

**Scientific Evidence of Commercial Artificial Intelligence
Products for Pulmonary Nodule Assessment on CT Scans:
A Systematic Review**

ELECTRONIC SUPPLEMENTARY MATERIAL

Table of Contents

Search Strategy.....	3
Medline	3
Embase	3
Web of Science.....	4
Cochrane	5
Google Scholar	6
Data extraction Form	7
Hierarchical model of Radiology AI Deployment and Assessment Rubric.....	8
Table S1: Overview of RADAR	8
Preferred Reporting Items for Systematic Reviews and Meta-Analyses	10
Fig. S1: Flow Diagram of Studies Selected for Review	10
Additional Supporting Information.....	11
Fig. S2: Geographic Distribution of Datasets in Included Studies	11
Table S2: Study Characteristics and Methodology	12
Adapted Quality Assessment Framework Based on QUADAS-2 and QUADAS-C.....	20
Domains with Signaling Questions.....	20
Fig. S3. Traffic-Light Plot of Risk of Bias Assessment	21

Search Strategy

Medline

(Solitary Pulmonary Nodule / OR Multiple Pulmonary Nodules / OR (((lung OR pulmonar*) ADJ6 (nodul* OR spot* OR small-mass* OR lesion*))) .ab,ti. OR ((lung OR pulmonar*) AND (nodul* OR spot* OR small-mass* OR lesion*)) .ti.) AND (exp Artificial Intelligence/ OR exp Machine Learning/ OR Bayes Theorem/ OR ((artificial* ADJ3 intelligen*) OR ((machine OR deep) ADJ3 learning) OR (Computer* ADJ3 heuristics*) OR (hyperheuristics*) OR (metaheuristics*) OR (multicriteria* ADJ3 decision* ADJ3 analysis*) OR (multiobjective* ADJ3 (optimization* OR algorithm*)) OR (neural* ADJ3 network*) OR (automatic* ADJ3 (speech* OR pattern*) ADJ3 recognition*) OR (back* ADJ3 propagation*) OR (Bayesian* ADJ3 learning*) OR (classification* ADJ3 (algorithm OR binary OR multiclass OR multilabel)) OR (classifier*) OR (computer* ADJ3 heuristics*) OR (cross-validation*) OR (data-mining*) OR ((feature OR algorithm OR edge*) ADJ3 detection*) OR (feature* ADJ3 (extraction OR learning OR ranking OR selection OR fusion*)) OR (fuzzy* ADJ3 system*) OR (hidden* ADJ3 Markov* ADJ3 model*) OR (iterative* ADJ3 closest* ADJ3 point*) OR (k-nearest* ADJ3 neighbo*) OR (kernel* ADJ3 method*) OR (knowledge* ADJ3 discovery*) OR (learning* ADJ3 algorithm*) OR (least* ADJ3 absolute* ADJ3 shrinkage* ADJ3 selection* ADJ3 operator*) OR (Markov* ADJ3 model*) OR (memristor*) OR (molecular* ADJ3 docking*) OR (multicriteria* ADJ3 decision* ADJ3 analysis*) OR (multifactor* ADJ3 dimensionality* ADJ3 reduction*) OR (network* ADJ3 learning*) OR (online* ADJ3 analytical* ADJ3 processing*) OR (outlier* ADJ3 detection*) OR (perceptron*) OR (radial* ADJ3 basis* ADJ3 function*) OR (random* ADJ3 forest*) OR (recursive* ADJ3 feature* ADJ3 elimination*) OR (recursive* ADJ3 partitioning*) OR (relevance* ADJ3 vector* ADJ3 machine*) OR (rough* ADJ3 set*) OR (support* ADJ3 vector* ADJ3 machine*) OR ((recognition* OR detection* OR classification* OR segmentation* OR diagnos*) ADJ3 (algorithm* OR network* OR computer-aided* OR automatic*)) OR bayesian*) .ab,ti. OR Al.ti.) AND (computer assisted tomography/ OR ((compute* ADJ3 tomogra*) OR ct) .ab,ti.) NOT (exp animals/ NOT humans/) AND english.la.

Embase

('lung nodule'/de OR 'multiple pulmonary nodules'/de OR 'lung lesion'/de OR (((lung OR pulmonar*) NEAR/6 (nodul* OR spot* OR small-mass* OR lesion*))) :Ab,ti OR ((lung OR pulmonar*) AND (nodul* OR spot* OR small-mass* OR lesion*)) :ti) AND ('artificial intelligence'/exp OR 'machine learning'/exp OR 'pattern recognition'/exp OR 'bayes theorem'/de OR ((artificial* NEAR/3 intelligen*) OR ((machine OR deep) NEAR/3 learning) OR (Computer* NEAR/3 heuristics*) OR (hyperheuristics*) OR (metaheuristics*) OR (multicriteria*

NEAR/3 decision* NEAR/3 analysis*) OR (multiobjective* NEAR/3 (optimization* OR algorithm*)) OR (neural* NEAR/3 network*) OR (automatic* NEAR/3 (speech* OR pattern*) NEAR/3 recognition*) OR (back* NEAR/3 propagation*) OR (Bayesian* NEAR/3 learning*) OR (classification* NEAR/3 (algorithm OR binary OR multiclass OR multilabel)) OR (classifier*) OR (computer* NEAR/3 heuristics*) OR (cross-validation*) OR (data-mining*) OR ((feature OR algorithm OR edge*) NEAR/3 detection*) OR (feature* NEAR/3 (extraction OR learning OR ranking OR selection OR fusion*)) OR (fuzzy* NEAR/3 system*) OR (hidden* NEAR/3 Markov* NEAR/3 model*) OR (iterative* NEAR/3 closest* NEAR/3 point*) OR (k-nearest* NEAR/3 neighbo*) OR (kernel* NEAR/3 method*) OR (knowledge* NEAR/3 discovery*) OR (learning* NEAR/3 algorithm*) OR (least* NEAR/3 absolute* NEAR/3 shrinkage* NEAR/3 selection* NEAR/3 operator*) OR (Markov* NEAR/3 model*) OR (memristor*) OR (molecular* NEAR/3 docking*) OR (multicriteria* NEAR/3 decision* NEAR/3 analysis*) OR (multifactor* NEAR/3 dimensionality* NEAR/3 reduction*) OR (network* NEAR/3 learning*) OR (online* NEAR/3 analytical* NEAR/3 processing*) OR (outlier* NEAR/3 detection*) OR (perceptron*) OR (radial* NEAR/3 basis* NEAR/3 function*) OR (random* NEAR/3 forest*) OR (recursive* NEAR/3 feature* NEAR/3 elimination*) OR (recursive* NEAR/3 partitioning*) OR (relevance* NEAR/3 vector* NEAR/3 machine*) OR (rough* NEAR/3 set*) OR (support* NEAR/3 vector* NEAR/3 machine*) OR ((recognition* OR detection* OR classification* OR segmentation* OR diagnos*) NEAR/3 (algorithm* OR network* OR computer-aided* OR automatic*)) OR bayesian*):ab,ti OR Al:ti AND ('computer assisted tomography'/exp OR ((compute* NEAR/3 tomogra*) OR ct):Ab,ti) NOT [conference abstract]/lim NOT ([animals]/lim NOT [humans]/lim) AND [english]/lim

