

Predictive value of multiple imaging predictive models for spread through air spaces of lung adenocarcinoma: A systematic review and network meta-analysis

CONG LIU¹, YU-FENG WANG², PENG WANG², FENG GUO³, HONG-YING ZHAO⁴,
QIANG WANG⁴, ZHI-WEI SHI⁵ and XIAO-FENG LI⁵

¹Department of Minimally Invasive Oncology, Xuzhou New Health Hospital (The Xuzhou Hospital Affiliated to Jiangsu University); Departments of ²Nuclear Medicine, ³Medical Oncology, ⁴Radiotherapy and ⁵Radiology, Xuzhou Cancer Hospital (The Xuzhou Hospital Affiliated to Jiangsu University), Xuzhou, Jiangsu 221000, P.R. China

Received July 1, 2023; Accepted January 3, 2024

DOI: 10.3892/ol.2024.14255

Abstract. Spread Through Air Spaces (STAS) is involved in lung adenocarcinoma (LUAD) recurrence, where cancer cells spread into adjacent lung tissue, impacting surgical planning and prognosis assessment. Radiomics-based models show promise in predicting STAS preoperatively, enhancing surgical precision and prognostic evaluations. The present study performed network meta-analysis to assess the predictive efficacy of imaging models for STAS in LUAD. Data were systematically sourced from PubMed, Embase, Scopus, Wiley and Web of Science, according to the Cochrane Handbook for Systematic Reviews of Interventions and A Measurement Tool to Assess systematic Reviews 2. Using Stata software v17.0 for meta-analysis, surface under the cumulative ranking area (SUCRA) was applied to identify the most effective diagnostic method. Quality assessments were performed using Cochrane Collaboration's risk-of-bias tool and publication bias was assessed using Deeks' funnel plot. The analysis encompassed 14 articles, involving 3,734 patients, and assessed 17 predictive models for STAS in LUAD. According to comprehensive analysis of SUCRA, the machine learning (ML)_Peri_tumour model had the highest accuracy (56.5), the Features_computed tomography (CT) model had the highest sensitivity (51.9) and the positron emission tomography (pet)_CT model had the highest specificity (53.9). ML_Peritumour model had the highest predictive performance. The accuracy was as follows: ML_Peritumour vs. Features_CT [relative risk (RR)=1.14; 95% confidence interval (CI), 0.99-1.32]; ML_Peritumour

vs. ML_Tumour (RR=1.04; 95% CI, 0.83-1.30) and ML_Peritumour vs. pet_CT (RR=1.04; 95% CI, 0.84-1.29). Comparative analyses revealed heightened predictive accuracy of the ML_Peritumour compared with other models. Nonetheless, the field of radiological feature analysis for STAS prediction remains nascent, necessitating improvements in technical reproducibility and comprehensive model evaluation.

Introduction

Lung cancer, particularly lung adenocarcinoma (LUAD), is a major contributor to cancer-associated mortality worldwide, accounting for approximately 11.4% of all global cancer cases and 18.0% of cancer-related deaths. This prominence is potentially linked to its unique patterns of invasion. (1). Aside from infiltration of myofibroblast stroma, lymphovascular and pleural invasion, spread through air spaces (STAS) has emerged as an invasion pattern in LUAD (2). It was identified initially by Kadota *et al* (3) and recognized as a distinct form of tumor spread in the 2015 World Health Organization classification (3). STAS is characterized by presence of micropapillary clusters, solid nests or individual cells in lung parenchyma air spaces beyond the tumor margin (4). The current diagnostic methodology for STAS is analysis of pathological specimens obtained from lung tissues excised during surgical procedures in patients (5). It is found in 14.8-56.4% of LUAD cases and is associated with lower survival rates and a worse prognosis compared with STAS-negative tumors. Therefore, identification of STAS can provide key information for the clinical treatment of patients with LUAD (6,7). Reports indicate a significant risk of local and distant recurrence in STAS-positive cases treated with sublobar resection (3,8), whereas patients who undergo lobectomy have no increased recurrence risk. Thus, early detection of STAS is of clinical importance.

Radiomics, the conversion of radiographic images into quantifiable information, offers the potential to improve diagnosis, prognosis and the development of predictive models (9-11). Previous advancements in predicting STAS status in LUAD using radiomics methods reported promising results (12,13).

Correspondence to: Professor Xiao-Feng Li, Department of Radiology, Xuzhou Cancer Hospital (The Xuzhou Hospital Affiliated to Jiangsu University), 131 Huancheng Road, Xuzhou, Jiangsu 221000, P.R. China
E-mail: lxf5818@163.com

Key words: predictive model, lung adenocarcinoma, spread through air spaces, network meta-analysis

However, bridging the gap between radiomics as a research tool and its clinical implementation presents challenges, including technical reproducibility, clinical validity, quantification and cost-effectiveness. There is also notable heterogeneity in previous studies, with lack of comprehensive evaluation of the performance of radiomics in predicting STAS in LUAD (14). Identifying factors affecting the predictive performance of radiomics is key for its clinical use. Several radiomics models employing computed tomography (CT), magnetic resonance imaging and positron emission tomography (PET)/CT have been developed for predicting STAS, showing diverse performance and indicating methodological variability (15,16). However, to date, there are no relevant network meta-analyses to evaluate the predictive value of these models, to the best of our knowledge. Therefore, the present study aimed to assess the risk of bias and methodological quality and to perform a network meta-analysis (NMA) to evaluate the effectiveness of radiomics models in predicting preoperative STAS in LUAD. This may be valuable for clinicians, radiologists and researchers in the field of LUAD diagnosis and treatment.

Materials and methods

Protocol and registration. The present review was performed in accordance with AMSTAR 2 (17). The methods and protocol for the present study were pre-registered, in accordance with standard procedures, in the International Platform of Registered Systematic Review and Meta-analysis Protocols (registration no. 202390105; DOI: 10.37766/inplasy2023.9.0105).

Retrieval strategy. A comprehensive literature search was performed using the following key terms: 'Risk factor', 'predictive', 'spread through air spaces', 'lung adenocarcinoma' and 'nomograms'. This search used the PubMed (pubmed.ncbi.nlm.nih.gov/), Embase (embase.com/), Scopus (https://www.scopus.com/), Wiley (https://onlinelibrary.wiley.com/) and Web of Science (https://www.webofscience.com/wos/) databases, with a cut-off date of May 1, 2023. The references of included studies were also systematically reviewed to obtain potentially relevant publications (Table I).

Inclusion and exclusion criteria. The inclusion criteria included the following: i) Study focuses on patients who have been diagnosed with LUAD and who exhibit STAS; ii) objective of the study is to develop a predictive model to accurately identify the presence of STAS in patients with LUAD. Tumor STAS was defined as tumor cells (micropapillary structures, solid nests, or single cells-spreading within air spaces in the lung parenchyma beyond the edge of the main tumor (3). The present study selected studies that included patients who underwent segmental or lobar resection.

The exclusion criteria were as follows: i) Predictive models that were not constructed based on radiological features; ii) no clear inclusion and exclusion criteria; iii) reviews and lecture-type literature; iv) literature for which the full text could not be obtained and v) literature for which data could not be extracted.

Literature screening. The initial screening of titles and abstracts was independently conducted by two researchers (CL and PW)

using Cochrane handbook's guidelines for systematic reviews of interventions (18), adhering to the predefined inclusion and exclusion criteria. Discrepancies or uncertainty about article inclusion were resolved through discussion or consultation with a third reviewer (XL).

Quality evaluation of literature. The quality of articles was assessed by two independent reviewers using Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) (19), which is a tool for assessing quality of diagnostic studies, focusing on 'risk-of-bias' and 'applicability concerns'. The risk-of-bias was assessed across four domains: Patient selection, index test, reference standard and flow and timing. The applicability was evaluated for the first three domains and rated as 'yes', 'no' or 'unclear', with 'yes' denoting low risk, 'no' indicating high risk and 'unclear' suggesting insufficient information. In cases of disagreement, a third reviewer was used for resolution. A Measurement Tool to Assess systematic Reviews was used for a stringent quality assessment (14).

