-automatic	ML	RF	345	158	–	Training-Validation split
Jiang et al. 2019[101]	PET/CT	Non-contrast CT	Semi-automatic	ML	SVM	–	–	–	10-fold cross-validation
Jiang et al. 2022[53]	CT	Non-contrast CT	Manual	ML	SVM	514	178	–	Training-Validation split
Kawazoe et al. 2023[45]	CT	Non-contrast CT	Semi-automatic	ML	SVM LR LGBM	120	44	–	Training-Validation split
Kawazoe et al.	CT	Non-contrast CT	Semi-automatic	ML	SVM	120	52	–	Training-Validation

2023[102]					LR				split
Koyasu et al. 2020[103]	PET/CT	Non-contrast CT	Manual	ML	RF XGBoost	–	–	–	10-fold cross-validation
Le et al. 2021[54]	CT	Non-contrast CT	Manual	ML	XGBoost	143	18	–	Training-Validation split
Li et al. 2018#[58]	CT	Non-contrast CT	Manual	DL ML	RF	810	200	–	Training-Validation split
Li et al. 2018[31]	CT	Contrast-enhanced	Semi-automatic	ML	SVM	–	–	–	3-fold cross-validation
Li et al 2019#[57]	PET/CT	Non-contrast CT	Manual	ML	Boosting ML scheme	115	–	–	10-fold cross-validation
Li et al. 2019[104]	CT	Non-contrast CT	Manual	ML	LR	236	76	–	Training-Validation split
Li et al. 2020[56]	CT	Non-contrast CT	Manual	ML	LR SVM	326	112	–	Training-Validation split

					RF NB Neural network				
Li et al. 2021[105]	PET	–	Semi-automatic	ML	SVM	50	25	–	Training-Validation split
Li et al. 2022[55]	PET/CT	Non-contrast CT	Manual	ML	LR	125	54	–	Training-Validation split
Liu et al. 2016[106]	CT	Non-contrast CT	Semi-automatic	Classical statistical model	–	–	–	–	–
Liu et al. 2020#[39]	CT	Contrast-enhanced	Semi-automatic	ML	LR	210	53	–	Training-Validation split
Liu et al. 2020[22]	PET/CT	Non-contrast CT	Manual	ML	XGBoost	111	37	–	Training-Validation split
Liu et al. 2022[59]	CT	Non-contrast CT	Manual	ML	LR DT RF	296	50	–	Training-Validation split

					SVM				
Liu et al. 2023[107]	PET/CT	Non-contrast CT	Manual	ML	LR DT RF SVM	–	–	–	10-fold cross-validation
Lu et al. 2020#[46]	CT	Non-contrast CT	Manual	ML	LR	83	–	21	Training-Validation split
Lu et al. 2020[108]	CT	Non-contrast CT	Semi-automatic	ML	KNN Bagging SVM RF	105	228	–	Training-Validation split
Lu et al. 2022[25]	CT	Non-contrast CT	Manual	ML	DT AdaBoost NB RF LR SVM XGBoost	140	61	–	Training-Validation split

					KNN				
Ma et al. 2020[73]	CT	Contrast-enhanced	Manual	ML	SVM	98	42	–	Training-Validation split
Mei et al. 2018[40]	CT	Non-contrast CT	Manual	Classical statistical model	–	–	–	–	–
Mu et al. 2020[85]	PET/CT	Non-contrast CT	Manual	DL	–	429	187	65	Training-Validation split
Nair et al. 2021[17]	PET/CT	Contrast-enhanced	Manual	ML	LR	–	–	–	LOOCV
Ninomiya et al. 2021[60]	CT	Contrast-enhanced	Manual	ML	SVM	99‡	99‡	95	Training-Validation split
Ninomiya et al. 2023[61]	CT	Contrast-enhanced	Not specified	ML	SVM	92	62	–	Training-Validation split
Omura et al. 2023[20]	CT	Contrast-enhanced	Automatic	ML	RF	–	–	–	Training-Validation split

Ríos Velázquez et al. 2017[62]	CT	Contrast + Non-contrast CT	Semi-automatic	ML	RF	353	352	–	Training-Validation split
Rossi et al. 2021[63]	CT	Non-contrast CT	Manual	ML	SVM	–	109	61	Training-Validation split
Ruan et al. 2022[109]	PET/CT	Non-contrast CT	Manual	ML	SVM	70	30	–	Training-Validation split
Shao et al. 2022[90]	CT	Non-contrast CT	Semi-automatic	DL	–	–	–	–	Training-Validation split
Shiri et al. 2020[18]	CT low dose CTD PET/CT	Contrast-enhanced	Manual Automatic[§]	ML	SVM KNN DT QDA MLP SGD LR NB GNB	82	68	–	10-fold cross-validation

					RF AdaBoost Bagging				
Shiri et al. 2022[88]	PET/CT	Non-contrast CT	Manual Automatic[§]	ML	RF	–	–	–	Training-Validation split
Song et al. 2021[42]	CT	Not specified	Manual Automatic	DL ML	SVM	528	137	–	Training-Validation split
Song et al. 2020[74]	CT	Non-contrast CT	Automatic	ML	DT	268	67	–	Training-Validation split
Trivizakis et al. 2021[110]	CT	Not specified	Not specified	DL ML	KNN DT RBF-GPC RBF-SVM Linear SVM Polynomial SVM Sigmoid SVM	–	–	–	5-fold cross-validation

Tu et al. 2019[64]	CT	Non-contrast CT	Not specified	ML	LR	243	161	–	Training-Validation split
Wang et al. 2019[21]	CT	Contrast-enhanced	Manual	ML	SVM	41	–	–	Training-Validation split
Wang et al. 2021[43]	CT	Non-contrast CT	Manual	DL	–	882	125	255	Training-Validation split
Wang et al. 2022#[65]	CT	Non-contrast CT	Manual	DL ML	LASSO	–	–	–	Training-Validation split¶
Wang et al. 2022##[47]	PET/CT	Non-contrast CT	Semi-automatic	ML	LR	180	78	–	Training-Validation split
Weng et al. 2021[66]	CT	Non-contrast CT	Semi-automatic	ML	LR	210	91	–	Training-Validation split
Wu et al. 2020[26]	CT	Contrast-enhanced	Manual	ML	LR	–	–	–	10-fold cross-validation
Xiao et al.	PET/CT	Non-contrast CT	Manual	DL	RF	121	29	–	Training-Validation

2023[91]				ML					split
Yamazaki et al. 2022[111]	CT	Non-contrast CT	Semi-automatic	ML	RF	–	–	–	–
Yang et al. 2020#[27]	CT	Contrast-enhanced	Semi-automatic	ML	LASSO	130	40	–	Training-Validation split
Yang et al. 2020[112]	PET/CT	Non-contrast CT	Semi-automatic	ML	RF	139	35	–	Training-Validation split
Yang et al. 2022#[67]	CT	Contrast + Non-contrast CT	Manual	ML	LR RF SVM GBT NB	327	66	19	Training-Validation split
Yang et al. 2022[23]	PET/CT	Non-contrast CT	Semi-automatic	ML	SVM DT RF	218	95	–	Training-Validation split
Yang et al.	CT	Contrast-enhanced	Manual	ML	LR	176	74	–	Training-Validation

2022[48]									split
Yip et al. 2017[89]	PET	–	Manual	Classical statistical model	–	–	–	–	–
Zhang et al. 2018[28]	CT	Non-contrast CT	Manual	ML	LR	140	40	–	Training-Validation split
Zhang et al. 2020#[70]	PET/CT	Non-contrast CT	Manual	ML	RF SVM LR	–	–	–	10-fold cross-validation
Zhang et al. 2020[113]	PET/CT	Non-contrast CT	Semi-automatic	ML	LR	175	73	–	Training-Validation split
Zhang et al. 2020##[68]	CT	Non-contrast CT	Semi-automatic	DL ML	RF SVM	638	71	205	Training-Validation split
Zhang et al.	CT	Contrast-enhanced	Semi-automatic	ML	LASSO	–	–	–	Training-Validation

2021[41]									split
Zhang et al. 2021[69]	CT	Non-contrast CT	Manual	ML	DT LR SVM Multivariate analysis for C-R-R model	294	126	–	Training-Validation split
Zhang et al. 2023[114]	PET	–	Manual	ML	SVM RF LR AdaBoost	–	–	–	10-fold cross-validation
Zhao et al. 2019[86]	CT	Non-contrast CT	Manual	DL ML	LR	348	116	116	Training-Validation split
Zhao et al. 2019[115]	CT	Non-contrast CT	Manual	ML	LR	322	315	–	Training-Validation split
Zhao et al. 2022[71]	PET/CT	Non-contrast CT	Semi-automatic	ML	LR	65	23	–	Training-Validation split

Zhu et al. 2022[32]	CT	Non-contrast CT	Semi-automatic	ML	LASSO RF SVM	875	217	–	Training-Validation split
Zhu et al. 2021[29]	CT	Contrast-enhanced	Manual	ML	SVM KNN RF LR	159	40	–	Training-Validation split
Zuo et al. 2023[87]	PET/CT	Non-contrast CT	Manual	ML	LGBM XGBoost RF LR	410	170	180	Training-Validation split

*In studies in which PET/CT was performed, only details about contrast were provided for PET acquisition. Consequently, it was assumed that CT scans were non-contrast enhanced.

†Not specified but inferred from the methodology and results.

‡Number of cases for training and validation sets not specified; only a total number for both cohorts provided.

§Manual segmentation for PET images; automatic segmentation for CT images.

¶80% Training-Validation split.

CT, computed tomography; CTD, contrast-enhanced diagnostic quality; DL, deep learning; DT, decision tree; ERBB2, v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2;

GBM, gradient boosted machine; GBT, gradient boosting tree; GNB, Gaussian Naives Bayes; GPC, Gaussian processes classification; LASSO, least absolute shrinkage and selection operator; LBP, local binary pattern; LGBM, Light gradient boosted machine; LOOCV, leave-one-out cross-validation; LR, logistic regression; ML, machine learning; MLP, multilayer perceptron; NB, Naive Bayes; KNN, K-nearest neighbors; PET, positron emission tomography; QDA, quadratic discriminant analysis; RBF, radial basis function; RF, random forest; SGD, stocastic gradient descendent; SVM, support vector machine; TP53, tumor suppressor protein 53.

Table 2. Clinical characteristics of the studies ($N = 89$) included in the systematic review. For those studies with the same name for the first author and published the same year, a hashtag was added to unequivocally indicate those that were included in the different meta-analyses and consequently, that are represented in the forest plots.