Web of Science

TS=((((lung OR pulmonar*) NEAR/5 (nodul* OR spot* OR small-mass* OR lesion*))) AND (((artificial* NEAR/2 intelligen*) OR ((machine OR deep) NEAR/2 learning) OR (Computer* NEAR/2 heuristics*) OR (hyperheuristics*) OR (metaheuristics*) OR (multicriteria* NEAR/2 decision* NEAR/2 analysis*) OR (multiobjective* NEAR/2 (optimization* OR algorithm*)) OR (neural* NEAR/2 network*) OR (automatic* NEAR/2 (speech* OR pattern*) NEAR/2 recognition*) OR (back* NEAR/2 propagation*) OR (Bayesian* NEAR/2 learning*) OR (classification* NEAR/2 (algorithm OR binary OR multiclass OR multilabel)) OR (classifier*) OR (computer* NEAR/2 heuristics*) OR (cross-validation*) OR (data-mining*) OR ((feature OR algorithm OR edge*) NEAR/2 detection*) OR (feature* NEAR/2 (extraction OR learning OR ranking OR selection OR fusion*)) OR (fuzzy* NEAR/2 system*) OR (hidden* NEAR/2 Markov* NEAR/2 model*) OR (iterative* NEAR/2 closest* NEAR/2 point*) OR (k-nearest* NEAR/2 neighbo*) OR (kernel* NEAR/2 method*) OR (knowledge* NEAR/2 discovery*) OR (learning* NEAR/2 algorithm*) OR (least* NEAR/2 absolute* NEAR/2 shrinkage* NEAR/2 selection* NEAR/2 operator*) OR (Markov* NEAR/2 model*) OR (memristor*) OR (molecular* NEAR/2

Eur Radiol (2026) Paramasamy J, Leliveld A, Groen J-W et al.

docking*) OR (multicriteria* NEAR/2 decision* NEAR/2 analysis*) OR (multifactor* NEAR/2 dimensionality* NEAR/2 reduction*) OR (network* NEAR/2 learning*) OR (online* NEAR/2 analytical* NEAR/2 processing*) OR (outlier* NEAR/2 detection*) OR (perceptron*) OR (radial* NEAR/2 basis* NEAR/2 function*) OR (random* NEAR/2 forest*) OR (recursive* NEAR/2 feature* NEAR/2 elimination*) OR (recursive* NEAR/2 partitioning*) OR (relevance* NEAR/2 vector* NEAR/2 machine*) OR (rough* NEAR/2 set*) OR (support* NEAR/2 vector* NEAR/2 machine*) OR ((recognition* OR detection* OR classification* OR segmentation* OR diagnos*) NEAR/2 (algorithm* OR network* OR computer-aided* OR automatic*)) OR bayesian*)) AND (((compute* NEAR/2 tomogra*) OR ct))) NOT DT=(Meeting Abstract OR Meeting Summary) AND LA=(english)

Cochrane

(((((lung OR pulmonar*) NEAR/6 (nodul* OR spot* OR small NEXT mass* OR lesion*)))Ab,ti OR ((lung OR pulmonar*) AND (nodul* OR spot* OR small NEXT mass* OR lesion*)):ti) AND (((artificial* NEAR/3 intelligen*) OR ((machine OR deep) NEAR/3 learning) OR (Computer* NEAR/3 heuristics*) OR (hyperheuristics*) OR (metaheuristics*) OR (multicriteria* NEAR/3 decision* NEAR/3 analysis*) OR (multiobjective* NEAR/3 (optimization* OR algorithm*)) OR (neural* NEAR/3 network*) OR (automatic* NEAR/3 (speech* OR pattern*) NEAR/3 recognition*) OR (back* NEAR/3 propagation*) OR (Bayesian* NEAR/3 learning*) OR (classification* NEAR/3 (algorithm OR binary OR multiclass OR multilabel)) OR (classifier*) OR (computer* NEAR/3 heuristics*) OR (cross NEXT validation*) OR (data NEXT mining*) OR ((feature OR algorithm OR edge*) NEAR/3 detection*) OR (feature* NEAR/3 (extraction OR learning OR ranking OR selection OR fusion*)) OR (fuzzy* NEAR/3 system*) OR (hidden* NEAR/3 Markov* NEAR/3 model*) OR (iterative* NEAR/3 closest* NEAR/3 point*) OR (k NEXT nearest* NEAR/3 neighbo*) OR (kernel* NEAR/3 method*) OR (knowledge* NEAR/3 discovery*) OR (learning* NEAR/3 algorithm*) OR (least* NEAR/3 absolute* NEAR/3 shrinkage* NEAR/3 selection* NEAR/3 operator*) OR (Markov* NEAR/3 model*) OR (memristor*) OR (molecular* NEAR/3 docking*) OR (multicriteria* NEAR/3 decision* NEAR/3 analysis*) OR (multifactor* NEAR/3 dimensionality* NEAR/3 reduction*) OR (network* NEAR/3 learning*) OR (online* NEAR/3 analytical* NEAR/3 processing*) OR (outlier* NEAR/3 detection*) OR (perceptron*) OR (radial* NEAR/3 basis* NEAR/3 function*) OR (random* NEAR/3 forest*) OR (recursive* NEAR/3 feature* NEAR/3 elimination*) OR (recursive* NEAR/3 partitioning*) OR (relevance* NEAR/3 vector* NEAR/3 machine*) OR (rough* NEAR/3 set*) OR (support* NEAR/3 vector* NEAR/3 machine*) OR ((recognition* OR detection* OR classification* OR segmentation* OR diagnos*) NEAR/3 (algorithm* OR network* OR computer NEXT aided* OR automatic*)) OR bayesian*):ab,ti) AND (((compute* NEAR/3 tomogra*) OR ct):Ab,ti)

Google Scholar

"lung|pulmonary nodule|nodules" "artificial intelligence"|"machine|deep learning" "computed tomography"

Data extraction Form

This data extraction form is designed to capture key details regarding the study characteristics, AI system specifications, methodological data, and main outcomes of the included studies.

The first section includes **Study Characteristics and AI System Details**, capturing essential information such as the study title, year of publication, primary author(s), source, vendor involvement, and the country of data origin. It also provides details about the type of AI system used, specifically whether it was a Computer-Aided Detection (CADe) or Computer-Aided Diagnosis (CADx) system, along with the manufacturer's name, software used, and its version.

Study Characteristics and AI System Details

Study	Year	Author	Source				Vendor involvement	Country
			CADe/ CADx	Manufacturer	Software	Version		

The **Methodological Data** section includes details about the study design (e.g., cross-sectional, cohort) and the study population or database. It specifies the reference standard used for comparison (e.g., histology or clinical follow-up) and the sample size. Additionally, it includes information about the index test, whether the study was single-center or multi-center, and whether it involved a screening setting. This section also captures whether the study evaluated results per lesion or per patient, as well as the number of nodule cases (cases) and non-nodule cases (controls). Finally, it provides the number of readers involved in evaluating the study.

