Review Article:

The efficacy of artificial intelligence-based imaging modalities for detection of emphysema: A systematic review

Mohamed K Seyam

Department of Physical Therapy and Health Rehabilitation, College of Applied Medical Sciences, Majmaah University, Al Majmaah 11952, Saudi Arabia

Received on Feb 09, 2025, accepted for publication on Aug 24, 2025, dx.doi.org/10.5455/mjhs.2025.04.013

Corresponding Author:
Mohamed K Seyam. Department of Physical Therapy and Health Rehabilitation, College of Applied Medical Sciences, Majmaah University, Al Majmaah 11952, Saudi Arabia; m.seyam@mu.edu.sa

OPEN ACCESS Distributed under Creative Commons CC BY-NC-ND 4.0

Abstract:

Emphysema detection has traditionally relied on clinical assessments and imaging modalities, such as computed tomography (CT) scans. Recently, artificial intelligence (AI)-based imaging modalities have been explored for their potential to enhance detection accuracy and efficiency. I conducted a systematic review to assess the efficacy of AI-based imaging modalities in detecting emphysema. Multiple databases were searched using tailored MeSH phrases, keywords, and Boolean operators. Each study was evaluated for quality using the QUADAS-AI tool and the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to assess the certainty of the evidence. Ten studies were included, that demonstrated varied efficacy of AI in the detection and quantification of emphysema. AI tools were generally found to correlate well with clinical pulmonary function tests and were able to accurately quantify emphysema from imaging studies. Some studies reported that AI was comparable or superior to radiologists in detecting emphysema, particularly in the context of lung cancer screenings. However, effectiveness varied with different AI approaches and settings. AI also showed potential in differentiating severity levels of chronic obstructive pulmonary disease (COPD) and outperforming traditional densitometry methods in texture analysis. AI-based imaging modalities show promise in enhancing the detection and management of emphysema. They offer potential not only as supplementary tools to traditional methods but also, in some cases, as superior alternatives. The integration of AI into clinical practice could improve diagnostic accuracy and patient outcomes, although variability in AI performance across different contexts should be considered.

Key words:

AI imaging, emphysema detection, systematic review, QUADAS-AI, GRADE methodology, pulmonary function tests, COPD, texture analysis

Introduction:

Emphysema is a primary global health concern; it is a pathological hallmark of chronic obstructive pulmonary disease (COPD), which is defined by the permanent deterioration of alveolar structures other than the proximal alveolar ducts [1]. Spirometry and high-resolution computed tomography (HRCT) are two classic diagnostic modalities that are clinical pillars for identifying and measuring emphysematous alterations. However, emphysema's varied clinical presentations and subtle morphological features highlight the need for more sophisticated, accurate, and targeted diagnostic instruments [2].

Artificial intelligence (AI) has become a disruptive force in medical imaging in recent years, bringing new methods of image analysis that use deep learning frameworks and machine learning algorithms [3]. With the early and accurate identification of pathological diseases like emphysema, these AI-based imaging modalities can improve patient outcomes while increasing diagnostic accuracy and throughput. AI algorithms are very good at deciphering intricate patterns in imaging data, spotting minute details that would be difficult for humans to perceive, and producing quantitative evaluations that accurately represent the underlying pathophysiology [4-5].

The effectiveness of these cutting-edge systems in the particular context of emphysema diagnosis is still being actively researched despite the growing interest in using AI in medical imaging [6]. The performance of AI-based imaging modalities needs to be thoroughly assessed because it affects professional judgment, patient care, and the distribution of healthcare resources. A thorough evaluation of AI's diagnostic skills, measured against accepted imaging criteria, is necessary in the context of its integration into clinical practice to guarantee that the application of these technologies is consistent with evidence-based medicine [6-8].

When diagnosing emphysema, low-density regions on CT scans are made more visible with the minimum intensity projection (MinIP) technique, making it easier to spot early and subtle symptoms of the disease [9-10]. This study investigated the applicability of MinIP as a specialized augmentation method for an unsupervised deep learning (DL) algorithm to automate the low-dose computed tomography (LDCT) scan emphysema diagnosis process [11-12].

Given how quickly AI is developing in medical imaging, a thorough review of the available data is required to determine the practical effectiveness of AI-based modalities for emphysema detection. Therefore, this systematic study aims to review the efficacy of artificial intelligence-based im-

aging modalities for detection of emphysema.

Materials and Methods:

Eligibility criteria:

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [13] shaped the methodological framework of this systematic review. The methodology of PECO (Population, Exposure, Comparator, Outcome) was carefully developed. Patients without constraints on age, gender, or severity of the disease who were clinically suspected of having emphysema or who had been diagnosed with COPD made up the group of interest. The exposure was examined using AI-based imaging modalities, such as convolutional neural networks (CNNs), combined with different imaging techniques like HRCT. The standard imaging method of diagnosing emphysema mainly involved radiologists' visual evaluation and measurements of lung density metrics, which served as the comparison. However, considering our review's scope and objectives, a comparator was not deemed mandatory. The previously mentioned diagnostic accuracy measures comprised the primary outcomes of interest. The inclusion and exclusion criteria devised for this review are represented in Table 1.

Table 1: Inclusion and exclusion criteria devised for this review.

Criteria	Description
Inclusion	
Study Design	Peer-reviewed articles reporting original research, including randomized controlled trials (RCTs), cohort studies, case-control studies, and cross-sectional studies.
Population	Studies involving patients with a clinical suspicion or established diagnosis of emphysema/COPD.
Intervention	Research utilizing AI-based imaging modalities for the detection of emphysema.
Comparators	Studies that included a comparator group assessed with traditional imaging techniques for emphysema.
Outcomes	Articles reporting on measures of diagnostic accuracy such as sensitivity, specificity, AUC, predictive values, likelihood ratios, and diagnostic odds ratios.
Exclusion	
Study Design	Editorials, letters, comments, conference abstracts, reviews, and meta-analyses.
Population	Studies exclusively involving animal models or pediatric populations.
Intervention	Research focusing on other pulmonary disorders without a specific focus on emphysema, or studies examining non-AI based imaging modalities.
Comparators	Studies lacking a traditional imaging comparator.
Outcomes	Studies not reporting specific diagnostic accuracy measures.

Database search protocol:

In order to find pertinent papers for our review, the databases of PubMed, Embase, Web of Science, Scopus, IEEE Xplore, Cochrane Library, and Google Scholar were searched. MeSH phrases, keywords, and Boolean operators, inclusion date between 2006 till 2023 specific to each platform's features and indexing terms were used to access and query each database, as shown in Table 2.