Author-Year	Target oncogene mutation	Design	Total of patients	Sex		Age Mean/median	Histology	Smoking status		TNM stage	Treatment
				Female	Male			Current /former smoker	Non-smoker		
Agüloğlu et al. 2022[92]	EGFR ALK	Unicentric	189	59	130	62/–	NSCLC	130	59	Stages I-IV	Naïve
Aerts et al. 2016[34]	EGFR	Unicentric	47	–	–	–/–	NSCLC	–	–	–	Naïve + post-treatment images
Agazzi et al. 2021[93]	EGFR ALK	Unicentric	84	39	45	–/63	ADC	57	27	–	Naïve
Aide et al. 2022[30]	EGFR	Unicentric	109	34	75	–/66	ADC	96	13	Stages II-IV	Naïve

Chang et al. 2021#[49]	EGFR	Unicentric	583	305	278	–/62	ADC	229	354	Stages I-III	Naïve
Chang et al. 2021##[72]	ALK	Unicentric	526	272	254	–/58.25	ADC	202	324	Stages I-IV	Naïve
Chen et al. 2021[94]	EGFR	Unicentric	223	109	114	64.63/–	NSCLC	55	168	Stages I-IV	Naïve
Chen et al. 2022[95]	EGFR	Unicentric	233	105	128	57.5/–	ADC	65	168	Stages I-IV	Naïve
Choe et al. 2021[96]	ALK	Unicentric	503	273	230	62.5/–	ADC	200	303	Stages I-IV	Not specified
Dang et al. 2021[19]	EGFR	Not specified	118	55	63	63.82/–	ADC, SCC	–	–	Stages I-III	No treatment*
Digumarthy et al. 2019[97]	EGFR	Unicentric	93	50	43	60/–	ADC, SCC	61	32	–	Naïve
Dong et al. 2022[33]	EGFR	Multicentric	132	64	68	58.8/–	NSCLC	42	90	Stages I-III	Naïve
Dong et al.	EGFR	Multicentric	525	250	275	– /65.5	NSCLC	373	152	–	Not specified

2021[50]	KRAS										
Feng et al. 2022[36]	EGFR	Multicentric	168	–	–	–/–	NSCLC	–	–	–	Not specified†
Gao et al. 2023[51]	EGFR	Unicentric	515	264	251	64/–	ADC	175	–	Stages I-IV	Naïve
Hao et al. 2022[98]	ALK	Unicentric	193	102	91	54.26/–	NSCLC	49	144	Stages II and IV	Naïve
He et al. 2022[99]	EGFR	Multicentric	758	317	441	55.6/–	NSCLC	358	400	Stages I-IV	Naïve
Hong et al. 2020[24]	EGFR	Unicentric	201	94	107	58.12/–	ADC	64	137	Stages I-IV	Naïve
Huang et al. 2018[35]	EGFR	Unicentric	46	–	–	–/–	NSCLC	–	–	–	Naïve + post-treatment images
Huang et al. 2022[44]	EGFR	Unicentric	195	72	123	61.14 –	NSCLC	127	68	–	No treatment*
Huang et al. 2022[37]	EGFR	Unicentric	1074	–	–	–/–	NSCLC	–	–	–	Not specified

Huo et al. 2022[52]	EGFR	Unicentric	608	272	336	61.7/–	ADC	0	335	Stages II and IV	Naïve
Hou et al. 2021[100]	EGFR	Unicentric	206	120	86	–/59	ADC, SCC, ASC‡	57	–	Stages I-IV	Naïve
Jia et al. 2019[38]	EGFR	Unicentric	503	249	254	–/60.5	ADC	80	423	Stages I-IV	Not specified†
Jiang et al. 2019[101]	EGFR	Unicentric	80	32	48	64/62.5	NSCLC	21	59	–	Naïve
Jiang et al. 2022[53]	EGFR	Unicentric	692	–	–	59/–	ADC	–	–	–	Naïve
Kawazoe et al. 2023[45]	EGFR	Unicentric	164	75	89	70.24/–	ADC	102	62	Stages I-IV	No treatment§
Kawazoe et al. 2023[102]	EGFR	Unicentric	172	77	95	70.76/–	ADC	107	65	Stages I-IV	Naïve
Koyasu et al. 2020[103]	EGFR	Unicentric	138	54	84	67.8/–	ADC, SCC	–	–	–	Not specified

Le et al. 2021[54]	EGFR KRAS	Multicentric	161	50	111	68.05/–	ADC, NSCLC NOS, SCC	61	100	–	Naïve
Li et al. 2018#[58]	EGFR	Unicentric	1010	457	553	–/63	ADC	262	748	Stages I-IV	Naïve
Li et al. 2018[31]	EGFR	Unicentric	51	19	32	58.1/–	ADC	24	27	Stages I-III	Not specified[†]
Li et al 2019#[57]	EGFR	Unicentric	115	62	53	–/63	NSCLC	36	79	Stages II and IV	Naïve
Li et al. 2019[104]	EGFR	Unicentric	312	164	148	Freq./Freq.[¶]	ADC, SCC	109	203	Stages II and IV	Naïve
Li et al. 2020[56]	EGFR	Multicentric	438	–	–	61.31/–	ADC	–	–	–	Naïve
Li et al. 2021[105]	EGFR	Unicentric	75	45	30	62/–	Lung cancer[**]	34	41	–	Not specified
Li et al. 2022[55]	EGFR	Unicentric	179	103	76	61.51/59.5	ADC	65	114	–	Naïve
Liu et al. 2016[106]	EGFR	Unicentric	298	172	126	–/60	ADC, Others	136	162	Stages II and IV	Naïve

Liu et al. 2020#[39]	EGFR	Unicentric	263	121	142	62.5/–	ADC	31	232	–	Not specified†
Liu et al. 2020[22]	EGFR	Unicentric	148	63	85	–/61.2	ADC	–	–	Stages II-IV	Naïve
Liu et al. 2022[59]	EGFR	Multicentric	346	141	205	66.69/–	ADC, SCC, LCC, PSC	225	121	–	Naïve
Liu et al. 2023[107]	EGFR	Unicentric	115	62	53	–/62.75	ADC	36	79	Stages I-IV	Naïve
Lu et al. 2020#[46]	EGFR	Unicentric	104	64	40	58.27/–	ADC	30	74	Stages I-IV	No treatment*
Lu et al. 2020[108]	EGFR	Multicentric	228††	85††	120††	67.94/–	ADC, SCC, NOS	–	–	Stages 0-IV	Not specified
Lu et al. 2022[25]	EGFR	Unicentric	201	99	102	64.81/–	ADC	84	117	Stages III-IV	Naïve
Ma et al. 2020[73]	ALK	Unicentric	140	87	53	54.19/–	ADC	45	95	Stages II and IV	Naïve
Mei et al. 2018[40]	EGFR	Unicentric	296	144	152	58.56/–	ADC	86	210	–	Not specified†

Mu et al. 2020[85]	EGFR	Multicentric	681	303	378	61,83/–	ADC, SCC	315	366	Stages I-IV	Naïve
Nair et al. 2021[17]	EGFR	Unicentric	50	18	32	–/–	NSCLC	35	15	–	Naïve
Ninomiya et al. 2021[60]	EGFR	Multicentric	194	74	120	–/67	NSCLC	128	66	Stages I-IV	Not specified
Ninomiya et al. 2023[61]	EGFR	Multicentric	154	86	68	–/67	Lung cancer	73	81	Stages I-IV	Not specified
Omura et al. 2023[20]	EGFR	Unicentric	99	65	34	66/–	ADC	41	–	Stages I-II	Naïve
Ríos Velázquez et al. 2017[62]	EGFR KRAS	Multicentric	763	459	304	65/–	ADC	548	215	Stages I-IV	Not specified
Rossi et al.	EGFR	Multicentric	170	–	–	–/–	ADC	110	30	–	Naïve

2021[63]											
Ruan et al. 2022[109]	EGFR	Unicentric	100	42	58	– / 64.5	NSCLC	33	67	Stages I-IV	Naïve
Shao et al. 2022[90]	EGFR	Unicentric	1096	–	–	58.26/–	NSCLC	–	–	–	Naïve
Shiri et al. 2020[18]	EGFR KRAS	Unicentric	150	–	–	69.1 –	ADC, SCC, NOS‡	–	–	–	Not specified
Shiri et al. 2022[88]	EGFR KRAS	Multicentric	136	–	–	–/–	ADC, SCC, NOS	–	–	–	Not specified
Song et al. 2021[42]	EGFR	Multicentric	665	336	329	Freq./Freq.¶	ADC	334	331	Stages II and IV	Not specified‡‡
Song et al. 2020[74]	ALK	Unicentric	335	196	139	57 / –	ADC	103	232	Stages I-IV	Naïve

Trivizakis et al. 2021[110]	EGFR	Unicentric	112	–	–	–/–	ADC, SCC	–	–	–	Not specified
Tu et al. 2019[64]	EGFR	Unicentric	404	211	193	59.95/–	NSCLC	114	290	Stages II and IV	Naïve
Wang et al. 2019[21]	EGFR	Unicentric	51	35	16	58.45/–	ADC	9	42	Stages 0-II	Not specified
Wang et al. 2021[43]	EGFR PD-L1	Unicentric	1262	642	620	57.7/–	ADC, SCC, Others‡	452	749	Stages I-IV	Not specified§§
Wang et al. 2022#[65]	EGFR PD-L1	Unicentric	3629	1674	1955	59.29/–	ADC, SCC, Others	1413	1981	Stages I-IV	Naïve
Wang et al. 2022##[47]	KRAS	Unicentric	258	78	180	62.35/–	NSCLC	166	92	–	No treatment*
Weng et al. 2021[66]	EGFR	Unicentric	301	145	156	64.95/–	NSCLC	110	191	–	Naïve

Wu et al. 2020[26]	EGFR	Unicentric	67	29	38	56.35/–	ADC, SCC	34	33	Stages III-IV	Naïve
Xiao et al. 2023[91]	EGFR	Unicentric	150	59	91	–/58	NSCLC	64	86	–	Not specified
Yamazaki et al. 2022[111]	EGFR	Unicentric	478	190	288	Freq./Freq.[¶]	ADC, SCC, Others[‡]	–	–	Stages II and IV	Naïve
Yang et al. 2020#[27]	EGFR	Unicentric	253	155	98	–/62	ADC	105	148	Stages III-IV	Naïve
Yang et al. 2020[112]	EGFR	Unicentric	174	81	93	61.72/–	ADC	59	115	Stages II and IV	Naïve
Yang et al. 2022#[67]	EGFR	Unicentric	412	223	189	62/–	ADC, SCC	105	307	–	Naïve
Yang et al. 2022[23]	EGFR	Unicentric	313	164	149	59.21/–	ADC	105	208	Stages II-IV	Naïve
Yang et al. 2022[48]	EGFR	Unicentric	250	–	–	56.35 / –	ADC	–	–	–	Treated with TKIs[¶¶]

Yip et al. 2017[89]	KRAS	Unicentric	348	214	134	–/65	ADC, NSCLC NOS, SC. Not available for 1 patient‡	286	62	Stages I-IV	Naïve
Zhang et al. 2018[28]	EGFR	Unicentric	180	46	134	59.7/–	ADC, SCC, Others	119	61	Stages III-IV	Naïve
Zhang et al. 2020#[70]	EGFR	Unicentric	173	58	115	60.8/–	ADC SCC, LCC, NSCLC-NOS	–	–	Stages I-IV	Naïve
Zhang et al. 2020[113]	EGFR	Unicentric	248	113	135	62.23/–	ADC	117	131	Stages I-IV	Naïve
Zhang et al. 2020##[68]	EGFR	Unicentric	914	493	421	59.79/–	ADC	–	–	–	Naïve

Zhang et al. 2021[41]	EGFR KRAS ERBB2 TP53	Unicentric	134	56	78	63.6/–	ADC, SCC, ASC	28	106	–	Not specified
Zhang et al. 2021[69]	EGFR	Unicentric	420	201	219	57.43/56.5	ADC	147	273	–	Naïve
Zhang et al. 2023[114]	EGFR	Unicentric	115	–	–	–/–	NSCLC	–	–	–	Naïve
Zhao et al. 2019[86]	EGFR	Unicentric	579	334	245	60.1/–	ADC	–	–	Stages 0-IV	Not specified
Zhao et al. 2019[115]	EGFR	Unicentric	637	368	269	59.9/–	ADC	49	588	–	Naïve
Zhao et al. 2022[71]	EGFR	Unicentric	88	39	49	64.23/–	ADC	31	57	Stages II and IV	Naïve
Zhu et al. 2022[32]	EGFR	Unicentric	1092	648	442	59.59/–	ADC	–	–	Stages I-III	Naïve

Zhu et al. 2021[29]	EGFR TP53	Unicentric	199	86	113	Freq./Freq.¶	ADC	94	105	Stages III-IV	Naïve
Zuo et al. 2023[87]	EGFR	Multicentric	767	372	395	–/62.04	ADC	–	–	Stages I-IV Others (34 patients)	Not specified

*Patients were excluded if treated with RT or chemotherapy, but targeted therapy is not specified.