Additionally, the methodological quality of studies was appraised using the Cochrane Handbook's risk-of-bias assessment tool (RevMan v.5.3.5, The Cochrane Collaboration) (20), covering six aspects (selection, performance bias, detection, attrition, reporting bias and other bias), which were categorized as 'yes', 'no' or 'unclear' to indicate the level of bias.

Data extraction. The data extracted primarily encompassed the following aspects: i) Characteristics of the included literature, such as author information, publication date, country of origin, predictive models and regression methods employed and predictive factors investigated; ii) details of the study subjects, such as sample size, sex distribution and tumor stage (according to the 8th edition of the AJCC staging standards) (21), with all participants having undergone surgery and iii) evaluation of effect indicators.

Statistical analysis. The effectiveness of various predictive models were evaluated based on their accuracy, sensitivity (SEN), and specificity (SPE). Predictive models were categorized according to their unique features for NMA to assess performance in predicting STAS. This NMA, conducted using Stata software (version 17.0; StataCorp LP) within a Bayesian framework, used the Markov Chain Monte Carlo Subset Simulation (22) in accordance with the PRISMA NMA guidelines (23). A nodal approach for quantifying and clarifying concordance between direct and indirect comparisons was adopted. The consistency criterion for the NMA was $P > 0.05$. Network diagrams visually represented diagnostic methods, with nodes symbolizing each method and lines representing direct comparisons. The size of nodes and thickness of lines corresponded to the number of studies. To detect possible publication bias in selected studies, funnel plots were constructed for each measure of diagnostic efficiency, employing symmetry criteria as a key validation technique. Statistical heterogeneity was evaluated using I^2 statistic, a measure in meta-analytical methods. This quantifies the proportion of the total variation in study estimates due to heterogeneity rather than chance. An I^2 value of 0% indicates no observed heterogeneity, whilst higher values suggest increasing heterogeneity, with guidelines typically considering 25, 50 and 75% as low, moderate and

Table I. Search strategy.

Database	Search query	Number of results
PubMed	(spread through air spaces [Mesh] OR STAS [Mesh] OR spread through air spaces [Title/Abstract] OR STAS [Title/Abstract]) AND (Lung cancer [Mesh] OR Lung adenocarcinoma [Mesh] OR Adenocarcinoma of Lung [Mesh] OR Lung cancer [Title/Abstract] OR Lung adenocarcinoma [Title/Abstract] OR Adenocarcinoma of Lung [Title/Abstract]) AND (Risk factor [Mesh] OR Prediction [Mesh] OR Nomograms [Mesh] OR Risk factor [Title/Abstract] OR Prediction [Title/Abstract] OR Nomograms [Title/Abstract])	64
Embase	((STAS)/br OR (('spread through air spaces'):ti)) AND ((Adenocarcinoma of Lung)/br OR ((Lung adenocarcinoma)/br) OR ((Lung cancer)/br)) AND ((prediction)/br OR ((Risk factor)/br) OR ((Nomograms)/br))	120
Scopus	(TITLE-ABS-KEY (stas) OR TITLE-ABS-KEY (spread AND through AND air AND spaces) AND (TITLE-ABS-KEY (lung AND cancer) OR TITLE-ABS-KEY (Adenocarcinoma AND of AND Lung) OR TITLE-ABS-KEY (Lung AND adenocarcinoma)) AND (TITLE-ABS-KEY (risk AND factor) OR TITLE-ABS-KEY (Prediction) OR TITLE-ABS-KEY (Nomograms))	81
Wiley	'STAS OR spread through air spaces' anywhere and 'Lung cancer OR Lung adenocarcinoma OR Adenocarcinoma of Lung' anywhere and 'prediction OR Risk factor OR Nomograms' anywhere	177
Web of science	((TS=(spread through air spaces)) OR TS=(STAS) OR TI=(STAS) OR AB=(STAS)) AND (TS=(Lung cancer) OR TS=(Adenocarcinoma of Lung) OR TS=(Lung adenocarcinoma) OR TI=(Lung adenocarcinoma)) AND (TS=(Prediction) OR TS=(Risk factor) OR TS=(Nomograms))	112

STAS, spread through air spaces.

high heterogeneity, respectively (18). Additionally, to ascertain the relative superiority of one method, the level of certainty for predictive models was quantified. This assessment was performed using surface under the cumulative ranking curve (SUCRA), forest plots and league tables.

Subgroup diagnostic meta-analyses were performed using Stata to assess relative predictive efficiency of composite models. The effectiveness of these models was evaluated using the area under the curve (AUC) derived from the summary receiver operating characteristics (sROC). Additionally, the Fagan plot was utilized to quantify the overall discriminatory power of a diagnostic test (24).

Results

Selection and characteristics of literature. Literature review was performed using 565 articles. Subsequent to this, a meticulous screening process was undertaken to ensure the relevance and quality of sources. This entailed the removal of 249 articles due to duplication. Further scrutiny, focusing on titles and abstracts, led to the exclusion of an additional 288 articles that were not pertinent. The remaining pool of 28 articles was subjected to a more rigorous evaluation, which included accessibility of full-text versions and the feasibility of data extraction. This process led to the disqualification of 14 articles, leaving a final count of 14 articles (Fig. 1).

These 14 articles, collectively encompassing data from 3,734 participants, were exclusively focused on patients diagnosed with STAS in LUAD. The predictive models were categorized into the following four distinct types based on their methodological approaches and radiological characteristics: i) Models developed using logistic regression analysis to screen CT features (Features_CT); ii) models using machine learning (ML) techniques to screen tumor radiological characteristics (ML_Tumour); iii) models applying ML for the screening of both tumor and peritumor radiological features (ML_Peritumour) and iv) models that used logistic regression analysis for screening of PET/CT features (pet_CT).

In the process of tumor and peritumor segmentation, the open-source software 3D Slicer v4.8.1 (slicer.org) was used. Moreover, all studies used pathological findings as a benchmark, forming a control group against which predictive models were evaluated. The data enabling direct comparative analysis in outcomes were also assessed (Table II) (5,15,16,25-35).

Quality assessment and publication bias. In the present study involving 14 articles and 17 predictive models, NMA was performed using Stata and the QUADAS-2 tool was used to assess quality, risk of bias and applicability of the articles. The inter-rater reliability (κ -agreement) between the reviewers was 0.87. A high risk of bias was detected in a few articles (3/14) in terms of patient selection (1/14) and reference standards