Table 2: Search strings utilised across the different databases.

Database	Search string
PubMed	("emphysema"[MeSH Terms] OR "emphysema"[Title/Abstract]) AND ("artificial intelligence"[MeSH Terms] OR "AI"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "neural networks"[Title/Abstract] OR "deep learning"[Title/Abstract]) AND ("imaging"[MeSH Terms] OR "computed tomography"[Title/Abstract] OR "magnetic resonance imaging"[Title/Abstract] OR "radiography"[Title/Abstract])
Embase	('emphysema'/exp OR emphysema) AND ('artificial intelligence'/exp OR 'machine learning'/exp OR 'deep learning'/exp OR 'neural networks'/exp) AND ('diagnostic imaging'/exp OR 'computed tomography'/exp OR 'magnetic resonance imaging'/exp OR 'radiography'/exp)
Web of Science	TS=(emphysema) AND TS=(("artificial intelligence" OR AI OR "machine learning" OR "neural networks" OR "deep learning")) AND TS=(("imaging modalities" OR "computed tomography" OR CT OR "magnetic resonance imaging" OR MRI OR radiography))
Scopus	TITLE-ABS-KEY (emphysema) AND TITLE-ABS-KEY (("artificial intelligence" OR AI OR "machine learning" OR "neural networks" OR "deep learning")) AND TITLE-ABS-KEY ((imaging OR "computed tomography" OR CT OR "magnetic resonance imaging" OR MRI OR radiography))
IEEE Xplore	("Abstract":emphysema) AND ("Abstract":("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural networks")) AND ("Abstract":(imaging OR "computed tomography" OR "magnetic resonance imaging" OR "radiography"))
Cochrane Library	(MeSH descriptor: [Emphysema] this term only) AND (MeSH descriptor: [Artificial Intelligence] this term only OR MeSH descriptor: [Machine Learning] this term only OR MeSH descriptor: [Neural Networks (Computer)] this term only OR MeSH descriptor: [Deep Learning] this term only) AND (MeSH descriptor: [Diagnostic Imaging] this term only OR MeSH descriptor: [Tomography, X-Ray Computed] this term only OR MeSH descriptor: [Magnetic Resonance Imaging] this term only OR MeSH descriptor: [Radiography] this term only)

Variables selected and analyzed:

A team of reviewers carried out the data extraction; they were all well-trained in using the data extraction form, which was tested on several papers and improved as necessary. To reduce subjectivity and bias in the process, the reviewers worked independently. Any disagreements were settled by conversation or, if arbitration was required, by a third reviewer. The data extraction form included the following items:

1. **Study characteristics:** Information such as the author(s), year of publication, country of origin, study design,

sample size, and duration of follow-up were recorded.

2. **Population details:** Data about participant demographics (age, sex, ethnicity), clinical presentation, diagnostic criteria for emphysema, and severity of disease were extracted.
3. **AI-based imaging modalities:** Specifics regarding the AI algorithms used (e.g., type of machine learning or deep learning model), training and validation of the models, imaging techniques employed, and any pre-processing steps were noted.
4. **Outcomes:** Primary outcome measures, including sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratios, diagnostic odds ratio, and area under the receiver operating characteristic (ROC) curve, were extracted. Information regarding secondary outcomes, such as inter-reader variability, was also included.

Risk of bias and certainty bias assessment:

Each included study was evaluated using the QUADAS-AI tool [14], which guided several domains, including patient selection, index test, reference standard, flow and scheduling, and other AI-specific variables, including the setup and validation of the AI model. Concurrently, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach [15] was used to assess the certainty in the evidence for each significant outcome assessed across the included papers.

Results:

Study selection process:

In the initial stage of the study selection procedure, 426 entries were discovered throughout all evaluated databases, as illustrated in (figure 1). After 59 duplicate records were removed before screening, 367 records were left to be evaluated. Forty-two items were deleted because viewing the entire content was not currently possible. Consequently, an attempt was made to obtain 325 reports. It was not possible to extract 39 of these records for further examination. Every one of the 286 remaining reports was carefully assessed for eligibility. Fifty-seven publications violated the PICO criteria: 36 were deemed off-topic, 44 had merely case reports, 24 featured animal research, 47 featured scoping reviews, 39 featured seminar pieces, and the final 29 featured editorials. When these criteria were applied, ten papers [16–25] were deemed suitable for evaluation.