†CT scans acquired prior surgery; no information on prior treatments.

‡Mainly adenocarcinoma cases.

§Patients did receive target treatment, but no information on the administration of other treatments (immunotherapy and/or chemotherapy) is specified.

¶These studies provide age data as frequencies establishing an age threshold.

**Inferred that NSCLC patients were included as it is specified that 17 patients had 19Del and 20 cases had L858R mutation; EGFR mutations are very rare in SCLC.

††In this study, there are 23 patients with no information about sex.

‡‡Image acquired 3 months before PCR; no information about treatments.

§§CT images acquired within 1 month before pathological diagnosis.

¶¶ Imaging-proven progression on first- or second-generation TKIs; patients underwent chest contrast-enhanced CT at the time of confirmed progression, and the interval between CT and confirmed progression was within 3 days.

ADC, adenocarcinoma; ALK, anaplastic lymphoma kinase; ASC, adenosquamous carcinoma; CT, computed tomography; EGFR, epidermal growth factor receptor; ERBB2, v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2; Freq. , frequency; KRAS, Kirsten rat sarcoma viral oncogene homologue; LCC, large cell lung carcinoma; NOS, not otherwise specified; NSCLC, non-small cell lung cancer; PCR, polymerase chain reaction; PD-L1, programmed death ligand 1; PSC, pulmonary sarcomatoid carcinoma; RT, radiotherapy; SCC, Squamous cell carcinoma; SCLC, small-cell lung cancer; TKI , tyrosine kinase inhibitor; TP53, tumor suppressor protein 53.

FIGURES

Figure 1. PRISMA flowchart. AI, artificial intelligence; CLAIM , Checklist for Artificial Intelligence in Medical Imaging; CT , computed tomography; MRI , magnetic resonance imaging.

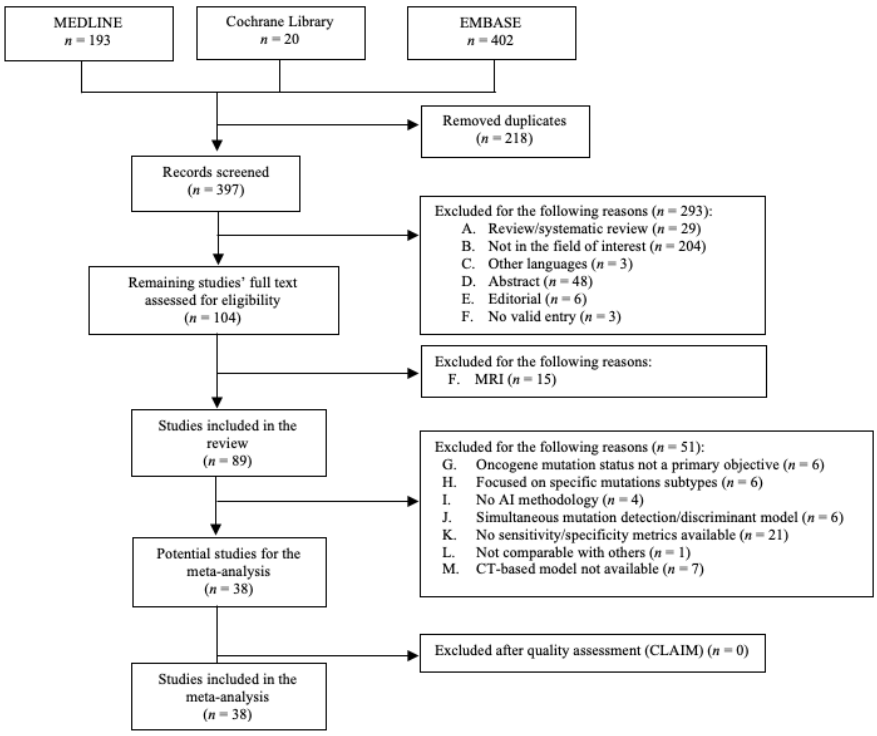


Figure 2. Forest plots of the included studies developing radiomics models using machine learning and/or deep learning methods for the prediction of EGFR mutation status. Numbers are estimated with 95% CIs in brackets and indicated by horizontal lines. For those studies with the same name for the first author and published the same year, a hashtag was added to unequivocally tag them as done in Tables 1 and 2 and in the reference list. EGFR, epidermal growth factor receptor; CI, confidence interval; DOR, diagnostic odds ration; FPR, false positive rate.

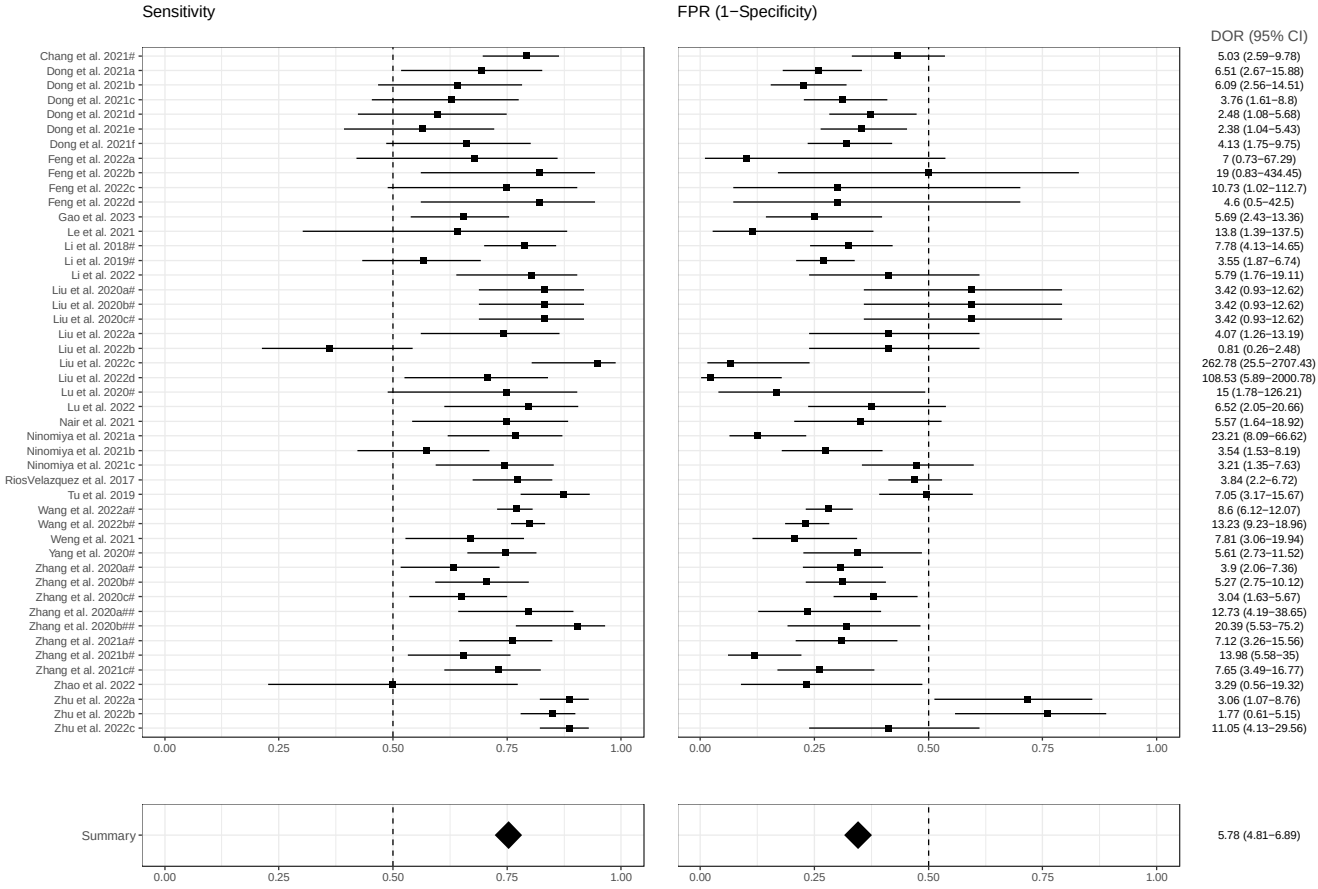
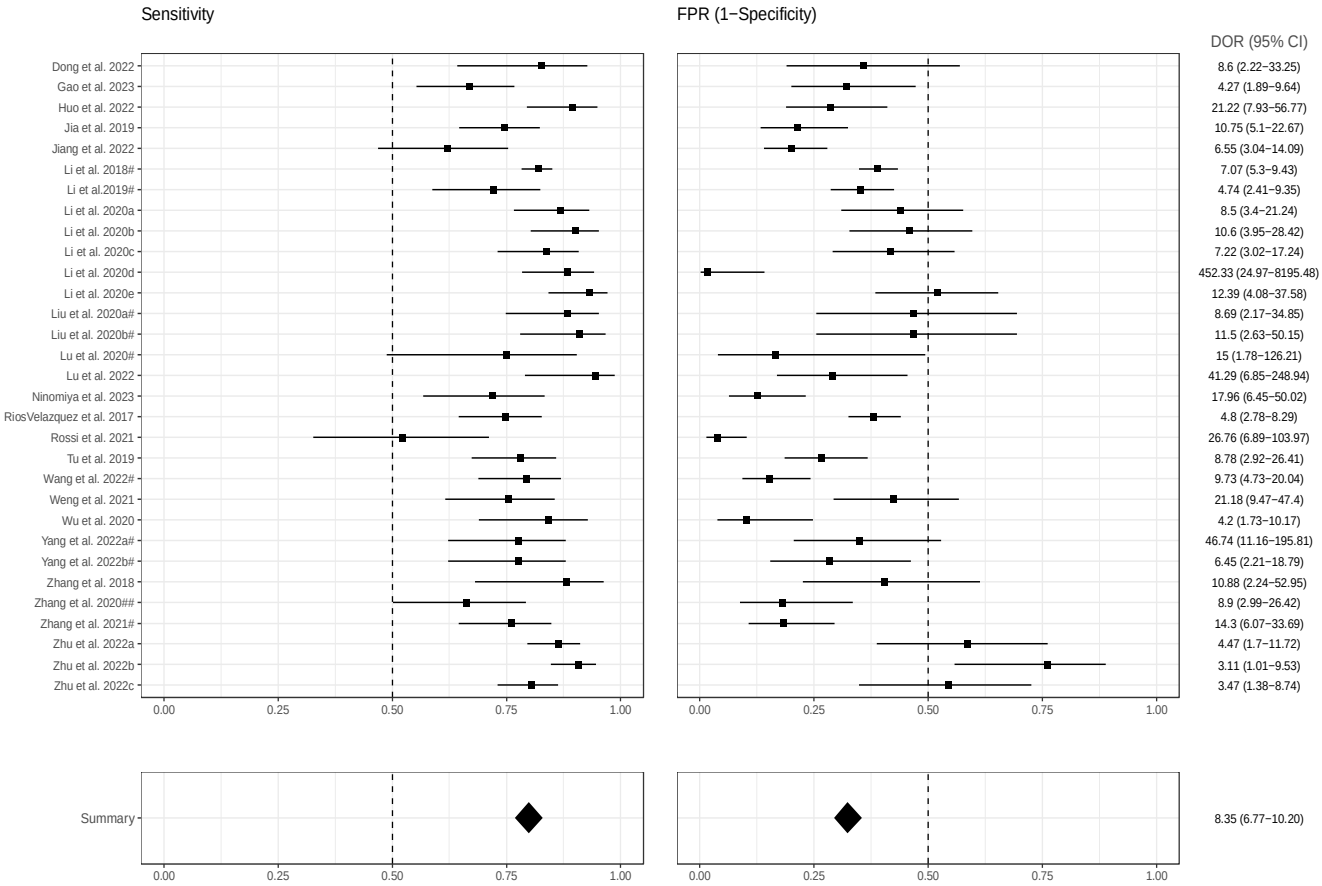
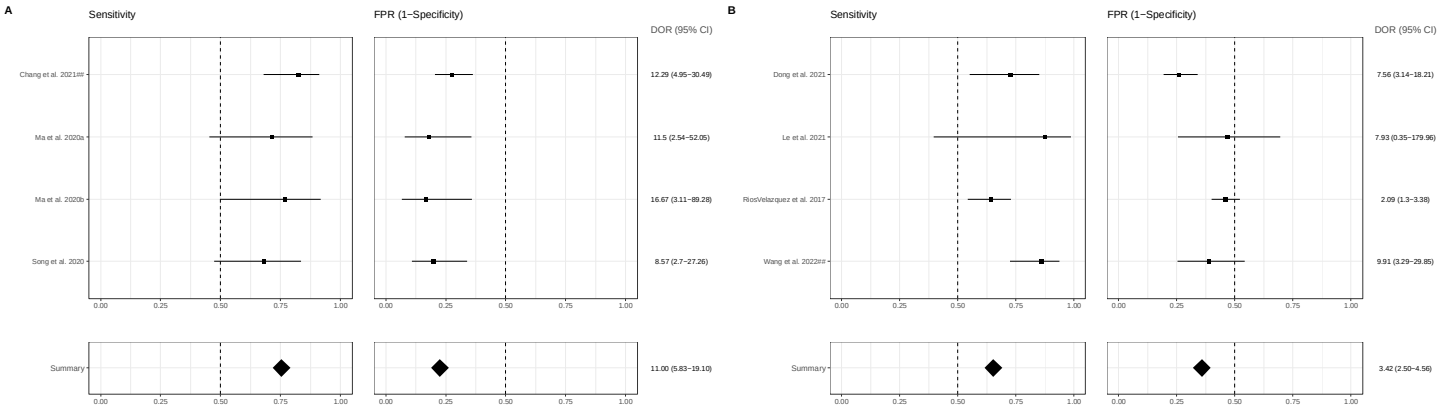