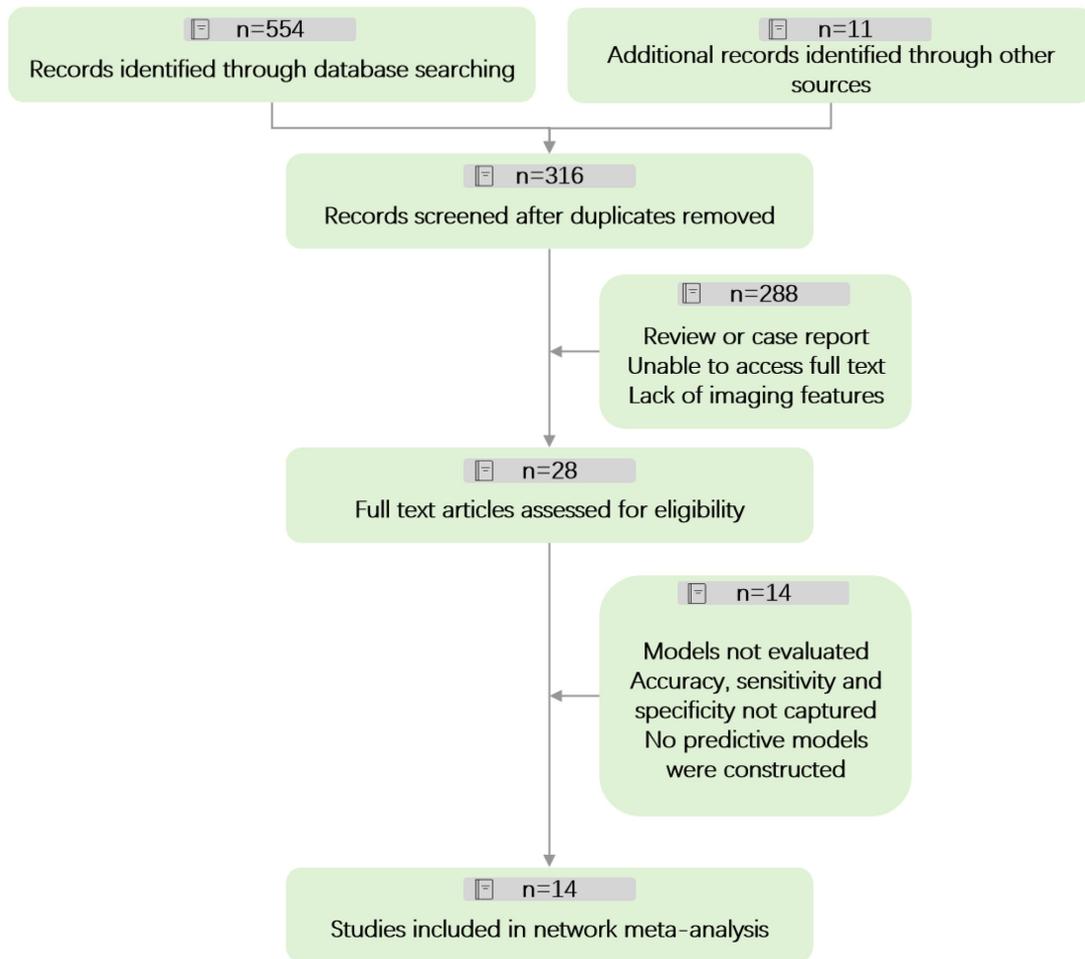


Figure 1. Comprehensive flow diagram of the literature selection process. The flowchart represents systematic screening and selection methodology in compliance with established meta-analysis protocols. The number inside each box reflects the cumulative count of studies at each sequential stage.

(2/14); however, the overall quality of the publications was satisfactory (Fig. 2).

NMA. NMA evaluated the relative risk (RR) values and 95% confidence intervals (CI) across different predictive models in terms of accuracy, SEN and SPE for STAS in LUAD.

Pairwise meta-analysis. NMA graph illustrates the comparative accuracy, SEN and SPE of predictive models (Fig. 3). Notably, CT_feature model group encompassed the largest sample size, followed by the ML_Peritumour model. Specifically, two studies directly compared the CT_feature and ML_Peritumour model, and one study contrasted the CT_feature model with ML_Tumour model (Fig. 3). Furthermore, the comprehensive evaluation of the included studies spanned all domains. Potential publication bias was assessed using funnel plots (Fig. 4). The roughly symmetric distribution suggested a negligible presence of publication bias or other forms of bias within the studies. This symmetry bolstered the reliability of findings.

Accuracy. Using the SUCRA, the accuracy of several predictive models for STAS was evaluated. Models ranked in descending order of accuracy were as follows: Control (100.0%); ML_Peritumour (56.5%); ML_Tumour (41.8%); pet_CT (41.4%)

and Features_CT (10.3%; Fig. 5A). A detailed two-by-two comparative analysis is presented in Table IIIA, highlighting the predictive efficacy of these models. ML_Peritumour model demonstrated superior accuracy, particularly compared with Features_CT (RR=1.14; 95% CI, 0.99-1.32), ML_Tumour (RR=1.04; 95% CI, 0.83-1.30) and pet_CT (RR=1.04; 95% CI, 0.84-1.29). A heterogeneity test revealed I^2 value of 20.4%. Consistently, the forest plot demonstrated the highest predictive accuracy for ML_Peritumour model (Fig. 6A).

SEN. SEN for different predictive models for STAS, derived from the SUCRA was as follows: Control (99.9%); Features_CT (51.9%); ML_Peritumour (49.9%); ML_Tumour (42.8%); and pet_CT (5.5%; Fig. 5B). Table IIIB shows a comparative league table for a two-by-two analysis of these models. Features_CT model exhibited superior SEN, especially compared with ML_Peritumour (RR=1.00; 95% CI, 0.89-1.13), ML_Tumour (RR=1.02; 95% CI, 0.88-1.18) and pet_CT (RR=1.17; 95% CI, 0.97-1.40). The heterogeneity test indicated an I^2 of 17.8%. Additionally, forest plot highlighted the superior predictive SEN of the Features_CT model (Fig. 6B).

SPE. SPE of different predictive models for STAS, ascertained using the SUCRA, was as follows: Control (99.7%); pet_CT (53.9%); ML_Peritumour (48.0%); ML_Tumour (42.7%); and

Table II. Characteristics of studies included in the meta-analysis.

First author/s, year	Country of origin	Number of patients	Sex (male/female)	Tumor stage	STAS (+)	Predictive model	Regression method	(Refs.)
Bassi <i>et al.</i> , 2022	Italy	149	85/64	I-III	98	CT features Tumor radiomics	Logistic ML	(25)
Qi <i>et al.</i> , 2021	China	216	160/56	I-III	56	CT features Peritumoral and tumoral radiomic features	Logistic ML	(26)
Liao <i>et al.</i> , 2022	China	256	122/134	I	85	Peritumoral and tumoral radiomic features	ML	(5)
Chen <i>et al.</i> , 2022	China	327	131/196	I-III	113	Peritumoral and tumoral radiomic features	ML	(27)
Kim <i>et al.</i> , 2018	Korea	276	129/147	I-III	92	CT features	Logistic	(28)
Qin <i>et al.</i> , 2022	China	503	201/302	I-III	241	CT features	Logistic	(29)
Chen <i>et al.</i> , 2022	China	85	Unknown	I-III	13	CT features	Logistic	(15)
Qi <i>et al.</i> , 2020	China	190	112/78	I	47	CT features	Logistic	(30)
Zhang <i>et al.</i> , 2020	China	762	276/486	I	83	CT features	Logistic	(31)
Han <i>et al.</i> , 2022	China	395	207/188	I	169	Tumor radiomics	ML	(32)
Takehana <i>et al.</i> , 2022	Japan	339	160/179	I	95	CT features Peritumoral and tumoral radiomic features	Logistic ML	(33)
Wang <i>et al.</i> , 2020	China	121	60/67	I	51	18F FDG-PET/CT	Logistic	(34)
Nishimori <i>et al.</i> , 2022	Japan	52	22/30	I	19	18F FDG-PET/CT	Logistic	(16)
Falay <i>et al.</i> , 2021	Turkey	63	41/22	I-III	33	18F FDG-PET/CT	Logistic	(35)

STAS, spread through air spaces; CT, computed tomography; 18F FDG-PET/CT, 2-deoxy-2-[fluorine-18]fluoro-D-glucose-positron emission tomography/CT; ML, machine learning.

Table III. League tables for predictive models.

