INNOVAPATH™
The Premier Journal Bridging Discovery & Innovation

Review Article
https://doi.org/10.63501/pj8sbm22

Artificial Intelligence in Pulmonary Medicine and Critical Care: *A Comprehensive Review of Current Applications, Challenges, and Future Directions*

Gurjot Garcha MD[1], Michael Scott, DO[2], Sohail Rao, MD, MA, DPhil[3], Rameez Rao, MD[4]

[1]Orlando Health, 1900 Don Wickman Drive, Clermont, FL 34711

[2]Oklahoma Pulmonary Care, 1245 N Birch Ave, Broken Arrow, OK 74012

[3]INNOVACORE™ Center for Research & Biotechnology, 6918 Camp Bullis Road, San Antonio, TX 78256.

[4]Orlando Health, 52 W Underwood Street, FL 32806.

[4]Corresponding Author Email: rameez.rao@orlandohealth.com

ABSTRACT:

Background: Artificial intelligence (AI) and machine learning (ML) technologies are rapidly transforming pulmonary medicine and critical care, offering unprecedented opportunities for improved diagnosis, prognosis, and patient management. The intensive care unit (ICU) generates vast amounts of real-time data, creating an ideal environment for AI applications, while pulmonary medicine benefits from advanced imaging analysis and predictive modeling.

Objective: This comprehensive review examines the current landscape of AI applications in pulmonary medicine and critical care, evaluating their clinical utility, performance metrics, implementation challenges, and future directions.

Methods: A systematic literature search was conducted in PubMed, Web of Science, Scopus, and the IEEE databases for studies published between 2020 and 2025, focusing on AI/ML applications in respiratory disease diagnosis, mechanical ventilation optimization, sepsis prediction, sleep medicine, pulmonary function test interpretation, and chronic respiratory disease management.

Results: AI systems demonstrate significant promise across multiple domains: lung cancer detection achieves over 90% sensitivity with reduced false positives by up to 30%; Acute Respiratory Distress Syndrome (ARDS) prediction models show pooled sensitivity of 0.81 and specificity of 0.82; sepsis prediction algorithms can alert clinicians up to 12 hours before clinical onset with AUC of 0.83; sleep apnea diagnosis using deep learning achieves 97% sensitivity for moderate-to-severe cases; and AI-based pulmonary function test interpretation achieves 86.6% diagnostic accuracy compared to 65.8% for pulmonologists. The FDA has authorized more than 950 AI-enabled medical devices, with pathology and radiology as major application areas.

Discussion: Despite promising performance metrics, significant barriers to clinical implementation persist. Model transparency remains a critical concern, as many algorithms function as "black boxes" unsuitable for high-stakes medical decisions. Validation studies are predominantly single-center and retrospective, limiting generalizability across diverse populations and healthcare settings. Regulatory frameworks are evolving, with the FDA's Predetermined Change Control Plans pathway and the EU AI Act establishing new compliance requirements. Workflow integration challenges, including alarm fatigue and IT infrastructure demands, must be addressed. Ethical considerations regarding data privacy, algorithmic bias, and liability for AI-assisted decisions require ongoing attention. Importantly, the deployment of AI must maintain a qualified human clinician as the final arbitrator of all diagnostic and therapeutic choices, ensuring that AI augments rather than replaces physician expertise. Future directions include multimodal data integration, federated learning for privacy-preserving research, and explainable AI approaches to enhance clinician trust within a supervised deployment framework.

Conclusions: AI technologies offer transformative potential for pulmonary medicine and critical care, though successful implementation requires addressing challenges including model transparency, workflow integration, regulatory compliance, and validation across diverse populations. Multi-center prospective trials are needed to establish clinical benefit and patient outcomes.

KEYWORDS: *artificial intelligence; machine learning; deep learning; pulmonary medicine; critical care; mechanical ventilation; sepsis; ARDS; lung cancer; sleep apnea; COPD; interstitial lung disease; pulmonary function tests; spirometry; clinician decision support*

INTRODUCTION:

The integration of artificial intelligence (AI) into healthcare represents one of the most significant technological advances of the 21st century, with pulmonary medicine and critical care emerging as particularly promising domains for these technologies. The evolution of AI in healthcare, introduced in the early 1970s, has accelerated dramatically, with an exponential rise in AI-related scholarly publications since 2014 (Karthika *et al*., 2024). The United States Food and Drug Administration (FDA) and European Conformité Européenne have approved more than 300 AI-based software and medical devices, most of which are intended for pulmonary imaging applications.

The intensive care unit (ICU) represents both an ideal testing ground and a significant challenge for AI applications in medicine. Modern ICUs generate enormous volumes of continuous data from physiological monitors, laboratory systems, imaging studies, and electronic health records (EHRs), creating fertile ground for predictive model development, enhanced decision support tools, and personalized care approaches (Nikolaidis *et al*., 2025). However, the complexity and heterogeneity of critically ill patients present unique challenges for algorithm development and validation.

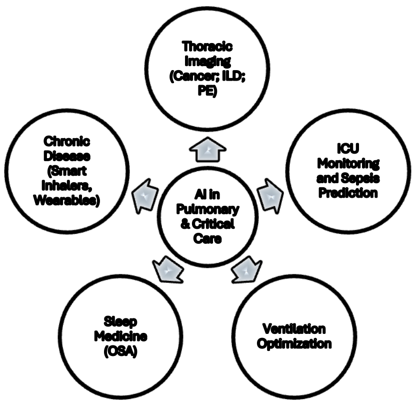
Chronic respiratory diseases, including chronic obstructive pulmonary disease (COPD) and asthma, affect hundreds of millions of people globally and represent a significant healthcare burden. An estimated 262 million people worldwide are affected by asthma, causing approximately half a million deaths annually, while COPD remains the third leading cause of death globally (Chen *et al*., 2025). AI-driven technologies, particularly machine learning (ML) and deep learning (DL), offer transformative potential to improve diagnostic accuracy, personalize treatment, predict exacerbations, and enhance patient self-management.