Figure 3. Forest plots of the included studies developing combined models (radiomics + clinical data) using machine learning and/or deep learning methods for the prediction of EGFR mutation status. Numbers are estimates with 95% CIs in brackets and indicated by horizontal lines. For those studies with the same name for the first author and published the same year, a hashtag was added to unequivocally tag them as done in Tables 1 and 2 and in the reference list. EGFR, epidermal growth factor receptor; CI, confidence interval; DOR, diagnostic odds ration;



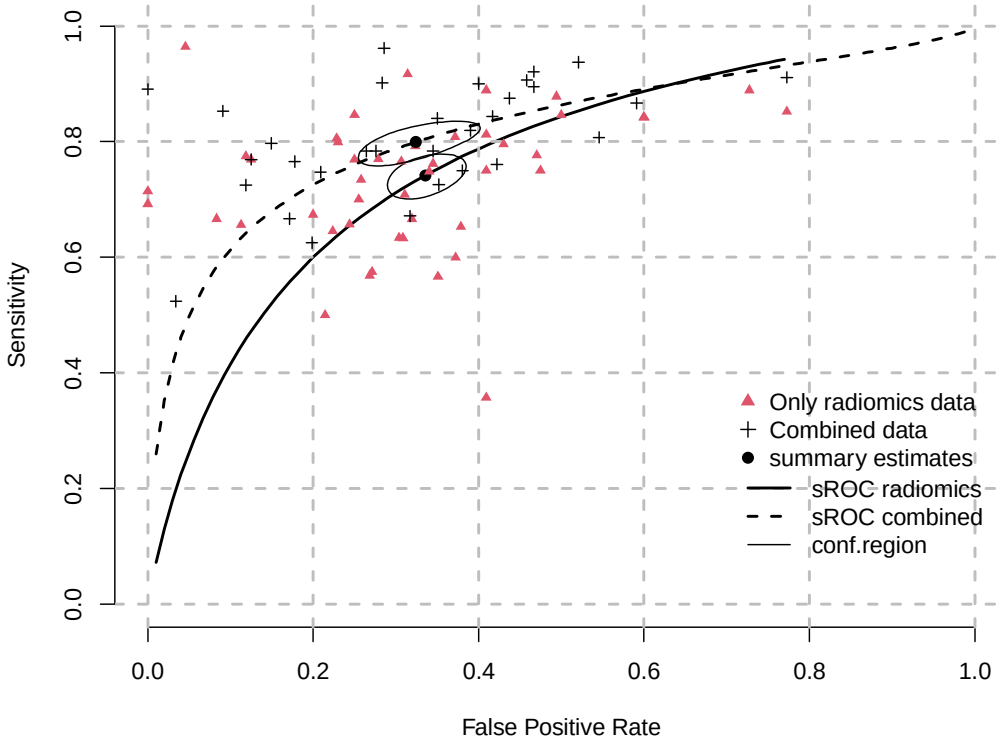
FPR, false positive rate.

Figure 4. Forest plots of the included studies developing radiomics models using machine learning and/or deep learning methods for the prediction of **A)** ALK and **B)** KRAS mutation status. Numbers are estimates with 95% CIs in brackets and indicated by horizontal lines. For those studies with the same name for the first author and published the same year, a hashtag was added to unequivocally tag them as done in Tables 1 and 2 and in the reference list. ALK, anaplastic lymphoma kinase; CI, confidence interval; DOR, diagnostic odds ration; FPR, false

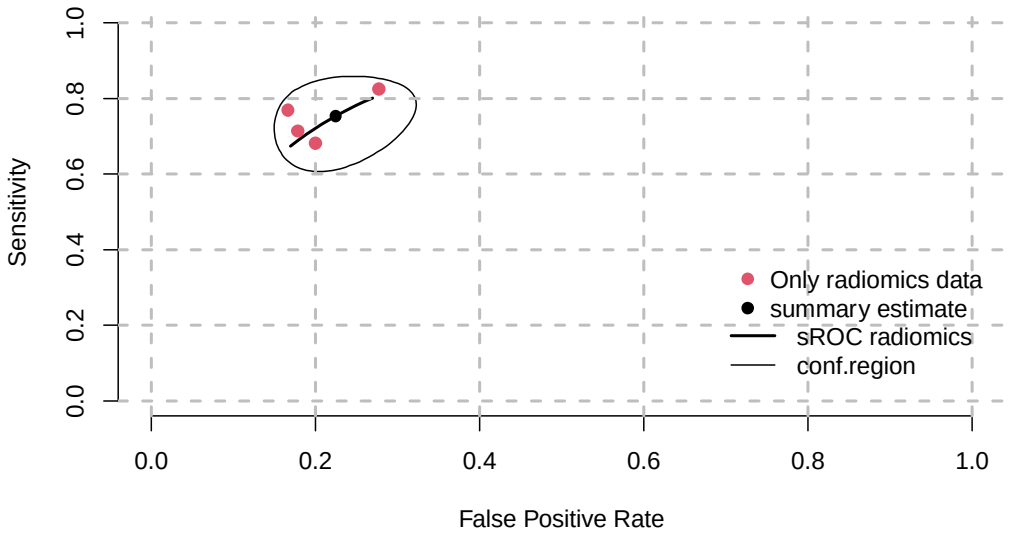


positive rate; KRAS, Kirsten rat sarcoma viral oncogene homologue.

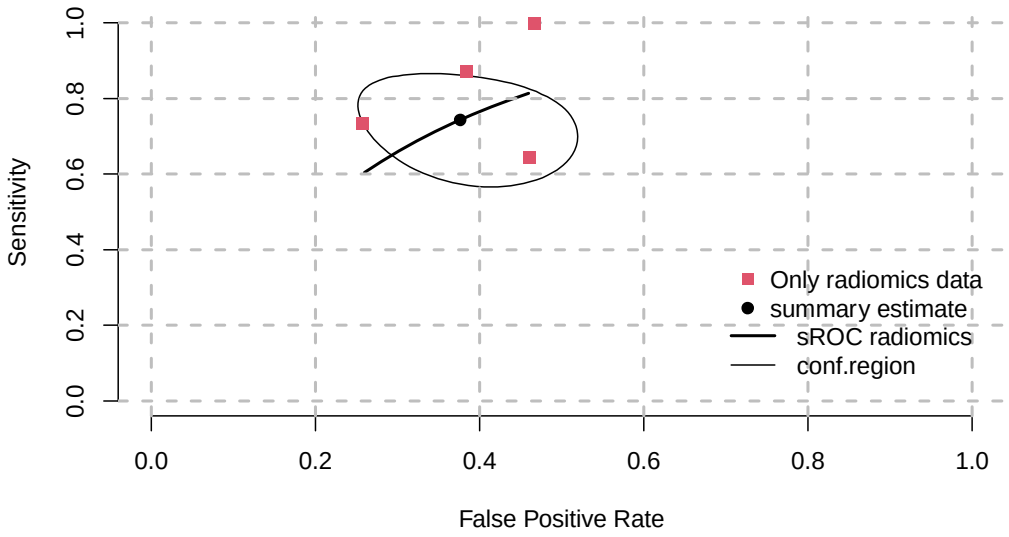
Supplementary Figure S1. Hierarchical sROC curves of included studies for the comparative performance of radiomics models and combined models (radiomics + clinical data) using machine learning and/or deep learning methods for the prediction of EGFR mutation status (n = 24 and n = 23 studies, respectively). EGFR, epidermal growth factor receptor.



Supplementary Figure S2. Hierarchical sROC curve of included studies for the performance of radiomics models for the prediction of ALK mutation status (*n* = 3). ALK, anaplastic lymphoma kinase.



Supplementary Figure S3. Hierarchical sROC curve of included studies for the performance of radiomics models for the prediction of KRAS mutation status ($n = 4$). KRAS, Kirsten rat sarcoma viral oncogene homologue.



Supplementary tables

Supplementary Table S1. Search strategy applied for the qualitative analysis (systematic review).

Databases	Search strategy
MEDLINE (via Pubmed)	("radiomics"[TIAB] OR "radiomic"[TIAB] OR "texture analysis"[TIAB]) AND ("lung neoplasms"[MESH] OR "lung cancer"[TIAB] OR "NSCLC"[TIAB] or "non-small cell lung cancer"[TIAB] OR "lung adenocarcinoma"[TIAB]) AND ("mutational status" OR "mutation" OR "molecular subtype" OR "ALK"[TIAB] OR "anaplastic lymphoma kinase"[TIAB] OR "BRAF"[TIAB] OR "EGFR"[TIAB] OR "Epidermal growth factor receptor"[TIAB] OR "ERRB2"[TIAB] OR "Receptor, ErbB-2"[MESH] OR "HER2"[TIAB] OR "KRAS"[TIAB] OR "Kirsten rat sarcoma virus"[TIAB] OR "Proto Oncogene Proteins c met"[TIAB] OR "NTRK"[TIAB] OR "ROS"[TIAB] OR "c-ros"[TIAB])
COCHRANE LIBRARY	("radiomics" OR “radiomic” OR "texture analysis") AND ("lung neoplasms" OR "lung cancer" OR “NSCLC” OR “non-small cell lung cancer”) AND ("mutational status" OR "mutation" OR "molecular subtype" OR "ALK" OR "anaplastic lymphoma kinase" OR “BRAF” OR “EGFR” OR “ERRB2” OR "Receptor, ErbB-2" OR “HER2” OR “KRAS” OR “Kirsten rat sarcoma virus” OR “Proto Oncogene Proteins c met” OR “NTRK” OR “ROS” OR “c-ros”)
EMBASE	('radiomics':ab,ti OR 'radiomics'/exp OR 'radiomic':ab,ti OR 'texture analysis':ab,ti) AND ('lung cancer'/exp OR 'lung cancer':ab,ti OR 'NSCLC'/exp OR 'NSCLC':ab,ti OR 'non small cell lung cancer'/exp OR 'non small cell lung cancer':ab,ti OR 'lung adenocarcinoma'/exp OR 'lung adenocarcinoma':ab,ti) AND

	('mutational status':ab,ti OR ('mutational' NEAR/2 'status') OR 'mutation':ab,ti OR 'mutation'/exp OR 'molecular subtype':ab,ti OR ('molecular' NEAR/2 'subtype') OR 'ALK':ab,ti OR 'ALK gene'/exp OR 'anaplastic lymphoma kinase':ab,ti OR 'anaplastic lymphoma kinase'/exp OR 'BRAF':ab,ti OR 'BRAF gene'/exp OR 'EGFR':ab,ti OR 'EGFR gene'/exp OR 'Epidermal growth factor receptor':ab,ti OR 'Epidermal growth factor receptor gene'/exp OR 'ERRB2':ab,ti OR 'ERRB2 gene'/exp OR 'epidermal growth factor receptor 2'/exp OR 'epidermal growth factor receptor 2':ab,ti OR 'HER2':ab,ti OR 'KRAS':ab,ti OR 'KRAS gene'/exp OR 'Kirsten rat sarcoma virus':ab,ti OR 'Kirsten rat sarcoma virus'/exp OR 'Proto Oncogene Proteins c met' OR 'MET':ab,ti OR 'MET gene'/exp OR 'NTRK':ab,ti OR 'NTRK gene'/exp OR 'c ros oncogene 1':ab,ti OR 'ROS1':ab,ti OR 'ROS1 gene'/exp)

Supplementary Table S2. Quality assessment results obtained after CLAIM evaluation.