Methodological data

Study Design	Setting		Study Population/ Database	Reference Standard	Sample size			Index test	
	Single-Center/ Multi-Center	Screening			per-lesion/ per-patient	Case	Control		#readers

The **Main Outcomes and Outcomes per RADAR-Level** section summarizes the key findings of each study, focusing on the performance of the AI system in terms of sensitivity, specificity, and false positive rate per scan. This section also captures whether the study reported outcomes based on the RADAR levels (1-6), assessing the AI system's diagnostic accuracy, clinical integration, and broader therapeutic impact.

Main Outcomes and Outcomes per RADAR-level

Key findings			RADAR					
Sensitivity	Specificity	FP/scan	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6

Hierarchical model of Radiology AI Deployment and Assessment Rubric

The Radiology AI Deployment and Assessment Rubric (RADAR) is a hierarchical efficacy model that outlines the various levels of impact AI tools can have on clinical practice. This model builds upon the original six-level framework established by Fryback and Thornbury in 1991, designed to evaluate the contributions of diagnostic imaging to patient management across distinct levels of efficacy. In the RADAR model, this framework is adapted to focus specifically on the implications of AI technologies in clinical practice by incorporating an additional seventh level. The first six levels of RADAR are interconnected and operate within the same hierarchical dimension addressing various aspects of efficacy. Level 1 examines the technical efficacy of AI systems, emphasizing their capability to accurately process and analyze imaging data. Level 2 addresses diagnostic accuracy by evaluating the algorithm's performance in detecting and characterizing lung nodules. Level 3 encompasses diagnostic thinking efficacy, exploring how AI influences clinical decision-making processes. Level 4 pertains to therapeutic efficacy, scrutinizing how AI informs treatment planning and management strategies. Level 5 focuses on patient outcome efficacy, analyzing the effects of AI-assisted diagnoses on patient health and quality of life. Level 6 assesses societal efficacy, evaluating the overall benefits and cost-effectiveness of AI implementations within healthcare systems.

The newly introduced 7th level operates in a separate dimension and emphasizes the necessity of appraising AI systems within their specific local contexts. This level examines how the insights and findings from the first six levels translate to and are applied within local hospital settings, ensuring that the implementation of AI technologies is contextually relevant and effective. **Table S1: Overview of RADAR** provides an overview of the seven levels as outlined in the RADAR model. Due to the focus of this systematic review on hierarchical efficacy, only the initial six levels of the RADAR framework are considered in this review.

Table S1: Overview of RADAR - Hierarchical Levels of Efficacy for AI in Radiology. This table presents the RADAR framework, adapted from Fryback and Thornbury's imaging efficacy model, to evaluate the development and implementation of radiology AI systems. The first six levels operate within a hierarchical dimension, addressing interconnected aspects of AI performance, clinical utility, patient outcomes and broader healthcare impact. The seventh level, operating in a separate contextual dimension, emphasizes the appraisal of AI systems within specific local hospital settings. For each level, three illustrative examples are provided to demonstrate its relevance to AI applications in radiology, with a particular focus on lung nodule detection and management.

Level	Description	Clinical Examples
-------	-------------	-------------------

Level 1: Technical Efficacy	Assessment of AI algorithm performance in processing CT images.	1. Algorithm's ability to process the CT images. 2. Does the algorithm work when using CTs with different protocols. 3. Computation time for analysis.
Level 2: Diagnostic Accuracy Efficacy	Evaluation of accuracy in evaluating lung nodules on CT.	1. Sensitivity and specificity of the AI model in detecting lung nodules. 2. Predictive values for positive and negative findings in clinical cases. 3. Identifying malignant nodules out of all nodules.
Level 3: Diagnostic Thinking Efficacy	Impact of AI on clinician decision-making regarding lung nodules.	1. Change in clinicians' probability estimates for malignancy before and after AI evaluation. 2. Influence of AI suggestions on differential diagnoses. 3. Clinicians' confidence in diagnosis post-AI assessment.
Level 4: Therapeutic Efficacy	Influence of AI findings on treatment strategies for patients with lung nodules.	1. Changes in management plans based on AI results (e.g., referral for biopsy). 2. Percentage of patients avoiding unnecessary procedures due to AI assessments. 3. Adjustments to treatment protocols based on AI-detected features.
Level 5: Patient Outcome Efficacy	Effect of AI-assisted diagnoses on patient health outcomes.	1. Improvement in patient survival rates with AI-guided interventions. 2. Changes in quality-adjusted life years (QALYs) for patients with AI-assisted treatment plans. 3. Assessment of morbidity rates post-AI diagnosis.
Level 6: Societal Efficacy	Broader impact of AI implementations on healthcare systems and costs.	1. Cost-effectiveness analysis of AI integration in lung cancer screening programs. 2. Evaluation of public health outcomes resulting from widespread AI adoption. 3. Benefit-cost analyses of AI tools across different populations.
Level 7: Local Efficacy	<i>Emphasis on the valuation of AI systems within the local context, assessing how prior valuations translate locally.</i>	1. <i>Evaluation of budget impact analysis to determine the financial feasibility of AI implementations in specific healthcare settings.</i> 2. <i>Application of multi-criteria decision analysis to assess stakeholder preferences regarding AI tool adoption.</i> 3. <i>Prospective monitoring of AI systems to ensure ongoing effectiveness and alignment with local healthcare needs.</i>

Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other sources

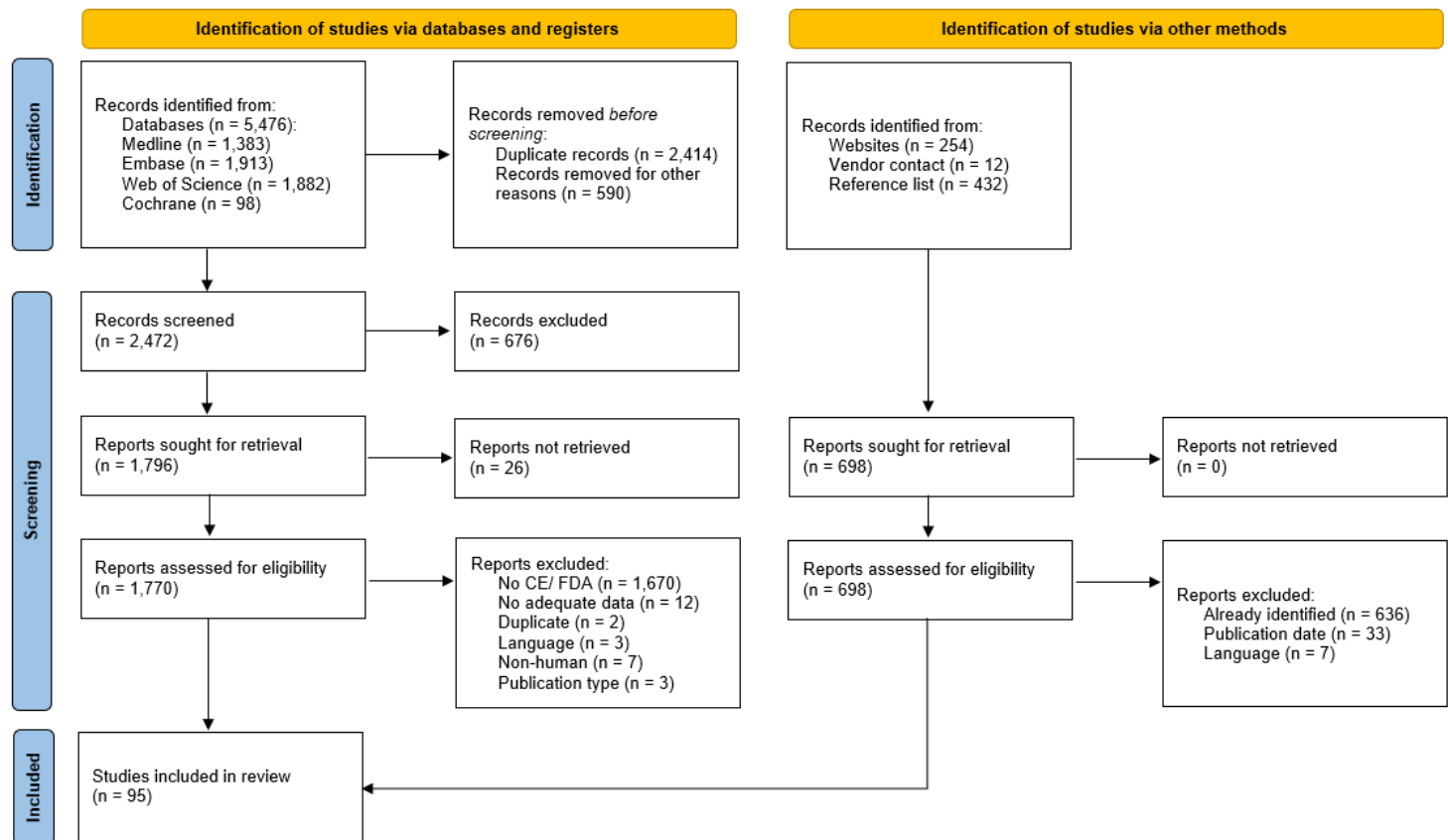


Fig. S1: Flow Diagram of Studies Selected for Review: The diagram illustrates the selection process for studies included in this review. CE = Conformité Européenne, FDA = Food and Drug Administration.

Additional Supporting Information

Fig. S2: Geographic Distribution of Datasets in Included Studies: This world map displays a heatmap of the geographic distribution of datasets used in the included studies. The intensity of the color represents the concentration of data sources, with the highest concentrations in the United States and China, where lung cancer screening studies are most commonly conducted.

Geographic Distribution of Data Sources in AI-based Lung Nodule Evaluations

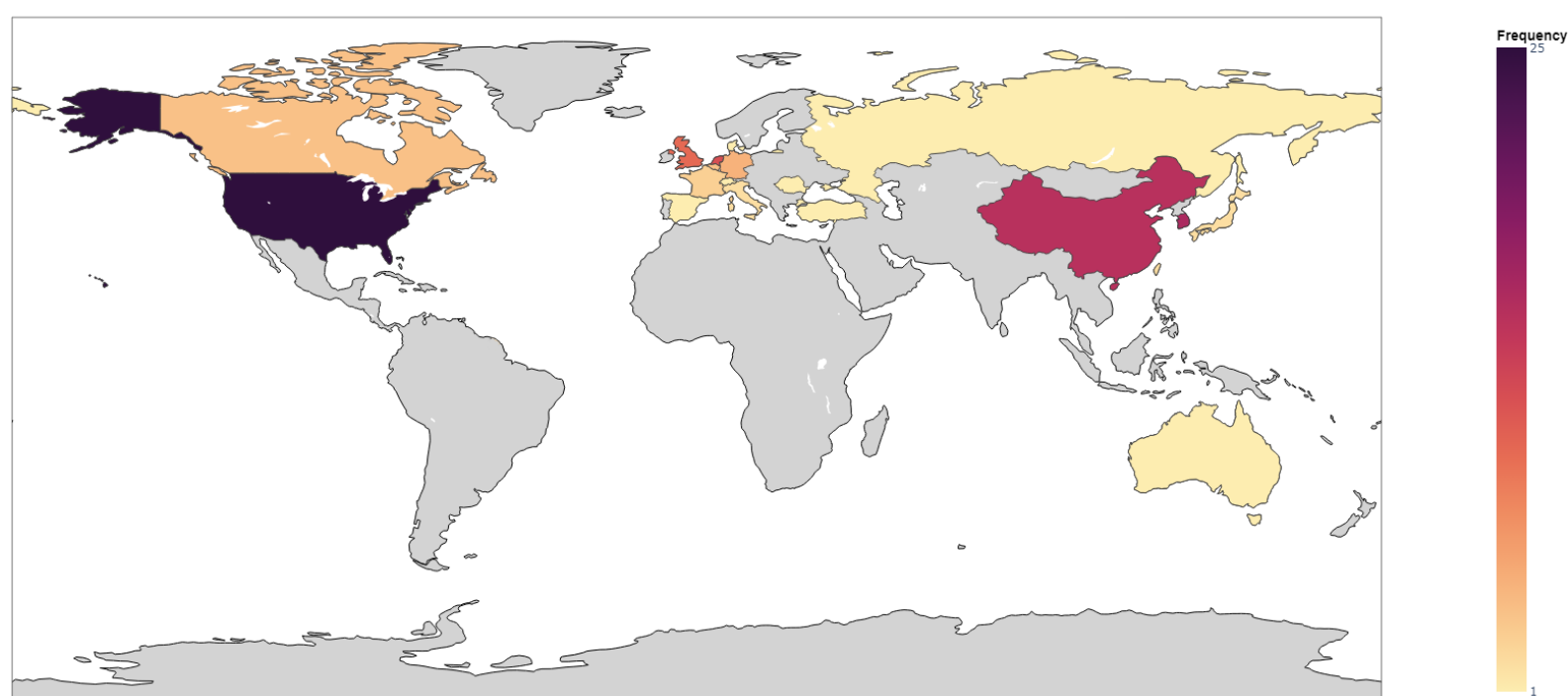


Table S2: Study Characteristics and Methodology: This table provides an overview of key study characteristics, including the author, study design, and setting. It details the study population or database used, the reference standard (RS) applied, and the RADAR assessment level assigned to each study. Additionally, the table includes information on the manufacturer and AI software utilized, as well as whether the study was conducted in a single-center or multi-center setting.