This comprehensive review examines the current state of AI applications across the spectrum of pulmonary medicine and critical care, from automated imaging analysis and disease classification to predictive modeling, pulmonary function test interpretation, and treatment optimization. We evaluate the evidence supporting various AI modalities, discuss implementation

https://doi.org/10.63501/pj8sbm22

challenges, and provide insights into future directions for this rapidly evolving field (Figure 1).

Figure 1: Overview of Artificial Intelligence Applications in Pulmonary Medicine and Critical Care.



Abbreviations: ILD: Interstitial Lung Disease; PE: Pulmonary Embolism; OSA; Obstructive Lung Disease

METHODS:

A comprehensive literature search was conducted across multiple electronic databases, including PubMed/MEDLINE, Web of Science, Scopus, IEEE Xplore, and the Cochrane Library. The search encompassed articles published from January 2020 through December 2025, with particular emphasis on recent systematic reviews, meta-analyses, and original research studies. Search terms included combinations of: "artificial intelligence," "machine learning," "deep learning," "pulmonary medicine," "respiratory care," "critical care," "intensive care unit," "mechanical ventilation," "sepsis," "ARDS," "lung cancer," "COPD," "asthma," "sleep apnea," "interstitial lung disease," "pulmonary function test," "spirometry," and "digital health."

Studies were included if they: (1) evaluated AI/ML algorithms for diagnosis, prognosis, or management of pulmonary or critical care conditions; (2) reported performance metrics such as sensitivity, specificity, accuracy, or area under the receiver operating characteristic curve (AUC); (3) were published in peer-reviewed journals in English; and (4) involved adult patient populations. Exclusion criteria included case reports, conference abstracts without full text, studies with fewer than 50 participants, and articles that focused solely on theoretical AI concepts without clinical application.

Data extracted from included studies encompassed: study design, patient population characteristics, AI/ML methodology employed, input data types, primary outcomes, performance metrics, validation approaches, and identified limitations. Given the heterogeneity of AI applications and outcome measures, a narrative synthesis approach was employed, with quantitative data summarized where meta-analytic pooling was appropriate.

RESULTS:

The key findings of AI applications in pulmonary medicine and critical care are summarized in Table 1, and detailed performance metrics from individual studies are presented in Table 2. A comparative visualization of AI model performance across applications is provided in Figure 2.

Table 1: Summary of AI Applications in Pulmonary Medicine and Critical Care

Application Domain	Key Performance Metrics	Primary AI Methods	Clinical Readiness
Lung Cancer Detection	Sensitivity >90%; FP reduction 30%; AUC 0.94-0.97	CNN, Deep Learning, Ensemble Models	FDA-cleared devices available
Sepsis Prediction	AUC 0.83; 12-hour advance prediction; Accuracy 0.91-0.96	Random Forest, XGBoost, LSTM	Limited clinical implementation
ARDS Prediction	Sensitivity 0.81; Specificity 0.82; AUC 0.88	ML Classifiers, Neural Networks	Research/Validation phase
Mechanical Ventilation	Weaning prediction AUC >0.80; R-C detection 99.4%	Reinforcement Learning, RNN	Emerging applications
Sleep Apnea Diagnosis	Sensitivity 97%; Accuracy 88-91%; R^2 0.92-0.96	Deep Learning (OxiNet), XGBoost	Commercial products available
PFT Interpretation	AI accuracy 86.6% vs 65.8% pulmonologists; 100% pattern recognition	KNN, Decision Trees, Neural Networks, XAI	Commercial software (ArtiQ.PFT)

https://doi.org/10.63501/pj8sbm22

COPD Management	High accuracy lung sound classification	CNN, ANN, Wearable Integration	Smart inhalers available
Interstitial Lung Disease	IPF diagnosis; 3-year survival prediction	CNN, RadImageNet, CALIPER	FDA-cleared (Fibresolve)
Pulmonary Embolism	High specificity and NPV; Automated detection	DCNN, NLP	FDA-approved (Aidoc)
Asthma Management	26% reduction in inhaler errors	ML with smart inhalers	Commercial platforms

Abbreviations: AUC, area under the curve; CNN, convolutional neural network; DCNN, deep convolutional neural network; FP, false positive; KNN, K-nearest neighbors; LSTM, long short-term memory; ML, machine learning; NLP, natural language processing; NPV, negative predictive value; PFT, pulmonary function test; R-C, resistance-compliance; RNN, recurrent neural network; XAI, explainable artificial intelligence.

AI in Lung Cancer Detection and Screening:

Lung cancer remains the leading cause of cancer-related deaths worldwide, accounting for 2.5 million new cases and 18.7% of total cancer deaths globally. AI-driven lung cancer screening has demonstrated remarkable diagnostic performance, achieving over 90% sensitivity compared to 70-80% with traditional methods, while reducing false positives by up to 30% (Salvatore *et al*., 2025). Deep learning algorithms, particularly convolutional neural networks (CNNs), have demonstrated the ability to detect smaller nodules than human radiologists and significantly reduce both false positives and false negatives (Table 1).

A systematic review of AI performance in lung cancer detection on CT thorax by Tan *et al*. (2025) evaluated 14 studies, with seven in the detection subgroup and eight in the classification subgroup. The introduction of lung cancer screening programs is expected to increase imaging volumes, placing greater demands on radiologists. AI models have the potential to detect and classify pulmonary nodules without compromising diagnostic performance, thereby addressing workforce constraints.