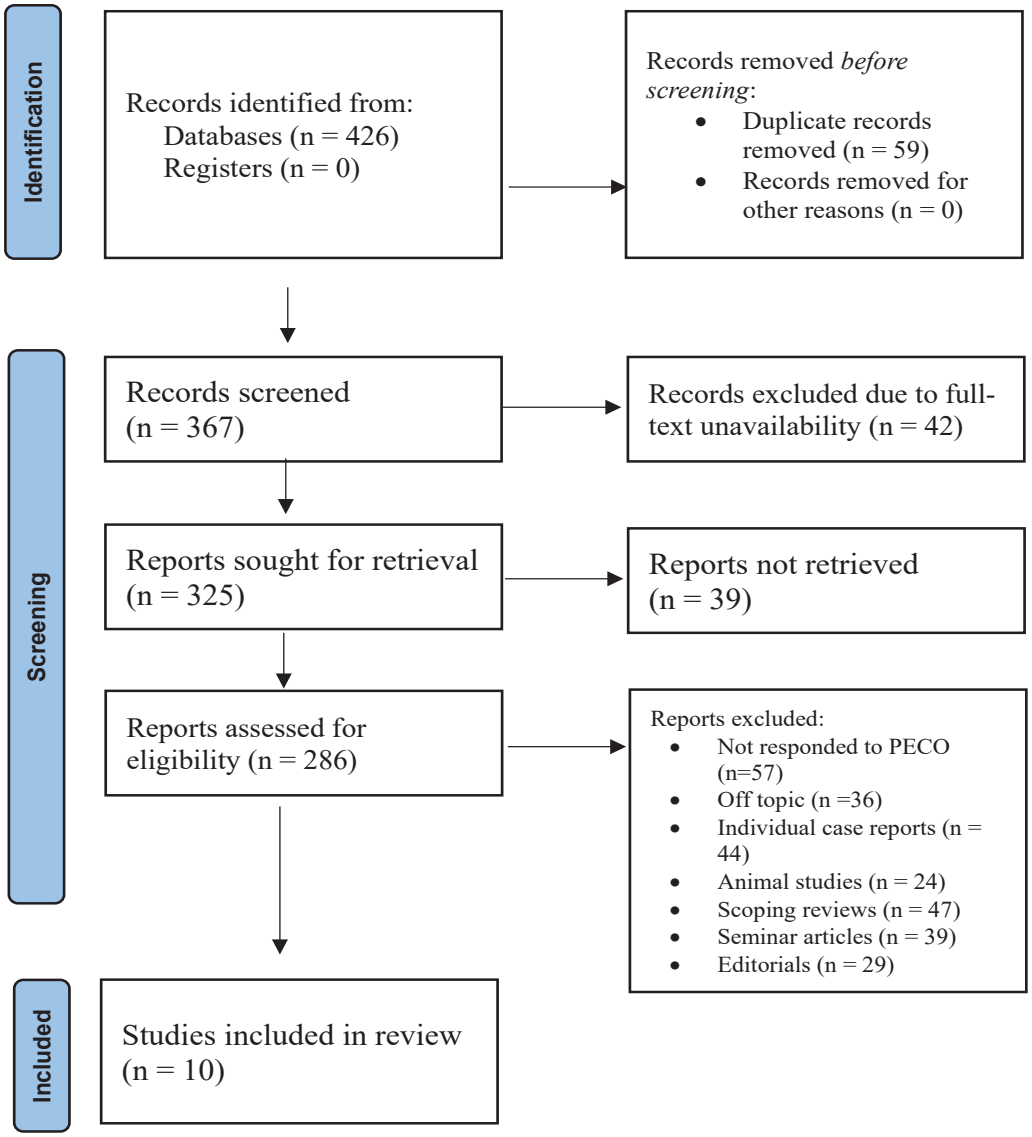


Figure 1: PRISMA protocol representation of the article selection process for the review

Assessed levels of bias:

Evaluating a subset of studies using the QUADAS-AI algorithm reveals that the bias risk varies between domains (figure 2). Research on patient selection by Binder et al. [16], Fischer et al. [18], Kumarasamy et al. [20], and Saad et al. [22] showed moderate bias, pointing to possible problems with participant selection that might affect results. While Fischer et al. [18], Kumarasamy et al. [20], Mendoza et al. [21], Saad et al. [22], and Xu et al. [25] showed moderate risk, suggesting potential inconsistencies in test applications, Binder et al. [16] and Campo et al. [17] reported low-risk in the Index Test domain, indicating reliable AI algorithm application.

Except for three research (Campo et al., Humphries et al., and Xu et al.) that

reported moderate risk, most studies for the Reference Standard achieved low risk, indicating trusted validation techniques. While some moderate risks were in Binder et al. [16], Humphries et al. [19], Mendoza et al. [21], and Sørensen et al. [23,] these risks were typically low and suggested potential timing and sequence problems.

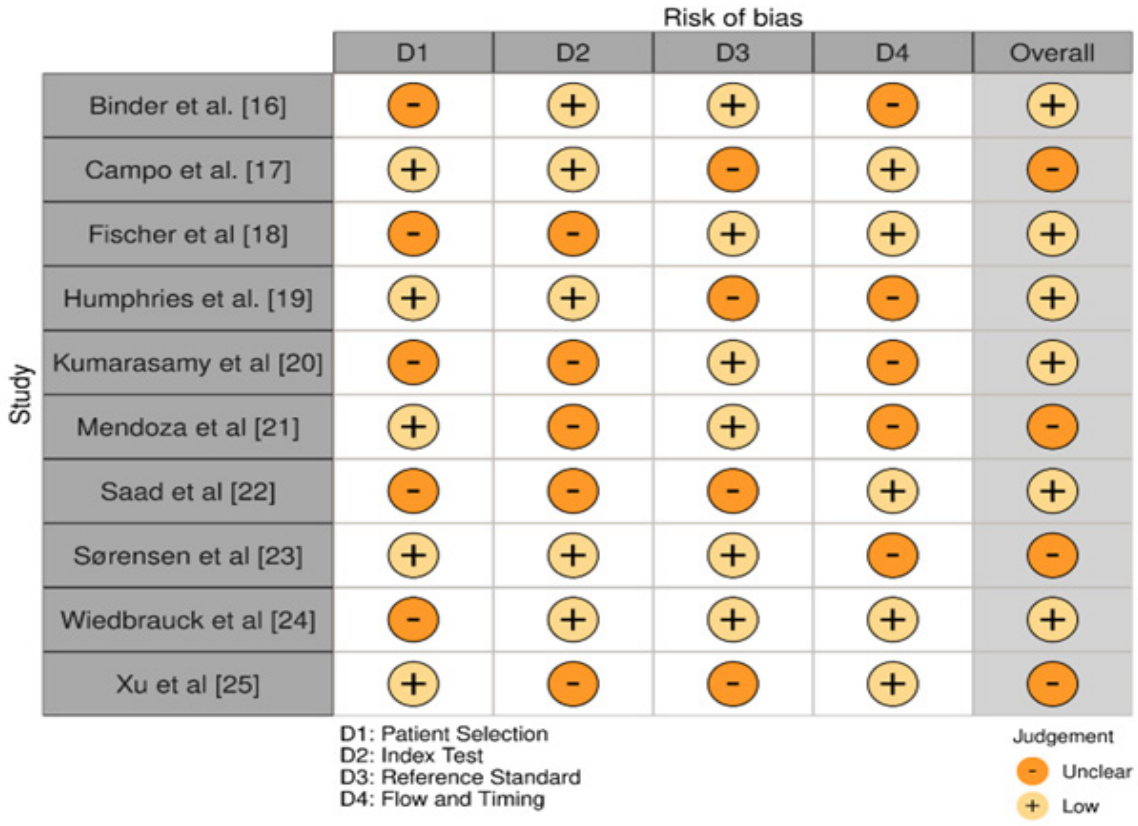


Figure 2: Bias assessment using the QUADAS-AI tool.