Author			Study Design	Setting	Study Population/ Database	RS	RA DA R
	Manufacturer	Software		Single/ Multi			
Bogoni et al. (2012)	Siemens Healthcare, Erlangen, Germany	Lung CAD	Cross-Sectional	Multi-Center	Early Diagnostic Research Network study	Findings labeled as LN by ≥ 3 readers.	2; 3
Zhao et al. (2012)	Siemens Healthcare, Erlangen, Germany	Lung CAD	Cross-Sectional	Multi-Center	NELSON	Findings by CAD and unaided RADs were reviewed by 2 reviewers.	2
Kim et al. (2013)	Siemens Healthcare, Erlangen, Germany	LungCAR E	Cross-Sectional	Single-Center	Patients with pulmonary GGNs identified on low-dose, thin-section chest CT.	CAD segmentations were reviewed by two reviewers.	1
Godoy et al. (2013)	Siemens Healthcare, Erlangen, Germany	Lung CAD	Cross-Sectional	Single-Center	Patients with GGO nodules ≥ 5 mm.	Findings of all readers were reviewed by one adjudicator.	2; 3
Scholten et al. (2015, 1)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Multi-Center	NELSON	Two readers identified all nodules, adjudicator was consulted for discrepancies.	1
Scholten et al. (2015, 2)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Multi-Center	NELSON	Two readers identified all nodules, in case of discrepancy the case was excluded.	1
Cohen et al. (2016, 1)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Single-Center	Patients whom GGNs were surgically resected.	Two radiologists measured all nodules.	1
Cohen et al. (2016, 2)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Single-Center	Patients whom GGNs were surgically resected.	<i>Software unenhanced vs enhanced CT</i>	1
Ritchie et al. (2016)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cohort	Multi-Center	PanCan	PanCan study radiologists together with the FU outcomes.	3
Jacobs et al. (2016)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Multi-Center	LIDC/IDRI	Two phase process: Four radiologists independently read all cases, then reviewed their own marks with anonymized annotations from the other three.	1; 2
Den Harder et al. (2016)	Philips	IntelliSpace Portal - Lung Nodule Assessment	Cross-Sectional	Single-Center	Patients aged ≥ 50 years scheduled for follow-up chest CT for known pulmonary nodules.	Two readers identified all solid nodules.	1; 2
Setio et al. (2017)	MeVis Medical Solutions AG Deephealth	Veolity NR	Cross-Sectional	Multi-Center	LIDC/IDRI	Preannotated LIDC-IDRI dataset with annotations from four readers.	2
Cohen et al. (2017)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Single-Center	Patients whom SSNs were surgically resected.	Two readers identified all nodules. <i>CAD in FBP vs MBIR</i>	1
Liang et al. (2017)	Siemens Healthcare, Erlangen, Germany	syngo.via Lung CAD	Cross-Sectional	Single-Center	Patients from the Mount Sinai lung cancer screening database with solid nodules < 30 mm.	Two readers assessment.	1
Garzelli et al. (2018)	Coreline Soft, Co. Ltd, Seoul, Korea	AVIEW LCS	Cross-Sectional	Single-Center	Patients who had undergone surgical resection for lung adenocarcinomas and	Two readers identified all nodules.	1

					preinvasive lesions that manifested as subsolid nodules on thin- section CT.		
Kim et al. (2018)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Single-Center	Patients whom SSNs were surgically resected.	Two readers identified all nodules.	1
Silva et al. (2018, 1)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Multi-Center	MILD	Findings by CAD were reviewed by 2 reviewers.	1; 2; 4
Aissa et al. (2018)	Riverain Technologies, Miamisburg, OH, USA	ClearRead CT	Cohort	Single-Center	Patients with malignant melanoma who underwent chest CT.	-	2; 3; 4
Lo et al. (2018)	Riverain Technologies, Miamisburg, OH, USA	ClearRead CT	Cross-Sectional	Multi-Center	NLST	Three readers identified all nodules. Majority vote for discrepancies.	2; 3
Takahashi et al. (2018)	Siemens Healthcare, Erlangen, Germany	Lung CAD	Cross-Sectional	Single-Center	Patients who underwent standard-dose CT for nodule evaluation, excluding those with pulmonary diseases, prior radiation, thoracic surgery, or >10 nodules.	One reader identified all nodules.	2; 3
Silva et al. (2018, 2)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cohort	Multi-Center	MILD	Readers identified all nodules and two reviewers assessed all CAD marks.	2; 3; 4; 5
Wagner et al. (2019)	Riverain Technologies, Miamisburg, OH, USA	ClearRead CT	Cross-Sectional	Single-Center	Patients who received a lung CT for the clinical indication of tumor and/or pulmonary nodule with histologically confirmed nodules.	Two readers identified all nodules based on the common report.	1; 2; 3
Mozaffary et al. (2019)	Siemens Healthcare, Erlangen, Germany	Lung CAD	Cross-Sectional	Single-Center	Emergency patients who underwent computed tomography pulmonary angiography (CTPA) for suspected pulmonary embolism (PE)	Findings by CAD were reviewed by 1 reviewer.	3
Liu et al. (2019)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Multi-Center	Patients who underwent thin-section chest CT scans for screening and in-patient purposes.	Two readers identified all nodules based on the common report. A reviewer was consulted for discrepancies.	2; 3; 4
Tammemagi et al. (2019)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Multi-Center	PanCan	Findings by CAD were reviewed by reviewer.	2
Wan et al. (2020)	Riverain Technologies, Miamisburg, OH, USA	ClearReadCT	Cross-Sectional	Single-Center	Patients who underwent CT-guided localization of lung nodules followed by video-assisted thoracic surgery and pathology, excluding those with prior thoracic surgery, nodules >2 cm, or metastases.	Pathologically proven lung nodules.	2
Baldwin et al. (2020)	Optellum	Lung Cancer Prediction	Cohort	Multi-Center	IDEAL	Pathologically proven lung nodules and stability assessments by radiologists.	2
Massion et al (2020)	Optellum	Lung Cancer Prediction	Cohort	Multi-Center	Patients with a solid pulmonary nodule 5–30 mm, excluding those with more than 5 nodules.	Pathologically proven lung nodules and stability assessments by radiologists.	2
Kozuka et al. (2020)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients suspected of lung cancer aged 20 years or older, who underwent CT examination, excluding those with severe pulmonary diseases and postoperative complications.	Two readers identified all nodules based on the common report. A reviewer was consulted for discrepancies.	2; 3