N	**Study**	**Score**			**Mean score**	**Cut-off**
		Reviewer 1 (A.J.P.)	**Reviewer 2 (F.B.B.)**	**Reviewer 3 (A.P.P.)**		
1	Chang et al. 2021[1]	28	24	27	26	17
2	Chang et al. 2021[2]	26	25	26	26	17
3	Dong et al. 2021[3]	23	22	19	21	18
4	Dong et al. 2022[4]	25	22	21	23	18
5	Feng et al. 2022[5]	20	22	23	22	18.5
6	Gao et al. 2023[6]	22	22	20	21	17.5
7	Huo et al. 2022[7]	22	25	25	24	17.5
8	Jia et al. 2019[8]	19	19	20	19	17
9	Jiang et al. 2022[9]	24	23	22	23	16.5
10	Le et al. 2021[10]	22	20	21	21	17.5
11	Li et al. 2018[11]	25	28	22	25	19
12	Li et al. 2019[12]	23	22	21	22	17
13	Li et al. 2020[13]	25	23	22	23	19
14	Li et al. 2022[14]	21	21	21	21	17

15	Liu et al. 2020[15]	22	25	23	23	17
16	Liu et al. 2022[16]	24	23	22	23	17
17	Lu et al. 2020[17]	27	30	26	28	17.5
18	Lu et al. 2022[18]	22	25	21	23	17.5
19	Ma et al. 2020[19]	24	26	22	24	17
20	Nair et al. 2021[20]	21	22	19	21	17.5
21	Ninomiya et al. 2021[21]	21	22	21	21	17.5
22	Ninomiya et al. 2023[22]	21	22	22	22	17
23	Rios Velazquez et al. 2017[23]	19	23	20	21	18
24	Rossi et al. 2021[24]	20	22	19	20	17.5
25	Song et al. 2020[25]	27	28	27	27	19
26	Tu et al. 2019[26]	19	21	20	20	17
27	Wang et al. 2022[27]	23	26	26	25	18
28	Wang et al. 2022[28]	24	22	22	23	19
29	Weng et al. 2021[29]	24	25	24	24	17.5
30	Wu 2020[30]	20	22	21	21	17
31	Yang 2020[31]	22	23	23	23	17
32	Yang 2022[32]	19	18	18	18	17
33	Zhang 2018[33]	26	27	23	25	17
34	Zhang 2020[34]	19	19	19	19	17

35	Zhang 2020[35]	22	23	23	23	17
36	Zhang 2021[36]	27	26	26	26	17
37	Zhao 2022[37]	23	24	23	23	19
38	Zhu 2022[38]	22	22	20	21	17.5

Supplementary Table S3. Studies included in the different meta-analyses conducted.

N	EGFR		ALK	KRAS
	Radiomics models	**Combined models**	**Radiomics models**	**Radiomics models**
1	Chang et al. 2021[1]	Dong et al. 2022[4]	Chang et al. 2021[2]	Dong et al. 2021[3]
2	Dong et al. 2021[3]	Gao et al. 2023[6]	Ma et al. 2020[19]	Le et al. 2021[10]
3	Feng et al. 2022[5]	Huo et al. 2022[7]	Song et al. 2020[25]	RiosVelazquez et al. 2017[23]
4	Gao et al. 2023[6]	Jia et al. 2019[8]		Wang et al. 2022[28]
5	Le et al. 2021[10]	Jiang et al. 2022[9]		
6	Li et al. 2018[11]	Li et al. 2018[11]		
7	Li et al. 2019[12]	Li et al. 2019[12]		
8	Li et al. 2022[14]	Li et al. 2020[13]		

9	Liu et al. 2020[15]	Liu et al. 2020[15]		
10	Liu et al. 2022[16]	Lu et al. 2020[17]		
11	Lu et al. 2020[17]	Lu et al. 2022[18]		
12	Lu et al. 2022[18]	Ninomiya et al. 2023[22]		
13	Nair et al. 2021[20]	Rios Velazquez et al. 2017[23]		
14	Ninomiya et al. 2021[21]	Rossi et al. 2021[24]		
15	RiosVelazquez et al. 2017[23]	Tu et al. 2019[26]		
16	Tu et al. 2019[26]	Wang et al. 2022[27]		
17	Wang et al. 2022[27]	Weng et al. 2021[29]		
18	Weng et al. 2021[29]	Wu 2020[30]		
19	Yang 2020[31]	Yang 2022[32]		
20	Zhang 2020[34]	Zhang 2018[33]		

21	Zhang 2020[35]	Zhang 2020[35]		
22	Zhang 2021[36]	Zhang 2021[36]		
23	Zhao 2022[37]	Zhu 2022[38]		
24	Zhu 2022[38]			

ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor; KRAS, Kirsten rat sarcoma viral oncogene homologue.

Supplementary Table S4. Type of models (radiomic model/deep learning or combined [radiomic features + clinical variables]) developed in the studies for EGFR prediction and the radiomics/clinical features included. EGFR, epidermal growth factor receptor.

Study	Models	Radiomic features	Clinical variables
Chang et al. 2021[1]	Radiomic	ShortRunLowGreyLevelEmphasis_AllDirection_offset1_SDH Percentile85 OneVoxelVolume Flatness ShortRunEmphasis_AllDirection_offset_SD HaralickCorrelation_AllDirection_offset4_SD Zone Percentage GLCM_Entropy_AllDirection_offset7_SD Correlation_AllDirection_offset7_SD CT_GLCMEntropy_AllDirection_offset1_SD HaralickCorrelation_angle135_offset7 LongRunHighGreyLevelEmphasis_angleO_offset 1 ShortRunLowGreyLevelEmphasis_AllDirection_offset7_SD	N/A

		HaralickCorrelation_AllDirection_offset1_SD SurfaceVolumeRatio	
Dong et al. 2021[3]	Deep learning	Not specified	N/A
Dong et al. 2022[4]	Combined	wavelet-HLL_GLCM_MaximumProbability wavelet-LLL_GLCM_MaximumProbability original_GLCM_SumEntropy) log-sigma-1-0-mm-3D_GLCM_MaximumProbability wavelet-LHL_firstorder_Kurtosis wavelet-LLL_firstorder_Skewness log-sigma-2-0-mm-3D_firstorder_Kurtosis original_shape_Sphericity wavelet-LHL_GLSZM_LargeAreaHighG	Smoking status Histological type
Feng et al. 2022[5]	Radiomic	Skewness.7_firstorder_wavelet-LHL SmallAreaHighGrayLevelEmphasis.7_GLSZM_wavelet-LHL HighGrayLevelZoneEmphasis.12_GLSZM_wavelet-HHH 90Percentile_firstorder_original Variance.4_firstorder_square Range.4_firstorder_square	

		GrayLevelVariance.26_GLSZM_wavelet-LHH JointAverage.11_GLCM_wavelet-HLH MeanAbsolute Deviation.4_firstorder_square RobustMeanAbsoluteDeviation.4_firstorder_square GrayLevelNonUniformity.32_GLSZM_wavelet-LLH | GrayLevelNonUniformity.1_girlm_original GrayLevelNonUniformity.4_girlm_logarithm GrayLevelNonUniformity.16_girlm_squareroot HighGrayLevelRunEmphasis.5_girlm_squareroot GrayLevelNonUniformityNormalized.21_GLSZM_wavelet-LLH LowGrayLevelRunEmphasis girlm_original LowGrayLevelRunEmphasis.5_girlm_squareroot GrayLevel Variance.32_GLSZM_wavelet-LLH Minimum.4_firstorder_square SmallArealowGrayLevelEmphasis.11_GLSZM_wavelet-HLH SmallArealowGrayLevelEmphasis.12_GLSZM_wavelet-HHH Mean.12_firstorder_wavelet-HHH SmallArealowGrayLevelEmphasis.9_GLSZM_wavelet-HLL Imc2.12 GLCM wavelet-HHH ADC	N/A
Gao et al. 2023[6]	Radiomic Combined	original_firstorder_Kurtosis original_firstorder_Median original_firstorder_Skewness log-sigma-1-0-mm-3D_firstorder_Energy	CEA Sex (male) Nodule type (sub-

		log-sigma-4-0-mm-3D_GLDM_DependenceVariance wavelet-LHL_GLRLM_LongRunLowGrayLevelEmphasis wavelet-HLL_firstorder_Energy	solidity)
Huo et al. 2022[7]*	Combined	137 features (not specified)	Age Sex (female) Non-smokers Clinical stage (I-II)
Jia et al. 2019[8]	Combined	94 features (not specified)	Sex Smoking history
Jiang et al. 2022[9]	Combined	Skewness Minimum Kurtosis Variance Minimum 10th percentile	Age Sex

		SumSquare SizeZoneNonUniformity HighGrayLevelZoneEmphasis ZoneVariance LargeDependence HighGrayLevelEmphasis LargeDependenceHighGrayLevel Emphasis DependenceEntropy	Smoking Tumor Family history
Le et al. 2021[10]	Radiomic	wavelet-LLLfirstorderEnergy wavelet-LHHGLSZMGrayLevelNonUniformityNormalized wavelet-HHLGLDMSmallDependenceLowGratLevelEmphasis wavelet-HLHGLCM_MCC wavelet-HLHGLSZMSmallAreaLowGrayLevelEmphasis wavelet-HHHGLCMjointEnergy wavelet-HHHGLRLMGrayLevelNonUniformityNormalized	N/A
Li et al. 2018[11]	Radiomic Combined	338 features (not specified)	Sex Smoking status