Hendrix *et al*. (2023) developed a deep learning-based AI system that demonstrated sensitivity of 96.9% for detecting primary lung cancers and 92.0% for metastases, with clinically acceptable false-positive rates, in external validation (Table 2). Kashyap *et al*. (2025) further advanced the field with an ensemble deep learning model for automated lung tumor detection and segmentation on CT scans, potentially reducing the labor-intensive nature of manual delineation while decreasing physician variability.

Table 2: AI Algorithm Performance Metrics Across Key Studies

Study/Algorithm	Application	Sensitivity	Specificity	AUC/Accuracy	Sample Size/Validation
Hendrix *et al*. (2023)	Lung Cancer Detection	96.9%	—	0.94	External (*n*=100)
OxiNet (2023)	OSA Diagnosis	99.8%*	—	—	*n*=12,923, 6 databases
NAVOY Sepsis	Sepsis Prediction	—	—	>qSOFA	Prospective validation
Nemati *et al*.	Sepsis (12h prior)	—	—	0.83	ICU cohort
Xiong meta-analysis	ARDS Prediction	0.81	0.82	0.88	33 studies pooled
Apnea Interact Xplainer	OSA (4-level)	97%**	—	0.74-0.81	*n*=15,807, 7 cohorts
Topalovic *et al*. (2019)	PFT Interpretation	—	—	100%† / 82%‡	*n*=6,000 interpretations
Saad *et al*. (2025)	PFT Interpretation	—	—	86.6%‡	*n*=440 prospective
Das *et al*. (2023)	PFT + XAI	—	—	+10.4%‡	*n*=78 pulmonologists
Gompelmann (2025)	PFT for ILD	—	—	86%‡	*n*=60, 25 physicians
Random Forest (meta)	Sepsis ICU	—	—	0.91	Multiple studies
XGBoost (meta)	Sepsis ICU	—	—	0.96	Multiple studies

**Moderate-to-severe OSA detection. **Early sleep apnea detection using oximetry only. †Pattern interpretation accuracy. ‡Diagnostic accuracy. Abbreviations: AUC, area under the curve; ICU, intensive care unit; ILD, interstitial lung disease; OSA, obstructive sleep apnea; PFT, pulmonary function test; qSOFA, quick Sequential Organ Failure Assessment.*

https://doi.org/10.63501/pj8sbm22

AI in Critical Care and Sepsis Prediction:

Sepsis remains one of the leading causes of mortality worldwide, characterized by a complex and heterogeneous clinical presentation. Despite advances in patient monitoring and biomarkers, early detection in the ICU is often hampered by incomplete data and diagnostic uncertainty. AI has emerged as a powerful tool for early sepsis detection, with various algorithms demonstrating the ability to predict sepsis hours before clinical manifestation (Papaioannou *et al*., 2025).

The NAVOY sepsis algorithm, tested by Persson *et al*. (2024), demonstrated the ability to detect patients at high risk of developing sepsis within 3 hours and achieved superior predictive performance compared with existing early warning scoring systems, including SOFA, qSOFA, MEWS, and NEWS2. Nemati *et al*. developed an AI model that predicted sepsis 12 hours before onset with an AUC of 0.83, representing a significant improvement over traditional methods (Table 2).

A systematic review and meta-analysis examining the prediction of early-onset sepsis in the ICU found that random forest (RF) and extreme gradient boosting (XGB) were the most frequently used algorithms, achieving high accuracy (0.911 for RF, 0.957 for XGB) as shown in Table 2. However, while ML algorithms capable of detecting evolving sepsis earlier than rule-based methods have proliferated, few have been implemented in clinical practice and have not demonstrated improved outcomes (Table 1).

AI in Mechanical Ventilation Management:

Optimizing mechanical ventilation is a complex, high-stakes intervention requiring precise and continuous adjustments. This task is complicated by the heterogeneity of patient responses, due to variability in underlying causes, lung mechanics, and individual physiological characteristics. AI applications in mechanical ventilation primarily focus on predicting outcomes, including the need for intubation, respiratory complications, and readiness for weaning (Misseri *et al*., 2024).

Stivi *et al*. (2024) used AI to predict the success of weaning from mechanical ventilation in patients with respiratory failure, including those with ARDS, demonstrating the potential for personalized ventilator management. Liu et al. developed a two-stage AI system to predict optimal weaning timing, whereas reinforcement learning approaches using Markov Decision Processes have demonstrated performance superior to standard physician clinical care in patients receiving mechanical ventilation.

The Feasible Artificial Intelligence with Simple Trajectories for Predicting Adverse Catastrophic Events (FAST-PACE) model can predict the onset of cardiac arrest or acute respiratory failure from 1 to 6 hours before occurrence, providing critical lead time for intervention. Additionally, AI algorithms have demonstrated 99.4% accuracy in detecting resistance-compliance levels in ventilated patients, enabling more responsive ventilator adjustments (Table 1).

AI in Acute Respiratory Distress Syndrome:

ARDS is a critical challenge in intensive care, marked by high mortality (35-46%) and significant patient heterogeneity, which limits the effectiveness of conventional supportive therapies. A meta-analysis by Xiong *et al*. (2025) evaluating the accuracy of AI in predicting ARDS found pooled sensitivity of 0.81 (95% CI: 0.76-0.85) and pooled specificity of 0.82 (95% CI: 0.77-0.86), with an AUC of 0.88 (95% CI: 0.85-0.91) (Table 2; Figure 2).

Li *et al*. (2025) provided a comprehensive review of AI/ML applications in ARDS management, highlighting diverse applications including early prediction and diagnosis using multimodal data (EHR, imaging, ventilator waveforms), advanced prognostic assessment that outperforms traditional scoring systems, and precise identification of ARDS subtypes to guide personalized treatment. AI models have successfully distinguished between hyperinflammatory and hypoinflammatory ARDS phenotypes, which have different mortality rates and treatment responses.