A, Accuracy: $I^2=20.4%$; random effects results.

pet_CT	1.00 (0.78,1.29)	1.04 (0.84,1.29)	0.92 (0.75,1.11)	1.38 (1.16,1.64)
1.00 (0.78,1.28)	ML_Tumour	1.04 (0.83,1.30)	0.91 (0.75,1.11)	1.38 (1.15,1.66)
0.96 (0.77,1.19)	0.96 (0.77,1.20)	ML_Per_i_tumour	0.88 (0.76,1.01)	1.32 (1.16,1.50)
1.09 (0.90,1.33)	1.09 (0.90,1.33)	1.14 (0.99,1.32)	Features_CT	1.51 (1.37,1.66)
0.72 (0.61,0.86)	0.73 (0.60,0.87)	0.76 (0.67,0.86)	0.66 (0.60,0.73)	Control

B, SEN: $I^2=17.8%$; random effects results.

pet_CT	1.14 (0.92,1.42)	1.16 (0.96,1.41)	1.17 (0.97,1.40)	1.37 (1.16,1.61)
0.87 (0.70,1.08)	ML_Tumour	1.02 (0.86,1.20)	1.02 (0.88,1.18)	1.20 (1.04,1.37)
0.86 (0.71,1.04)	0.98 (0.83,1.16)	ML_Per_i_tumour	1.00 (0.89,1.13)	1.18 (1.06,1.30)
0.86 (0.71,1.03)	0.98 (0.85,1.13)	1.00 (0.89,1.12)	Features_CT	1.17 (1.09,1.26)
0.73 (0.62,0.86)	0.84 (0.73,0.96)	0.85 (0.77,0.94)	0.85 (0.79,0.92)	Control

C, SPE: $I^2=9.1%$; random effects results.

pet_CT	0.95 (0.67,1.35)	0.97 (0.72,1.29)	0.81 (0.62,1.06)	1.34 (1.06,1.70)
1.05 (0.74,1.50)	ML_Tumour	1.02 (0.75,1.39)	0.85 (0.64,1.13)	1.41 (1.09,1.84)
1.03 (0.77,1.39)	0.98 (0.72,1.34)	ML_Per_i_tumour	0.83 (0.69,1.01)	1.39 (1.17,1.65)
1.24 (0.95,1.62)	1.18 (0.89,1.56)	1.20 (0.99,1.46)	Features_CT	1.66 (1.46,1.90)
0.75 (0.59,0.94)	0.71 (0.54,0.92)	0.72 (0.61,0.86)	0.60 (0.53,0.68)	Control

SEN, sensitivity; SPE, specificity; CI, confidence interval; ML, machine learning; CT, computed tomography; pet, positron emission tomography.

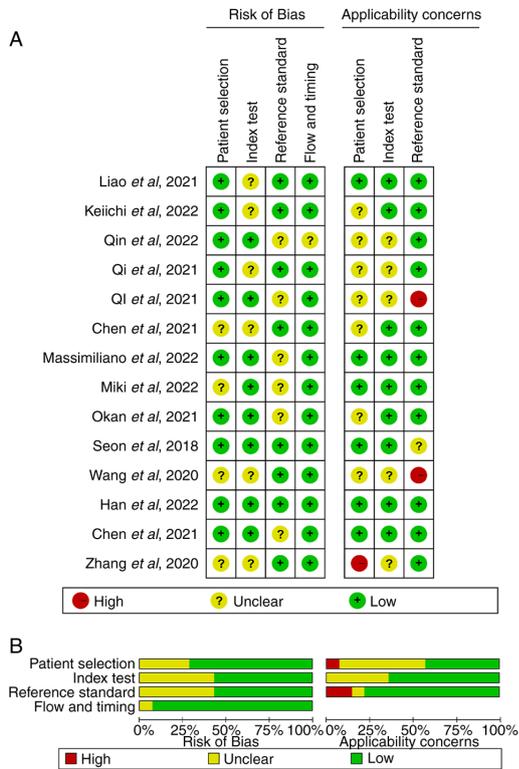


Figure 2. Bias risk in included studies based on Quality Assessment of Diagnostic Accuracy Studies-2 criteria. (A) Risk of bias assessment. (B) Quality assessment graph includes the risk of bias graph, the % of each rank in the quality assessment graph.

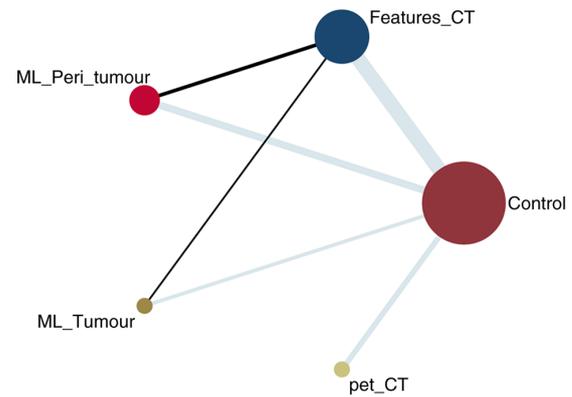


Figure 3. Network diagram of studies included in meta-analysis. The nodes represent distinct categories within the meta-analysis framework, with the size of each node proportional to the amount of data or number of studies. The edges indicate the strength of the interactions, with thickness representing strength of the evidence supporting the interaction between categories. ML_Per_i_tumour model was constructed by ML screening of tumor and peritumor radiological features. Features_CT model was constructed by screening CT features through logistic regression analysis. The control was established based on pathological findings utilized as the gold standard for comparison and validation. ML_Tumour model was constructed by screening tumor radiological features through ML. pet_CT model was constructed by screening PET/CT features by logistic regression analysis. ML, machine learning; CT, computed tomography; PET, positron emission tomography.

Features_CT (5.7%; Fig. 5C). A comprehensive league table in Table IIIC compares these models in a two-by-two format.