AI-based imaging and its correlation with clinical metrics:

The description of clinical trials [18, 20, 22-24] evaluating the effectiveness of AI-based imaging modalities in identifying and assessing emphysema is given in Table 3, along with a correlation between the results and clinical metrics. Comprising a sample size of 141 patients, Fischer et al. [18] used deep convolutional networks with adversarial training to analyze CT scans of the lungs. The Tiffeneau index,

a crucial spirometry parameter, averages 0.57 ± 0.13. AI measurement of emphysema using two distinct approaches yielded results of 9.96% ± 11.87% and 8.04% ± 10.32%. Strong negative correlations were seen in the study between the quantifications and the Tiffeneau index ($\rho = -0.86$ and $\rho = -0.85$, $p < 0.0001$).

Kumarasamy et al. [20] used SIEMENS AI Rad software to analyze ultra-low dose CT scan data from 308 patients, most of whom were male and between the ages of 65 and 79. The FEV1/FVC ratios and the FEV1 absolute values were among the spirometry measurements. While the AI emphysema score was only moderately correlated ($r = 0.335$, $p < 0.001$) with radiologist assessments, it was somewhat higher than the scores determined by radiologists (AUC = 0.718). Saad et al. [22] used Coreline Soft's AVIEW software on non-contrast CT scans to assess 80 cases with different COPD severities. The study classified the severity of COPD using the Tiffeneau index and discovered substantial differences in the D-value and the inspiratory LAA-950% between COPD severity groups and strong correlations between these measures and COPD severity.

Sørensen et al.'s [23] use of texture analysis and densitometry on CT scans involved 1915 Danish Lung Cancer Screening Trial participants. They looked at relationships with COPD diagnoses, correlations between AI measurements and anticipated values for FEV1%, and forecasts of lung function deterioration. Compared to densitometric measures, texture analysis demonstrated a more significant correlation with lung function and higher diagnostic accuracy for COPD stages. Using AI, Wiedbrauck et al. [24] examined 298 patients' low attenuation volume percentage (LAV%) on standard chest CT and spirometry testing, with a threshold of less than -950 HU. The FEV1/FVC ratio and LAV% showed a strong negative connection ($\boldsymbol{\rho} = -0.477$, $P < 0.001$), suggesting that higher LAV% is linked to more severe airway obstruction and higher visual emphysema grades based on spirometric criteria.

Table 3: Summary of studies evaluating AI-based imaging modalities for emphysema detection and correlation with clinical metrics.

Study ID	Patient sample size and demographics	AI modality employed	Relevant spirometry metrics	Key findings on AI efficacy	Correlation with clinical metrics	Conclusion drawn
Fischer et al [18]	141 patients (72 women, 69 men) aged 23-91 years	Deep convolutional networks with adversarial training for lung CT	Mean Tiffeneau index (TI): 0.57 ± 0.13	Mean emphysema quantification: 9.96% ± 11.87% and 8.04% ± 10.32% using two methods	Very strong Spearman correlation with TI: $\rho = -0.86$ and $\rho = -0.85$ ($p < 0.0001$)	AI demonstrates high accuracy in quantifying emphysema, strongly correlating with clinical pulmonary function tests.
Kumarasamy et al [20]	308 participants, 81% male, aged 65-79 years	SIEMENS AI Rad software for ultra-low dose CT	FEV1/FVC ratio: 0.69 (0.11); FEV1: 86.26 (20.44)	AI emphysema score: 4.82 (5.56); Radiologist score: 2.94 (4.2)	Moderate correlation with radiologist scores ($r = 0.335$, $p < 0.001$); AUC for AI: 0.718	AI is effective in detecting emphysema in LC screening, though slightly less effective than radiologist assessment.
Saad et al [22]	80 cases with varying degrees of COPD severity	Coreline Soft's AVIEW, non-contrast CT using density masks	Tiffeneau index used for COPD severity classification	Significant distinctions in Inspiratory LAA-950% and D-value among COPD severity groups	Significant relationship between Inspiratory LAA-950%, D-value and COPD severity	AI effectively differentiates between severity levels of COPD, validating its use in clinical settings.
Sørensen et al [23]	1915 patients from Danish Lung Cancer Screening Trial, baseline low-dose lung CT scans from 2004-2010	Texture analysis and densitometry on CT scans; evaluated individually and in combination	FEV1% predicted; association with mild to severe COPD diagnosis and lung function decline prediction	Texture showed a stronger association with lung function and higher diagnostic AUC for COPD stages compared to densitometric measures	Texture analysis demonstrated significant correlations with lung function and predictive capability for rapid lung function decline	AI-based texture analysis proves to be more effective than densitometry in assessing COPD severity and predicting lung function decline
Wiedbrauck et al [24]	298 patients, routine chest CT and spirometry tests	AI software quantifying low attenuation volume percentage (LAV%) with a threshold < -950 HU	FEV1/FVC ratio; used to indicate airway obstruction	Significant correlations between LAV% and FEV1/FVC; higher LAV% associated with higher visual emphysema grades	Strong negative correlation between LAV% and FEV1/FVC ratio ($\varrho = -0.477$, $P < 0.001$)	AI-based LAV% quantification effectively correlates with both spirometric and visual assessments of emphysema, supporting its clinical use

Advances in AI algorithms pertaining to emphysema detection:

Table 4 represents the papers [16-17, 19, 21, and 25] that investigate the effectiveness of AI imaging modalities for identifying and categorizing emphysema. Using CT scans split into patches and examined using a generative model with variational expectation-maximization, Binder et al. [16] examined participants from the COPDGene research. By classifying patches according to illness classifications that closely matched clinical disease subtypes, our method allowed for the unsupervised detection and clustering of patients. The study found that this method improves the knowledge and categorization of COPD subtypes by sophisticated AI techniques despite the lack of reported sensitivity and specificity.

Using CT scans and simulated X-rays, Campo et al. [17] examined a sizable dataset comprising 7,377 individuals. We used an 11-layer convolutional neural network (CNN) trained on 2,705 pictures to determine the low attenuation area percentage (LAA%). According to the study, the average error for emphysema percentages was 3.96 ± 4.60. At 3%, 5%, and 10% thresholds, respectively, the corresponding sensitivity and specificity were 67.96% and 80.18%, 78.36% and 76.90%, and 85.68% and 80.42%. Humphries et al. [19] employed deep learning methods incorporating LSTM and convolutional neural networks to study a sizable population from the COPDGene and ECLIPSE research. Comparing the algorithm's classification to conventional visual assessments, 34% of scans were deemed more severe and 13% less severe. The study emphasizes the importance of deep learning in improving diagnostic and prognostic evaluations in COPD.