Cho et al. (2020)	Monitor Corporation, Seoul, Korea	LuCAS	Cohort	Multi-Center	NLST	Two readers independently classified CT as positive or negative for lung nodules, also all CAD marks were reviewed. A reviewer was consulted for discrepancies.	2
Hwang et al. (2021, 1)	MeVis Medical Solutions AG, Bremen, Germany	Visia	Cross-Sectional	Multi-Center	K-LUCAS	Findings by CAD were reviewed by 1 reviewer.	1; 2
Jarnalo et al. (2021)	Deephealth	Veye	Cross-Sectional	Single-Center	Patients who had undergone chest CT excluding those with more than five nodules.	Two readers identified all nodules based on the common report and CAD marks. A reviewer was consulted for discrepancies.	2
Jacobs et al. (2021, 1)	Deephealth		Cross-Sectional	Multi-Center	DSB2017 (which is NLST; DLST; screening program at Lahey Hospital and Medical Center) PanCan	Preannotated DSB2017 and PanCan dataset	2
Blazis et al. (2021)	Deephealth	Veye	Cross-Sectional	Single-Center	Patients who had undergone chest CT.	One reader identified all nodules using CAD in regular clinical workflow. One reader reidentified all nodules. A reviewer was consulted for discrepancies.	1; 2
Li et al. (2021)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients with surgically resected and histologically proven lung nodules or masses, excluding those with diffuse pulmonary diseases and presurgical chemotherapy or radiotherapy treatment.	Two readers identified all nodules. A reviewer was consulted for discrepancies.	1; 2
Singh et al. (2021)	Riverain Technologies, Miamisburg, OH, USA	ClearRead CT	Cross-Sectional	Multi-Center	NLST	Two readers identified all nodules. A reviewer was consulted for discrepancies.	2; 3
Chamberlin et al. (2021)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients who underwent routine lung cancer screening.	Two readers identified all nodules and reviewed CAD marks.	2
Rueckel et al. (2021)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients who underwent whole-body CT scans in the emergency department.	Findings by CAD were reviewed by 1 reviewer and compared with the original report. An adjudicator was consulted for discrepancies.	2; 3
Park et al. (2021, 1)	VUNO	Med-LungCT AI	Cross-Sectional	Single-Center	Patients with contrast-enhanced chest CT who underwent curative resection of lung adenocarcinoma, with SSN sizes ranging from 6 mm to 3 cm, up to seven SSNs.	Two readers identified all nodules.	2
Hsu et al. (2021)	Riverain Technologies, Miamisburg, OH, USA	ClearReadCT	Cross-Sectional	Single-Center	Patients with histopathological proven malignant nodules and small pulmonary nodules (≤ 10 mm), excluding those with severe pulmonary diseases, >10 nodules, or postoperative complications.	Findings of all readers were reviewed by two reviewers.	2; 3
Han et al. (2021)	Optellum	Lung Cancer Prediction	Cross-Sectional	Multi-Center	Patients with solid lung nodules 5–15 mm in diameter, excluding those with more than 5 nodules.	Consensus of three readers' assessment.	2
Hwang et al. (2021, 2)	Coreline Soft, Co. Ltd, Seoul, Korea	A-view Lungscreen	Cohort	Multi-Center	K-LUCAS	Medical records.	1
Jacobs et al. (2021, 2)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cohort	Multi-Center	NLST		3
AutrassEAU et al. (2021)	Mindshare Medical	RevealAI-Lung	Cohort	Multi-Center	Patients with 15+ years of asbestos exposure for screening of pleuropulmonary lesions, and those referred for clinically indicated unenhanced chest CT.	One reader identified all nodules and medical records were checked for final diagnosis.	2

Ahn et al. (2021)	VUNO	Med-LungCT AI	Cross-Sectional	Single-Center	Patients who had curative resection of non-small cell lung cancer with subsolid lesions ≥ 6 mm on CT, and no history of radiation therapy or neoadjuvant chemotherapy.	Segmentations by CAD were reviewed by two reviewers.	1
Tsakok et al. (2021)	Optellum	Lung Cancer Prediction	Cohort	Single-Center	Patients with an incidental nodule of >5 – 15 mm with recommending investigation or follow up.	Pathologically proven lung nodules and stability assessments by radiologist.	2; 4
Heuvelmans et al. (2021)	Optellum	Lung Cancer Prediction	Cohort	Multi-Center	LUCINDA	Pathologically proven lung nodules and stability assessments by radiologist.	1; 2; 4
Çoruh et al. (2021)	HuiYing Medical Technology Co., Ltd., Beijing, China		Cohort	Single-Center	Patients diagnosed with pulmonary neoplasms through radiology and histopathology.	Pathologically proven lung nodules and stability assessments by radiologist.	1; 2; 4
Park et al. (2021, 2)	VUNO		Cohort	Single-Center	Patients with surgically resected lung lesions (SSNs) within the adenocarcinoma spectrum (AAH, AIS, MIA, IPA), excluding those with history of lung cancer or neoadjuvant chemotherapy.	Pathologically proven lung nodules and stability assessments by radiologist.	2
Park et al. (2022)	VUNO	Med-LungCT AI	Cohort	Multi-Center	NLST	Information provided by NLST regarding Lung-RADS	2; 3; 4
Lancaster et al. (2022)	Coreline Soft, Co. Ltd, Seoul, Korea	A-view Lungscreen	Cross-Sectional	Multi-Center	Moscow Lung Cancer Screening	Three readers assessment.	2
Diao et al. (2022)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Multi-Center	Patients aged between 40–69 years from a lung cancer screening program.	One reader assessed all CTs and reviewed all reader's and CAD findings.	2; 3
Kim et al. (2022)	Optellum	Lung Cancer Prediction Virtual Nodule Clinic	Cohort	Multi-Center	Patients with 5- to 30-mm IPNs from screening and diagnostic chest CT scans, excluding GGNs and calcified nodules.	Pathologically proven lung nodules and stability assessments by radiologist.	3; 4
Chetan et al. (2022)	Optellum	Lung Cancer Prediction	Cross-Sectional	Multi-Center	NLST	NLST radiologists' annotations.	2
Lee et al. (2022)	Coreline Soft, Co. Ltd, Seoul, Korea	AVIEW LCS	Cross-Sectional	Single-Center	Patients with lung metastasis from colorectal cancer, excluding those with extrathoracic metastases in >2 organs.	Findings by radiologists and AI-assisted user were reviewed (cross-examined).	2; 3
Hempe et al. (2022)	Deephealth	Veye	Cross-Sectional	Single-Center	Patients who had undergone chest CT.	- No RS for detection - Appropriate management recommendation: discrepant BTS grades were reevaluated, and a consensus grade was agreed upon in a consensus meeting.	3; 4
Murchison et al. (2022)	Deephealth	Veye	Cross-Sectional	Single-Center	Patients aged 50–74 years, current smokers, or with a smoking history and/or radiological evidence of pulmonary emphysema excluding those with evidence of diffuse pulmonary diseases.	Two readers identified all nodules. A reviewer was consulted for discrepancies.	1; 2; 3
Hu et al. (2022, 1)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients with extra-thoracic malignancies scheduled for chest CT to detect pulmonary metastasis or monitor treatment response were included, excluding minors, pregnant patients and those with known pulmonary diseases.	Three readers identified all nodules.	1; 2
Hu et al.	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients with extra-thoracic malignancies scheduled for chest CT to detect pulmonary metastasis or monitor	Findings identified by all three readers aided by CAD were included.	1; 3