ML has also enabled the identification of "recruitable" versus "non-recruitable" ARDS subtypes by integrating quantitative CT imaging with respiratory mechanics, thereby enabling personalized PEEP settings and responses to recruitment maneuvers. Furthermore, AI-driven analysis of ventilator waveforms has improved detection of patient-ventilator asynchrony, a common complication associated with worse outcomes.

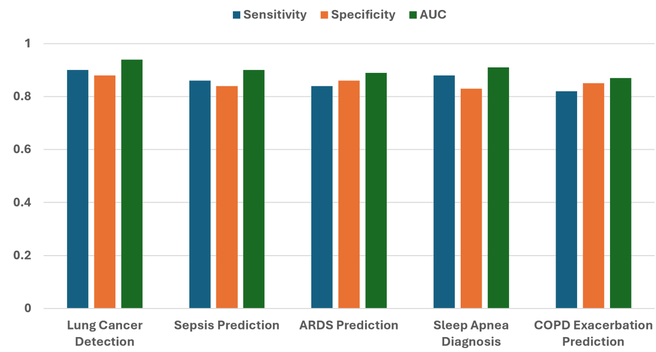
AI in Sleep Medicine:

Obstructive sleep apnea (OSA) affects approximately 1 billion adults globally and remains significantly underdiagnosed despite its association with cardiovascular diseases, metabolic disorders, and neurocognitive impairments. Traditional diagnosis through polysomnography (PSG) is resource-intensive and has limited accessibility. AI has emerged as a transformative tool for the diagnosis and management of OSA (Suliman *et al.*, 2025).

Nasifoglu *et al.* (2023) developed OxiNet, a deep learning model trained on 12,923 polysomnography recordings from six independent databases, achieving strong performance in OSA diagnosis from oximetry signals alone (Table 2). The system missed only 0.2% of all moderate-to-severe OSA patients compared to 21% for the best benchmark method. This represents a significant advancement in accessible, cost-effective screening.

Zhang *et al.* (2025) introduced the Apnea Interact Xplainer, a transparent AI system analyzing 15,807 polysomnography recordings from seven multi-ethnic cohorts, achieving accuracies of 0.738-0.810 for four-level severity classification with 99.8% accuracy within one severity grade (Table 2). Notably, the system achieved 97% sensitivity for early detection of sleep apnea using only oximetry signals, enabling practical home-based screening. ML techniques have achieved classification accuracy of 88-91% for predicting OSA severity using gradient boosting models (Kim *et al.*, 2023), as illustrated in Figure 2.

Figure 2: Comparative Performance of AI Models Across Pulmonary and Critical Care Applications.



Clinical Threshold (0.8)
Abbreviations: AUC: Area Under the Curve; ARDS: Acute Respiratory Distress Syndrome; COPD: Chronic Obstructive Pulmonary Disease.

AI in Chronic Obstructive Pulmonary Disease (COPD):

COPD represents a significant global health challenge, with the Global Initiative for Chronic Obstructive Lung Disease (GOLD) estimating substantial worldwide prevalence. Chen *et al.* (2025) conducted a systematic review of digital health technologies and AI algorithms in COPD, identifying applications across disease prediction, diagnostic accuracy, and patient management (Table 1).

AI applications in COPD include lung-sound classification using artificial neural networks with high accuracy to distinguish healthy from unhealthy patterns (Zhang *et al.*, 2025), voice-based COPD classification, and cough-sound analysis for early detection. Deep learning-based kernel adaptation has enhanced the quantification of emphysema on low-dose CT to predict long-term mortality (Park *et al.*, 2024).

ML models that integrate clinical, radiological, and spirometric data have shown promise in predicting COPD exacerbations, thereby enabling proactive intervention. Wearable technologies, including smartwatches and portable spirometers, can now capture real-time respiratory data, with AI

https://doi.org/10.63501/pj8sbm22

algorithms analyzing respiratory rate, oxygen saturation, and cough frequency to provide early warnings of exacerbations.

AI in Interstitial Lung Disease:

Interstitial lung diseases (ILDs), particularly idiopathic pulmonary fibrosis (IPF), present diagnostic challenges due to overlapping clinical, radiological, and pathological features. Wang *et al*. (2023) developed an AI system using RadImageNet-pretrained models to diagnose five types of ILD using multimodal data and predict 3-year survival rates.

The FDA granted De Novo clearance to Fibresolve in 2024, the first AI-enabled software specifically authorized for adjunctive diagnostic use in lung fibrosis (Uribe *et al*., 2025; Table 3).

Table 3: FDA-Authorized AI/ML Medical Devices in Pulmonary Medicine (2023-2025)

Device/Software	Manufacturer	Indication	Authorization Type/Date
Fibresolve	Imvaria	ILD/IPF diagnosis from CT	De Novo, 2024
Aidoc PE	Aidoc Medical	Pulmonary embolism detection on CTPA	510(k), 2018-ongoing
Viz.ai PE	Viz.ai	PE triage and notification	510(k), 2021
Riverain ClearRead CT	Riverain Technologies	Lung nodule detection	510(k), 2019
Optellum LCP	Optellum	Lung cancer prediction from CT	510(k), 2021
Qure.ai qXR	Qure.ai	Chest X-ray abnormality detection	510(k), 2020
Lunit INSIGHT CXR	Lunit	Chest radiograph AI analysis	510(k), 2019
Zebra Medical AI	Zebra Medical Vision	Multiple pulmonary findings	510(k), 2018-2020

Abbreviations: CT, computed tomography; CTPA, computed tomography pulmonary angiography; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis; PE, pulmonary embolism.