Using methods including ROI scaling, Local Binary Patterns (LBPs), Kernel Density Estimation (KDE), and the k-nearest neighbors (kNN) classifier, Mendoza et al. [21] examined 342 CT scans from the COPDGene project. In leave-one-out validation, the KDE had the highest mean success rate (0.664 ± 0.2175). Xu et al. [25] used multidetector row CT (MDCT) to evaluate 34 human volunteers and categorized them into various groups according to the severity of COPD as determined by pulmonary function tests (PFTs). Using a Bayesian classifier and 3-D texture data, the study showed improved sensitivity in identifying emphysematous alterations across different stages of COPD. For example, the sensitivity was 99% in 3-D and 72% in 2-D for the NN category (non-smokers with normal PFTs), while the specificity was 100% and 75% in the same range.

Table 4: Advances in AI algorithms for emphysema detection using imaging techniques.

Study ID	Population details	Imaging technique	AI/Algorithm details	Classification results	Sensitivity/Specificity	Conclusion drawn
Binder et al. [16]	Patients with COPD from the COPDGene study	CT scans, divided into patches	Generative model with variational EM, clusters patients and classifies patches by disease subtypes	Accurate alignment with clinical disease subtypes	Not specifically reported	Enhanced understanding and classification of COPD through unsupervised detection and patient clustering
Campo et al. [17]	7,377 subjects with simulated X-rays and CT scans	CT scans and simulated X-rays; LAA% calculation	CNN with 11 layers, trained on 2,705 images	Mean error in emphysema percentages: 3.96 ± 4.60	- 3% threshold: Sensitivity 67.96%, Specificity 80.18% - 5% threshold: Sensitivity 78.36%, Specificity 76.90% - 10% threshold: Sensitivity 85.68%, Specificity 80.42%	Demonstrated potential of CNNs for accurate emphysema assessment from X-ray and CT simulations, enhancing diagnostic accuracy
Humphries et al. [19]	COPDGene: 10,192 smoking individuals aged 45-80, excluding non-COPD respiratory conditions; ECLIPSE: 2,164 COPD patients (stages 2-4) and 582 controls	CT scans	Deep learning using convolutional neural network and LSTM, trained on 2,407 COPDGene participants, tested on 7,143 COPDGene and 1,962 ECLIPSE participants	Algorithm classified 34% of scans as more severe and 13% as less severe compared to visual assessments	Not quantified due to lack of specific sensitivity/specificity data	Deep learning enhances diagnostic accuracy and prognostic evaluations in COPD, showing potential for refined diagnostic and prognostic processes in chronic disease management
Mendoza et al [21]	342 CT scans from COPDGene, smokers of various GOLD levels (0-4, U) and 42 non-smokers	Standardized CT scans across different resolutions	ROI sizing (24.18x24.18 mm), LBPs, KDE, kNN (k=5) classifier	Leave-one-out validation; KDE highest mean success rate (0.664 ± 0.2175)	Not specified	Demonstrates that KDE's simpler approach can achieve comparable accuracy to more complex methods, supporting its use in clinical settings for emphysema detection
Xu et al [25]	34 humans on MDCT, classified into EC, MC, NC, NN, NS based on COPD severity from PFTs	Multidetector row CT (MDCT)	24 3-D texture features, Bayesian classifier, leave-one-out validation	Sensitivity in 3-D/2-D: EC (85%/71%), MC (90%/82%), NC (88%/50%), NN (100%/60%)	NN (2-class): Sensitivity 99%/72%, Specificity 100%/75%	3-D texture analysis enhances the ability to discriminate early and subtle emphysematous changes, critical for targeted interventions in COPD management.

GRADE inferences observed:

As per the GRADE criteria assessment, the risk of bias across these studies was considered low to moderate (figure 3). This classification reflects rigorous methodologies, although some studies could have further minimized biases related to study design or data handling. Inconsistency among findings was low, indicating that the studies generally agreed on the effectiveness of AI in emphysema-related tasks. Indirectness was also low, as the studies addressed the research question with appropriate patient populations and relevant clinical metrics. Imprecision was rated as moderate, suggesting that while the results are promising, some studies had wide confidence intervals or small sample sizes that could affect the reliability of the findings. No additional factors were identified that significantly affected the certainty of the evidence.

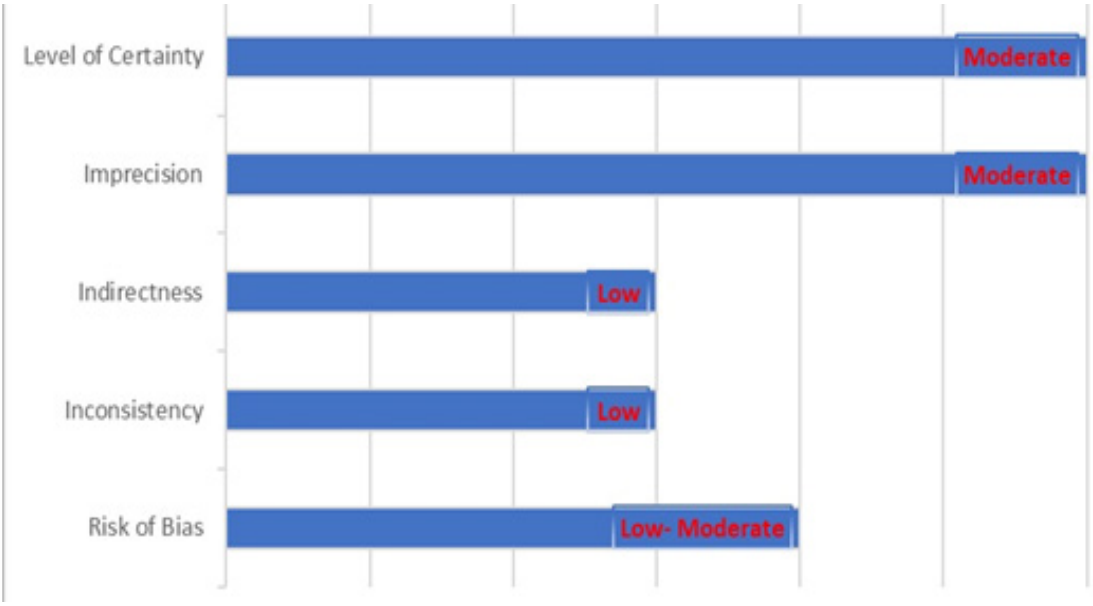


Figure 3: GRADE assessment observations of 10 reports for AI efficacy in emphysema quantification and diagnosis.