(2022, 2)					treatment response were included, excluding minors, pregnant patients and those with known pulmonary diseases.		
Jiang et al. (2022)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients in whom CECT and non-contrast ULD CT were performed.	Findings by CAD-aided radiologist on CECT were reviewed by 1 reviewer	1; 3
Guo et al. (2022)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients with BMI of ≥ 25 kg/m ² who underwent non-contrast-enhanced chest CT for the initial evaluation or follow-up of PNAs, excluding those with evidence of diffuse pulmonary diseases.	Findings by two CAD-aided radiologists on LDCT, adjudicator was consulted for discrepancies. .	2
Zhang et al. (2022)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Multi-Center	Patients who had undergone chest CT. Netherlands-China Big-3	Findings by two CAD-aided radiologists and findings in the initial report were reviewed by two radiologists.	3
Zhu et al. (2022)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients who underwent a low-dose non-contrast-enhanced chest CT scan of pulmonary nodules, excluding those with evidence of diffuse pulmonary diseases.	Four readers identified all nodules on adaptive iterative dose reduction 3D enhanced (eAIDR 3D)	1; 2
Mușeteșcu et al. (2022)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Single-Center	Patients diagnosed with rheumatoid arthritis.	Findings by CAD agreed by reviewer and findings identified by reader.	2; 3
Abadia et al. (2022)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients with at least one lung condition and one suspicious lung nodule on radiology reports.	One reader identified all nodules and findings by CAD were reviewed by a reviewer.	2; 3
Yacoub et al. (2022)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients who had undergone non-contrast chest CT.	Two readers identified all nodules, discrepancies were resolved by consensus discussions.	2
Kocher et al. (2022)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients diagnosed as having breast cancer.	One reader identified all nodules.	2
Gheysens et al. (2022)	Siemens Healthcare, Erlangen, Germany	MM Oncology	Cross-Sectional	Single-Center	Patients undergoing routine oncologic follow-up with thoraco-abdominal CT	Two readers identified all nodules on SDCT one time with and one time without CAD.	1; 3
Hall et al. (2022)	MeVis Medical Solutions AG, Bremen, Germany	Veolity	Cross-Sectional	Multi-center	Lung Screen Uptake Trial (LSUT)	Original reports were compared with radiographer's output, scans of not documented nodules in the original report were re-reviewed by one radiologist.	1; 3
Cavallo et al. (2023)	AIDOC		Cohort	Single-Center	Patients who had CT scans of the chest, abdomen, and pelvis obtained in the emergency setting.		
Zhang et al. (2023)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients with persistent GGNs, excluding those with severe organ dysfunction, radiotherapy or chemotherapy, lung metastases.	Findings by CAD were reviewed by 2 reviewers.	2
Zhu et al. (2023)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients who were admitted for healthcare screening or respiratory symptoms, excluding those with a history of prior respiratory treatment and severe pulmonary disease.	Findings by CAD were reviewed by 3 reviewers.	1; 2
Yoshida et al. (2023)	Riverain Technologies, Miamisburg, OH, USA	ClearReadCT	Cross-Sectional	Single-Center	Patients with 5-30 mm subsolid nodules on non-contrast chest CT, suspected to be adenocarcinomas, excluding those with more than 6 nodules.	Radiology reports and findings by CAD were reviewed by 2 reviewers.	1; 2
Smith et al. (2023)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients who had undergone SPECT/CT imaging.	One reader identified all nodules.	1; 2

Chamberlin et al. (2023)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients who had undergone PET/CT imaging for an oncological imaging purpose.	One reader identified all nodules.	2
Michelin et al. (2023)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cohort	Single-Center	Patients with a CT scan showing an incidental pulmonary nodule (6–30 mm) in an emergency context, excluding those with >3 nodules or cancer.	One reader identified all nodules and assessed medical records for final diagnosis.	1; 2; 3; 4
Adams et al. (2023)	Mindshare Medical	RevealAI-Lung	Cohort	Multi-Center	Patients from NLST, the Fred Hutchinson Lung Cancer Screening program, and real-world cases with any CT-detected lung nodule.	Pathologically proven lung nodules and stability assessments by radiologist.	2; 3
Weir-McCall et al. (2023)	Optellum	Lung Cancer Prediction	Cohort	Multi-Center	Patients with an indeterminate solitary pulmonary nodule (8–30 mm), excluding those with a history of malignancy or multiple nodules.	Pathologically proven lung nodules and stability assessments by radiologist.	1; 2
Paez et al. (2023)	Optellum	Lung Cancer Prediction	Cohort	Multi-Center	Patients with an IPN (6–30 mm) in largest axial diameter, either incidentally or through screening.	Pathologically proven lung nodules and stability assessments by radiologist.	2
Landy et al. (2023)	Optellum	Lung Cancer Prediction Convolutional Neural Network (LCP-CNN)	Cohort	Multi-Center	NLST	Pathologically proven lung nodules and stability assessments by radiologist.	3
Liu et al. (2024)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Single-Center	Patients with lung adenocarcinoma, excluding those with severe organ dysfunction and metastasis.	Findings by two CAD-aided radiologists and findings in the initial report.	2; 3
Yoo et al. (2024)	Monitor Corporation, Seoul, Korea	LuCAS	Cross-Sectional	Single-Center	First-time outpatients at a respiratory clinic scheduled for chest CT, excluding those with a BMI > 30 kg/m ² .	One reader identified all nodules.	2
Paramasamy et al. (2024)	Qure.ai	qCT	Cross-Sectional	Single-Center	Patients who had undergone chest CT, excluding those with > 10 nodules.	Radiology reports and findings by reader were compared, adjudicator was consulted for discrepancies.	2; 3
Santos et al. (2024)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients who had undergone chest CT.	Two readers identified all nodules.	2
Sourlos et al. (2024)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	ImaLife (Imaging in Lifelines)	Two reviewers reviewed discrepancies between findings for AI and reader.	2; 3
Peters et al. (2024)	Riverain Technologies, Miamisburg, OH, USA Siemens Healthcare, Erlangen, Germany	ClearReadCT AI-Rad Companion	Cross-Sectional	Multi-Center	Patients who had undergone chest CT with and without lung tumors.	Two readers identified all nodules.	1; 2; 3; 4
Mark et al. (2024)	Siemens Healthcare, Erlangen, Germany	AI-Rad Companion	Cross-Sectional	Single-Center	Patients with non-contrast chest CT, excluding those with severe pulmonary disease.	Findings by reporting radiologists and CAD were reviewed by one reviewer.	2; 3
Cho et al. (2024)	Coreline Soft, Co. Ltd, Seoul, Korea	AVIEW Lung Nodule CAD	Cohort	Single-Center	Patients with solid organ malignancies.	One reader reviewed medical records for final diagnosis and stability assessments.	2; 3