Li et al. 2019[12]	Radiomic Combined	CT_GGS_Gray Span CT_GGC_Gray Mean	Age Sex Smoking status Clinical stage Lesion location
Li et al. 2020[13]	Combined	12 features (not specified)	Sex Age Smoking status
Li et al. 2022[14]	Radiomic	3 features (not specified)	–
Liu et al. 2020[15]	Radiomic Combined	**RADIOMIC MODEL:** wavelet-HLH_GLDM_DependenceVariance wavelet-LHL_GLDM_LargeDependenceLowGrayLevelEmphasis logarithm_GLCM_InverseVariance square_GLDM_DependenceVariance wavelet-HLH_GLDM_LargeDependenceHighGrayLevelEmphasis	

		wavelet-HHH_GLCM_Id log-sigma-0-5-mm-3D_GLSZM_ZoneEntropy square_GLCM_Correlation original_GLCM_ClusterShade wavelet-LHH_GLDM_LargeDependenceHighGrayLevelEmphasis **COMBINED MODEL:** wavelet-HLH_GLDM_DependenceVariance custom_PatientSex logarithm_GLCM_InverseVariance square_GLCM_Correlation wavelet-HLL_firstorder_Kurtosis wavelet-LHL_GLRLM_LongRunLowGrayLevelEmphasis wavelet-HLL_firstorder_Median original_GLSZM_SizeZoneNonUniformityNormalized exponential_firstorder_Skewness	Age Sex Smoking history

		wavelet-LLH_GLCM_ClusterShade	
Liu et al. 2022[16]	Radiomic	Mean absolute deviation 60 Percentile area Convex Correlation Dissimilarity 5-1 Homogeneity 2 10-4 Homogeneity 2 -333-7 Information measure corr 1 8-1 Information measure corr 1 9-7 Information measure corr 1 2-4 Inverse diff norm 6-4 Inverse variance 8-4 Inverse variance 8-1 Max Probability 12-7 Max Probability	–

		-333 Run length nonuniformity	
Lu et al. 2020[17]	Radiomic Combined	original_GLSZM_SmallAreaHighGrayLevelEmphasis original_GLSZM_SmallAreaLowGrayLevelEmphasis original_GLDM_LowGrayLevelEmphasis log-sigma-1-0-mm-3D_GLCM_Cluster Prominence log-sigma-3-0-mm-3D_GLDM_DependenceNonUniformityNormalized wavelet-LLL_GLCM_InverseVariance wavelet-LLH_GLCM_Imc2 wavelet-HLL_firstorder_Mean wavelet-HLL_GLSZM_LowGrayLevelZoneEmphasis wavelet-HLL_GLDM_SmallDependenceHighGrayLevelEmphasis wavelet-HLH_GLSZM_SizeZoneNonUniformityNormalized wavelet-HHH_firstorder_Skewness wavelet-HHH_GLSZM_Size Zone Non Uniformity Normalized	Sex Smoking status Pathohistological subtype Vascular infiltration status
Lu et al. 2022[18]	Radiomic Combined	1269 features (not specified)	Age Sex Smoking status Stage of disease

			Serum level of tumor markers (CEA, CYFRA 21-1, SCC, Pro-GRP)
Nair et al. 2021[20†]	Radiomic	NGTDM_600_Complexity Glrl_Saggital_30_ShortRunEmphasis Glrl_Saggital_30_ShortRunHighGrayLevelEmphasis Glrl_Saggital_120_ShortRunHighGrayLevelEmphasis Glrl_Coronal_120_ShortRunHighGrayLevelEmphasis Glrl_Coronal_30_ShortRunEmphasis Glrl_Saggital_120_ShortRunEmphasis Glrl_Axial_30_ShortRunEmphasis Glrl_Coronal_120_ShortRunEmphasis FirstOrder_HistogramBin2	–
Ninomiya et al. 2021[21]	Radiomic	**BN MODEL:** b0_GLCM_Energy_45, b1/b0_GLSZM_ZSN_104	

		b1_GLCM_SumAverage_122 b0_GLRLM_Lrlge_97) **OI MODEL:** GLRLM_ ShortRunLowGrayLevelEmphasis GLSZM_ LowGrayLevelZoneEmphasis GLSZM_ShortZoneLowGrayEmphasis **WD MODEL:** GLSZM_ LowGrayLevelZoneEmphasis _LL	–
Ninomiya et al. 2023[22]	Combined	GLSZM_SmallAreaLowGrayLevelEmphasis GLSZM_LargeAreaEmphasis Hist.RootMeanSquared GLDM_DependenceVariance	Sex Smoking status
RiosVelazquez et al. 2017[23]	Radiomic Combined	imaging.Wavelet_LHH_GLCM_invDiffmomnor imaging.LoG_sigma_3_mm_3D_GLSZM_highIntensityLarteAreaEmp	

		imaging.Wavelet_LLL_GLCM_clusProm imaging.GLCM_maxProb imaging.Wavelet_LLL_stats_energy imaging.Wavelet_HLL_stats_var imaging.LoG_sigma_3_mm_3D_GLSZM_largeAreaEmphasis imaging.Wavelet_LLH_stats_range imaging.Wavelet_LHH_GLCM_clusProm imaging.Wavelet_LLL_GLSZM_highIntensityLarteAreaEmp imaging.Wavelet_HLH_GLSZM_lowIntensitySmallAreaEmp imaging.Wavelet_HHL_stats_energy imaging.Wavelet_LLH_stats_mean imaging.Stats_median imaging.Wavelet_HHL_GLCM_maxProb imaging.LoG_sigma_3_mm_3D_GLCM_clusProm imaging.Shape_spherDisprop imaging.Stats_kurtosis imaging.Wavelet_HHH_GLCM_correl1 imaging.LoG_sigma_3_mm_3D_rlgl_grayLevelNonuniformity	Stage Sex Smoking status Age Race
Rossi et al. 2021[24]	Combined	First order_90 Percentile First order_Entropy First order_Maximum First order_Median First order_Robust mean absolute deviation	Sex Smoking status

		First order_Root mean squared First order_Skewness First order_Uniformity GLCM_Correlation GLCM_Difference average GLCM_Difference entropy GLCM_InverseDifference GLCM_ InverseDifferenceMoment GLCM_ InverseDifferenceMomentNormalized GLCM_ InverseDifferenceNormalized GLCM_ InformationalMeasureCorrelation1 GLCM_ InformationalMeasureCorrelation2 GLCM_InverseVariance GLCM_JointEnergy GLCM_JointEntropy GLCM_MaximalCorrelationCoefficient	

		GLCM_MaximumProbability GLCM_SumEntropy GLDM_DependenceEntropy GLDM_DependenceNonUniformity GLDM_Dependence NonUniformityNormalized GLDM_DependenceVariance GLDM_GrayLevelNonUniformity GLDM_LargeDependenceEmphasis GLDM_SmallDependenceEmphasis GLRLM_GrayLevelNonUniformity GLRLM_GrayLevelNonUniformityNormalized GLRLM_RunEntropy GLRLM_RunPercentage GLRLM_ShortRunEmphasis GLSZM_GrayLevelNonUniformity GLSZM_GrayLevelNonUniformityNormalized	

		GLSZM_SizeZoneNonUniformityNormalized GLSZM_SmallAreaEmphasis GLSZM_ZoneEntropy GLSZM_ZoneVariance NGTDM_Coarseness	
Tu et al. 2019[26]	Radiomic Combined	X0_GLRLM_ RunLengthNon-Uniformity X4_H_median X0_GLCM_homogeneity1	Maximum diameter Location Sex
Wang et al. 2022[27]	Radiomic Combined‡	Not specified	Age Sex Tumor staging Number Size Past recurrence Medication status

Weng et al. 2021[29]	Combined	SmallAreaEmphasis LongRunHigh GreyLevelEmphasis_angle0_offset4 ClusterProminence_All Direction_offset7_SD InverseDifference Moment_All Direction_offset4_SD LowGreyLevel Run Emphasis_All Direction_offset4_SD LongRunLowGrey Level Emphasis_All Direction_offset7_SDCorrelation_angle0_offset7 std Deviation GLCM Energy_All Direction_offset4_SD	Smoking status Spiculation Air bronchogram CEA SCCA
Wu 2020[30]	Combined	Not specified	Smoking status Histological subtype
Yang 2020[31]	Radiomic	Not specified	–
Yang 2022[32]	Radiomic Combined	Nonwavelet-LHH_NGTDM_Strength wavelet-LHH_GLDM_DependenceEntropy wavelet-LLL_GLSZM_LargeAreaLowGrayLevelEmphasis wavelet-LLL_firstorder_Minimum	Sex Emphysema Interstitial lung disease

		wavelet-LLH_NGTDM_Contrast wavelet-LHH_NGTDM_Strength log-sigma-1-5-mm-3D_firstorder_Kurtosis wavelet-LHL_GLCM_ClusterShade wavelet-LHH_NGTDM_Strength wavelet-LLL_GLSZM_LargeAreaLowGrayLevelEmphasis wavelet-LLH_firstorder_Mean original_NGTDM_Contrast original_firstorder_Kurtosis log-sigma-1-5-mm-3D_firstorder_Kurtosis wavelet-LLL_NGTDM_Contrast original_GLCM_MaximumProbability wavelet-LLL_GLSZM_LargeAreaLowGrayLevelEmphasis	
Zhang 2018[33]	Combined	IIF.range IIF.Skewness W_{LLH}F.IF.mean_absulute_eviation	Histological subtype Sex Smoking status

		W_{LHH}F.IF.median W_{LLH}F.IF.mean W_{LLH}F.GLCM.variance GLRLM_ HighGrayLevelRunEmphasis	
Zhang 2020[34]	Radiomic	GLSZM_ HighGrayLevelZoneEmphasis GLDM_DependenceVariance GLSZM_ GreyLevelNon UniformityNormalized GLSZM_ZoneEntropy	–
Zhang 2020[35]	Radiomic Combined	784 features (not specified)	Sex Histopathological Subtype Age
Zhang 2021[36]	Radiomic Combined	fo_Skewness exp_GLRLM_ShortRunEmphasis exp_GLRLM_ShortRunHighGrayLevelEmphasis exp_GLDM_SmallDependenceEmphasis	Smoking history Bubble-like lucency Pleural attachment Pleural retraction

		grad_GLDM_DependenceEntropy LLH_fo_90P LLH_GLCM_SumEntropy LLL−fo_kurtosis LLL−GLCM_ClusterProminence LLL_GLSZM_GrayLevelNonUniformityNormalized LLL_GLSZM_GrayLevelVariance LLL_GLSZM_ZoneEntropy	
Zhao 2022[37§]	Radiomic	CT_Shape_Sphericity CT_GLRLM_ShortRunEmphasis CT_GLRLM_ShortRunHighGreyLevelEmphasis CT_NGLDM_Busyness CT_Glzlm_ShortZoneEmphasis	–
Zhu 2022[38]	Radiomic Combined	log_sigma_1.0_mm_3D_GLRLM_RunVariance wavelet_LLH_firstorder_RootMeanSquared log-sigma-2-0-mm-3D_GLCM_ClusterShade	Sex Age Emphysema

		wavelet_HHH_firstorder_Mean	Pathological subtype

*Combined model also included 14 CT features: location (peripheral), tumor size ≥3cm, subsolid density, spiculation, lobulation, air bronchogram , air space, necrosis , calcification (presence), vascular convergence sign, pleural retraction sign, pleural effusion, lymphatic metastasis and multiple pulmonary metastasis.

†Top 10 selected features. The maximum number of texture features included was determined by maximizing cross-validated accuracy. This value was not the same for each binary group or each machine learning model.

‡Note that this model includes radiomic features + deep features and clinical variables.

§Model 1.

BN, Betti numbers; CEA, carcinoembryonic antigen; CYFRA 21-1, fragment of cytokeratin sub-unit 19; GLCM, gray-level co-occurrence matrix; GLDM, gray-level dependence matrix; GLRLM, gray-level run-length matrix; GLSZM, gray-level size zone matrix; NGTDM, neighbouring gray tone difference matrix; OI, original image; Pro-GRP, pro-gastrin-releasing peptide; SCC, squamous cell carcinoma antigen; WD, wavelet decomposition.

Supplementary Table S5. Type of models (radiomic model or combined [radiomic features + clinical variables]) developed in the studies for ALK prediction and the radiomics/clinical features included. ALK, anaplastic lymphoma kinase.