Discussion:

The thematic threads identified across the studies that we included in our review[16-25], ranging from AI as a supplementary tool to its potential to outperform traditional methods, highlight the multifaceted roles AI can assume in the evolving landscape of pulmonary healthcare. Fischer et al. [18] proved that AI accurately measured emphysema, correlating closely with clinical pulmonary function testing. This study aligns with the findings of Wiedbrauck et al. [24], who also highlighted the correlation between standard clinical assessments and AI-based quantification (LAV%). This highlights the potential and dependability of AI to be seamlessly integrated into clinical workflows.

Fischer et al.[18], however, depart from Kumarasamy et al. [20], who proposed that AI was marginally less successful than radiologist assessments in lung cancer screening, notwithstanding its effectiveness. This suggests that although AI is quite accurate in quantification tasks, its performance compared to that of human specialists may change depending on the application. Furthermore, compared to Sørensen et al. [23], which highlighted AI's superiority over traditional densitometry in texture analysis, Fischer et al. focus more broadly on quantification rather than specific methodological advancements. This implies that while Fischer et al.'s [18] findings do not address comparative methodological efficacy, they do generally support the integration of AI into clinical practice.

Kumarasamy et al. [20] found that AI was marginally less effective than radiologists in detecting emphysema during lung cancer screenings. This contrasts with Saad et al. [22] and Sørensen et al. [23], where AI was demonstrated to effectively discriminate severity levels of COPD and surpass densitometry in texture analysis, suggesting a key diagnostic role in both circumstances. Kumarasamy et al.'s [20] findings chime more with Mendoza et al. [21], where a more straightforward AI method (KDE) achieved equivalent accuracy to more complicated methods, demonstrating that more straightforward or less effective AI tools can still be helpful in clinical contexts when used to supplement traditional methods. Additionally, the emphasis on AI as an additional tool in Kumarasamy et al. contrasts with the findings of Binder et al.[16] and Humphries et al. [19], both of which explain AI's potential in strengthening diagnostic and prognostic processes, suggesting a more central position for AI in COPD care.

Saad et al. [22] highlighted AI's effectiveness in differentiating between COPD severity levels, underscoring its utility in direct clinical application. In terms of proving AI's practical superiority in particular analytical tasks (severity differentiation vs. texture analysis), this work aligns with Sørensen et al. [23]. According to both research, AI can supplement or replace conventional techniques. Saad et al. [22], on the other hand, depart from Kumarasamy et al. [20], where AI was found to be less valuable than radiologists in certain situations, suggesting that the application of AI can vary depending on the clinical context. Compared with Xu et al. [25], which focused on the early detection of modest emphysematous alterations, Saad et al. [22] present evidence for AI's usefulness across a more extensive spectrum of disease severities, suggesting a new focus on AI application within pulmonary treatment.

According to recent studies, combining various diagnostic modalities may significantly improve the identification and tracking of COPD. Combining quantified CT imaging (specifically, texture-based radiomic features), demographic information, and spirometric measurements instead of using each modality independently can enhance the predictive accuracy of machine learning models for COPD progression, according to a population-based study conducted within the CanCOLD cohort [26].

Furthermore, using biomarkers representing lung irregularities characteristic of COPD and emphysema may enhance AI-based COPD diagnosis. For example, Chen et al. [27] used deep learning analysis of blood RNA-seq gene expression in conjunction with CT imaging to find patterns associated with pulmonary structural alterations and inflammatory processes called image-expression axes (IEAs). An emphysema-specific IEA was found in this study, which included 1223 patients from the COPDGene project. This IEA was connected with lower lung function, increased neutrophil counts, and radiographic evidence of central emphysema. On the other hand, more significant airways and higher white blood cell counts were associated with the airway-specific IEA. Significantly, those who scored highly on the IEA airway and emphysema indicators also had worse clinical outcomes, as shown by higher mMRC and SGRQ scores and shorter 6-minute walk test distances [27].

These results highlight the potential to further our understanding of the structural and inflammatory mechanisms underlying emphysema by combining blood transcriptomics, CT imaging, and phenotypic data in AI models. Sputum [28-29] and more biomarkers [30-31] could be added to improve AI's capacity to precisely diagnose emphysema, anticipate exacerbations, and better patient outcomes. These thorough diagnostic techniques are essential for putting the new treatable characteristic strategy in emphysema therapy into practice since they offer a more individualized treatment plan based on specific prognostic markers[32].

Wang et al.[33] highlighted that AI could reliably identify and characterize several subtypes of COPD, such as emphysema, which is consistent with our results of AI's strong quantification capabilities. Their research also complemented our overall observations on the diagnostic adaptability of AI by elaborating on the kinds of distinctive factors that AI might analyze, such as bronchial wall thickness and bronchiectasis, which were not particularly included in our review.

A one-dimensional fully convolutional

neural network was described by Estepar et al. [34]; it achieved high discrimination (AUC 0.91) by effectively differentiating between emphysema and other COPD subtypes from flow-volume curves. This result confirms what we observed from Sørensen et al. [23] on the superiority of AI over conventional techniques in texture analysis. However, Estepar et al.'s [34] emphasis on using AI for multiclass classification provides a more comprehensive picture of the possible uses of AI in pulmonary illness differential diagnosis, a complex topic that our review did not fully address.

In their discussion of AI's implications in emphysema management, Robertson et al. [35] emphasized the technology's ability to lower healthcare costs and increase diagnostic specificity, particularly in countries with limited resources. The results of our review, which highlighted the use of AI in clinical practice, somewhat supported this viewpoint. However, Robertson et al.[35] offered a more comprehensive perspective on the global applications of AI, particularly in low- and middle-income nations, and emphasized the significance of a multimodal strategy that integrates CT imaging, demographic information, and spirometry. This all-encompassing strategy for incorporating AI into diagnostic procedures supports our findings while also going beyond them by outlining a framework for widespread adoption and highlighting AI's contribution to increasing access and decreasing inequities.

Limitations:

A few limitations can be said to have affected the findings and the generalizability of our findings. First, the variation in AI approaches among the included studies was a significant obstacle to thoroughly synthesizing data. AI systems differed not just in their algorithmic architecture, which may be as simple as machine learning models or as complicated as deep learning networks, but also in terms of their training datasets' size, quality, and demographic representation. This unpredictability hampered our capacity to conduct a meta-analysis and reach a consensus regarding the best AI method for emphysema detection.