The system uses ML-based pattern recognition to assess CT scans for IPF diagnosis, identifying patterns associated with IPF that are differentiated from those of non-IPF ILDs. This represents a significant milestone in AI-assisted pulmonary diagnosis.

Quantitative CT analysis tools, including CALIPER (Computer-Aided Lung Informatics for Pathology Evaluation and Rating), have enabled the identification of novel imaging biomarkers. Pulmonary vessel-related structures, expressed as a percentage of total lung volume, emerged as the most robust predictor of mortality among all HRCT pattern scores (de la Orden Kett Morais *et al*., 2024).

AI in Asthma Management and Smart Inhalers:

Asthma affects over 262 million people globally, with poor outcomes frequently attributed to non-adherence to medication, improper inhaler technique, and inadequate engagement with healthcare services. Digital health technologies, including smart inhalers and AI-powered applications, have emerged as solutions to these challenges (Gaillard *et al*., 2024).

The smart inhalers market is projected to grow at a CAGR of 21.26% between 2025 and 2032, driven by the integration of Bluetooth-enabled sensors, AI analytics, and mobile applications. Smart inhalers track medication usage, assess inhalation technique, and provide personalized reminders. The InspirerMundi platform uses ML to monitor inhaler use and assess medication self-administration remotely (Chan *et al*., 2021).

AI-powered exacerbation prediction has shown particular promise. Lugogo *et al*. (2024) used data from electronic multi-dose dry powder inhalers with integrated sensors combined with clinical information to construct ML models using gradient-boosting trees that predicted impending exacerbations (Gogali *et al*., 2025). A randomized trial comparing personalized inhaler education delivered via smart spacers with usual care showed a 26.2% reduction in inhalation errors, whereas usual care showed a 14.6% increase (Table 1).

AI in Pulmonary Embolism Detection:

Pulmonary embolism (PE) remains a critical condition requiring timely detection. AI integration with CT pulmonary angiography (CTPA) has significantly advanced PE detection, enhancing diagnostic accuracy and efficiency. The FDA approval of the Aidoc AI model represents a significant advancement, demonstrating high specificity and negative predictive value in PE diagnosis (Naser *et al*., 2025; Table 3).

Li *et al*. (2025) reviewed the growing role of AI in PE diagnosis using CTPA, examining the capabilities of deep learning models, particularly DCNNs, for enhanced image-based detection, and the use of natural language processing to improve risk stratification using electronic health records. These tools have demonstrated the potential to reduce diagnostic delays in this time-sensitive condition. Additional FDA-authorized devices for pulmonary applications are listed in Table 3.

AI in Pulmonary Function Test Interpretation:

Pulmonary function tests (PFTs) are essential diagnostic tools for evaluating respiratory diseases, including spirometry, body plethysmography, and diffusing capacity measurements. However, interpreting PFTs remains challenging due to interrater variability, variation in guideline strategies, and the complexity of integrating multiple parameters with clinical context. AI has emerged as a promising solution to enhance the accuracy and consistency of PFT interpretation (Saad *et al*., 2025).

A landmark study by Topalovic *et al*. (2019) demonstrated that AI-based software outperformed pulmonologists in interpreting PFTs. In this multicenter study involving 120 pulmonologists from 16 European hospitals who evaluated 50 cases each (6,000 total interpretations), pulmonologists matched ATS/ERS guidelines for pattern interpretation in only 74.4% of cases with moderate inter-rater agreement (κ=0.67). More strikingly, pulmonologists achieved correct diagnoses in only 44.6% of cases with substantial variability (κ=0.35). In contrast, the AI software achieved 100% accuracy for PFT pattern interpretation and 82% diagnostic accuracy (Table 2; Topalovic *et al*., 2019)

Saad *et al*. (2025) conducted a retrospective-prospective study using 3,230 retrospective cases for algorithm training and 440 prospective cases for evaluation. The AI achieved 86.59% accuracy, compared with 65.82% for pulmonologists; Fleiss's kappa for inter-rater agreement among eight pulmonologists was only 0.46 (moderate agreement). Various algorithms, including Decision Trees, Support Vector Machines, K-Nearest Neighbors, Naïve Bayes, and Neural Networks, were evaluated, with K-Nearest Neighbors (K=7) and Decision Trees (maximum depth=4) achieving optimal performance.

The collaboration between explainable AI (XAI) and pulmonologists has demonstrated synergistic benefits. Das *et al*. (2023) conducted a multicenter intervention study showing that when pulmonologists used XAI-supported interpretation, preferential diagnostic accuracy increased by 10.4% in phase 1 (n=16 pulmonologists) and 5.4% in phase 2 (n=62 pulmonologists), with both improvements being highly significant ($p<0.001$ and $p<0.0001$, respectively). The XAI system provided Shapley values to explain AI predictions, increasing both diagnostic confidence and inter-rater agreement (Das *et al*., 2023).

AI has also demonstrated particular value in improving the early diagnosis of interstitial lung disease by interpreting PFTs. Gompelmann *et al*. (2025) demonstrated that AI-based decision support significantly improved ILD diagnosis in a multicenter study involving 25 physicians who evaluated 60 cases. The AI software (ArtiQ.PFT) matched clinical diagnoses in 86% of cases. In two cases with a mismatch, the AI-suggested diagnosis of ILD was subsequently confirmed by the chest clinic. Interpretation time decreased from 10.6±4.1 minutes to 5.6±5.6 minutes with AI support ($p<0.001$) (Gompelmann *et al*., 2025).

Beyond interpretation, AI has been applied to estimate lung volumes from spirometry alone and even from chest radiographs. Helgeson *et al*. (2025) developed ML algorithms using 121,498 PFTs from Mayo Clinic to estimate static lung volumes (TLC, RV, FRC) from spirometry measurements, potentially enabling assessment of restriction, hyperinflation, and air-trapping without body plethysmography (Helgeson *et al*., 2025). Ueda *et al*. (2024) developed a deep learning model trained on over 140,000 chest radiographs to estimate FVC and FEV1, with Pearson's correlation coefficient exceeding 0.90, offering an alternative for patients unable to perform spirometry (Ueda *et al*., 2024).