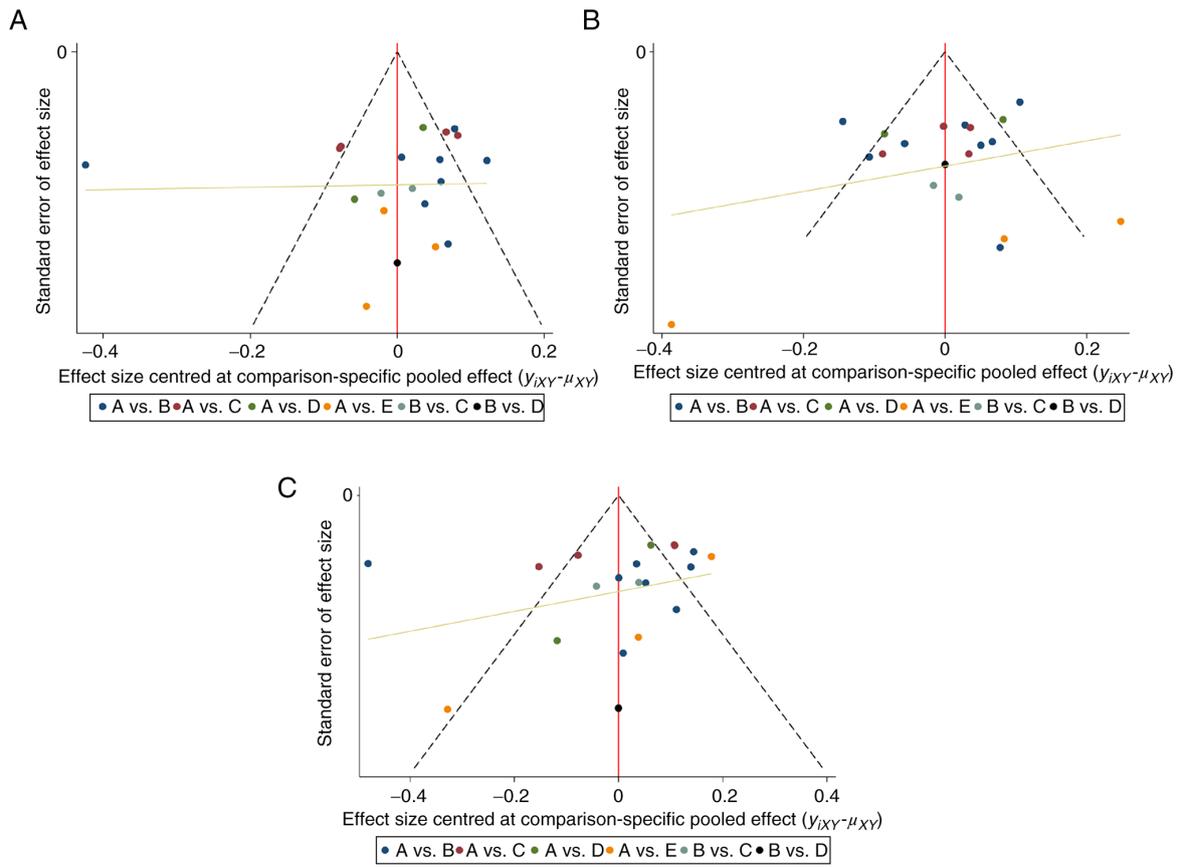


Figure 4. Funnel plots of network meta-analysis comparing predictive models. Plots for (A) accuracy, (B) sensitivity and (C) specificity represent predictive values within the network meta-analysis. Each data point represents a study. A, control; B, Features_CT; C, ML_Peritumour; D, ML_Tumour and E, pet_CT. ML, machine learning; CT, computed tomography; pet, positron emission tomography.

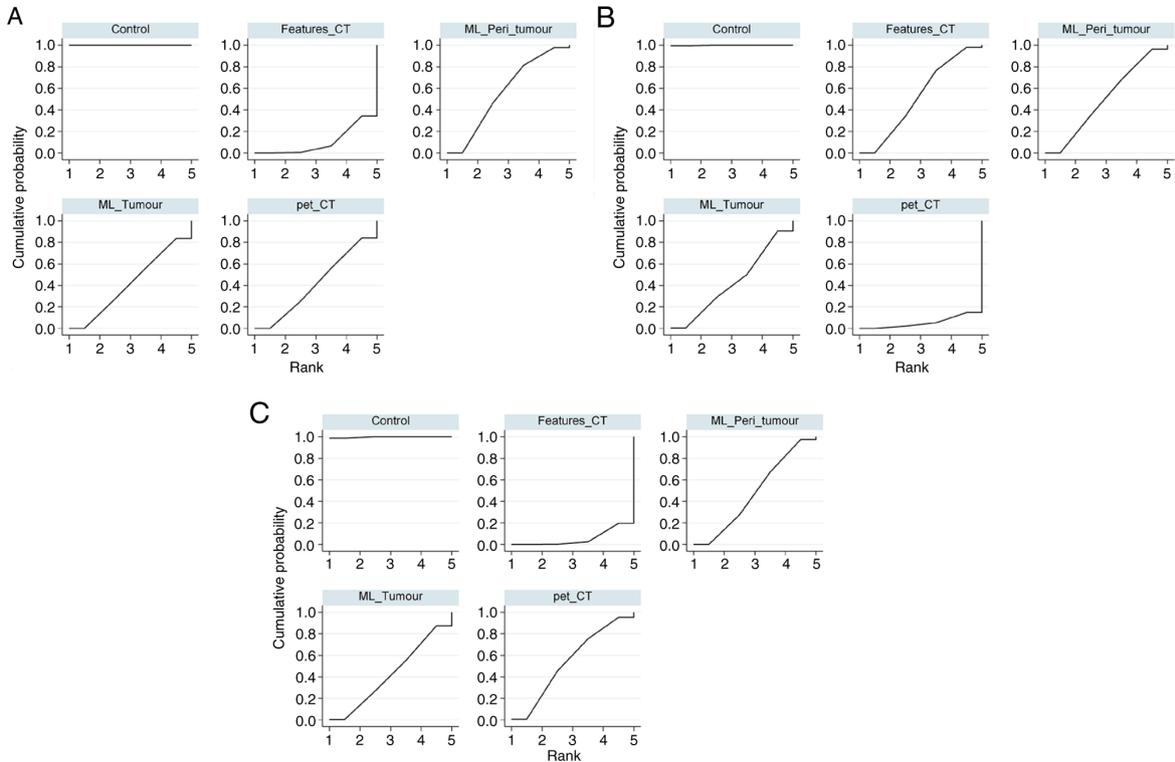


Figure 5. Surface under the cumulative ranking curve for predictive models. Plots for (A) accuracy, (B) sensitivity and (C) specificity display cumulative probability distributions for treatment in a rank-ordered fashion, with each panel corresponding to a unique treatment scenario. ML, machine learning; CT, computed tomography; pet, positron emission tomography.

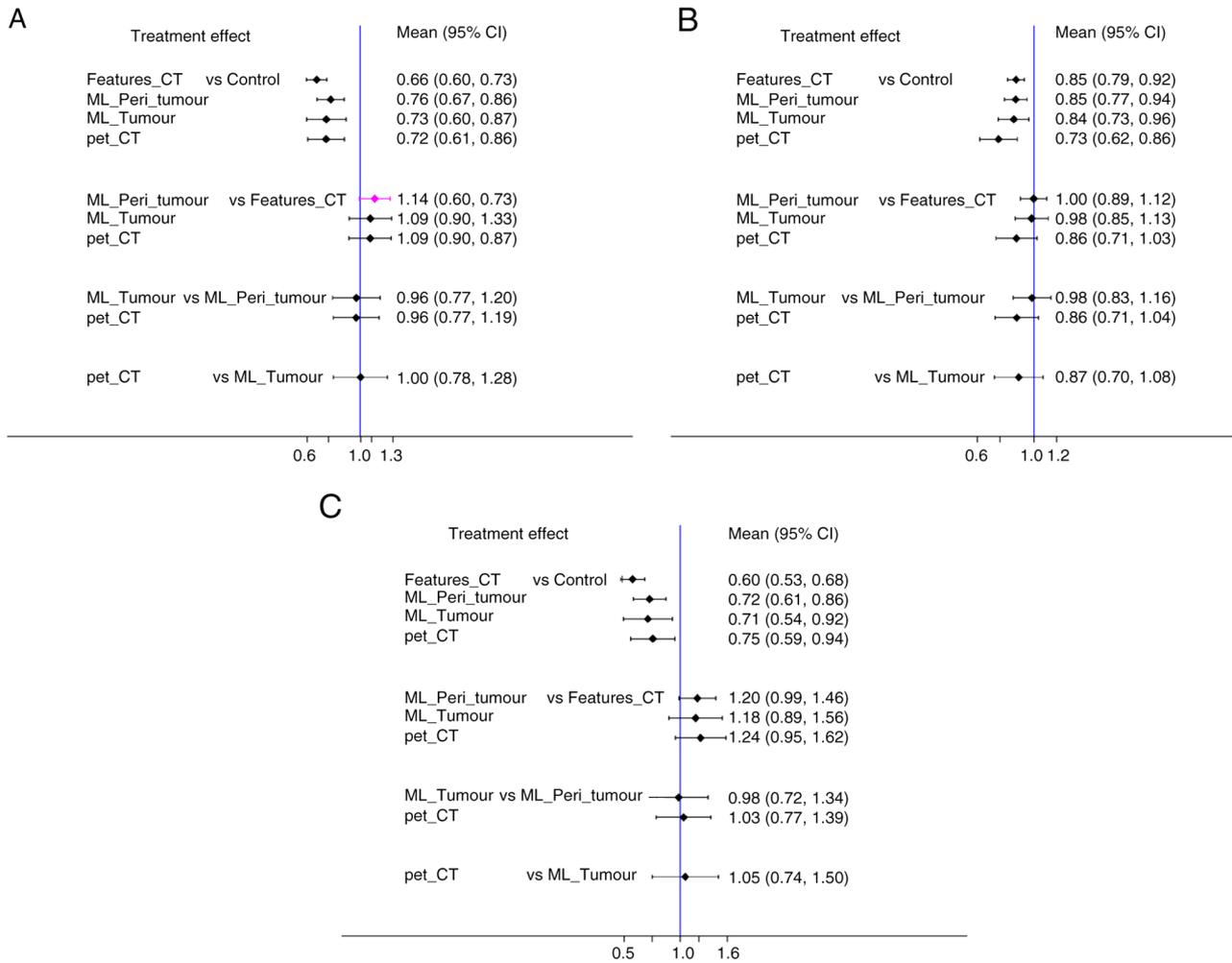


Figure 6. Forest plots for predictive models. Plots for (A) accuracy, (B) sensitivity and (C) specificity present pairwise meta-analysis results comparing the efficacy of imaging modalities. Features_CT, ML_Tumour, ML_Peritumour and pet_CT were compared with each other and a control group. ML, machine learning; CT, computed tomography; pet, positron emission tomography; CI, confidence interval.