Study	Models	Radiomic features	Clinical variables
Chang et al. 2021[2]	Radiomic	CT_uniformity CT_LongRunEmphasis_AllDirection_offset4_SD CT_HaraEntropy CT_GLCMEnergy_angle135_offset7 CT_LongRunHighGreyLevelEmphasis_angle45_offset1 CT_LongRunLowGreyLevelEmphasis_AllDirection_offset7_SD CT_Correlation_AllDirection_offset4_SD CT_Percentile70 CT_HaralickCorreltion_AllDirection_offset4_SD CT_LongRunLowGreyLevelEmphasis_AllDirection_offset4_SD CT_LongRunEmphasis_angle135_offset4 CT_LongRunHighGreyLevelEmphasis_angle90_offset4	–

		CT_LongRunLowGreyLevelEmphasis_AllDirection_offset1_SD CT_HaralickCorreltion_AllDirection_offset7_SD CT_ShortRunEmphasis_AllDirection_offset1_SD CT_LongRunHighGreyLevelEmphasis_angle0_offset1 CT_GLCMEntropy_angle90_offset1 CT_Percentile30 CT_LongRunEmphasis_angle90_offset4 CT_LongRunEmphasis_AllDirection_offset1_SD	
Ma et al. 2020[19]	Radiomic	**PRE-CONTRAST MODEL:** wavelet-LLL_GLCM_DifferenceVariance wavelet-LLH_firstorder_Median wavelet-LLH_NGTDM_Busyness wavelet-LHL_GLSZM_LargeAreaLowGrayLevelEmphasis wavelet-HHH_GLSZM_LargeAreaLowGrayLevelEmphasis wavelet-LHL_firstorder_Energy	–

		wavelet-HHL_firstorder_90Percentile wavelet-HHL_GLCM_JointEntropy wavelet-HHL_firstorder_Uniformity wavelet-HHL_firstorder_RobustMeanAbsoluteDeviation wavelet-LHH_GLDM_LargeDependenceLowGrayLevelEmphasis wavelet-HLH_firstorder_Median wavelet-LHL_GLDM_LargeDependenceLowGrayLevelEmphasis wavelet-HHL_GLCM_InverseDifference wavelet-HHL_firstorder_InterquartileRange wavelet-HHL_GLCM_MaximumProbabiblity wavelet-HHH_GLSZM_SmallAreaLowGrayLevelEmphasis wavelet-HHL_firstorder_Mean wavelet-HLL_GLCM_ClusterShade wavelet-HHL_GLSZM_SmallAreaLowGrayLevelEmphasis wavelet-LHH_GLCM_ MaximalCorrelationCoefficient wavelet-LLL_GLSZM_SizeZoneNonUniformityNormalized	

		wavelet-LLL_GLSZM_SmallAreaEmphasis wavelet-HHL_GLCM_ InverseDifferenceNormalized **POST-CONTRAST MODEL:** wavelet-LHH_GLDM_SmallDependenceHighGrayLevelEmphasis wavelet_HHL_GLSZM_GrayLevelNonUniformity wavelet-LLH_firstorder_Mean wavelet-LLH_GLSZM_HighGrayLevelZoneEmphasis wavelet-LLH_GLSZM_SmallAreaHighGrayLevelEmphasis wavelet-LLH_GLSZM_SmallAreaLowGrayLevelEmphasis wavelet-HHH_GLCM_MaximumProbability wavelet-LLL_GLDM_LargeDependenceLowGrayLevelEmphasis wavelet-HLL_GLDM_DependenceVariance wavelet-HHH_firstorder_Mean wavelet-HHH_GLDM_LowGrayLevelEmphasis wavelet-LLH_firstorder_90Percentile	

		wavelet-HHL_GLDM_DependenceVariance wavelet-HHH_GLCM_MaximalCorrelationCoefficient wavelet-HHH_NGTDM_Contrast wavelet-original_GLCM_InverseVariance wavelet-LLH_firstorder_Range wavelet-HHL_GLCM_MaximalCorrelationCoefficient wavelet-HLL_GLSZM_GrayLevelNonUniformityNormalized	
Song et al. 2020[25]	Radiomic Combined	**RADIOMIC MODEL:** Original_Firstorder_90Percentile Original_Firstorder_Entropy Original_Firstorder_Maximum Wavelet-LHH_Firstorder_10Percentile Wavelet-HLL_Firstorder_Median Wavelet-HHH_Firstorder_Mean LoG-sigma-1-0-mm-3D_Firstorder_Median	Age Sex Smoking history

		LoG-sigma-1-0-mm-3D_Firstorder_RootMeanSquared LoG-sigma-1-0-mm-3D_Firstorder_Minimum LoG-sigma-2-0-mm-3D_Firstorder_10Percentile LoG-sigma-3-0-mm-3D_Firstorder_90Percentile LoG-sigma-5-0-mm-3D_Firstorder_Skewness Original_ GLCM _ClusterShade Wavelet-LHH_ GLCM _Correlation Wavelet-LHL_ GLCM _InverseDifferenceNormalized Wavelet-HHH_ GLCM _InformationalMeasureofCorrelation1 LoG-sigma-1-0-mm-3D_GLCM _Autocorrelation LoG-sigma-2-0-mm-3D_GLCM _InverseVariance Original_GLSZM_ SmallAreaHighGrayLevelEmphasis Wavelet-HHH_ GLSZM _ SmallAreaHighGrayLevelEmphasis Wavelet-HLL_ GLSZM _ZoneEntropy Wavelet-HLH_ GLSZM _ZoneEntropy LoG-sigma-2-0-mm-3D_GLSZM _ZoneEntropy	Smoking index Clinical stage Distal metastasis Pathological invasiveness of the tumor

		LoG-sigma-3-0-mm-3D_GLSZM _ SmallAreaEmphasis LoG-sigma-3-0-mm-3D_GLSZM _ Size-ZoneNonUniformityNormalized LoG-sigma-5-0-mm-3D_GLSZM _ GrayLevelNonUniformityNormalized Wavelet-LHH_GLDM_ LargeDependenceHighGrayLevelEmphasis LoG-sigma-1-0-mm-3D_GLDM _ HighGrayLevelEmphasis LoG-sigma-3-0-mm-3D_GLRLM_RunPercentage LoG-sigma-4-0-mm-3D_GLRLM _ LongRunLowGrayLevelEmphasis **COMBINED MODEL:** Current smoker Stage I Male Local lymphadenopathy Pericardial effusion Left Lower Lobe lesion No cavity in the lesion	

		Lobulated margin No pleural retraction sign No local lymphadenopathy Wavelet-HHL_Firstorder_Kurtosis Wavelet-HLL_Firstorder_Median Wavelet-LHH_Firstorder_Skewness Wavelet-LLL_Firstorder_Minimum Wavelet-HLH_Firstorder_Median LoG-sigma-1-0-mm-3D_Firstorder_Minimum LoG-sigma-2-0-mm-3D_Firstorder_Minimum Wavelet-LLL_GLCM_ClusterShade Wavelet-LLH_GLCM _InformationalMeasureofCorrelation2 Wavelet-HLH_GLCM _InformationalMeasureofCorrelation2 Wavelet-HLH_GLCM_InformationalMeasureofCorrelation1 LoG-sigma-1-0-mm-3D_GLCM_InformationalMeasureofCorrelation1 LoG-sigma-3-0-mm-3D_GLCM_InformationalMeasureofCorrelation2	

		LoG-sigma-5-0-mm-3D_GLCM_InformationalMeasureofCorrelation2 Original_Shape_MajorAxisLength Wavelet-HLH_GLSZM_SizeZoneNon-Uniformity LoG-sigma-4-0-mm-3D_GLSZM_GrayLevelNonUniformityNormalized Wavelet-HLH_GLDM_ LargeDependenceHigh GrayLevelEmphasis Wavelet-HHH_GLDM_ LargeDependenceHigh GrayLevelEmphasis Original_GLRLM_ HighGrayLevelRunEmphasis	

GLCM , gray-level co-occurrence matrix; GLDM , gray-level dependence matrix; GLRLM , gray-level run-length matrix; GLSZM , gray-level size zone matrix; NGTDM , neighbouring gray tone difference matrix.

Supplementary Table S6. Type of models (radiomic model or combined [radiomic features + clinical variables]) developed in the studies for KRAS prediction and the radiomics/clinical features included. KRAS, Kirsten rat sarcoma viral oncogene homologue.

Study	Models	Radiomic features	Clinical variables
Dong et al. 2021[3]	Radiomic	Not specified	–
Le et al. 2021[10]	Radiomic	wavelet-LLHGLSZMLargeAreaEmphasis wavelet-LLLGLDMDependenceEntropy wavelet-LHHGLDMLargeDependenceLowGrayLevelEmphasis ori-firstorderkurtosis wavelet-HLHGLCMInverseVariance wavelet-HLLGLSZMSmallAreaHighGrayLevelEmphasis wavelet-LHHGLCMId wavelet-HHLGLCMDifferenceEntropy	–

		wavelet-LLLGLSZMGrayLevelNonUniformityNormalized wavelet-HHHGLCMDifferenceAverage wavelet-HHHGLDMDependenceEntropy	
RiosVelazquez et al. 2017[23]	Radiomic Combined	imaging.LoG_sigma_3_mm_3D_GLSZM_highIntensityLarteAreaEmp imaging.Wavelet_LHH_GLCM_clusProm imaging.Wavelet_LHH_GLCM_energy imaging.Wavelet_LLL_stats_energy imaging.Wavelet_LLL_stats_median imaging.LoG_sigma_3_mm_3D_GLSZM_largeAreaEmphasis imaging.Wavelet_HHH_GLSZM_lowIntensitySmallAreaEmp imaging.Wavelet_HHH_GLCM_correl1 imaging.LoG_sigma_3_mm_3D_GLCM_clusProm imaging.Wavelet_LLL_GLSZM_highIntensityLarteAreaEmp imaging.Wavelet_HHL_stats_energy imaging.Wavelet_HLL_stats_var imaging.Wavelet_HLH_GLSZM_lowIntensitySmallAreaEmp imaging.Wavelet_LHH_rlgl_GrayLevelNonuniformity imaging.LoG_sigma_3_mm_3D_GLSZM_lowIntensitySmallAreaEmp imaging.GLCM_clusShade imaging.Wavelet_LHH_GLCM_invDiffmomnor imaging.Wavelet_HLL_stats_min imaging.Wavelet_LLL_rlgl_longRunHighGrayLevEmpha imaging.Wavelet_LLH_stats_mean	Stage Sex Smoking status Age Race
Wang et al. 2022[28]	Radiomic	CT_square_GLSZM_SizeZoneNonUniformityNormalized CT_wavelet-LHH_GLDM_DependenceNonUniformityNormalized	–

		CT_wavelet-HHL_firstorder_Skewness CT_wavelet-HHL_GLDM_DependenceNonUniformityNormalized	

GLCM , gray-level co-occurrence matrix; GLDM , gray-level dependence matrix; GLRLM , gray-level run-length matrix; GLSZM , gray-level size zone matrix.

Supplementary Table S7. Results of the meta-regression analyzing the effects of age, type of segmentation (manual/semi-automatic/automatic), type of model (radiomics/combined [radiomic features + clinical data) and artificial intelligence methodology (machine learning/deep learning).