AI-based spirometry quality control represents another critical application. A systematic review identified six studies that applied AI to spirometry quality assurance, including the assessment of maneuver acceptability, usability, and error detection during test performance. AI software has demonstrated the ability to evaluate more than 90% of spirometry sessions as meeting quality standards and to facilitate quality control in clinical trials for asthma and COPD (Topole *et al*., 2023). In primary care settings, AI-assisted spirometry interpretation achieved 86% agreement with clinical diagnoses while reducing interpretation time by nearly half (Table 1).

DISCUSSION:

Clinical Implications and Translational Challenges:

The evidence presented in this review demonstrates the substantial potential of AI technologies across the spectrum of pulmonary medicine and critical care (Table 1; Figure 2). However, several crucial factors must be addressed for successful clinical translation. Model transparency remains a significant concern, as many ML algorithms operate as "black boxes," making it difficult for clinicians to understand the algorithms' reasoning (Papaioannou *et al*., 2025). This opacity is particularly problematic in the uncertainty-laden clinical environment of critical care, where treatment decisions may have life-or-death consequences.

The success of explainable AI in PFT interpretation demonstrates a promising model for human-AI collaboration. When pulmonologists were provided with AI predictions along with Shapley value explanations, diagnostic accuracy improved significantly while maintaining clinical autonomy. This approach, in which AI serves as a decision-support tool rather than a replacement for clinical judgment, may represent the optimal integration strategy for many pulmonary applications.

Workflow integration presents another substantial barrier. AI tools must seamlessly integrate into existing clinical workflows without adding cognitive burden or disrupting established care processes. The real-time delivery of predictions to bedside clinicians requires robust IT infrastructure and careful consideration of alarm fatigue. As Nikolaidis *et al*. (2025) noted, although AI tools have shown promise in many areas of medicine, their use in ICU settings remains limited and inconsistent, constrained by concerns around transparency, workflow integration, and clinician trust.

Validation and Generalizability:

A critical limitation across AI applications is the lack of external validation and demonstration of generalizability. Many published studies utilize single-center retrospective data, limiting their applicability to broader patient populations (Table 2). The Medical Information Mart for Intensive Care (MIMIC) database is frequently used, but models developed on this dataset may not perform equivalently across diverse healthcare settings and populations. Multi-center prospective randomized controlled trials remain notably scarce in the field of ICU AI applications.

Furthermore, AI models must be trained on large, diverse datasets to ensure adequate performance across diverse clinical contexts, ethnicities, and populations. The heterogeneity of conditions such as ARDS and sepsis creates additional challenges, as

https://doi.org/10.63501/pj8sbm22

models may perform differently across disease subtypes. Federated learning approaches offer potential solutions by enabling model training across institutions while preserving data privacy.

The Imperative for Clinician Oversight:

Despite the impressive performance metrics achieved by AI systems across pulmonary and critical care applications, the necessity of maintaining a qualified clinician as the final arbitrator in all diagnostic and therapeutic decisions cannot be overstated. This "clinician-in-the-loop" paradigm is not merely a regulatory requirement but a fundamental ethical and clinical imperative. AI algorithms, regardless of their sophistication, operate within the constraints of their training data and cannot fully account for the nuanced clinical context, patient preferences, comorbidities, and psychosocial factors that experienced clinicians integrate into their decision-making.

Several factors underscore the critical importance of human oversight. First, AI models are susceptible to dataset shift and may perform unpredictably when encountering patient populations or clinical scenarios underrepresented in their training data. A qualified physician can recognize when AI recommendations may be unreliable due to atypical presentations or rare conditions. Second, the "black box" nature of many deep learning algorithms means that even high-performing models may occasionally produce erroneous outputs for reasons that are not immediately apparent. Human clinicians serve as an essential safety net for identifying and correcting such errors before they affect patient care.

Third, medical decision-making frequently involves value judgments and trade-offs that extend beyond pattern recognition. Decisions regarding goals of care, risk tolerance, and treatment preferences require human empathy, communication skills, and ethical reasoning that AI systems cannot replicate. The physician's role in shared decision-making with patients and families remains irreplaceable. Fourth, accountability and liability for medical decisions must ultimately rest with licensed healthcare professionals who can be held to professional standards and who bear legal and ethical responsibility for patient outcomes.

The optimal model for AI deployment in pulmonary medicine and critical care (in fact, in all clinical settings) is therefore one of augmented intelligence rather than autonomous AI. In this framework, AI enhances human capabilities by processing large volumes of data, identifying patterns, flagging abnormalities, and providing decision support. At the same time, the qualified clinician retains authority over final diagnostic and therapeutic decisions. Regulatory bodies, including the FDA and professional societies, have increasingly emphasized this requirement, and healthcare institutions that implement AI tools must ensure that workflows preserve meaningful physician oversight rather than encouraging complacency with automation.

Regulatory and Ethical Considerations:

The regulatory landscape for AI in healthcare is evolving rapidly. The FDA has authorized more than 950 AI-enabled medical devices (Table 3), and the Predetermined Change Control Plans (PCCP) pathway, finalized in December 2024, provides a framework for adaptive AI systems. In Europe, the AI Act (Regulation EU 2024/1689), effective August 2024, classifies medical AI as high-risk and requires dual compliance with both the AI Act and the MDR/IVDR requirements by August 2027.