The pet_CT model showed enhanced SPE, particularly against ML_Peritumour (RR=1.03; 95% CI, 0.77-1.39), ML_Tumour (RR=1.05; 95% CI, 0.74-1.50) and Features_CT (RR=1.24; 95% CI, 0.95-1.62). The heterogeneity test yielded I^2 of 9.1%. The forest plot indicated the superior predictive SPE of the pet_CT model (Fig. 6C).

Subgroup diagnostic MA. Diagnostic MA scrutinized the predictive capabilities of the ML_Peritumour and Features_CT models. AUC of the sROC for the ML_Peritumour model was 0.86 (95% CI, 0.82-0.88), while for the Features_CT model it was 0.81 (95% CI, 0.77-0.84; Fig. 7A and B, respectively). Fagan plot analysis, which assessed the predictive potency of models, demonstrated the relative superiority of the ML_Peritumour model (Fig. 7C and D).

Discussion

The present MA evaluated predictive accuracy of several models for STAS in LUAD. Analyzing 14 studies encompassing 3,734 patients, four predictive models were assessed. Among these, the ML_Peritumour model, using ML to analyze tumor and peritumor radiographic features, was the

most effective. This model demonstrated superior performance in accuracy, SEN and SPE, evidenced by its SUCRA values of 56.5, 49.9 and 48.0, respectively. Furthermore, a diagnostic MA supported the efficacy of the ML_Peritumour model, indicating a pooled AUC of 0.86 (95% CI, 0.82-0.88).

Previous studies have substantially deepened understanding of STAS in LUAD, especially regarding its prediction via radiological and pathological features (36,37). Investigations into predictive CT characteristics for STAS in small-sized LUAD have reported that attributes such as consolidation tumor ratio, spiculation, satellites, ground glass ribbon sign, pleural attachment and unclear tumor-lung interface are effective predictors of STAS (30,38). This aligns with the present accuracy of the ML_Peritumour and other ML-based models, underscoring the significance of CT features in these models.

The evolving understanding of the association between tumor stromal cells and STAS, along with the role of stromal cells in STAS pathogenesis, is noteworthy. Advanced medical information technology, including three-dimensional space convolution and fuzzy neural networks, has demonstrated potential in enhancing diagnostic SEN and SPE for lung cancer, suggesting promising avenues for future STAS prediction models (39).

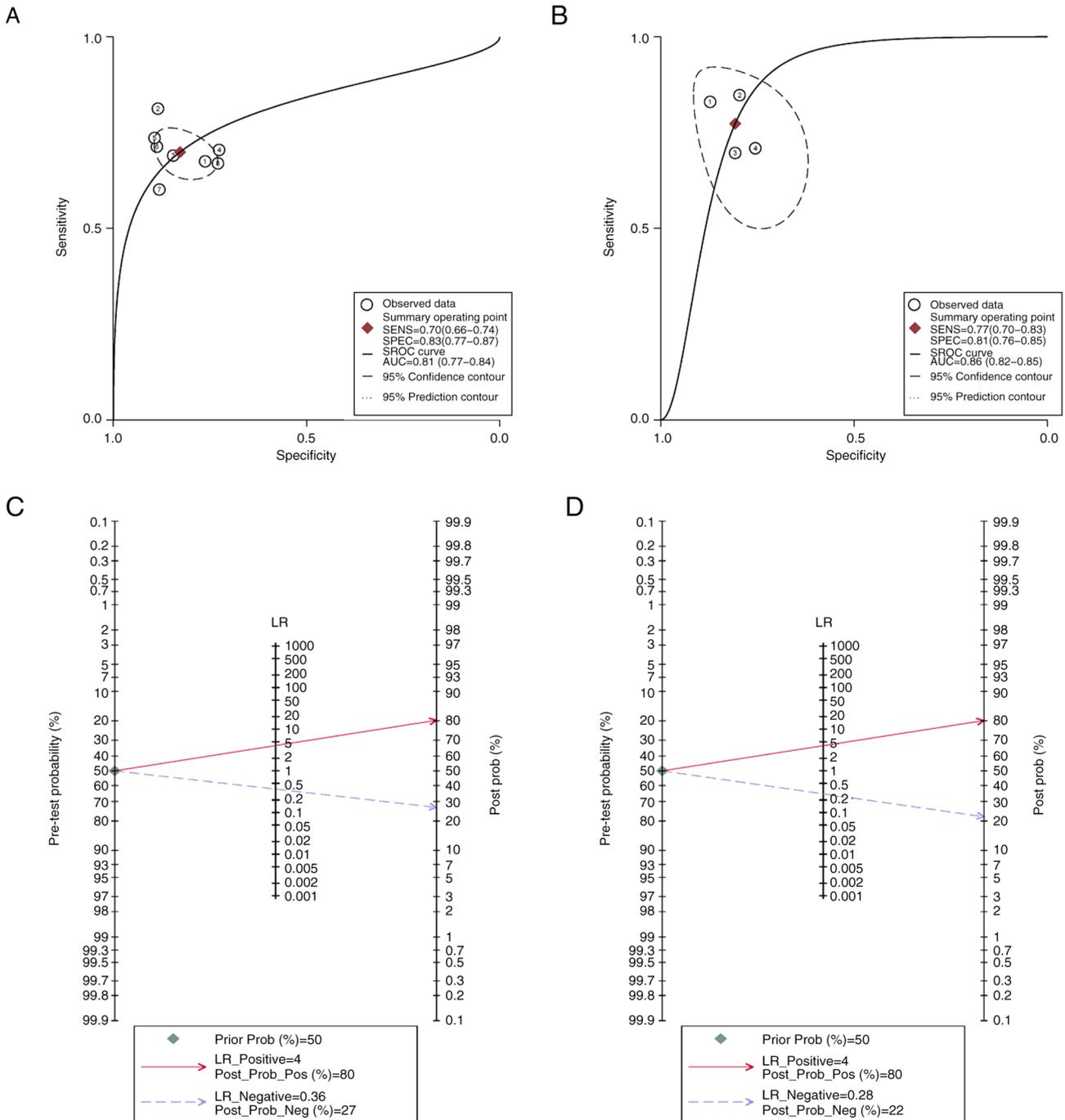


Figure 7. Subgroup diagnostic meta-analyses for predictive models. (A) Meta-analysis of Diagnostic Subgroups based on Features_CT models: AUC of SROC was 0.81 (0.77-0.84); (B) Meta-analysis of Diagnostic Subgroups based on ML_Peritumour models: AUC of SROC was 0.86 (0.82-0.85); (C) Fagan's nomogram for the assessment of post prob (intersection of pre-test prob and LR) based on Features_CT models; (D) Fagan's nomogram for the assessment of post prob (intersection of pre-test prob and LR) based on ML_Peritumour models. SROC, summary receiver operating characteristic; AUC, area under the curve; SENS, sensitivity; SPEC, specificity; LR, likelihood ratio; Prob, probability; Post, post-test; Pos, positive; Neg, negative.