AGE Fixed-effects coefficients					
	Estimate	**SE**	**z**	***p*-value**	**CI 95%**
tsens.(Intercept)	4.488	2.181	2.058	0.040	[0.214, 8.762]
tsens.AGE	-0.052	0.035	-1.483	0.138	[-0.121, 0.017]
tfpr.(Intercept)	-0.156	2.288	-0.068	0.946	[-4.640, 4.328]
tfpr.AGE	-0.012	0.037	-0.318	0.750	[-0.084, 0.061]
Variance components: between-studies Std. Dev and correlation matrix					
	SD		**tsens**		**tfpr**
tsens	0.395		-		0.819
tfpr	0.431		0.819		-
TYPE OF SEGMENTATION Fixed-effects coefficients					
	Estimate	**SE**	**z**	***p*-value**	**CI 95%**
tsens.(Intercept)	1.217	0.127	9.599	0.000	[0.968, 1.465]
tsens.SegmentationSemiautomatic	-0.086	0.241	-0.356	0.722	[-0.558, 0.387]

tsens.SegmentationUnknown	0.347	0.708	0.491	0.624	[-1.040, 1.734]
tfpr.(Intercept)	-0.984	0.133	-7.426	0.000	[-1.244, -0.725]
tfpr.SegmentationSemiautomatic	0.272	0.254	1.067	0.286	[-0.227, 0.770]
tfpr.SegmentationUnknown	0.397	0.697	0.569	0.569	[-0.970, 1.763]

Variance components: between-studies Std. Dev and correlation matrix

	SD	**tsens**	**tfpr**
tsens	0.465	-	0.827
tfpr	0.511	0.827	-

CONTRAST

Fixed-effects coefficients

	Estimate	**SE**	**z**	***p*-value**	**CI 95%**
tsens.(Intercept)	1.244	0.608	2.045	0.041	[0.052, 2.437]
tsens.Contrastcontrast-enhanced	0.151	0.650	0.232	0.817	[-1.124, 1.425]
tsens.Contrastnon-contrast CT	-0.096	0.620	-0.156	0.876	[-1.312, 1.119]
tfpr.(Intercept)	-0.619	0.638	-0.970	0.332	[-1.870, 0.632]
tfpr.Contrastcontrast-enhanced	-0.509	0.683	-0.745	0.456	[-1.848, 0.829]
tfpr.Contrastnon-contrast CT	-0.309	0.650	-0.476	0.634	[-1.584, 0.965]

Variance components: between-studies Std. Dev and correlation matrix

	SD	**tsens**	**tfpr**
tsens	0.468	-	0.889
tfpr	0.511	0.889	-

TYPE OF MODEL

Fixed-effects coefficients

	Estimate	**SE**	**z**	***p*-value**	**CI 95%**
tsens.(Intercept)	1.281	0.133	9.603	0.000	[1.020, 1.543]
tsens.Modelrad	-0.193	0.197	-0.980	0.327	[-0.579, 0.193]
tfpr.(Intercept)	-0.939	0.146	-6.413	0.000	[-1.226, -0.652]
tfpr.Modelrad	0.006	0.222	0.026	0.979	[-0.429, 0.441]

Variance components: between-studies Std. Dev and correlation matrix

	SD	**tsens**	**tfpr**
tsens	0.434	-	0.828
tfpr	0.524	0.828	-

AI METHODOLOGY

Fixed-effects coefficients

	Estimate	**SE**	**z**	***p*-value**	**CI 95%**
tsens.(Intercept)	1.215	0.232	5.2240	0.000	[0.761, 1.670]
tsens.TypeML	-0.022	0.257	-0.087	0.930	[-0.526, 0.481]

tfpr.(Intercept)	-0.898	0.257	-3.489	0.000	[-1.403, 0.394]
tfpr.TypeML	-0.052	0.285	-0.184	0.854	[-0.610, 0.505]

Variance components: between-studies Std. Dev and correlation matrix

	SD	**tsens**	**tfpr**
tsens	0.442	-	0.795
tfpr	0.520	0.795	-

AI , artifical intelligence; CI , confidence interval; ML , machine learning; rad, model including only radiomic features; SE, standard error; SD, Standard deviation; z, standard score in a gaussian distribution; tsens, logarithmic transformation of sensitivity; tfpr, logarithmic transformation of false positive rate.

REFERENCES

1 Chang C, Zhou S, Yu H et al. A clinically practical radiomics-clinical combined model based on PET/CT data and nomogram predicts EGFR mutation in lung adenocarcinoma. *Eur Radiol*. 2021; 31(8): 6259-6268.

2 Chang C, Sun X, Wang G et al. A machine learning model based on PET/CT radiomics and clinical characteristics predicts ALK rearrangement status in lung adenocarcinoma. *Front Oncol*. 2021; 11: 603882.

3 Dong Y, Hou L, Yang W et al. Multi-channel multi-task deep learning for predicting EGFR and KRAS mutations of non-small cell lung cancer on CT images. *Quant Imaging Med Surg*. 2021; 11(6): 2354-2375.

4 Dong Y, Jiang Z, Li C et al. Development and validation of novel radiomics-based nomograms for the prediction of EGFR mutations and Ki-67 proliferation index in non-small cell lung cancer. *Quant Imaging Med Surg*. 2022; 12(5): 2658-2671.

5 Feng Y, Song F, Zhang P et al. Prediction of EGFR mutation status in non-small cell lung cancer based on ensemble learning. *Front Pharmacol*. 2022; 13: 897597.

6 Gao J, Niu R, Shi Y et al. The predictive value of [(18)F]FDG PET/CT radiomics combined with clinical features for EGFR mutation status in different clinical staging of lung adenocarcinoma. *EJNMMI Res*. 2023; 13(1): 26.

7 Huo JW, Luo TY, Diao L et al. Using combined CT-clinical radiomics models to identify epidermal growth factor receptor mutation subtypes in lung adenocarcinoma. *Front Oncol*. 2022; 12: 846589.

8 Jia TY, Xiong JF, Li XY et al. Identifying EGFR mutations in lung adenocarcinoma by noninvasive imaging using radiomics features and random forest modeling. *Eur Radiol*. 2019; 29(9): 4742-4750.

9 Jiang M, Yang P, Li J et al. Computed tomography-based radiomics quantification predicts epidermal growth factor receptor mutation status and

efficacy of first-line targeted therapy in lung adenocarcinoma. *Front Oncol*. 2022; 12: 985284.

10 Le NQK, Kha QH, Nguyen VH et al. Machine learning-based radiomics signatures for EGFR and KRAS mutations prediction in non-small-cell lung cancer. *Int J Mol Sci*. 2021; 22(17).

11 Li XY, Xiong JF, Jia TY et al. Detection of epithelial growth factor receptor (EGFR) mutations on CT images of patients with lung adenocarcinoma using radiomics and/or multi-level residual convolutionary neural networks. *J Thorac Dis*. 2018; 10(12): 6624-6635.

12 Li X, Yin G, Zhang Y et al. Predictive power of a radiomic signature based on (18)F-FDG PET/CT images for EGFR mutational status in NSCLC. *Front Oncol*. 2019; 9: 1062.

13 Li S, Luo T, Ding C et al. Detailed identification of epidermal growth factor receptor mutations in lung adenocarcinoma: Combining radiomics with machine learning. *Med Phys*. 2020; 47(8): 3458-3466.

14 Li S, Li Y, Zhao M et al. Combination of (18)F-fluorodeoxyglucose PET/CT radiomics and clinical features for predicting epidermal growth factor receptor mutations in lung adenocarcinoma. *Korean J Radiol*. 2022; 23(9): 921-930.

15 Liu G, Xu Z, Ge Y et al. 3D radiomics predicts EGFR mutation, exon-19 deletion and exon-21 L858R mutation in lung adenocarcinoma. *Transl Lung Cancer Res*. 2020; 9(4): 1212-1224.

16 Liu Y, Zhou J, Wu J et al. Development and validation of machine learning models to predict epidermal growth factor receptor mutation in non-small cell lung cancer: a multi-center retrospective radiomics study. *Cancer Control*. 2022; 29: 10732748221092926.

17 Lu X, Li M, Zhang H et al. A novel radiomic nomogram for predicting epidermal growth factor receptor mutation in peripheral lung adenocarcinoma. *Phys Med Biol*. 2020; 65(5): 055012.

18 Lu J, Ji X, Wang L et al. Machine learning-based radiomics for prediction of epidermal growth factor receptor mutations in lung adenocarcinoma. *Dis Markers*. 2022; 2022: 2056837.

19 Ma DN, Gao XY, Dan YB et al. Evaluating solid lung adenocarcinoma anaplastic lymphoma kinase gene rearrangement using noninvasive radiomics biomarkers. *Onco Targets Ther*. 2020; 13: 6927-6935.

20 Nair JKR, Saeed UA, McDougall CC et al. Radiogenomic models using machine learning techniques to predict EGFR mutations in non-small cell lung cancer. *Can Assoc Radiol J*. 2021; 72(1): 109-119.

21 Ninomiya K, Arimura H, Chan WY et al. Robust radiogenomics approach to the identification of EGFR mutations among patients with NSCLC from three different countries using topologically invariant Betti numbers. *PLoS One*. 2021; 16(1): e0244354.

22 Ninomiya K, Arimura H, Tanaka K et al. Three-dimensional topological radiogenomics of epidermal growth factor receptor Del19 and L858R mutation subtypes on computed tomography images of lung cancer patients. *Comput Methods Programs Biomed*. 2023; 236: 107544.

23 Rios Velazquez E, Parmar C, Liu Y et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res*. 2017; 77(14): 3922-3930.

24 Rossi G, Barabino E, Fedeli A et al. Radiomic detection of EGFR mutations in NSCLC. *Cancer Res*. 2021; 81(3): 724-731.

25 Song L, Zhu Z, Mao L et al. Clinical, conventional ct and radiomic feature-based machine learning models for predicting ALK rearrangement status in lung adenocarcinoma patients. *Front Oncol*. 2020; 10: 369.

26 Tu W, Sun G, Fan L 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.

27 Wang C, Ma J, Shao J et al. Predicting EGFR and PD-L1 status in NSCLC patients using multitask ai system based on CT images. *Front Immunol*. 2022; 13: 813072.

28 Wang J, Lv X, Huang W et al. Establishment and optimization of radiomics algorithms for prediction of KRAS gene mutation by integration of NSCLC gene mutation mutual exclusion information. *Front Pharmacol*. 2022; 13: 862581.

29 Weng Q, Hui J, Wang H et al. Radiomic feature-based nomogram: a novel technique to predict EGFR-activating mutations for EGFR tyrosin kinase inhibitor therapy. *Front Oncol*. 2021; 11: 590937.

30 Wu S, Shen G, Mao J et al. CT radiomics in predicting EGFR mutation in non-small cell lung cancer: a single institutional study. *Front Oncol*. 2020; 10: 542957.

31 Yang C, Chen W, Gong G et al. Application of CT radiomics features to predict the EGFR mutation status and therapeutic sensitivity to TKIs of advanced lung adenocarcinoma. *Transl Cancer Res*. 2020; 9(11): 6683-6690.

32 Yang X, Liu M, Ren Y et al. Using contrast-enhanced CT and non-contrast-enhanced CT to predict EGFR mutation status in NSCLC patients-a radiomics nomogram analysis. *Eur Radiol*. 2022; 32(4): 2693-2703.

33 Zhang L, Chen B, Liu X 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.

34 Zhang M, Bao Y, Rui W et al. Performance of (18)F-FDG PET/CT radiomics for predicting EGFR mutation status in patients with non-small cell lung cancer. *Front Oncol*. 2020; 10: 568857.

35 Zhang B, Qi S, Pan X et al. Deep CNN model using CT radiomics feature mapping recognizes EGFR gene mutation status of lung adenocarcinoma. *Front Oncol*. 2020; 10: 598721.

36 Zhang G, Cao Y, Zhang J et al. Predicting EGFR mutation status in lung adenocarcinoma: development and validation of a computed tomography-based radiomics signature. *Am J Cancer Res*. 2021; 11(2): 546-560.

37 Zhao HY, Su YX, Zhang LH et al. Prediction model based on 18F-FDG PET/CT radiomic features and clinical factors of EGFR mutations in lung adenocarcinoma. *Neoplasma*. 2022; 69(1): 233-241.

38 Zhu H, Song Y, Huang Z et al. Accurate prediction of epidermal growth factor receptor mutation status in early-stage lung adenocarcinoma, using radiomics and clinical features. *Asia Pac J Clin Oncol*. 2022; 18(6): 586-594.