Another drawback was the use of retrospective studies, which contain biases in selection and may overstate the effectiveness of AI systems since instances are pre-selected. Our knowledge of the efficacy of these AI modalities in clinical practice has been distorted by the underrepresentation of prospective studies in the literature, which are more representative of real-world performance.

Additionally, most studies did not offer a thorough external validation of the AI models. Particularly in medical applications, the generalizability of AI discover-

ies is essential, and the absence of rigorous external validation raises questions about how well these AI systems work across various clinical contexts and demographics. If this kind of evaluation is lacking, AI system diagnostic accuracy may be overestimated.

Clinical recommendations:

Several recommendations for the integration and implementation of AI in pulmonary healthcare may be drawn from the thematic threads found in the evaluated papers. First, AI has proven to be quite accurate in measuring lung diseases like emphysema, and it strongly correlates with clinical pulmonary function testing. This implies that clinical procedures for quantifying lung illnesses can dependably incorporate AI. It is advised that to improve the accuracy and efficiency of diagnosis, healthcare providers consider implementing AI tools.

However, the effectiveness of AI in comparison to human experts, like radiologists, seems to rely on the particular use case. For instance, AI performed marginally worse than radiologists in lung cancer tests. This suggests that, especially in complicated diagnostic circumstances, AI must be integrated carefully and contextually to enhance human skill rather than take its place. In situations when AI tools may not yet be able to match human performance, healthcare systems should make sure that they are employed to augment radiologist assessments.

Furthermore, the results imply that less complex AI techniques can benefit and prosper in therapeutic contexts. This suggests that even simpler AI tools, which could be easier to use and less expensive, can have a significant impact when combined with conventional diagnostic techniques. In light of their unique requirements and available resources, healthcare organizations ought to think carefully about implementing both basic and sophisticated AI technologies.

Conclusion:

As per the results of our investigation, AI-based imaging modalities show promise for enhancing emphysema detection and quantification. AI was found to quantify emphysema from imaging data and correlate well with clinical pulmonary function tests. The efficiency of various AI approaches and clinical contexts varied, but in some cases, such as lung cancer screenings, some AI models were on par with or even better than traditional radiologist assessments.

Acknowledgments: The author extends their appreciation to the team of reviewers of the faculty members of Majmaah University for helping with this research and

is grateful to them for their support and assistance in this investigation.

Conflicts of Interest: The author declares no conflict of interest.

References:

1. Foreman KJ, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, et al. Forecasting life expectancy years of life lost and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. Lancet. 2018;392(10159):2052–2090. doi: 10.1016/S0140-6736(18)31694-5.
2. Diab N, Gershon AS, Sin DD, Tan WC, Bourbeau J, Boulet L-P, et al. Underdiagnosis and overdiagnosis of chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2018;198(9):1130–1139. doi: 10.1164/rccm.201804-0621CI.
3. Ruparel M, Quaife SL, Dickson JL, Horst C, Tisi S, Hall H, et al. Prevalence Symptom Burden and Underdiagnosis of Chronic Obstructive Pulmonary Disease in a Lung Cancer Screening Cohort. Ann Am Thorac Soc. 2020;17(7):869–878. doi: 10.1513/AnnalsATS.201911-857OC.
4. Hoidal JR, Niewoehner DE. Pathogenesis of emphysema. Chest. 1983;83(4):679–685. doi: 10.1378/chest.83.4.679.
5. Gonzalez J, Marín M, Sánchez-Salcedo P, Zulueta JJ: Lung cancer screening in patients with chronic obstructive pulmonary disease. Ann Transl Med 4(8), 2016. http://atm.amegroups.com/article/view/9852
6. Durawa A, Dziadziuszko K, Jelitto-Górska M, Szurowska E. Emphysema-The review of radiological presentation and its clinical impact in the LDCT screening era. Clin Imaging. 2020;64:85–91. doi: 10.1016/j.clinimag.2020.04.018.
7. Schilham AMR, Van Ginneken B, Gietema H, Prokop M. Local noise weighted filtering for emphysema scoring of low-dose CT images. IEEE Trans Med Imaging. 2006;25(4):451–463. doi: 10.1109/TMI.2006.871545.
8. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced Lung Cancer Mortality with Low-Dose Computed Tomographic Screening. N Engl J Med. 2011;365(5):395–409. doi: 10.1056/NEJMoa1102873
9. Ghonge NP, Chowdhury V. Minimum-intensity projection images in high-resolution computed tomography lung: Technology update. Lung India Off Organ Indian Chest Soc.

2018;35(5):439. doi: 10.4103/lungindia.lungindia_489_17.

10. Lan H, Nishitani H, Nishihara S, Ueno J, Takao S, Iwamoto S, et al. Using the MDCT thick slab MinIP method for the follow-up of pulmonary emphysema. J Med Invest. 2011;58(3–4):175–179. doi: 10.2152/jmi.58.175.
11. Remy-Jardin M, Remy J, Gosselin B, Copin MC, Wurtz A, Duhamel A. Sliding thin slab minimum intensity projection technique in the diagnosis of emphysema: Histopathologic-CT correlation. Radiology. 1996;200(3):665–671. doi: 10.1148/radiology.200.3.8756912.
12. Satoh S, Ohdama S, Shibuya H. Sliding thin slab minimum intensity projection imaging for objective analysis of emphysema. Radiat Med. 2006;24(6):415–421. doi: 10.1007/s11604-006-0045-y.
13. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021 Mar 29;372:n71. doi: 10.1136/bmj.n71. PMID: 33782057; PMCID: PMC8005924.
14. Sounderajah, V., Ashrafian, H., Rose, S. et al. A quality assessment tool for artificial intelligence-centered diagnostic test accuracy studies: QUADAS-AI. Nat Med 27, 1663–1665 (2021). https://doi.org/10.1038/s41591-021-01517-0
15. Balshem H., Helfand M., Schünemann H.J., Oxman A.D., Kunz R., Brozek J., Vist G.E., Falck-Ytter Y., Meerpohl J., Norris S., et al. GRADE guidelines: 3. Rating the quality of evidence. J. Clin. Epidemiol. 2011; 64:401–406. doi: 10.1016/j.jclinepi.2010.07.015
16. Binder P, Batmanghelich NK, Estepar RS, Golland P. Unsupervised Discovery of Emphysema Subtypes in a Large Clinical Cohort. Mach Learn Med Imaging. 2016 Oct;10019:180-187. doi: 10.1007/978-3-319-47157-0_22. Epub 2016 Oct 1. PMID: 28232959; PMCID: PMC5317320.
17. Campo MI, Pascau J, José Estépar RS. EMPHYSEMA QUANTIFICATION ON SIMULATED X-RAYS THROUGH DEEP LEARNING TECHNIQUES. Proc IEEE Int Symp Biomed Imaging. 2018 Apr;2018:273-276. doi: 10.1109/ISBI.2018.8363572. Epub 2018 May 24. PMID: 30450153; PMCID: PMC6239425.