Ethical considerations include data privacy, algorithmic bias, and the potential for AI to exacerbate healthcare disparities if not carefully implemented. Sensitive health data use necessitates ethical guidelines and reflective, representative datasets. Additionally, questions remain about liability when AI-assisted decisions lead to adverse outcomes, and the appropriate balance between AI recommendations and clinical judgment.

Future Directions:

Several emerging paradigms show promise for advancing AI in pulmonary and critical care medicine. Large Language Models (LLMs) and

https://doi.org/10.63501/pj8sbm22

agent-based AI systems may enable enhanced data integration and autonomous decision support. Graph Neural Networks and causal inference methods offer potential for understanding complex relationships in disease. Self-supervised learning could address the challenges posed by limited labeled medical data, while explainable AI (XAI) approaches may help overcome transparency barriers.

The integration of multimodal data, combining imaging, physiological signals, genomic information, and clinical notes, represents a promising direction (Figure 1). AI systems that can synthesize information across these diverse data types may better capture disease complexity and enable truly personalized medicine approaches.

CONCLUSIONS:

Artificial intelligence and machine learning technologies are poised to transform pulmonary medicine and critical care, offering improvements in diagnostic accuracy, predictive capability, and treatment personalization. The evidence demonstrates significant advances across multiple domains: lung cancer screening with sensitivity exceeding 90%; sepsis prediction hours before clinical onset; ARDS identification and phenotyping; sleep apnea diagnosis with near-perfect sensitivity; PFT interpretation with accuracy exceeding that of pulmonologists; and enhanced management of chronic respiratory diseases through digital health technologies (Tables 1-3; Figures 1-2).

However, successful clinical implementation requires addressing substantial challenges. Model transparency and interpretability must be improved to build clinician trust and enable appropriate clinical decision-making. Robust validation across diverse populations and healthcare settings is essential before widespread deployment. Regulatory frameworks continue to evolve to balance innovation with patient safety. Workflow integration strategies must minimize disruption while maximizing clinical benefit.

The future of AI in pulmonary and critical care medicine will likely involve increasingly sophisticated multimodal systems that integrate diverse data types, personalized treatment algorithms, and seamless clinical workflows. Multi-center prospective trials are urgently needed to demonstrate not only algorithmic performance but meaningful improvements in patient outcomes. With careful attention to implementation challenges, AI technologies offer the potential to significantly enhance respiratory care and improve outcomes for millions of patients worldwide.

REFERENCES

1. Chan, A. H. Y., Pleasants, R. A., Dhand, R., Guo, Y. X., Li, T. H., & Hew, M. (2021). Digital inhalers for asthma or chronic obstructive pulmonary disease: A scientific perspective. *Pulmonary Pharmacology & Therapeutics, 71*, 102104. https://doi.org/10.1016/j.pupt.2021.102104
2. Chen, Z., Hao, J., Sun, H., Li, M., Zhang, Y., & Qian, Q. (2025). Applications of digital health technologies and artificial intelligence algorithms in COPD: Systematic review. *BMC Medical Informatics and Decision Making, 25*(1), 77. https://doi.org/10.1186/s12911-025-02870-7
3. Das, N., Happaerts, S., Gyselinck, I., et al. (2023). Collaboration between explainable artificial intelligence and pulmonologists improves the accuracy of pulmonary function test interpretation. *European Respiratory Journal, 61*(5), 2201720. https://doi.org/10.1183/13993003.01720-2022
4. de la Orden Kett Morais, S. R., Felder, F. N., & Walsh, S. L. F. (2024). From pixels to prognosis: Unlocking the potential of deep learning in fibrotic lung disease imaging analysis. *British Journal of Radiology, 97*(1161), 1517-1525. https://doi.org/10.1093/bjr/tqae108
5. Gaillard, E. A., et al. (2024). Current technological advancement in asthma care. *Expert Review of Respiratory Medicine, 18*(7), 499-512. https://doi.org/10.1080/17476348.2024.2380067
6. Gogali, A., & Kostikas, K. (2025). Artificial intelligence in asthma. *Journal of Allergy and*