Masci et al. (2024)	Deephealth	Veye Lung Nodules	Cross-Sectional	Single-Center	Patients with an extra-thoracic primary malignancy who underwent pulmonary metastasectomy excluding those with known pulmonary diseases.	One radiologist reviewed all CTs together with the original radiology report and pathology report.	2; 3
Yoon et al. (2024)	Monitor Corporation, Seoul, Korea	LuCAS	Cohort	Single-Center	Patients who underwent pulmonary metastasectomy, excluding those with known pulmonary diseases.	Two reviewers reviewed CAD findings using electronic medical records and follow-up CT images.	2; 3
Gao et al. (2024)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cross-Sectional	Multi-Center	Patients from Beijing Science and Technology Project-Baseline Survey of Lung Cancer in Beijing and Study on Early Prevention and Treatment Strategies	A panel with at least three readers read all CTs.	1; 2
Xiong et al. (2024)	InferVision Medical Health, Beijing, China	InferRead CT Lung	Cohort	Single-Center	Patients with GGNs who underwent surgical resection excluding those with severe pulmonary diseases.	Pathologically proven lung nodules and stability assessments by radiologist.	1; 2
Quanyang et al. (2024)	HuiYing Medical Technology Co., Ltd., Beijing, China	Medical Imaging Assisted Diagnosis Platform	Cohort	Single-Center	Patients who underwent lung cancer screening at the National Cancer Center	Two radiologists assessed all nodules with a management strategy based on the National Comprehensive Cancer Network (NCCN) follow-up management guidelines. An adjudicator was consulted for discrepancies.	3; 4
Kim et al. (2024)	Optellum	Lung Cancer Prediction Virtual Nodule Clinic	Cohort	Single-Center	Patients without active lung cancer who underwent bronchoscopy biopsy for 8–30 mm solid or part-solid PNs.	Pathologically proven lung nodules and stability assessments by radiologist.	1; 2; 3
Kondrashova et al. (2024)	Coreline Soft, Seoul, Korea Siemens Healthineers, Forchheim, Germany	Aview ChestCTExplore	Cross-sectional	Multi-Center	HANSE	One reader identified all nodules, CAD reread all scans, a reviewer reviewed all included nodules.	1; 2; 4

Adapted Quality Assessment Framework Based on QUADAS-2 and QUADAS-C

The risk of bias and applicability concerns were evaluated as low, high, or unclear in four domains—patient selection, index test, reference standard, and flow and timing—using signaling questions.

Domains with Signaling Questions

DOMAIN 1: PATIENT SELECTION

1. Was a consecutive or random sample of patients enrolled?
2. Was a case-control design avoided?
3. Are the patients' inclusion/exclusion criteria as well as the study population described appropriately?
4. Was there a detailed description of the study population?
5. Was there a clear description of the data source?

DOMAIN 2: INDEX TEST(S)

1. Is the description of the AI method detailed?
2. Were the AI results by the human reader interpreted without knowledge of the results of the reference standard?

DOMAIN 3: REFERENCE STANDARD

1. Is the reference standard described?
2. Were the reference standard results interpreted without knowledge of the results of AI?

DOMAIN 4: FLOW AND TIMING

1. Was the interval between the AI deployment phase (aided AI phase) and reference standard (or unaided AI phase) described?
2. Did all patients receive a reference standard?
3. Did all patients receive the same reference standard?
4. Were all patients included in the analysis?

Fig. S3. Traffic-Light Plot of Risk of Bias Assessment: This figure presents a traffic-light plot illustrating the risk of bias assessment for each individual study across all evaluated domains. Green, yellow, and red indicators represent low, unclear, and high risk of bias, respectively.

	Risk of bias domains			
	D1	D2	D3	D4
Bogoni et al. (2012)				
Zhao et al. (2012)				
Godoy et al. (2013)				
Jacobs et al. (2016)				
Den Harder et al. (2016)				
Setio et al. (2017)				
Aissa et al. (2018)				
Lo et al. (2018)				
Takahashi et al. (2018)				
Silva et al. (2018, 2)				
Mozaffary et al. (2019)				
Liu et al. (2019)				
Kozuka et al. (2020)				
Cho et al. (2020)				
Hwang et al. (2021, 1)				
Jacobs et al. (2021)				
Blazis et al. (2021)				
Singh et al. (2021)				
Chamberlin et al. (2021)				
Rueckel et al. (2021)				
Park et al. (2021, 1)				
Hsu et al. (2021)				
Lee et al. (2022)				
Hempel et al. (2022)				
Hu et al. (2022, 1)				
Hu et al. (2022, 2)				
Jiang et al. (2022)				
Guo et al. (2022)				
Zhang et al. (2022)				
Zhu et al. (2022)				
Mușetescu et al. (2022)				
Abadia et al. (2022)				

Study	Yacoub et al. (2022)	-	X	-	X
	Kocher et al. (2022)	X	-	X	-
	Gheysens et al. (2022)	-	-	X	-
	Hall et al. (2022)	X	X	X	-
	Cavallo et al. (2023)	+	+	-	X
	Yoshida et al. (2023)	X	-	X	X
	Smith et al. (2023)	+	-	-	X
	Chamberlin et al. (2023)	X	+	-	X
	Santos et al. (2024)	X	-	+	-
	Yoo et al. (2024)	X	+	-	-
	Sourlos et al. (2024)	+	-	X	X
	Mark et al. (2024)	-	+	X	-
	Cho et al. (2024)	+	X	-	-
	Masci et al. (2024)	X	-	+	+
	Yoon et al. (2024)	+	-	-	X
	Ritchie et al. (2016)	-	X	+	+
	Silva et al. (2018, 1)	-	X	X	X
	Wagner et al. (2019)	X	+	-	?
	Jarnalo et al. (2021)	X	-	X	-
	Li et al. (2021)	X	-	-	-
	Murchison et al. (2022)	X	-	-	X
	Zhang et al. (2023)	X	?	X	X
	Zhu et al. (2023)	X	-	X	?
	Liu et al. (2024)	X	-	-	-
	Paramasamy et al. (2024)	-	+	-	?
	Peters et al. (2024)	-	X	+	-
	Kondrashova et al. (2024)	-	-	-	X
	Kim et al. (2013)	X	-	-	-
	Scholten et al. (2015, 1)	+	-	+	X
	Scholten et al. (2015, 2)	-	-	-	X
	Cohen et al. (2016, 1)	X	-	-	-
	Cohen et al. (2016, 2)	X	-	-	+
	Cohen et al. (2017)	X	-	+	-

Liang et al. (2017)				
Garzelli et al. (2018)				
Kim et al. (2018)				
Tammemagi et al. (2019)				
Wan et al. (2020)				
Baldwin et al. (2020)				
Massion et al (2020)				
Han et al. (2021)				
Hwang et al. (2021, 2)				
Jacobs et al. (2021, 2)				
Autrusseau et al. (2021)				
Ahn et al. (2021)				
Tsakok et al. (2021)				
Heuvelmans et al. (2021)				
Çoruh et al. (2021)				
Park et al. (2021, 2)				
Park et al. (2022)				
Lancaster et al. (2022)				
Diao et al. (2022)				
Kim et al. (2022)				
Chetan et al. (2022)				
Michelin et al. (2023)				
Adams et al. (2023)				
Weir-McCall et al. (2023)				
Paez et al. (2023)				
Landy et al. (2023)				
Gao et al. (2024)				
Xiong et al. (2024)				
Quanyang et al. (2024)				
Kim et al. (2024)				

Domains:
D1: Patient selection.
D2: Index test.
D3: Reference standard.
D4: Flow & timing.

Judgement
 High
 Some concerns
 Low
 No information