Another notable development is the association between fluorodeoxyglucose (FDG) metabolic tumor burden, measured by PET/CT and STAS. Studies using PET/CT metrics such as standardized uptake value and total lesion glycolysis have reported that LUAD with low FDG uptake is associated with a lower incidence of STAS, whilst subtypes with higher FDG uptake, such as solid predominant adenocarcinoma, show a higher incidence of STAS (34,35). Furthermore, integration of ML techniques for analyzing radiological data for tumor

and peritumor features has resulted in models with improved predictive accuracy for STAS. A study by Liao *et al* (5) involving 256 patients, integrated tumor radiomic signature (TRS) with peritumoral radiomic signatures (PRS) and developed an effective gross radiomic signature model. Particularly, TRS combined with the PRS-15 mm model exhibited substantial predictive accuracy, achieving an AUC of 0.854 in the development and 0.870 in the validation cohort. The focus on the peritumoral environment represents

a notable advancement over prior research (28), which predominantly concentrated on the primary tumor alone. The success of the ML_Peritumour model in the present study highlights the potential of merging radiomic features from both tumor and peritumor regions, providing a more comprehensive approach to STAS prediction. This is relevant since STAS, typically found at the tumor periphery, may be more accurately predicted using preoperative CT images of tumor margins (26,33).

The present analysis revealed that the majority of radiomics studies on STAS prediction were in early or intermediate stages of research. The rigorous design of these studies is vital for validating the feasibility of radiological approaches. The present study identified limitations, including a lack of reproducibility analysis, internal validation and comprehensive performance evaluation of models. Notably, none of the included studies performed phantom or test-retest analyses for validating feature robustness (40,41) and only three studies addressed calibration, which is a key metric for evaluating prediction consistency with actual outcomes (5,15,32). To advance clinical application and practicality of radiomics, attention must be paid to external validation, cost-effectiveness and availability of open data. Validation with data from other institutions or different time periods is key for confirming model generalizability (42). However, only one study in the present analysis validated radiomic signatures with external data (15). Furthermore, a lack of open data and code availability, essential for assessing reproducibility, was a common limitation across studies (43). In the present NMA, included studies encompassed two different regression methods for constructing predictive models: ML (6/14) and binary logistic regression (8/14). ML focuses on the accuracy of the final model, while binary logistic regression also pays attention to metrics such as the odds ratio for each variable (44). The present research did not reveal any heterogeneity between different types of regression method. However, the comprehensive analysis indicated that the ML_Peritumour model held greater value in predicting STAS, potentially due to its consideration of radiomic signatures in the peritumoral region. Nevertheless, use of models developed using diverse regression methods is a limitation of the present study. To enhance robustness and comprehensiveness of results of the present study, the incorporation of additional studies for detailed subgroup analyses is key. Additionally, 8/14 studies reviewed focused on the occurrence of STAS in patients with stage I LUAD. However, the MA did not demonstrate inter-study heterogeneity in this aspect. There is need for more research to evaluate the predictive efficiency of radiomic models for STAS in early-stage LUAD, where accurate diagnosis is of paramount importance. Assessing STAS status prior to developing a surgical plan is key, as it significantly influences the selection of surgical strategies.

In conclusion, the present study is the first NMA to integrate several predictive models, to the best of our knowledge. The findings underscore the superior predictive efficacy of tumor and peritumor-based radiological models. Nonetheless, research in radiological features for STAS prediction is in its early stages, and significant enhancements are needed, particularly in technical reproducibility and comprehensive model evaluation. The reliability of these studies requires

experimental verification due to limited external validation. Additionally, the scarcity of direct model comparisons in the analyzed studies, primarily relying on indirect comparisons, may affect quality assessment results, underscoring the need for more direct model comparisons in future. The ethnographic and geographic applicability of these findings, primarily contributed to by researchers in Asia, also needs further validation.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Clinical Medicine Science and Technology Development Foundation of Jiangsu University (grant no. JLY2021082) and Xuzhou Science and Technology Bureau (grant no. KC23229).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

XL designed the study and revised the manuscript. CL performed statistical analysis and wrote the manuscript. CL and PW performed the literature review and quality assessment. PW, YW, FG and ZS interpreted the data. CL and XL confirm the authenticity of all the raw data. HZ and QW were involved in the study design and critically revised the manuscript. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Zhuo Y, Feng M, Yang S, Zhou L, Ge D, Lu S, Liu L, Shan F and Zhang Z: Radiomics nomograms of tumors and peritumoral regions for the preoperative prediction of spread through air spaces in lung adenocarcinoma. *Transl Oncol* 13: 100820, 2020.
2. Blaauwgeers H, Flieder D, Warth A, Harms A, Monkhurst K, Witte B and Thunnissen E: A prospective study of loose tissue fragments in non-small cell lung cancer resection specimens: An alternative view to 'spread through air spaces'. *Am J Surg Pathol* 41: 1226-1230, 2017.
3. Kadota K, Nitadori JI, Sima CS, Ujiie H, Rizk NP, Jones DR, Adusumilli PS and Travis WD: Tumor spread through air spaces is an important pattern of invasion and impacts the frequency and location of recurrences after limited resection for small stage I lung adenocarcinomas. *J Thorac Oncol* 10: 806-814, 2015.

4. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, *et al*: The 2015 world health organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol* 10: 1243-1260, 2015.
5. Liao G, Huang L, Wu S, Zhang P, Xie D, Yao L, Zhang Z, Yao S, Shanshan L, Wang S, *et al*: Preoperative CT-based peritumoral and tumoral radiomic features prediction for tumor spread through air spaces in clinical stage I lung adenocarcinoma. *Lung Cancer* 163: 87-95, 2022.
6. Terada Y, Takahashi T, Morita S, Kashiwabara K, Nagayama K, Nitadori JI, Anraku M, Sato M, Shinozaki-Ushiku A and Nakajima J: Spread through air spaces is an independent predictor of recurrence in stage III (N2) lung adenocarcinoma. *Interact Cardiovasc Thorac Surg* 29: 442-448, 2019.
7. Niu Y, Han X, Zeng Y, Nanding A, Bai Q, Guo S, Hou Y, Yu Y, Zhang Q and Li X: The significance of spread through air spaces in the prognostic assessment model of stage I lung adenocarcinoma and the exploration of its invasion mechanism. *J Cancer Res Clin Oncol* 149: 7125-7138, 2023.
8. Eguchi T, Kameda K, Lu S, Bott MJ, Tan KS, Montecalvo J, Chang JC, Rekhtman N, Jones DR, Travis WD and Adusumilli PS: Lobectomy is associated with better outcomes than sublobar resection in spread through air spaces (STAS)-positive T1 lung adenocarcinoma: A propensity score-matched analysis. *J Thorac Oncol* 14: 87-98, 2019.
9. Ji GW, Zhang YD, Zhang H, Zhu FP, Wang K, Xia YX, Zhang YD, Jiang WJ, Li XC and Wang XH: Biliary tract cancer at CT: A radiomics-based model to predict lymph node metastasis and survival outcomes. *Radiology* 290: 90-98, 2019.
10. Autorino R, Gui B, Panza G, Boldrini L, Cusumano D, Russo L, Nardangeli A, Persiani S, Campitelli M, Ferrandina G, *et al*: Radiomics-based prediction of two-year clinical outcome in locally advanced cervical cancer patients undergoing neoadjuvant chemoradiotherapy. *Radiol Med* 127: 498-506, 2022.
11. Cong H, Peng W, Tian Z, Vallières M, Chuanpei X, Aijun Z and Benxin Z: FDG-PET/CT radiomics models for the early prediction of locoregional recurrence in head and neck cancer. *Curr Med Imaging* 17: 374-383, 2021.
12. Jiang C, Luo Y, Yuan J, You S, Chen Z, Wu M, Wang G and Gong J: CT-based radiomics and machine learning to predict spread through air space in lung adenocarcinoma. *Eur Radiol* 30: 4050-4057, 2020.
13. Li C, Jiang C, Gong J, Wu X, Luo Y and Sun G: A CT-based logistic regression model to predict spread through air space in lung adenocarcinoma. *Quant Imaging Med Surg* 10: 1984-1993, 2020.
14. Liu A, Sun X, Xu J, Xuan Y, Zhao Y, Qiu T, Hou F, Qin Y, Wang Y, Lu T, *et al*: Relevance and prognostic ability of Twist, Slug and tumor spread through air spaces in lung adenocarcinoma. *Cancer Med* 9: 1986-1998, 2020.
15. Chen Y, Jiang C, Kang W, Gong J, Luo D, You S, Cheng Z, Luo Y and Wu K: Development and validation of a CT-based nomogram to predict spread through air space (STAS) in peripheral stage IA lung adenocarcinoma. *Jpn J Radiol* 40: 586-594, 2022.
16. Nishimori M, Iwasa H, Miyatake K, Nitta N, Nakaji K, Matsumoto T, Yamanishi T, Yoshimatsu R, Iguchi M, Tamura M and Yamagami T: 18F FDG-PET/CT analysis of spread through air spaces (STAS) in clinical stage I lung adenocarcinoma. *Ann Nucl Med* 36: 897-903, 2022.
17. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E and Henry DA: AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 358: j4008, 2017.
18. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ and Welch VA: *Cochrane handbook for systematic reviews of interventions* version 6.4 (updated August 2023). *Cochrane*, 2023.
19. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA and Bossuyt PM; QUADAS-2 group: QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 155: 529-536, 2011.
20. Wang SR, Li QL, Tian F, Li J, Li WX, Chen M, Sang T, Cao CL and Shi LN: Diagnostic value of multiple diagnostic methods for lymph node metastases of papillary thyroid carcinoma: A systematic review and meta-analysis. *Front Oncol* 12: 990603, 2022.
21. Kutob L and Schneider F: Lung cancer staging. *Surg Pathol Clin* 13: 57-71, 2020.
22. Heinecke A, Tallarita M and De Iorio M: Bayesian splines versus fractional polynomials in network meta-analysis. *BMC Med Res Methodol* 20: 261, 2020.
23. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP, *et al*: The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. *Ann Intern Med* 162: 777-784, 2015.
24. Kapor S, Rankovic MJ, Khazaei Y, Crispin A, Schüler I, Krause F, Lussi A, Neuhaus K, Eggmann F, Michou S, *et al*: Systematic review and meta-analysis of diagnostic methods for occlusal surface caries. *Clin Oral Investig* 25: 4801-4815, 2021.
25. Bassi M, Russomando A, Vannucci J, Ciardiello A, Dolcianni M, Ricci P, Pernazza A, D'Amati G, Terracciano CM, Faccini R, *et al*: Role of radiomics in predicting lung cancer spread through air spaces in a heterogeneous dataset. *Transl Lung Cancer Res* 11: 560-571, 2022.
26. Qi L, Li X, He L, Cheng G, Cai Y, Xue K and Li M: Comparison of diagnostic performance of spread through airspaces of lung adenocarcinoma based on morphological analysis and perinodular and intranodular radiomic features on chest CT images. *Front Oncol* 11: 654413, 2021.
27. Chen LW, Lin MW, Hsieh MS, Yang SM, Wang HJ, Chen YC, Chen HY, Hu YH, Lee CE, Chen JS, *et al*: Radiomic values from high-grade subtypes to predict spread through air spaces in lung adenocarcinoma. *Ann Thorac Surg* 114: 999-1006, 2022.
28. Kim SK, Kim TJ, Chung MJ, Kim TS, Lee KS, Zo JI and Shim YM: Lung Adenocarcinoma: CT features associated with spread through air spaces. *Radiology* 289: 831-840, 2018.
29. Qin L, Sun Y, Zhu R, Hu B and Wu J: Clinicopathological and CT features of tumor spread through air space in invasive lung adenocarcinoma. *Front Oncol* 12: 959113, 2022.
30. Qi L, Xue K, Cai Y, Lu J, Li X and Li M: Predictors of CT morphologic features to identify spread through air spaces preoperatively in small-sized lung adenocarcinoma. *Front Oncol* 10: 548430, 2020.
31. Zhang Z, Liu Z, Feng H, Xiao F, Shao W, Liang C, Sun H, Gu X and Liu D: Predictive value of radiological features on spread through air space in stage cIA lung adenocarcinoma. *J Thorac Dis* 12: 6494-6504, 2020.
32. Han X, Fan J, Zheng Y, Ding C, Zhang X, Zhang K, Wang N, Jia X, Li Y, Liu J, *et al*: The value of CT-based radiomics for predicting spread through air spaces in stage IA lung adenocarcinoma. *Front Oncol* 12: 757389, 2022.
33. Takehana K, Sakamoto R, Fujimoto K, Matsuo Y, Nakajima N, Yoshizawa A, Menju T, Nakamura M, Yamada R, Mizowaki T and Nakamoto Y: Peritumoral radiomics features on preoperative thin-slice CT images can predict the spread through air spaces of lung adenocarcinoma. *Sci Rep* 12: 10323, 2022.
34. Wang XY, Zhao YF, Yang L, Liu Y, Yang YK and Wu N: Correlation analysis between metabolic tumor burden measured by positron emission tomography/computed tomography and the 2015 World Health Organization classification of lung adenocarcinoma, with a risk prediction model of tumor spread through air spaces. *Transl Cancer Res* 9: 6412-6422, 2020.
35. Falay O, Selçukbiricik F, Tanju S, Erus S, Kapdağlı M, Cesur E, Yavuz Ö, Bulutay P, Firat P, Mandel NM and Dilege Ş: The prediction of spread through air spaces with preoperative 18F-FDG PET/CT in cases with primary lung adenocarcinoma, its effect on the decision for an adjuvant treatment and its prognostic role. *Nucl Med Commun* 42: 922-927, 2021.
36. Wang S, Shou H, Wen H, Wang X, Wang H, Lu C, Gu J, Xu F, Zhu Q, Wang L and Ge D: An individual nomogram can reliably predict tumor spread through air spaces in non-small-cell lung cancer. *BMC Pulm Med* 22: 209, 2022.
37. Toki MI, Harrington K and Syrigos KN: The role of spread through air spaces (STAS) in lung adenocarcinoma prognosis and therapeutic decision making. *Lung Cancer* 146: 127-133, 2020.
38. Sun F, Huang Y, Yang X, Zhan C, Xi J, Lin Z, Shi Y, Jiang W and Wang Q: Solid component ratio influences prognosis of GGO-featured IA stage invasive lung adenocarcinoma. *Cancer Imaging* 20: 87, 2020.
39. Bai S, Wang Z, Sun Z and Liu Z: Study on the relationship between lung cancer stromal cells and air cavity diffusion based on an image acquisition system. *Contrast Media Mol Imaging* 2022: 2492124, 2022.
40. Li Y, Reyhan M, Zhang Y, Wang X, Zhou J, Zhang Y, Yue NJ and Nie K: The impact of phantom design and material-dependence on repeatability and reproducibility of CT-based radiomics features. *Med Phys* 49: 1648-1659, 2022.

41. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ and Kattan MW: Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology* 21: 128-138, 2010.
42. Sauerbrei W, Taube SE, McShane LM, Cavenagh MM and Altman DG: Reporting recommendations for tumor marker prognostic studies (REMARK): An abridged explanation and elaboration. *J Natl Cancer Inst* 110: 803-811, 2018.
43. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, Topol EJ, Ioannidis JPA, Collins GS and Maruthappu M: Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 368: m689, 2020.
44. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY and Van Calster B: A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 110: 12-22, 2019.



Copyright © 2024 Liu et al. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.