18. Fischer AM, Varga-Szemes A, van Assen M, Griffith LP, Sahbaee P, Sperl JI, Nance JW, Schoepf UJ. Comparison of Artificial Intelligence-Based Fully Automatic Chest CT Emphysema Quantification to Pulmonary Function Testing. AJR Am J Roentgenol. 2020 May;214(5):1065-1071. doi: 10.2214/AJR.19.21572. Epub 2020 Mar 4. PMID: 32130041.
19. Humphries SM, Notary AM, Centeno JP, Strand MJ, Crapo JD, Silverman EK, Lynch DA; Genetic Epidemiology of COPD (COPDGene) Investigators. Deep Learning Enables Automatic Classification of Emphysema Pattern at CT. Radiology. 2020 Feb;294(2):434-444. doi: 10.1148/radiol.2019191022. Epub 2019 Dec 3. PMID: 31793851; PMCID: PMC6996603.
20. Kumarasamy, C. & Bennett, K. & Adler, B. & Murray, C. & DeKlerk, N. & Brims, F.. (2022). EP01.03-003 AI Detection of Emphysema in an Ultra-Low Dose CT Lung Cancer Screening Program. Journal of Thoracic Oncology. 17. S172. 10.1016/j.jtho.2022.07.285.
21. Mendoza CS, Washko GR, Ross JC, Diaz AA, Lynch DA, Crapo JD, Silverman EK, Acha B, Serrano C, Estépar RS. EMPHYSEMA QUANTIFICATION IN A MULTI-SCANNER HRCT COHORT USING LOCAL INTENSITY DISTRIBUTIONS. Proc IEEE Int Symp Biomed Imaging. 2012:474-477. doi: 10.1109/ISBI.2012.6235587. PMID: 23743800; PMCID: PMC3670097.
22. Saad, M.M., Bayoumy, A.A., EL-Nisr, M.M. et al. Assessment of artificial intelligence-aided chest computed tomography in diagnosis of chronic obstructive airway disease: an observational study. Egypt J Radiol Nucl Med 54, 97 (2023). https://doi.org/10.1186/s43055-023-01043-8
23. Sørensen L, Nielsen M, Petersen J, Pedersen JH, Dirksen A, de Bruijne M. Chronic Obstructive Pulmonary Disease Quantification Using CT Texture Analysis and Densitometry: Results From the Danish Lung Cancer Screening Trial. AJR Am J Roentgenol. 2020 Jun;214(6):1269-1279. doi: 10.2214/AJR.19.22300. Epub 2020 Apr 7. PMID: 32255690.
24. Wiedbrauck D, Karczewski M, Schoenberg SO, Fink C, Kayed H. Artificial Intelligence-Based Emphysema Quantification in Routine Chest Computed Tomography: Correlation With Spirometry and Visual Emphysema Grading. J Comput Assist Tomogr. 2023 Dec 18. doi: 10.1097/RCT.0000000000001572. Epub ahead of print. PMID: 38110294.

25. Xu Y, Sonka M, McLennan G, Guo J, Hoffman EA. MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. IEEE Trans Med Imaging. 2006 Apr;25(4):464-75. doi: 10.1109/TMI.2006.870889. PMID: 16608061.
26. Makimoto K,Hogg JC,Bourbeau J,Tan WC,Kirby M. CT imaging with machine learning for predicting progression to COPD in individuals at risk. Chest. 2023;164(5): 1139-1149. doi: https://doi.org/10.1016/j.chest.2023.06.008
27. Chen J,Xu Z,Sun L,et al. Deep learning integration of chest computed tomography imaging and gene expression identifies novel aspects of COPD. Chronic Obstr Pulm Dis. 10(4): 355-368. doi: https://doi.org/10.15326/jcopdf.2023.0399
28. Schaefer R,Khona M,Fiete IR. No free lunch from deep learning in neuroscience: a case study through models of the entorhinal-hippocampal circuit. bioRxiv. 2023:1-16. doi: https://doi.org/10.1101/2022.08.07.503109
29. Tang LYW,Coxson HO,Lam S,Leipsic J,Tam RC,Sin DD. Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. Lancet Digit Health. 2020;2(5):e259-e267. doi: https://doi.org/10.1016/S2589-7500(20)30064-9
30. Esther CR Jr,O'Neal WK,Anderson WH,et al. Identification of sputum biomarkers predictive of pulmonary exacerbations in COPD. Chest. 2022;161(5):1239-1249. doi: https://doi.org/10.1016/j.chest.2021.10.049
31. Xiong Y,Ba X,Hou A,et al. Automatic detection of mycobacterium tuberculosis using artificial intelligence. J Thorac Dis. 2018;10(3):1936-1940. doi: https://doi.org/10.21037/jtd.2018.01.91
32. Agusti A,Bel E,Thomas M,et al. Treatable traits: toward precision medicine of chronic airway diseases. Eur Respir J. 2016;47(2):410-419. doi: https://doi.org/10.1183/13993003.01359-2015
33. Wang F, Li S, Gao Y, Li S. Computed tomography-based artificial intelligence in lung disease—Chronic obstructive pulmonary disease. MedComm – Future Med. 2024; 3:e73. doi:10.1002/mef2.73
34. Estépar RSJ. Artificial Intelligence in COPD: New Venues to Study a Complex Disease. Barc Respir Netw Rev. 2020 May-Dec;6(2):144-160. doi: 10.23866/BRNRev:2019-0014. PMID: 33521399; PMCID: PMC7842269.

35. Robertson NM, Centner CS, Siddharthan T. Integrating Artificial Intelligence in the Diagnosis of COPD Globally: A Way Forward. Chronic Obstr Pulm Dis. 2024 Jan 25;11(1):114-120. doi: 10.15326/jcopdf.2023.0449. PMID: 37828644; PMCID: PMC10913925.