Clinical Immunology: In Practice, 13(9), 2397-2404. https://doi.org/10.1016/j.jaip.2025.05.009
7. Gompelmann, D., Gysan, M. R., Desbordes, P., et al. (2025). AI-powered evaluation of lung function for diagnosis of interstitial lung disease. *Thorax, 80*(7), 445-450. https://doi.org/10.1136/thorax-2024-221537
8. Helgeson, S. A., Quicksall, Z. S., Johnson, P. W., Lim, K. G., Carter, R. E., & Lee, A. S. (2025). Estimation of static lung volumes and capacities from spirometry using machine learning: Algorithm development and validation. *JMIR AI, 4*, e65456. https://doi.org/10.2196/65456
9. Hendrix, W., et al. (2023). Deep learning for the detection of benign and malignant pulmonary nodules in non-screening chest CT scans. *Communications Medicine, 3*, 159. https://doi.org/10.1038/s43856-023-00388-5
10. Karthika, M., Sreedharan, S., Shevade, M., Mathew, J., & Ray, S. (2024). Artificial intelligence in respiratory care. *Frontiers in Digital Health, 6*, 1502434. https://doi.org/10.3389/fdgth.2024.1502434
11. Kashyap, M., Wang, X., Panjwani, N., et al. (2025). Automated deep learning-based detection and segmentation of lung tumors at CT. *Radiology, 314*(1), e233029. https://doi.org/10.1148/radiol.233029
12. Kim, S. Y., et al. (2023). Application of various machine learning techniques to predict obstructive sleep apnea syndrome severity. *Scientific Reports, 13*, 6379. https://doi.org/10.1038/s41598-023-33170-7
13. Li, L., Peng, M., Zou, Y., Li, Y., & Qiao, P. (2025). AI in CTPA pulmonary embolism detection. *Frontiers in Medicine, 12*, 1514931. https://doi.org/10.3389/fmed.2025.1514931
14. Li, R., Yue, J., Lu, H., et al. (2025). Artificial intelligence and machine learning in acute respiratory distress syndrome management: Recent advances. *Frontiers in Medicine, 12*, 1597556. https://doi.org/10.3389/fmed.2025.1597556
15. Misseri, G., Piattoli, M., Cuttone, G., Gregoretti, C., & Bignami, E. G. (2024). Artificial intelligence for mechanical ventilation: A transformative shift in critical care. *Therapeutic Advances in Pulmonary and Critical Care Medicine, 19*, 29768675241298918. https://doi.org/10.1177/29768675241298918
16. Naser, A. M., Vyas, R., Morgan, A. A., et al. (2025). Role of artificial intelligence in the diagnosis and management of pulmonary embolism: A comprehensive review. *Diagnostics, 15*(7), 889. https://doi.org/10.3390/diagnostics15070889
17. Nasifoglu, H., et al. (2023). Deep learning for obstructive sleep apnea diagnosis based on single-channel oximetry. *Nature Communications, 14*, 4881. https://doi.org/10.1038/s41467-023-40604-3
18. Nikolaidis, K., et al. (2025). Machine learning and artificial intelligence in intensive care medicine: Critical recalibrations from rule-based systems to frontier models. *Journal of Clinical Medicine, 14*(12), 4026. https://doi.org/10.3390/jcm14124026
19. Papaioannou, V., et al. (2025). AI advances in ICU with an emphasis on sepsis prediction: An overview. *Machine Learning and Knowledge Extraction,* 7(1), 6. https://doi.org/10.3390/make7010006
20. Park, H., Hwang, E. J., & Goo, J. M. (2024). Deep learning-based kernel adaptation enhances quantification of emphysema on low-dose chest CT for predicting long-term mortality. *Investigative Radiology, 59*(4), 278-286. https://doi.org/10.1097/RLI.0000000000001003
21. Persson, I., Macura, A., Becedas, D., & Sjövall, F. (2024). Early prediction of sepsis in intensive care patients using the machine learning algorithm NAVOY sepsis. *Journal of Critical Care, 80*, 154400. https://doi.org/10.1016/j.jcrc.2024.154400
22. Saad, T., Pandey, R., Padarya, S., Singh, P., Singh, S., & Mishra, N. (2025). Application of artificial intelligence in the interpretation of pulmonary function tests. *Cureus, 17*(4), e82056. https://doi.org/10.7759/cureus.82056
23. Salvatore, M., et al. (2025). New perspectives on lung cancer screening and artificial intelligence. *Cancers, 17*(5), 826. https://pmc.ncbi.nlm.nih.gov/articles/PMC11943706/
24. Stivi, T., Padawer, D., Dirini, N., Nachshon, A., Batzofin, B., & Ledot, S. (2024). Using artificial

intelligence to predict mechanical ventilation weaning success in patients with respiratory failure, including those with acute respiratory distress syndrome. *Journal of Clinical Medicine, 13*(5), 1505. https://doi.org/10.3390/jcm13051505

25. Suliman, M. A., et al. (2025). Artificial intelligence in obstructive sleep apnea: Transforming diagnosis and management. *Respiratory Medicine, 243*, 108100. https://doi.org/10.1016/j.rmed.2025.108100
26. Tan, J., et al. (2025). A systematic review of AI performance in lung cancer detection on CT thorax. *Healthcare, 13*(13), 1510. https://doi.org/10.3390/healthcare13131510
27. Topole, E., Biondaro, S., Montagna, I., et al. (2023). Artificial intelligence-based software facilitates spirometry quality control in asthma and COPD clinical trials. *ERJ Open Research, 9*, 00292-2022. https://doi.org/10.1183/23120541.00292-2022
28. Topalovic, M., Das, N., Burgel, P. R., et al. (2019). Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. *European Respiratory Journal, 53*(4), 1801660. https://doi.org/10.1183/13993003.01660-2018
29. Ueda, D., Matsumoto, T., Yamamoto, A., et al. (2024). A deep learning-based model to estimate pulmonary function from chest x-rays: Multi-institutional model development and validation study in Japan. *The Lancet Digital Health, 6*(8), e547-e555. https://doi.org/10.1016/S2589-7500(24)00113-4
30. Uribe, J., Kalra, J., Muelly, M., Reicher, J., & Kheir, F. (2025). Clinical experience with the first FDA-authorized artificial intelligence tool in interstitial lung disease and idiopathic pulmonary fibrosis [Abstract]. *American Journal of Respiratory and Critical Care Medicine, 211*, A1709. https://doi.org/10.1164/ajrccm.2025.211.Abstracts.A1709
31. Wang, X., et al. (2023). Interstitial lung disease diagnosis and prognosis using an AI system integrating longitudinal data. *Nature Communications, 14*, 2272. https://doi.org/10.1038/s41467-023-37720-5
32. Xiong, Y., Gao, Y., Qi, Y., et al. (2025). Accuracy of artificial intelligence algorithms in predicting acute respiratory distress syndrome: A systematic review and meta-analysis. *BMC Medical Informatics and Decision Making, 25*(1), 44. https://doi.org/10.1186/s12911-025-02869-0
33. Zhang, J., et al. (2025). Transparent artificial intelligence-enabled interpretable and interactive sleep apnea assessment across flexible monitoring scenarios. *Nature Communications, 16*, 8147. https://doi.org/10.1038/s41467-025-62864-x
34. Zhang, W., Li, X., Liu, Q., et al. (2025). High-accuracy lung sound classification for pulmonary disease diagnosis using an artificial neural network. *Frontiers in Bioengineering and Biotechnology, 13*, 1583416. https://doi.org/10.3389/fbioe.2025.1583416